UNIVERSITY OF NATIONAL AND WORLD ECONOMY

The Spirit Makes The Power

11TH INTERNATIONAL CONFERENCE ON APPLICATION OF INFORMATION AND COMMUNICATION TECHNOLOGY AND STATISTICS IN ECONOMY AND EDUCATION

ICAICTSEE – 2021

November 25 – 26th, 2021

University of National and World Economy

Sofia, Bulgaria

CONFERENCE PROCEEDINGS

UNDER THE AUSPICES OF:

ifip

the leading edge of information technology

PUBLICATIONS INDEXED BY:

EBSCO HOST

AND

ProQuest

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

University of National and World Economy (UNWE), Sofia, Bugaria & School of Electronics Engineering, Vellore Institue of Technology, Vellore, India

Invite You to

International Conference on Application of Information and Communication Technology and Statistics in Economy and Education

ICAICTSEE-2021

Virtual Conference on

RECENT TRENDS IN BIOMEDICAL TECHONOLOGY

25th to 26th November

MS Teams

Early bird registration

20th November, 2021

Registration Fee Rs 250/-

Important Dates:

- Submission of Abstracts/Full Paper - 17th November
- Submission of Final Camera Ready Papers - January 16th 2022
- ICAICTSEE-2021 Conference - November 25th - 26th, 2021

For more Details, scan below

IEEE Signal Processing Society

TAM The AI & ML Club

VIT MAS

Indexed in

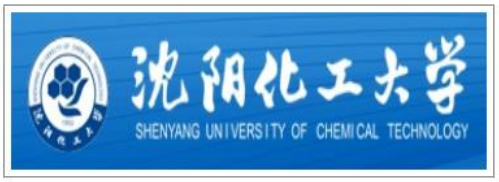
ProQuest

EBSCO HOST

For more queries contact: enquiry.icaictsee@vit.ac.in

COORGANIZING UNIVERSITIES AND INSTITUTIONS

Shimane University, Japan



Shenyang University of Chemical Technology, China

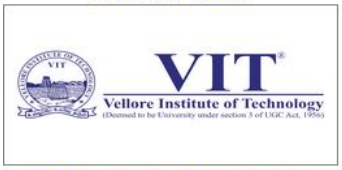
University of Genoa, Italy

Ukrainian Federation of Informatics, Ukraine

National Changhua University of Education, Taiwan

University Danubius of Galati, Romania

Kharkiv National University of Radioelectronics, Ukraine



Vellore Institue of Technology, India

Tomar Politechnic Institute, Portugal

Volga State University of Technology, Russian Federation

MB University, Serbia

Bohdan Khmelnytsky National University of Cherkasy, Ukraine

Kharkiv State University of Food Technology and Trade, Ukraine

Korkyt Ata Kyzylorda University, Kazakhstan

Bashkir State University, Russian Federation

Kyiv National University of Culture and Arts, Ukraine

Southern Federal University, Russian Federation

Gomel State University named after Francisk Skorina, Republic of Belarus

Artifex University of Bucharest, Romania

Academia de Studii Economice din Moldova

COOPERATED BY

CREATIVITY AND HARMONY SOCIETY

PROCEEDINGS OF THE

11TH INTERNATIONAL CONFERENCE ON APPLICATION OF INFORMATION AND COMMUNICATION TECHNOLOGY AND STATISTICS IN ECONOMY AND EDUCATION

ICAICTSEE-2021

EDITORS:

PROF. DR. DIMITER G. VELEV

DEPT. OF INFORMATION TECHNOLOGIES AND COMMUNICATIONS
UNIVERSITY OF NATIONAL AND WORLD ECONOMY
UNSS – STUDENTSKI GRAD, 1700 SOFIA, BULGARIA
DGVELEV@UNWE.BG

ASSOC. PROF. DR. PLAMENA ZLATEVA

INSTITUTE OF ROBOTICS
BULGARIAN ACADEMY OF SCIENCES, SOFIA, BULGARIA

PROF. DR. SIVAKUMAR R.

SCHOOL OF ELECTRONICS ENGINEERING
VELLORE IINSTITUTE OF TECHNOLOGY
VELLORE, TAMIL NADU, INDIA

ISSUED FOR PUBLICATION: SEPTEMBER 15TH, 2023

SIZE: JIS-B5

QUIRES: 31,8

PAGES: 509

VOLUME: 100

ISSN 2367-7635 (PRINT)

ISSN 2367-7643 (ONLINE)

PUBLISHING COMPLEX – UNWE, SOFIA, BULGARIA

CONFERENCE COMMITTEES

CONFERENCE CHAIR

Dimiter G. **Velev**	University of National and World Economy	***Bulgaria***

PROGRAM COMMITTEE

César Correa **Arias**	Universidad de Guadalajara	***Mexico***
Saji **Baby**	GEO Environmental Consultation	***Kuwait***
Gaetano **Cascini**	Politecnico di Milano	***Italy***
Olena **Chaikovska**	Kiev National University of Culture and Arts	***Ukraine***
Luisa **Chimenz**	University of Genoa	***Italy***
Sonia **Chipeva**	University of National and World Economy	***Bulgaria***
Sérgio Manuel Serra da **Cruz**	Universidade Federal Rural do Rio de Janeiro	***Brazil***
Tadeusz **Czachorski**	Institue of Theoretical and Applied Informatics, PAS	***Poland***
Linda **Daniela**	University of Latvia	***Latvia***
Vladimir **Dimitrov**	Sofia State University	***Bulgaria***
Marco **Ferrante**	University of Genoa	***Italy***
Marin **Galabov**	University of National and World Economy	***Bulgaria***
Carlos Arturo Torres **Gastelu**	Universidad Veracruzana	***Mexico***
Vasil **Georgiev**	Sofia State University	***Bulgaria***
Valentin **Goev**	University of National and World Economy	***Bulgaria***
Igor **Grebennik**	Kharkiv National University of Radio Electronics	***Ukraine***
Chih-Hsuan **Huang**	Hubei University of Economics	***China***
Aytekin **Isman**	Sakarya University	***Turkey***
Kalinka **Kaloyanova**	Sofia State University	***Bulgaria***
Madjid Malikovich **Karimov**	Tashkent University of Information Technologies	***Uzbekistan***
Sergii **Kavun**	Kharkiv University of Technology STEP	***Ukraine***
Hee Dong **Kim**	Hankuk University of Foreign Studies	***S.Korea***
Peter **Kokol**	University of Maribor	***Slovenia***
Özgen **Korkmaz**	Amasya University	***Turkey***
George **Kovacs**	Computer and Automation Institutre, NAS	***Hungary***
Oleksiy E. **Kovalenko**	Inst. of Mathematical Machines and Systems Problems, NASU	***Ukraine***
Patrick **Letouze**	Universidade Federal do Tocantins	***Brazil***
Violeta **Manevska**	St. Kliment Ohridski University - Bitola	***Macedonia***
Emanuel-Stefan **Marinescu**	Danubius University Galati	***Romania***
Liviu-Mihail **Marinescu**	Danubius University Galati	***Romania***
Célio Gonçalo Cardoso **Marques**	Polytechnic Institute of Tomar	***Portugal***
Pedja M. **Milosavljevic**	University of Nis	***Serbia***

Dear Colleagues,

*On behalf of the Conference Committees I have the honor and pleasure to thank you all for your true participation in the work of the **11th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE-2021)**, http://icaictsee.unwe.bg/, which took place on November 25-26th, 2021 at the University of National and World Economy (UNWE), Sofia, Bulgaria.*

*For a first time **ICAICTSEE** shared experience with the **Recent Trends in Biomedical Technology Conference**, co-organized by the School of Eelectronics Engineering, Vellore Institute of Technology, Vellore, India.*

Due to the COVID-19 pandemic, both conferences were conducted virtually.

I would like specially to express my gratitude to all of the program committee members for their genuine support without which it would never had happened.

Due to the current state of scientific development in all spheres of human activity, the constant knowledge and skill actualization of the academia and researchers in the field of Information and Communication Technology (ICT) is an obligatory necessity, especially when the world is in a global economic and financial crisis. The definition of long-term scientific research tasks in this area has a strategic importance, which is even truer for young researchers, lecturers and Ph.D. students.

The conference aims were conducting analysis of the current problems and presenting results of the ICT application in different areas of economy, education and related areas of scientific knowledge; outlining the existing possibilities for the application of modern ICT tools, methods, methodologies and information systems in economy and education; discussing advanced and emerging research trends with a long-term importance in the field of ICT application in economy and education.

The conference has established an academic environment that fosters the dialogue and exchange of ideas between different levels of academic and research communities.

The conference outcome is 61 published research papers (33 from foreign participants), the explosion of fresh ideas and the establishment of new professional relations.

Dimiter G. Velev
ICAICTSEE-2021 Chair

CONTENTS

Pulmonary Embolism Detection using Convoluted Neural Networks

Shubham Kumar, Tanuja Swathi Seeram, Sivakumar R

School of Electronics Engineering, Vellore Institute of Technology, Vellore , India.
shubham.kumar2019bec@vitstudent.ac.in, tanujaswathi.seeram2019@vitstudent.ac.in
rsivakumar@vit.ac.in

Abstract. The paper reviews CTPA (Computed Tomography Pulmonary Angiography) which is a comprehensive imaging system that has benefited numerous patients in diagnosing a life-threatening condition like as PE (Pulmonary embolisms). PEs are among the diagnoses that are frequently delayed, despite the fact that prompt identification and treatment are critical for minimizing the damage that PEs can cause. We describe a Deep-Learning model based on Convolutional Neural Networks (CNNs) that can detect and categories Pulmonary Embolisms from CTPA scans in this paper. The research makes use of a dataset offered by the Radiological Society of North America (RSNA), which is made up of scans taken from numerous hospitals around North America and made public as part of a Kaggle Competition on the same issue. Our model extracts features using Xception, a cutting-edge CNN, and then uses these features to identify all of the metrics connected with the studies using transfer learning. This model can classify most metrics with an accuracy of more than 90%, with the exception of heart-related measures, where it only achieves an accuracy of 80%.

Keywords. Convolutional Neural Networks, Computed Tomography Pulmonary Angiography, Pulmonary Embolisms, Kaggle, RSNA STR Pulmonary Embolism Detection, Image Preprocessing, Deep Learning.

1. Introduction

A pulmonary embolism (PE) is a blockage in a lung artery that occurs suddenly. It usually occurs when a blood clot breaks away and travels through the bloodstream to the lungs, and it can be fatal, especially if the clot is large or there are many of them. To make a diagnosis, the healthcare professional will need the patient's medical history, a physical examination, and blood and imaging tests. This is a time-consuming procedure that can lead to over-diagnosis. The most frequent form of medical imaging technology used to evaluate patients with suspected PE is CT Pulmonary Angiography (CTPA)[4]. Hundreds of pictures from CT scans must be reviewed in detail to find blockages in the pulmonary arteries. As a result, there is a gap in the market where Machine Learning might be used to detect PE symptoms based on clot identification on CT scans. This is what we hope to accomplish with this project.

The goal of this study is to create and test an end-to-end deep learning model capable of detecting PE using the RSNA's CTPA dataset. Over 7200 CTPA pictures and accompanying labels for training have been collected from real patients across North America and are included in the dataset. Our pipeline receives a pre-processed version of these photos that includes the required pixels in predetermined ranges thanks to a technique called windowing, which is discussed in more detail in section IV. The pipeline uses state-of-the-art convolutional neural networks to extract characteristics from this image, which are subsequently used for categorization across nine labels that correlate to each image on the image and study level. The

output is a binary classification for each label, which is then compared to the targets and losses and accuracies calculated.

2. Literature Survey

Machine learning (ML) is one of the most effective methods for integrating, analyzing, and forecasting big, heterogeneous data sets [15]. One of the most critical concerns in Artificial Intelligence systems is medical diagnosis via image processing and machine learning [1].

Early work on applying ML to the problem of PE diagnosis concentrated on using clinical factors or perfusion imaging as inputs to Neural Networks rather than CTPA, which resulted in limited generalization [18]. Other early approaches to automatic detection relied on extensive feature engineering and pre-processing without external validation, and were thus costly in more than one way [13]. By directly creating the learning architectures, Deep Learning eliminates or decreases the requirement for human feature engineering on images, allowing more freedom in designing the end-to-end pipeline for this automated process.

The use of CTPA imaging for PE diagnosis has become more popular in recent years. This is advantageous since CTPA imaging is the most routinely used imaging method for PE tests, and data collecting allows definite diagnosis to be made utilizing various learning methods. Furthermore, the developing field of applying Deep Learning techniques to medical applications allows for the steady improvement of a collaborative process between doctors and machines by delegating rudimentary diagnosis to automated DL pipelines and further inspection to clinicians. Even if the model isn't perfect, it will undoubtedly cut down on diagnosis time by allowing clinicians to better understand which paths they should investigate further. Instead of working with a full CTPA scan, the work in [16] substantially pre-processes aspects of pictures by segmentation and vessel-alignment, which is then provided as an input to a 3D CNN, whereas the work in [21] sub-divides each CTPA into small cubes. Our method involves substantially less feature engineering, as it simply requires windowing and performing rudimentary transformations to extract the key features from the pictures, and it works on whole CT scans rather than subdividing them. As a result, if our findings are improved, they will make a stronger argument for integrating with existing medical diagnostic pipelines.

3. Dataset and Preprocessing

We used the dataset given by the Radiological Society of North America (RSNA) [14] for this project, which is divided into three parts:

- A total of 1,790,594 training pictures with Labels are included in this training set, which includes of 7279 studies.
- There are 650 investigations in this test dataset, which amount to 146,853 unlabeled photos.
- A proprietary dataset containing 1517 research that Kaggle uses to assess submissions.

Multiple photos are sorted into series in a study. As a result, identifiers are assigned to each image:

- StudyInstanceUID: This is a one-of-a-kind identifier for the study that the image belongs to.
- SeriesInstanceUID: Identifies the image's place in the study's series.
- SOPInstanceUID: An image's unique identifier, defining the slice of a patient's data relevant to the chest (in the context of CT scans, there are several images taken at different locations and angles from the patient body called slices).

There are 14 labels for each training image, divided into three levels:

- Labels for each image at the image-level The sole image-level label is the one that indicates whether or not a PE is present on the image.
- Labels for Exam Levels: These labels are used to detect labels corresponding to PE characteristics, such as the right valve to left valve ratio (greater than or less than 1), the location of the PE (left, right, centre), the nature of the PE (acute, chronic, or both), and a label indicating whether the study was conclusive or not.
- Informational Labels: These labels are used to indicate whether the radiologists saw any motion or contrast issues in the study, if there were any artefacts, or if the diagnosis was something other than PE.

We used the information from [10] to narrow down the labels that needed to be predicted, eventually settling on 9:

- acute_and_chronic_pe,
- central_pe,
- chronic_pe,
- indeterminate,
- leftsided_pe,
- pe_present_on_image,
- rightsided_pe,
- rv_lv_ratio_gte_1,
- rv_lv_ratio_lt.

The photos are saved in the Digital Imaging and Communications in Medicine (DICOM) format [2], which is an international standard for digital medical image sharing, storage, and transmission. The study includes a series of 2D picture slices of various patients. These images are created by delivering Iodine to the patient and contrast to the chest, as discussed in [19].

As a result, the raw pixels contain information from several instances of the exam phases and through various tissues. Tissue density is measured in Hounsfield Units (HU) [9], using a scale that ranges from -1000 HU (for air) to 0 HU (for water). On disc, CT scan images may be represented differently than in memory. This is primarily due to the fact that DICOM pictures may contain negative values, although they are often stored as an unsigned integer. We may convert the values to HU units by using the intercept and slope values from the DICOM data format (which are provided by the device manufacturer). To extract the crucial information from the images, we windowed the CT numbers, which limited the pixels to a range of values, allowing improved contrast for the lungs.

Windowing is commonly done by defining a centre point and a length of the window for filtering the pixels, as described in [5] and [20].

We discovered that the majority of HU values were in the range of -1000 and 0, indicating that most of the region of an image is filled with air and water, in our exploratory data analysis using [7] as a reference. To segregate the lung tissue and provide a good contrast for the PE to be visible, we performed windowing three times on each slice with three different pairs of (centre, width): (40, 400), (100, 700), and (-600, 1500) in accordance with the architecture described in IV. As illustrated in 1, these were stacked to form a 3-channel Image that could be used by our feature extractor.

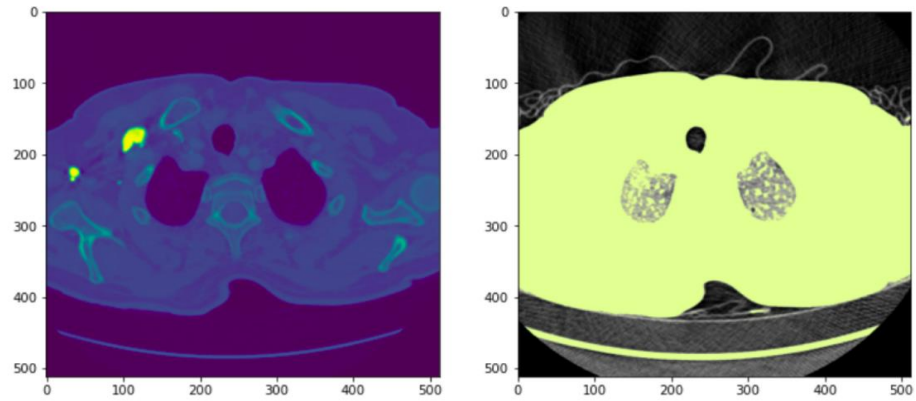


Fig. 1. A comparison of the original DICOM image (left) and the windowed 3-channel image (right)

We applied further augmentations on the photos after windowing to boost the performance of our workflow. To better detect the PE, we applied random 8-degree rotations in both directions and adjusted the contrast of the image. We used a random horizontal flip to images (with a 50% chance of flipping the slice image) to enable for better detection of the PE location and prediction of ventricular ratios. Section V delves deeper into the enhancements.

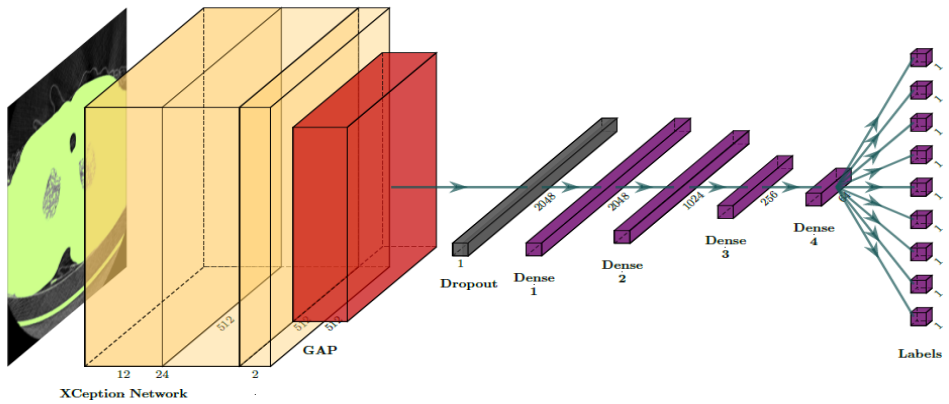


Fig. 2. For categorization, the image is sent into the Xception network, which extracts latent features, which are then averaged using a Global Average Pooling Layer. Following that, a 25% dropout is performed, followed by four back-to-back fully connected lay.

4. Methodology

The architecture is made up of two parts: (1) a convolutional base for extracting the image's latent features, and (2) a classifier for predicting the image's labels. We used the Xception network [3], a state-of-the-art network developed by Google, for the convolutional base and eliminated its top fully connected layer. Xception's feature extraction foundation is made up of 36 convolutional layers grouped into 14 modules, all of which have linear residual components save the first and last. The data is routed through three different processes: The entry flow is made up of four convolutional layers with batch-normalization and residual

connections, the middle flow is made up of three separable convolutions that are repeated eight times, and the exit flow is made up of two normalised convolutional layers. The Xception feature extractor produces a 16x16x2048 vector of latent features as its output. A Global Average Pooling (GAP) Layer is used to relay this information. This layer reduces the dimensionality of the data by averaging the 16x16 values, resulting in a 2048-length feature vector. This feature vector is then passed through a dropout layer, which sets 25% of the 2048 values to zero at random. The advantage of this dropout layer is that it forces the classifier to learn more robust weights that are not impacted by losing random Neurons, which helps stabilise the network and reduces over-fitting, as discussed in [8].

The feature vector is then passed through four dense, fully linked layers that extract additional particular features after the dropout. The first dense layer learns the right parameters to reduce the feature vector to half its original size, i.e. from 2048 to 1024. The feature vectors are reduced to 1/4 of their original size in the second and third dense layers, i.e. from 1024 to 256 and subsequently to 64. This 64-length vector is supplied into a nine-output multi-output configuration, each of which corresponds to one label. In contrast to the preceding layers, which were all activated by ReLu, each of these output layers is activated by a sigmoid function, and they aggregate the 64 features to output the probability of one label.

The Xception network is a Keras module that was pre-trained on the Imagenet Dataset [6]. During training, we made its layers untrainable. For the other layers, we utilized a test-train validation split of 0.8 and a learning rate of 0.001. The Adam optimizer [12] was employed in the model, which creates an exponential moving average of the gradient and the squared gradient, then decays them at rates that are set internally. The batch size for each training was 8, and we utilized a binary cross-entropy loss function to ensure the gradients decay proportionate to mistakes because the output was split into nine probability values between 0 and 1.

Figure 2 depicts the architecture. We utilized convolutional blocks to depict the three regions of the network, scaled according to the duration of the flows and the input size, because we couldn't entirely accommodate the Xception network.

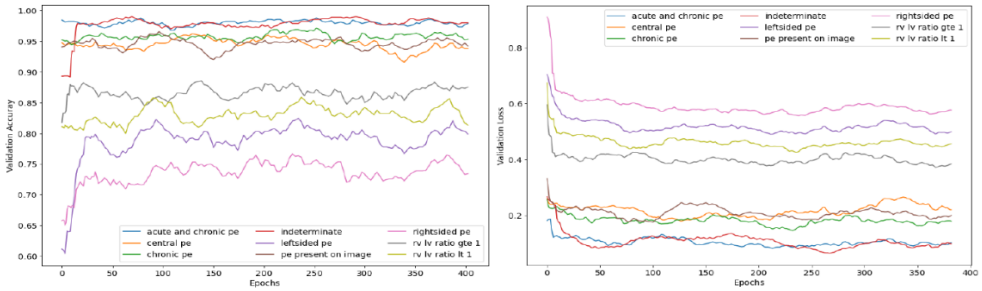


Fig. 3. Validation Accuracy for all metrics (Left). Validation Losses for all metrics (Right).

5. Experiments and Results

Because the data for this competition was roughly 1 TB, and our systems were not capable of handling or training it, our experimental pipeline was severely confined to the metrics and time constraints of kernels on Kaggle. As a result, our model was trained on Kaggle GPUs for a maximum of 4 or 5 hours per week (40 hours per week), giving us the ability to train at most 400 epochs per run. As measurement measures, we used the loss over-training and validation, as well as the classification accuracy of our model. We kept track of the accuracy and saved checkpoints to see if the model could improve its accuracy as the

training progressed. For all labels, accuracy reached a saturation point near the conclusion of training on the validation set in all experiments. We decided to utilise Binary Cross-Entropy loss while training our model because the problem was split into 9 binary classification problems, and we assessed the accuracy between the predicted labels and the validation set labels.

Our baseline was the vanilla architecture described in section IV, which performed well on the validation set, with a loss of 1.85199 and an average accuracy of roughly 89.6 percent, equally weighted for all measures. We attempted a few different additions to see if we could beat the baseline. The first change was to increase the dense layers, but this resulted in lower performance on the validation set and a tendency to overfit. As a feature extractor, we switched from Xception to Efficentnets [17], but this resulted in noisy outputs. We attempted applying numerous augmentations to the dataset with a progressive learning rate to make it work with efficient-nets, but we were still unable to enhance the validation accuracy. Then we tried augmentations like Mixup [23] and Cutmix [22], but they only served to lower our training metrics. As a result, we chose to keep with our current architecture and continue our trials with it. We opted to work with a single validation split of 80 percent per batch of training and test out hyper-parameter tweaking on this pipeline because the amount of training data was not an issue. We initially investigated the effect of learning rate: we reduced learning rate to 0.0001 and 0.0005, which reduced accuracies across all metrics by up to 5% while raising the loss to around 2.33 on the validation set. The generalisation performance on the validation set was reduced by 1.5 percent when the learning rate was increased to 0.005. As a result, we continued to learn at a pace of 0.001. The next step was to include random cropping in the prepossessing phase, resulting in a 320x320 cropped image from the 512x512 originals. However, the outputs became more noisy as a result of this. On this pipeline, we also tried cut-mix and mixup, but the results were sub-optimal. Rotation and windowing, which we attempted with a number of angles and windows before deciding on a combined random rotation of 8 degrees and 3-channel window stated in III, enhanced individual performances. We added the random horizontal flip augmentation to the image to boost the performance on the location of PE and the ventricle ratios, and it was this modification that finally allowed us to beat the baseline.

Figure 4 depicts the loss over the course of the training as well as during validation. Finally, the total loss on the validation set was 1.79224, and all parameters except chronic PE lost less than the baseline when compared to the baseline. This model's unweighted mean accuracy was 92 percent, compared to 89 percent for the baseline, while individual accuracy improved on all measures except chronic PE, where this model predicted a validation accuracy of 98 percent compared to 99 percent for the baseline. The validation losses and accuracies for each metric across the whole training period are shown in Figure 3. A running mean of 20 epochs was used to smooth these graphs. Throughout all of our runs, one consistent finding was that the model consistently underperformed on the ventricle ratios. We were unable to increase the performance of this statistic due to the time constraints.

Kaggle uses the private dataset to assess the generalisation performance. The normal technique is to train the model and then use it to build a submissions file using the predictions set. This submissions file is included with the kernel that will be run by Kaggle on the private dataset to determine scores. On the study and image-level labels, Kaggle employs a weighted log loss score for this competition. The loss L_{ij} for study i and label j is determined by multiplying the target of the label y_{ij} by the log of the probability predicted by the model for that label, as shown in Equation 1. The weight w_{ij} of the label is used to calculate the loss. In [11], the weights are provided.

$$L_{ij} = -w_j \times \left[y_{ij} \times \log(p_{ij}) + (1 - y_{ij}) \times \log(1 - p_{ij})\right] \quad (1)$$

The loss calculation for image-level labels is shown in Equation 2, where w is 0.07361963 and qi is the proportion of positive to negative pe images in the dataset. The exam is denoted by the letter I and the letter k is the picture identification for the exam.

$$L_{ik} = -[w \times p_i] \times [y_{ik} \times \log(p_{ik}) + (1 - y_{ik}) \times \log(1 - p_{ik})] \quad (2)$$

One of the restrictions for entering this competition was that notebooks must not utilise the internet, and because our notebook downloaded the Imagenet weights from the internet, we were unable to submit a complete entry. To produce an unofficial and private submission, we worked around it by first training our model independently, then retrieving the network file and re-uploading it as a dataset. We were able to see how well our model performed on the private dataset as a result of this. There are two types of scores that are calculated: The public score is derived using about 28% of the data, whereas the private score is calculated using 72% of the test dataset. The secret score of our model was 0.899, whereas the public score was 0.924. As a result, while our model's generalisation performance was not as good as the top entries, it wasn't far behind in terms of performance. We were unable to extract data during the runs because these tests are primarily executed on the Kaggle backend.

Fig. 4. Loss observed on Training and Validation Sets

6. Conclusion and Future Work

We described an end-to-end technique for detecting PE from chest CT scans in this paper. The model consisted of using the Xception CNN to extract latent features from the images, on which we applied a classifier to predict the 9 labels of the classification job, and we worked with data provided by the Radiological Society of North America. These labels were predicted by breaking down the multi-category classification problem into nine binary classification units, which were then used to assess the model's accuracy. On the validation test set, the model was able to achieve an accuracy of 92 percent. While this is not yet good enough for direct medical use, we provide an architecture that can be fine-tuned to improve categorization quality. There are two areas where further improvements can be made, as identified: To increase accuracy, the feature extractor can be ensembled with RNN units, which treat the collection of images in a study as a series, and the data can be further pre-processed to get a better balance between the labels.

References

1. Al-Ayyoub, Mahmoud, Ghaith Husari, Omar Darwish, and Ahmad Alabed-alaziz. "Machine learning approach for brain tumor detection." In Proceedings of the 3rd International Conference on Information and Communication Systems, pp. 1-4. 2012.

2. Bidgood Jr, W. D., S. C. Horii, F. W. Prior, and D. E. Van Syckle. "Understanding and using DICOM, the medical image communication standard." J Am Med Inform Assoc 4 (1997): 199Y212.
3. Chollet, Francois. "Deep learning with separable convolutions." arXiv Prepr. arXiv1610 2357 (2016).
4. CTPA. https : / /radiopaedia.org/ articles / ct - pulmonaryangiogram- technique.
5. Das, M. J., and L. B. Mahanta. "Lung Segmentation from CT Images: Impact of Different Window Settings on the Accuracy of Segmentation." Journal of Emerging Technologies and Innovative Research 5, no. 12 (2018): 189-193.
6. Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A large-scale hierarchical image database." In 2009 IEEE conference on computer vision and pattern recognition, pp. 248-255. Ieee, 2009.
7. EDA for Data. https : / /www. kaggle . com / nitindatta / pulmonary-embolism-dicom-preprocessing-eda.
8. Hinton, Geoffrey E., Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R. Salakhutdinov. "Improving neural networks by preventing co-adaptation of feature detectors." arXiv preprint arXiv:1207.0580 (2012).
9. Hounsfield, GODFREY N. "Computed medical imaging. Nobel lecture, Decemberr 8, 1979." Journal of computer assisted tomography 4, no. 5 (1980): 665-674.
10. Insights on Labels. https://www.kaggle.com/prk007/ insights-from-tabular-and-image-data.
11. Kaggle Evaluation Criterion. https://www.kaggle.com/ c / rsna - str - pulmonary - embolism-detection /overview/ evaluation.
12. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).
13. Liang, Jianming, and Jinbo Bi. "Computer aided detection of pulmonary embolism with tobogganing and mutiple instance classification in CT pulmonary angiography." In Biennial International Conference on Information Processing in Medical Imaging, pp. 630-641. Springer, Berlin, Heidelberg, 2007.
14. Radiological Society of North America. https : / /www. rsna.org/.
15. Ravì, Daniele, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, and Guang-Zhong Yang. "Deep learning for health informatics." IEEE journal of biomedical and health informatics 21, no. 1 (2016): 4-21.
16. Tajbakhsh, Nima, Michael B. Gotway, and Jianming Liang. "Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 62-69. Springer, Cham, 2015.
17. Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In International Conference on Machine Learning, pp. 6105-6114. PMLR, 2019.
18. Tourassi, Georgia D., Carey E. Floyd, H. Dirk Sostman, and R. Edward Coleman. "Acute pulmonary embolism: artificial neural network approach for diagnosis." Radiology 189, no. 2 (1993): 555-558.
19. Weon, Byung Mook, Jung Ho Je, Yeukuang Hwu, and Giorgio Margaritondo. "Phase contrast X-ray imaging." International journal of nanotechnology 3, no. 2-3 (2006): 280-297.
20. windowing explained. https : / / www . kaggle . com / redwankarimsony / ct - scans - dicom - files - windowing - explained.
21. Yang, Xin, Yi Lin, Jianchao Su, Xiang Wang, Xiang Li, Jingen Lin, and Kwang-Ting Cheng. "A two-stage convolutional neural network for pulmonary embolism detection from CTPA images." IEEE Access 7 (2019): 84849-84857.
22. Yun, Sangdoo, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. "Cutmix: Regularization strategy to train strong classifiers with localizable features." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 6023-6032. 2019.
23. Zhang, Hongyi, Moustapha Cisse, Yann N. Dauphin, and David Lopez-Paz. "mixup: Beyond empirical risk minimization." arXiv preprint arXiv:1710.09412 (2017).

COVID 19 Detection and Classification using Machine Learning

Dipanjan Acharya, Krishna Prasad M, Sivakumar R

Department of Sensor and Technology, Vellore Institute of Technology, Vellore, India
dipanjan.acharya2019@vitstudent .ac.in, krishnaprasad.m2019@vitstudent.ac.in, rsivakumar@vit.ac.in

Abstract. COVID 19 disease is causing a pandemic worldwide. As of now, reverse transcription-polymerase chain reaction (RT-PCR) test is the de facto test used to examine the existence of the ribonucleic acid (RNA) of SARS-CoV-27. RTPCR is one of the viable approaches to test for the disease. It is being used to test people for COVID 19 in most countries. Despite high specificity of testing with RT-PCR, this method has very low sensitivity. There is also high variability in efficacy as different methods of sampling and time of symptom occurence differ from person to person. The testing technique proposed as an essential tool for COVID-19 screening in this paper is especially effective as a complementary method to RT-PCR. COVID 19 infects mainly the lungs and alveoli. Lung CT scan images are being used. Therefore, it is essential for us to develop a machine learning code that can analyze lung computerized tomography images to detect presence of COVID 19. This is what we are trying to achieve by this project.

Keywords. COVID, CT images, Lung CT scan images, detection, Machine Learning

1. Introduction

About COVID 19

Coronavirus belongs to a large family of viruses which infect humans and animals alike. The symptoms associated with an infection of COVID 19 are similar to headaches, fatigues, cough and common cold. Since COVID 19 has never been manifested and recorded in humans before, there is no innate immunity in humans against the virus, Hence anyone can be infected.

The virus can thrive airborne for several hours or on surface for 14 days or more. Covid 19 impacts the lungs and the alveoli in our body. Therefore, in most cases chest/lung CT scan is appropriate choice for Covid 19 detection.

Some concerns we may have in our project:

- Chest CT and X-rays are inefficient in distinguishing between COVID-19 and other respiratory infections, like Tuberculosis.
- Majority of patients suffering from COVID-19 may have normal chest CTs or x-rays
- As COVID-19 is highly contagious, usage of imaging equipments on COVID-19 patients pose a serious health hazard for other patients and healthcare providers.

Our response

Our study shows that chest CT is sensitive and sufficiently specific for the diagnosis of COVID-19. But its sensitivity is still subject to revision. As we learn more about COVID 19 and coronavirus we will be able to use CT scan more efficiently. CT scans focus on the detection of hazy, patchy, “ground glass” white spots in the lung, a tell-tale sign of Covid-19.

Therefore, more we learning about the virus and how it infects the body we will be able to use the CT scan imaging technique with greater efficiency.

We do not intend to compare RT PCR test with that of lung CT scan based COVID-19 detection but hope that Lung CT scan imaging technique can be used to supplement RT PCR or other diagnostic test findings. Tests like RT PCR may take some time to get results and a CT scan conduct during this time can confirm findings of the diagnostic test. Patients with severe symptoms can obtain benefits from imaging because of assessment of the seriousness of the disease. With the additions of lab tests, thorough medical history and a physical exam, CT scans and x rays can be helpful for devising a treatment plan for a patient.

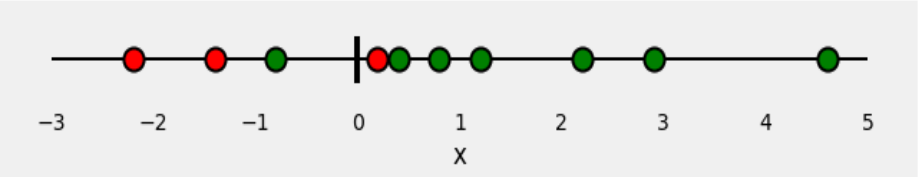
Lastly as we learn more about diagnosis using CT better and efficient equipment keeping in mind covid prevention guidelines and better testing problem can be developed.

2. Theory

Binary Cross Entropy is a type of Binary Classification. If a model is fit in this classification, it will predict the probability of the model belonging to one of the two classes. Loss function is used to evaluate the probability. A high loss indicates to a bad prediction and a low loss means that the prediction is good.

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

In a typical binary classification like our example, the loss function is the binary cross-entropy / log loss. where y is the label (1 means green and 0 means red.) and p(y) is the predicted probability of the point (how green is the point) for all N points.



Expected outcome

- Final Loss and accuracy is calculated
- A confusion matrix is obtained which gives us the True Positive (TP), False Negative, True Negative (TN) and False Positive (FP)
- Secondary parameters such as Sensiti vity can be obtained from the TP, TN, FP and FN values
- Graphs of Validation Accuracy and Model Accuracy can also be plotted
- Classification of any given image into Covid or Non Covid Image (Most Important Result)

Definitions

Pooling – In this operation, a two-dimensional filter slides over each channel of the feature map and sums up the features lying within the region covered by the filter.

Dense - A layer that is connected with its preceding layer in a highly convoluted manner. The neurons of one layer are connected to every neuron of its preceding layer.

The Dropout layer periodically sets an input unit to 0 with a fixed frequency rate at

each step for the entire timespan of the training time, which helps in prevention of overfitting.

Activation Function – This function defines how the weighted sum of the input is transformed into an output from a node or nodes in a layer of the network

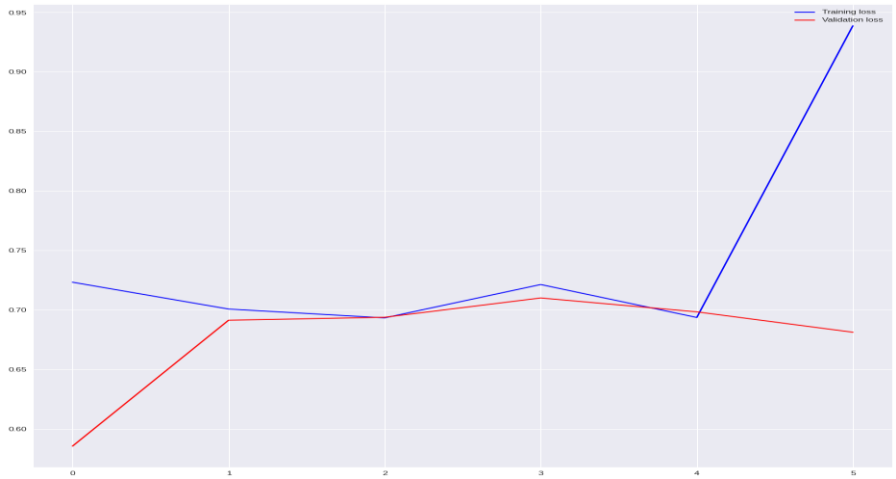
Flattening – It is the conversion of 2D data into an 1D array for sending the 1D array as input to the next layer

3. Algorithm

- Import required libraries such as numpy, matplotlib, pandas
- Categorize the dataset into training, validation and test datasets (70%. 20% and 10% respectively)
- Training set is for training the model
- Validation is for checking if our model is functioning properly (like mock test)
- Test set is for testing the model (like final test)
- We randomize the order of the dataset for better learning as overfitting causes the model to become rigid and categorize images not included in the dataset randomly. So this makes the results unreliable if randomization is not employed
- We use Flatten, Activation, Dense, Pooling functions to construct the CNN architecture
- A final prediction is made by the model after it has completed learning

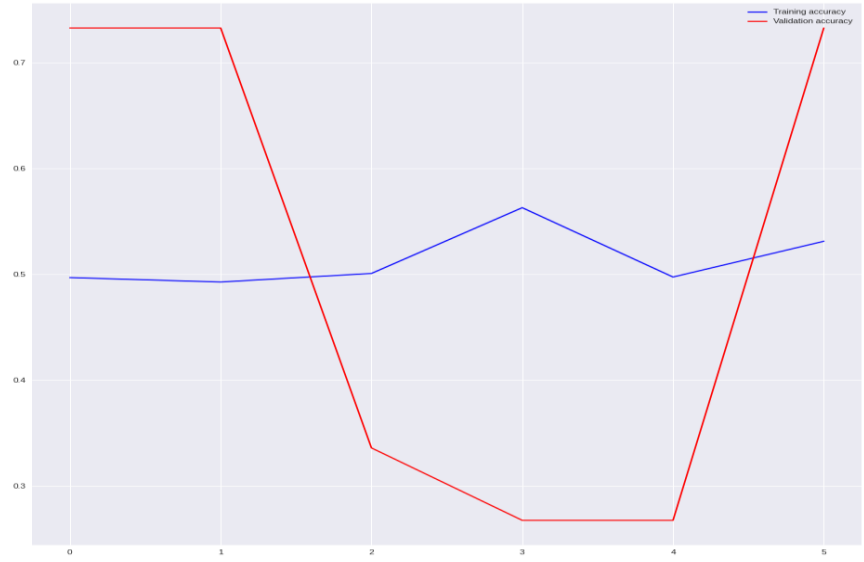
4. Result

Training and Validation Loss Graph (1st one)



This graph shows the loss obtained from the execution of the formulated code. As Binary Cross Entropy works on the process of binary classification and loss function is used to predict the probability, the loss function graph is an important indicator of the machine performance

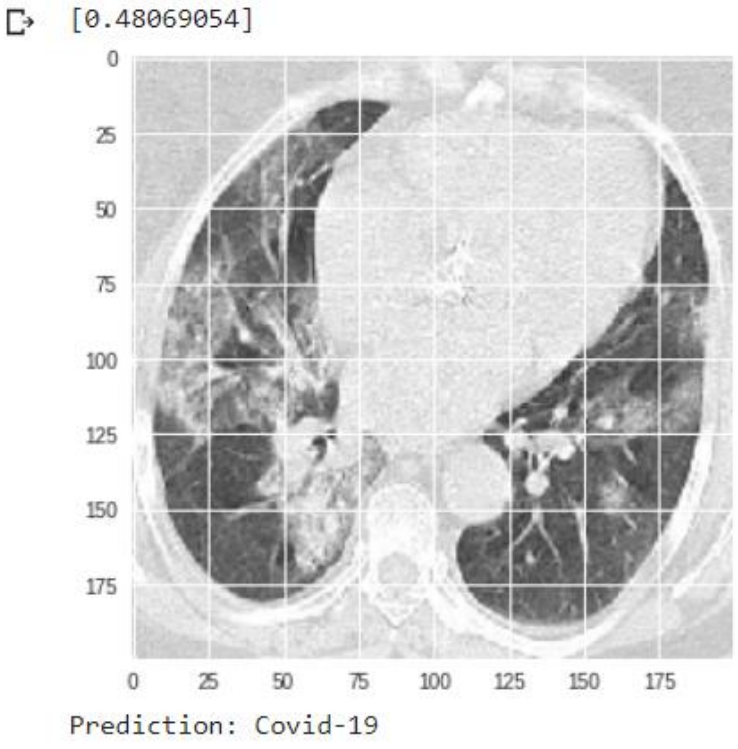
Training and Validation Accuracy Graphs (2nd one)



This graph shows the accuracy of the test and validation sets.

Test accuracy refers to the accuracy with which the model is able to learn from the data provided and validation accuracy is used to know about the reciprocation prediction rate of the machine. Validation data is not used to train the data.

Model Prediction and Probability



This is the final result which we obtain as result. Any picture we obtain, the machine will predict as to whether it is a COVID or non-COVID image.

5. Conclusion

In this project we were able to create a python code for detecting COVID 19 using CT scan images using the technique of binary cross entropy.

We obtained the CT Scan images from the database and using the technique mentioned above, we were able to confirm the presence or absence of COVID

References

1. Thoracic computerized tomography scans in one hundred eighteen orthopedic patients during the COVID-19 pandemic: identification of chest lesions; added values; help in managing patients; burden on the computerized tomography scan department (Jacques Hernigou & François Cornil & Alexandre Poignard & Said El Bouchaibi & Jean Mani & Jean François Naouri & Patrick Younes4 & Philippe Hernigou) International Orthopedics (2020)
2. A Rapid, Accurate and Machine-Agnostic Segmentation and Quantification Method for CT-Based COVID-19 Diagnosis (Longxi Zhou, Zhongxiao Li, Juexiao Zhou, Haoyang Li, Yupeng Chen, Yuxin Huang, Dexuan Xie, Lintao Zhao, Ming Fan , Shahrukh Hashmi, Faisal Abdelkareem , Riham Eiada , Xigang Xiao, Lihua L)
3. 3D Deep learning based detection of COVID-19 virus using lung Computed Tomography (Abdullah Aman Khan , Sidra Shafiq , Rajesh Kumar, Jay Kumar , Amin Ul Haq)
4. Rapid COVID 19 Diagnosis using Deep Learning of CT Scan (Hamed Tabrizchi, Amir Mosavi, Akos Szabo-Gali, Imre Felde, Laszlo Nada)
5. Generation of COVID-19 Chest CT Scan Images using Generative Adversarial Networks (Prerak Mann, Sahaj Jain, Saurabh Mittal, Aruna Bhat)
6. COVID CT-Net: Predicting Covid-19 From Chest CT Images Using Attentional Convolutional Network (Shakib Yazdani, Shervin Minaee, Rahele Kafieh, Narges Saeedizadeh and Milan. Sonka)
7. Automated Detection of Lung Cancer Using CT Scan Images (Ariful Hoque;A.K.M. Ashek Farabi;Fahad Ahmed;Md. Zahidul Islam)
8. Dataset Link
   https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset
   https://blog.radiology.virginia.edu/covid-19-and-imaging/
   https://www.cochrane.org/CD013639/INFECTN_how-accurate-chest-imagingdiagnosing-covid-19
   https://www.statnews.com/2020/04/16/ct-scans-alternative-to-inaccuratecoronavirus-tests/
   https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-avisual-explanation-a3ac6025181a

Breast Cancer Detection using Thermography

Ridita Garg, S. S Nithin Raja, Sivakumar R

School of Electronics Engineering, Vellore Institute of Technology, Vellore, India
ridita.garg2019@vitstudent.ac.in, nithinraja.ss2019@vitstudent.ac.in, rsivakumar@vit.ac.in

Abstract. The paper surveys thermal breast images utilized for malignant growth location—an altogether harmless and non-contact imaging strategy. The technique can build the early analysis pace of breast cancer. The profound learning-based technique distinguishes breast malignant growth with a 93.95% exactness rate. In the course of recent years, a few procedures have been proposed for this reason, for example, mammography, which is often utilized for breast cancer growth determination. Be that as it may, bogus up-sides of mammography can happen when the patient is analyzed decidedly by another method. Also, the expected results of utilizing mammography might urge patients and doctors to search for other analytic strategies. Our survey investigated infrared advanced imaging, which accepts that a fundamental thermal correlation between a healthy breast and a breast with malignant growth generally shows an expansion in thermal movement in the precancerous tissues and the regions encompassing creating cancer disease. In this paper, we utilized SVM and KNN Model.

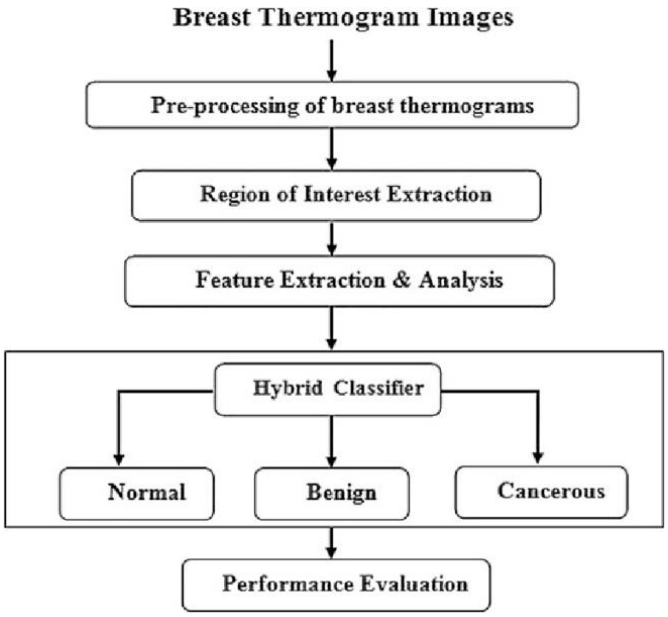
Keywords. Breast Cancer, Thermography, Machine learning, SVM, KNN, Mammography

1. Introduction

The human body naturally manages the creation, growth, and death of the cells in its tissues. Once this process starts to work abnormally, and the cells are not dying at the rate they should, we see an increase in the ratio of cell growth to cell death, which is a direct cause of cancer. Breast cancer occurs when cells in the breast divide and grow without reasonable control. It is a well-known disease around the world; in the USA, one in eight women will be diagnosed with breast cancer in her lifetime, and more than 40,000 is dying every year in USA. Thermal breast images are used for cancer detection. An entirely non-invasive and noncontact imaging technique is used. The method can increase the early diagnosis rate of breast cancer. Deep learning based method detects breast cancer with 93.95% accuracy rate.

The main objective of this project is to help doctors analyze the huge datasets of cancer data and find patterns with the patient's data and that cancer data available. With this analysis we can predict whether the patient might have breast cancer or not. There are two types of cancers. Malignant and Benign. Malignant cancers are cancerous. These cells keep dividing uncontrollably and start affecting other cells and tissues in the body. They spread to all other parts of the body and it is hard to cure this type of cancer. Benign cancer is non-cancerous.

Unlike malignant, this tumor does not spread to other parts of the body and hence is much less risky that malignant. In many cases, such tumors don't really require any treatment. In our paper we use KNN and SVM models to classify and predict if the cancerous cells are malignant or benign using numerous parameters.



2. Definition of thermography

Thermography is a test that uses an infrared camera to detect heat patterns and blood flow in body tissues. Digital infrared thermal imaging (DITI) is the type of thermography that's used to diagnose breast cancer. DITI reveals temperature differences on the surface of the breasts to diagnose breast cancer.

Mammography (taking X-ray pictures of the breasts) is the most effective breast cancer screening method and the only method proven to increase the chance of survival through earlier detection

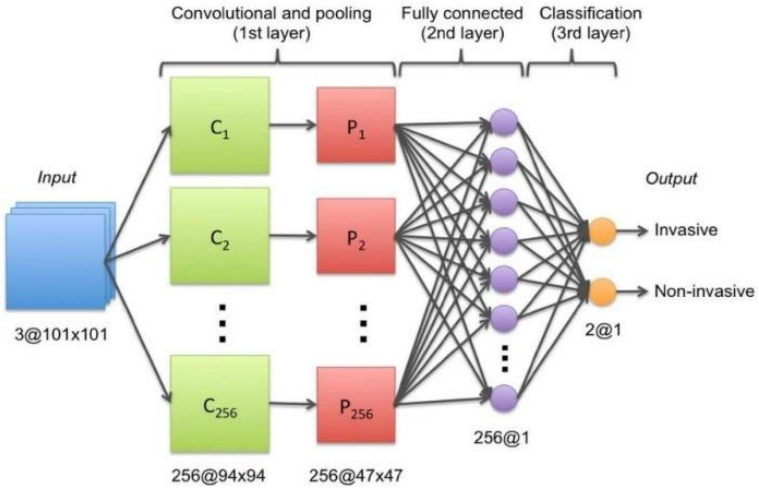
3. Previous techniques and comments

Based on several articles from 2002–2010, the authors of previous related research Kosus et al., 2010 undertook a review that illustrates the limitations of screen-film mammography (SFM). For example, a large number of false positives occur with this technique, with a rate between 4% and 34%. There is a considerable amount of interest in image processing through a network of neurons. To be more precise, several researchers have explored the detection and assessment of severity of cancer in this manner shows how, thanks to ConvNet, we can classify cancer as invasive or non-invasive. In addition, some authors have demonstrated the power of the neural network model (NN) and its variants (Recurrent Neural Network (RNN), Deep Neural Network (DNN), etc.). They also stressed the importance of solving the problem of cancer of the breast through innovative techniques in computer science.

4. Dataset

The utilized dataset for the research/analysis was acquired from the Research Data Base (DMR) containing frontal thermo gram images, with a resolution of 640 × 480 pixels. The dataset contains images of middle aged individuals these images include breasts of different shapes and sizes, such as medium, wide, and asymmetric breasts. 1) ID number 2) Diagnosis (M = malignant, B = benign) 3-32) Ten real-valued features are computed for each cell

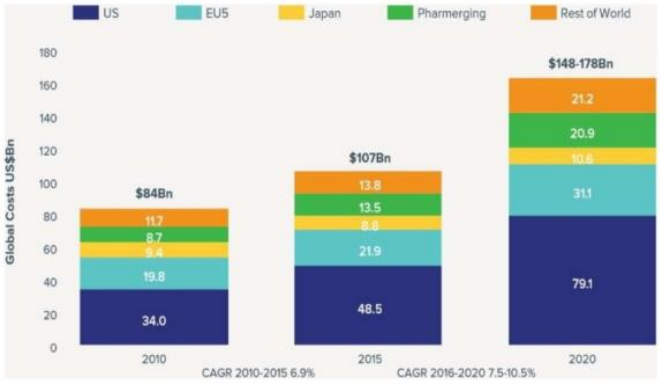
nucleus: a) radius (mean of distances from center to points on the perimeter) b) texture (standard deviation of gray-scale values) c) perimeter d) area e) smoothness (local variation in radius lengths) f) compactness (perimeter^2 / area - 1.0) g) concavity (severity of concave portions of the contour) h) concave points (number of concave portions of the contour) I) symmetry j) fractal dimension ("coastline approximation" - 1) The mean, standard error and "worst" or largest (mean of the three largest values) of these features were computed for each image, resulting in 30 features. For instance, field 3 is Mean Radius, field 13 is Radius SE, and field 23 is Worst Radius. Class distribution: 357 benign, 212 malignant



5. Economic aspects of breast cancer

Breast cancer is the most common cancer in women, accounting for 23% of all newly diagnosed oncology cases. 85% of families in which one of the parents will have cancer are without the help of family, friends, non-profit organizations, or even debt.

Emerging gene or epigenetic therapies that directly influence the genetic information responsible for tumor growth will be even more expensive Oncology cost growth is expected to be between 7.5% and 10.5% per year by 2020, when global oncology costs will exceed $150 billion. Give these factors, the need to find a new method of detection or achieve even a small improvement in current processes is more than evident.



6. Proposed work

The review clearly shows the possibilities to map fields to search for specific quantitative models of sustainability. These models have to be based on in-depth qualitative analysis of the concepts that make up the modern understanding of corporate sustainability. The discussed above is a model of this kind.

In our research paper we aim to find the accuracy of the machine learning models which are used in detecting the types and classifying the given cancer image. We are using a KNN model supported with SVM model. A brief flow diagram is shown below.

Image processing:

The images we obtained from the data base was converted into a numerical equivalent format with all 12 parameters. So that the ML model would easily process the data.

Dataset Preprocessing:

The data then is randomly split into training and testing set and fed to the ML models.

KNN classifier:

The KNN classifier play the most important role of classifying and identifying the various parameters of breast cancer which helps us in the detection of the type of cancer. It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.

SVM

In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is a number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well and thus it helps the KNN classifier to be more accurate.

Performance evaluation:

In our model we have attained an accuracy of 93.5% in the classification of the types of cancer.

Advantages of KNN:

1. No Training Period: KNN is called Lazy Learner (Instance based learning). It does not learn anything in the training period. It does not derive any discriminative function from the training data. In other words, there is no training period for it. It stores the training dataset and learns from it only at the time of making real time predictions. This makes the KNN algorithm much faster than other algorithms that require training e.g. SVM, Linear Regression etc.

2. Since the KNN algorithm requires no training before making predictions, new data can be added seamlessly which will not impact the accuracy of the algorithm.

3. KNN is very easy to implement. There are only two parameters required to implement KNN i.e. the value of K and the distance function (e.g. Euclidean or Manhattan etc.)

Code:

```
Import sys
import scipy
import numpy
import matplotlib
import pandas
import sklearn
```

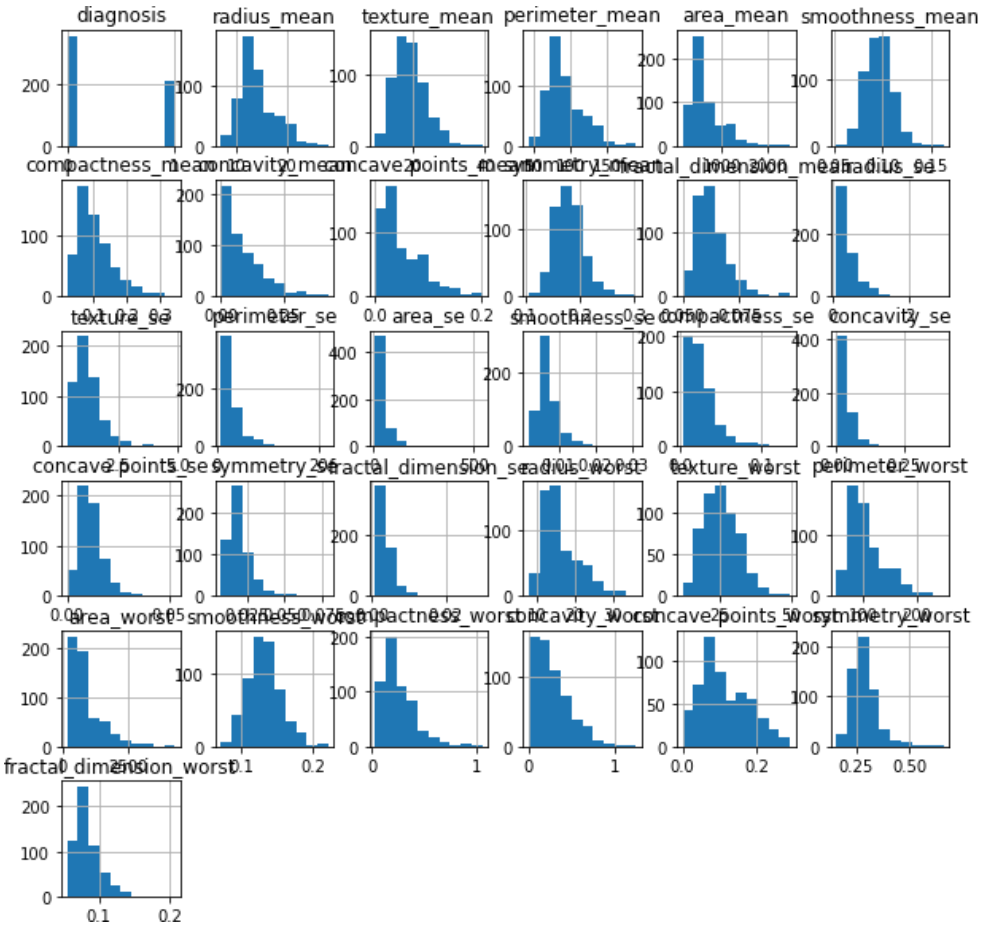
```
import numpy as np
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn import model_selection
from sklearn.metrics import classification_report
from sklearn.metrics import accuracy_score
from pandas.plotting import scatter_matrix
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
!pip install shap
import shap
shap.initjs
url = "DMR.data"
names = ['id',
'clump_thickness','uniform_cell_size','uniform_cell_shape','marginal_adhesion','single_epith
elial_size','bare_nuclei','bland_chromation','normal_nucleoli','nutises','class']
df = pd.read_csv("data.csv")
#preprocessing the data
df.replace('?',-99, inplace=True)
print(df.axes)
df.drop(['id'],1,inplace = True)
#print the shape of the dataset
print(df.shape)
encode={"M":1,"B":0}
df['diagnosis']=df['diagnosis'].map(encode)
df.isnull().sum()
df.head()
df.columns
df.drop("Unnamed: 32",axis=1,inplace=True)
#Dataset visualisation
print(df.loc[6]
#Dataset Visualisation
print(df.loc[6])
print(df.describe())
#plot histogram for each varibale
df.hist(figsize=(10,10))
plt.show()
scatter_matrix(df,figsize=(18,18))
plt.show()
X_train, X_test, y_train, y_test =
model_selection.train_test_split(df.drop(['diagnosis'],axis=1),df['diagnosis'],test_size=0.2)
#testing options
seed = 42
scoring = 'accuracy'
models = []
models.append(('KNN', KNeighborsClassifier(n_neighbors=5)))
```

```
models.append(('SVM',SVC()))
#evaluate each model in turn
results = []
names = []
for name, model in models:
kfold = model_selection.KFold(n_splits=10,random_state=seed,shuffle=True)
cv_results=model_selection.cross_val_score(model,X_train,y_train,cv=kfold,scoring
=scoring)
results.append(cv_results)
names.append(name)
msg="%s: %f (%f)" %(name,cv_results.mean(),cv_results.std())
print(msg)
#make prediction on validation dataset
for name, model in models:
model.fit(X_train,y_train)
predictions = model.predict(X_test)
print(name)
print(accuracy_score(y_test,predictions))
print(classification_report(y_test,predictions))
#Accuracy is the ration of correctly predicted observation to the total observations
pg. 9
#prediction (false positive) ratio of correctly predicted positive observations to the
total predicted positive observations
#Recall (sensitivity): (false negative) ratio of correctly predicted positive observations
to all observations in actuall class - yes
#f1-score: is the weighted average of prediction and recall. Therfore, this score takes
both positive and negative
from sklearn.metrics import confusion_matrix
sns.heatmap(confusion_matrix(y_test,predictions),annot=True)
clf = KNeighborsClassifier()
clf.fit(X_train,y_train)
accuracy = clf.score(X_test,y_test)
print(accuracy)
prob=clf.predict_proba(X_test)
svm = SVC(probability=True)
svm.fit(X_train,y_train)
accuracy = svm.score(X_test,y_test)
print(accuracy)
prob_svm=svm.predict_proba(X_test)
plt.figure(figsize=(8,8))
plt.plot(y_test - prob[:,1],linestyle=' ',marker="o",label="KNN")
plt.plot(y_test -prob_svm[:,1],linestyle=" ",marker = "+",label="SVM")
plt.legend()
def sample_feature_importance(idx, type='condensed'):
if type == 'condensed':
return shap.plots.force(shap_values[idx])
elif type == 'waterfall':
return shap.plots.waterfall(shap_values[idx])
else:
```

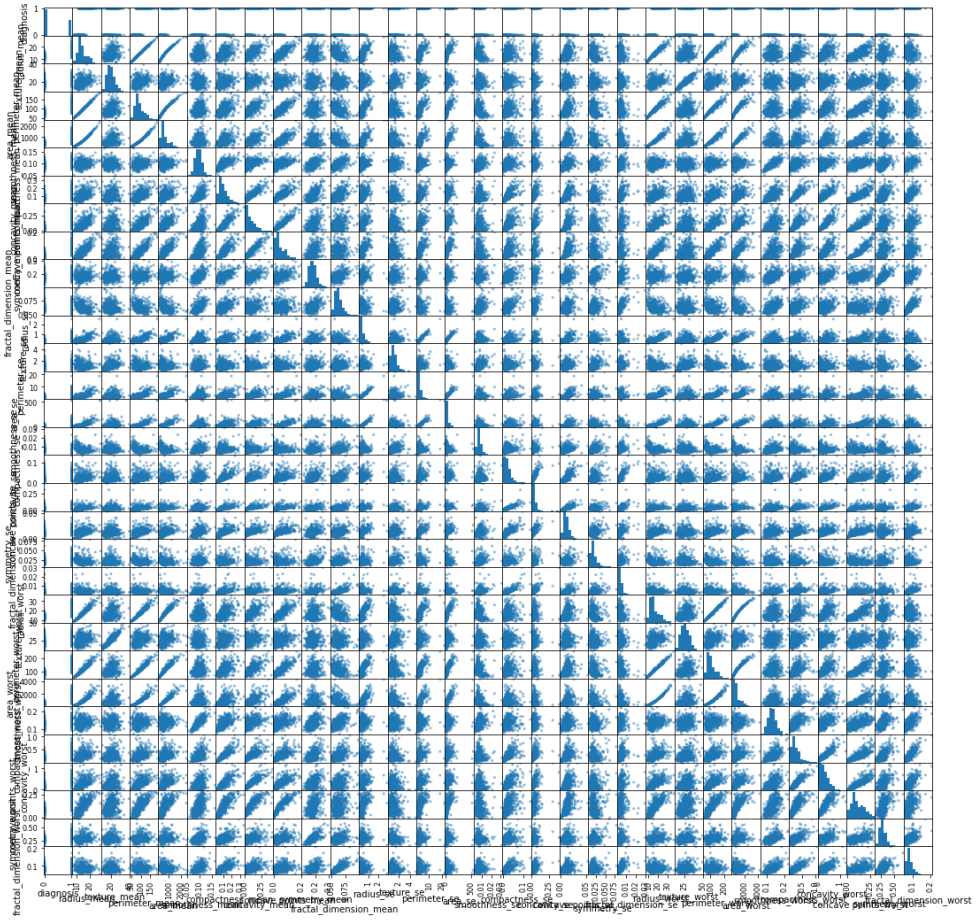
return "Return valid visual ('condensed', 'waterfall')"
shap.plots.waterfall(shap_values[56])
shap.plots.waterfall(shap_values[78])

7. Results and discussion

Histograms of various parameters:



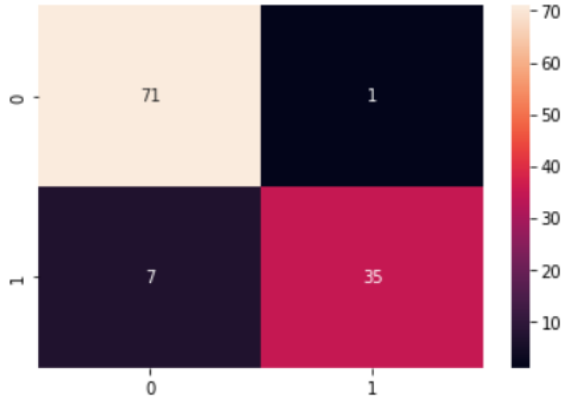
Scatter matrix



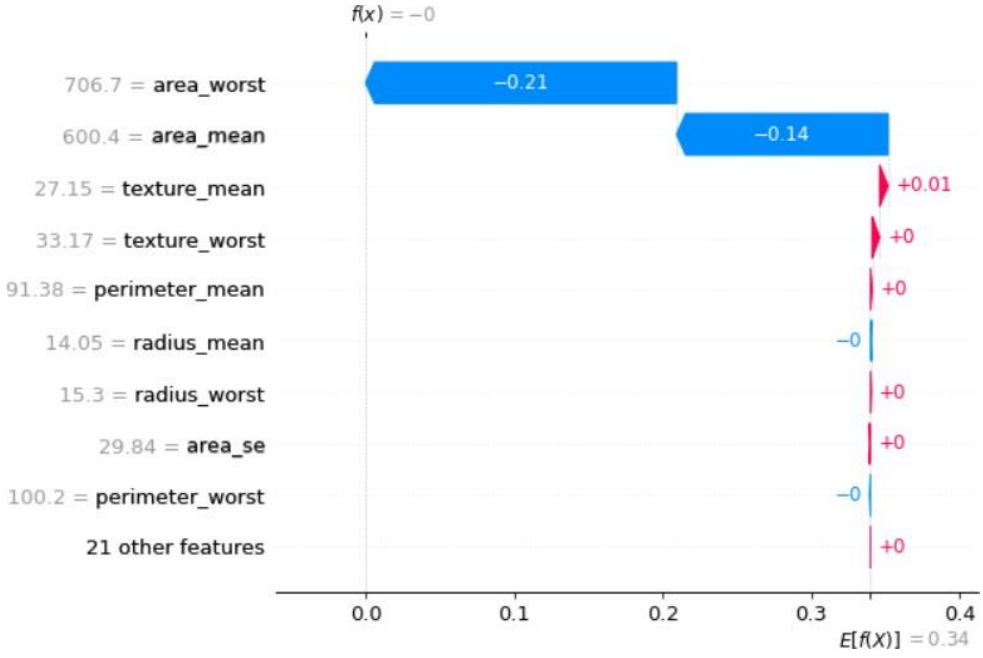
Accuracy of the models:

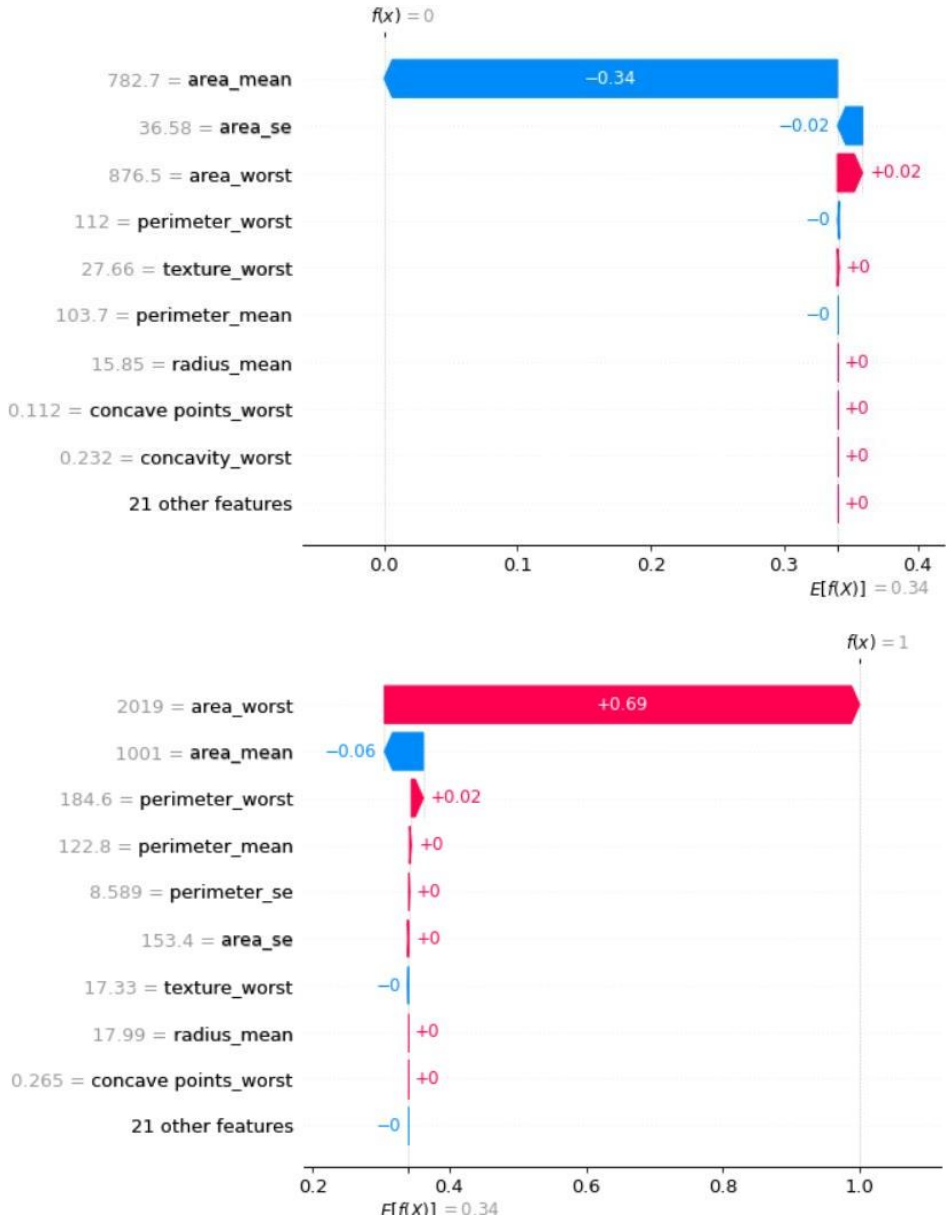
```
KNN
0.9385964912280702
              precision    recall  f1-score   support

           0       0.96      0.94      0.95        72
           1       0.91      0.93      0.92        42

    accuracy                           0.94       114
   macro avg       0.93      0.94      0.93       114
weighted avg       0.94      0.94      0.94       114

SVM
0.9298245614035088
              precision    recall  f1-score   support

           0       0.91      0.99      0.95        72
           1       0.97      0.83      0.90        42

    accuracy                           0.93       114
   macro avg       0.94      0.91      0.92       114
weighted avg       0.93      0.93      0.93       114
```

Confusion matrix:



Waterfall plots for major impact parameters:





We are able to achieve an accuracy of 93.08% using the SVM and KNN models in the prediction of cancer.

8. Conclusion

Our review of the literature explored infrared digital imaging, which assumes that a basic thermal comparison between a healthy breast and a breast with cancer always shows an increase in thermal activity in the precancerous tissues and the areas surrounding developing breast cancer. In this paper, we used SVM and KNN Model. The novel contribution of this

paper is the production of a comparative study of several breast cancer detection data sets using deep learning models.

References

1. Breast Cancer Detection Through Gabor Filter Based Texture Features Using Thermo grams Images https://ieeexplore.ieee.org/document/8703342
2. Optimization of skin cooling by computational modelling for early thermo graphic detection of breast cancer https://www.researchgate.net/publication/328663471_Optimization_of_skin_cooling_by_computational_modeling_for_early_thermographic_detection_of_breast_cancer
3. Patient/Breast-Specific Detection of Breast Tumor Based on Patients' Thermo grams, 3D Breast Scans, and Reverse Thermal Modelling https://www.mdpi.com/2076-3417/11/14/6565
4. Breast Cancer Detection in Thermal Infrared Images Using epresentation Learning and Texture Analysis Methods https://www.mdpi.com/2079-9292/8/1/100
5. Breast Cancer detection from Thermo grams Using Feature Extraction and Machine Learning Techniques https://ieeexplore.ieee.org/document/9033713
6. SPAER: Sparse Deep Convolutional Auto encoder Model to Extract Low Dimensional Imaging Biomarkers for Early Detection of Breast Cancer Using Dynamic Thermography https://www.mdpi.com/2076-3417/11/7/3248
7. Hybrid methods for feature extraction for breast masses classification https://www.researchgate.net/publication/319617463_Hybrid_methods_for_feature_extraction_for_breast_masses_classification
8. Thermal Imaging Based Breast Cancer Analysis Using BEMD and Uniform RLB https://ieeexplore.ieee.org/document/8082268
9. Feasibility test of Dynamic Cooling for detection of small tumors in IR thermo graphic breast imaging https://www.researchgate.net/publication/335944091_Feasibility_test_of_Dynamic_Cooling_for_detection_of_small_tumors_in_IR_thermographic_breast_imaging
10. Breast cancer diagnosis using thermography and convolutional neural networks https://www.sciencedirect.com/science/article/abs/pii/S0306987719313659

Joint Reconstruction of MRI - PET

Sivakumar R, Boppana Hemanth, Mekala Radha Krishna Vardhan, N N Sai Deepak Reddy

Department of Sensor and Technology, Vellore Institute of Technology, Vellore, India
rsivakumar@vit.ac.in, boppana.hemanth2019@vitstudent.ac.in, mekala.radhakrishna2019@vitstudent.ac.in, saideepak.reddynn2019@vitstudent.ac.in.

Abstract. The combined positron emission tomography (PET) and magnetic resonance imaging (MRI) scanners obtain the functional PET and anatomical or functional MRI data concurrently. As the data of both techniques are more likely to show structures that are alike, we have utilized this by joint reconstruction of PET and MRI. This can be achieved by using MATLAB functions. Our results shown the improved image quality and resolution of joined PET and MRI image. By this information we can help doctors to view both anatomical and functional data of a body, it reduces the effort and time of doctors as in the case of taking both PET and MRI separately. In disparity to currently used PET/CT systems, PET/MRI offers not only enhanced soft-tissue contrast and decreased the levels of ionizing radiation, but also a plenty of MRI-specific information such as functional, spectroscopic and diffusion tensor imaging. The fusion of these two excellent diagnostic imaging techniques into a single scanner improves the diagnostic accuracy by easing the accurate registration of molecular aspects and metabolic alterations of the diseases with exact correlation to anatomical findings and morphological information. PET/MRI is a very useful diagnostic technique for oncological imaging and for use in cancer screening in the decades to come due to the relatively less radiation exposure compared to PET/CT and the very high soft tissue resolution of MRI.

Keywords. Positron emission tomography, magnetic resonance imaging, joint reconstruction.

1. Introduction

MRI: Magnetic resonance imaging (MRI) may be a medical imaging technique that uses a magnetic flux and computer-generated radio waves to generate detailed images of the organs and tissues in human body. Most MRI machines are massive, tube-shaped magnets. When you sprawl inside an MRI machine, the magnetic flux temporarily reorganizes water molecules in your body. Radio waves cause these organized atoms to generate indistinct signals, which are utilized to create cross-sectional MRI images — like pieces in a loaf of bread. The MRI machine also can generate 3D images which will be observed from distinct angles.

PET: A positron emission tomography (PET) scan is an imaging technique which will assist in disclosing the metabolic or biochemical function of your tissues and organs. The PET scan utilizes a radioactive drug (tracer) to point out both normal and abnormal metabolic activity. A PET scan can often identify the abnormal metabolism of the tracer in diseases before the disease appears on other imaging tests, like computed tomography (CT) and magnetic resonance imaging (MRI).

The tracer is most frequently injected into a vein inside your hand or arm. The tracer will then accumulate into areas of your body that have higher levels of metabolic or

biochemical activity, which frequently pinpoints the situation of the disease. The PET images are typically merged with CT or MRI and are called PET-CT or PET-MRI scans.

2. Joint MRI PET scan

A PET/MRI checkup is a 2 in 1 scan that combines images from a positron emission tomography (PET) checkup and a magnetic resonance imaging (MRI) checkup. This new mongrel technology harnesses the strengths of PET and MRI to produce some of the most largely detailed pictures of the interior of your body . Doctors will use these images for diagnosing medical conditions and plan their treatment. For example, PET/MRI scans of the brain are useful for Alzheimer's disease, epilepsy, and brain tumors.

Advantages of PET/MRI:

1) Two tests can be taken at once
2) Radiation dose is reduced
3) Image quality is enhanced
4) Good comprehensive diagnosis

Working

MRI scans use a robust magnetic flux to supply detailed images of internal structures of the body. They can also provide information about how well these structures are functioning. PET scans use tracers to spotlight abnormalities that indicate disease. Until now, scientists couldn't integrate PET and MRI for simultaneous scanning because MRI's powerful magnets interfered with the imaging detectors on the PET scanner. PET and MRI scans are conducted separately, and therefore the separate images later merged. That merger, however, requires a posh computer process.

The whole procedure generally take 2 to 3 hours. The PET/MRI scan itself takes about an hour, depending on the number of scans. Your child may need extra recovery time after sedation.

3. MATLAB commands used

(1) A = Imread(filename) interprets the image from the file described by filename, deriving the format of the file from its constituents. If filename may be a multi-image file, then Imread interprets the fundamental image within the file.

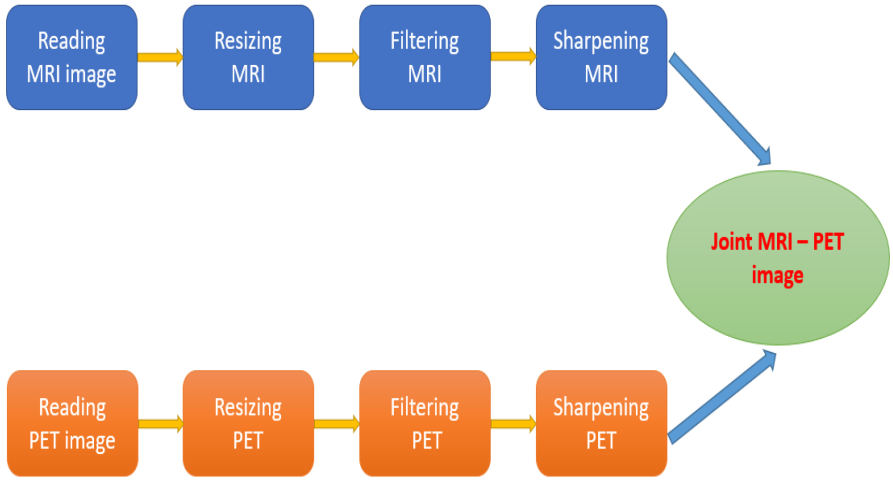
(2) A = Imshow(I) shows the grayscale representation I in a figure. Imshow utilizes the default display scope for the image data type and enhances the figure, axes, and image object attributes for image display.

(3) B = imresize(A,scale) remits image B that's scale times the proportions of image A. The input image A are frequently a grayscale, RGB, binary, or else a categorical image. If A has quite two proportions, then imresize only rescales the primary two proportions. If scale is in the middle of 0 and 1, then B is lesser than A. If scale is larger than 1, then B is greater than A. By default, imresize utilizes bicubic interpolation.

(4) B = imsharpen(A) enhances the grayscale or true color (RGB) input image A by making use of the unsharp masking technique.

(5) C = imfuse(A,B) generates a combined image from two different images, A and B. If A and B are different dimensions, imfuse pads the lesser dimensions with zeros in order that both images are an identical size in prior to creating the composite image. The output, C, may be a numeric matrix consisting of a merged version of images A and B.

4. Flow chart for joint MRI PET reconstruction



5. Methodology

a) Fusing image: We use this imaging fusion method to gather all the data from different images and fuse it as a single image, so that the final image is more informative, accurate and precise and contains all information from multiple source images.

b) Resizing image: We rescale the image so the that all the images can be processed equally and the output image of desired size can be obtained without magnification. This resizing of an image can be done by simply using imresize vector and specifying the number of rows and columns required in the output image.

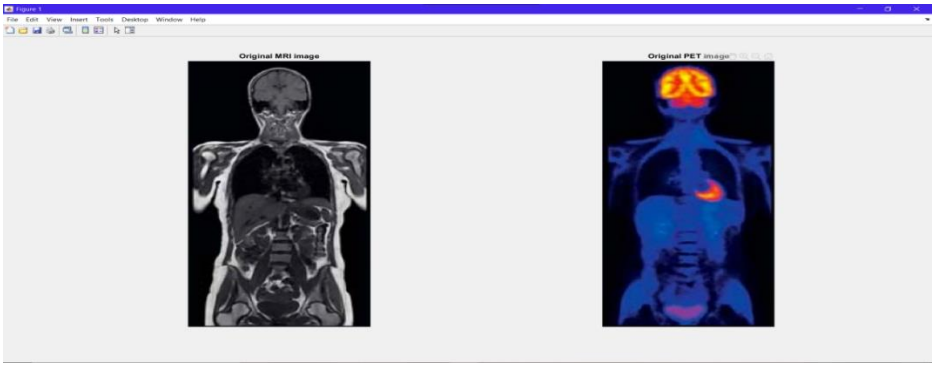
c) Filtering image: Filtering the image is one of the vital steps in any image processing system. The noise present in the image or any unwanted data can be removed during filtering. Wiener filter is best suited for smoothening noise and removing additive noise simultaneously inverting the blurring.

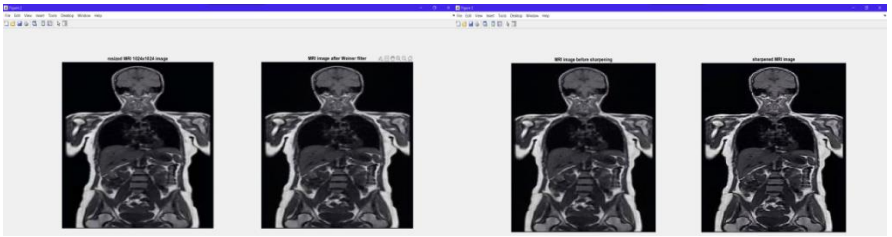
d) Sharpening image: Sharpening is a technique that enhances and highlights the edges and fine details in an image. For the biomedical imaging edges can be a useful information so edge preservation, improving local contrast, sharpening and image enhancement is to be done for fine details.

6. Simulation results

For images of patient suffering from Alzheimer's disease and epilepsy:

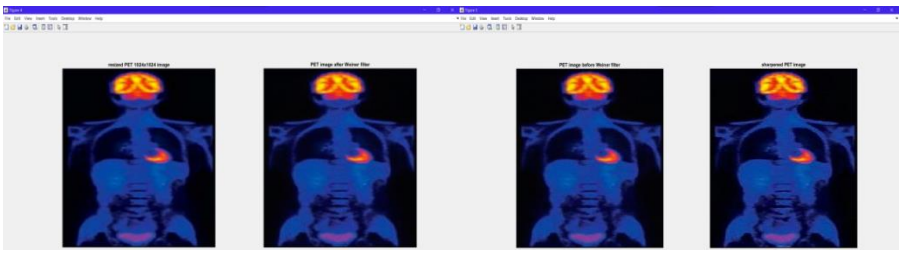
Original MRI and PET images





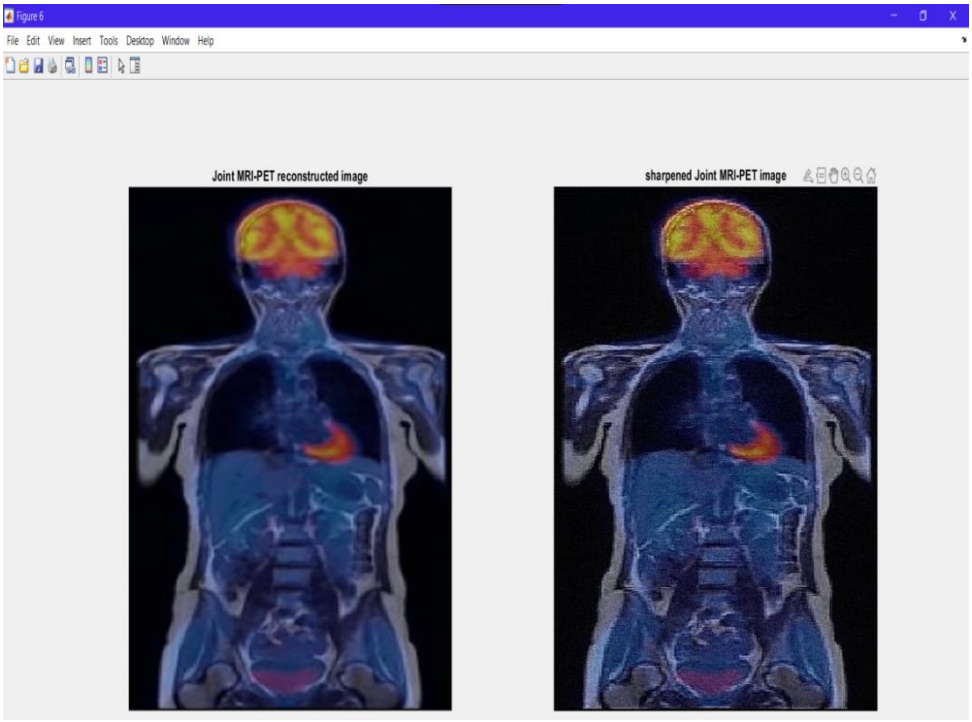
(a) (b)

Fig. (a) MRI image resized to 1024x1024 pixel size and MRI image after filtering using wiener filter. Fig. (b) MRI image before sharpening and after sharpening using unsharp masking sharpening operator.



(c) (d)

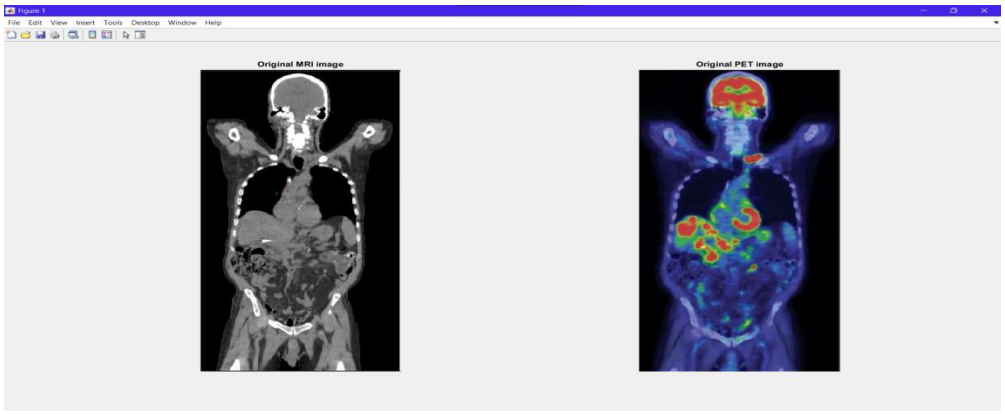
Fig. (c) PET image resized to 1024x1024 pixel size and PET image after filtering using wiener filter. Fig. (d) PET image before sharpening and after sharpening using unsharp masking sharpening operator.



(e)

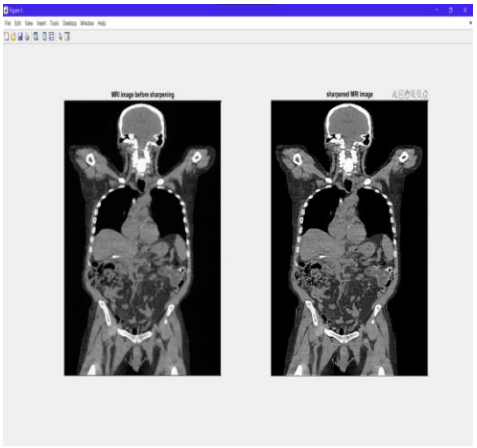
Fig. (e) Joint reconstructed MRI-PET image and sharpened joint MRI-PET image.

1) For images of patient suffering from liver cancer:



(f)

Fig. (f) Original MRI and PET images.



(g) (h)

Fig. (g) MRI image resized to 1024x1024 pixel size and MRI image after filtering using wiener filter. Fig. (h) MRI image before sharpening and after sharpening using unsharp masking sharpening operator.

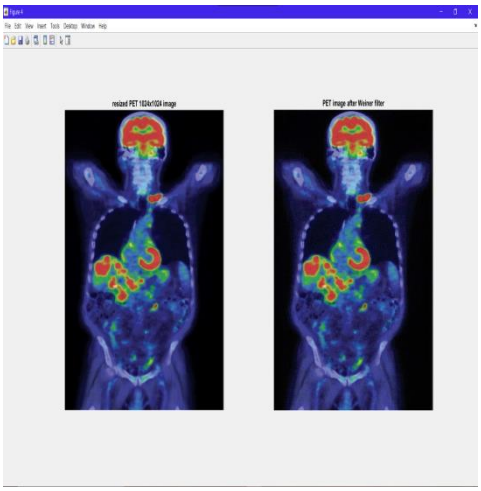


Fig. (i) PET image resized to 1024x1024 pixel size and PET image after filtering using wiener filter.

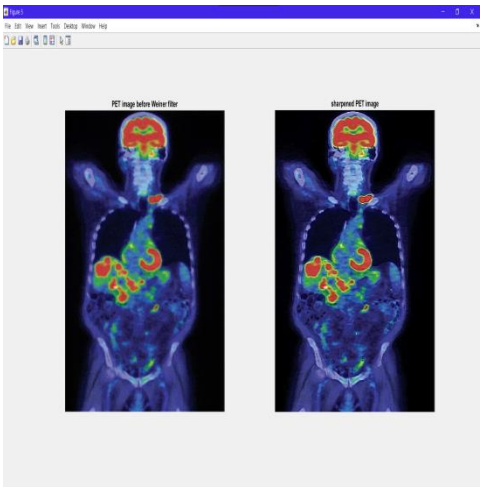
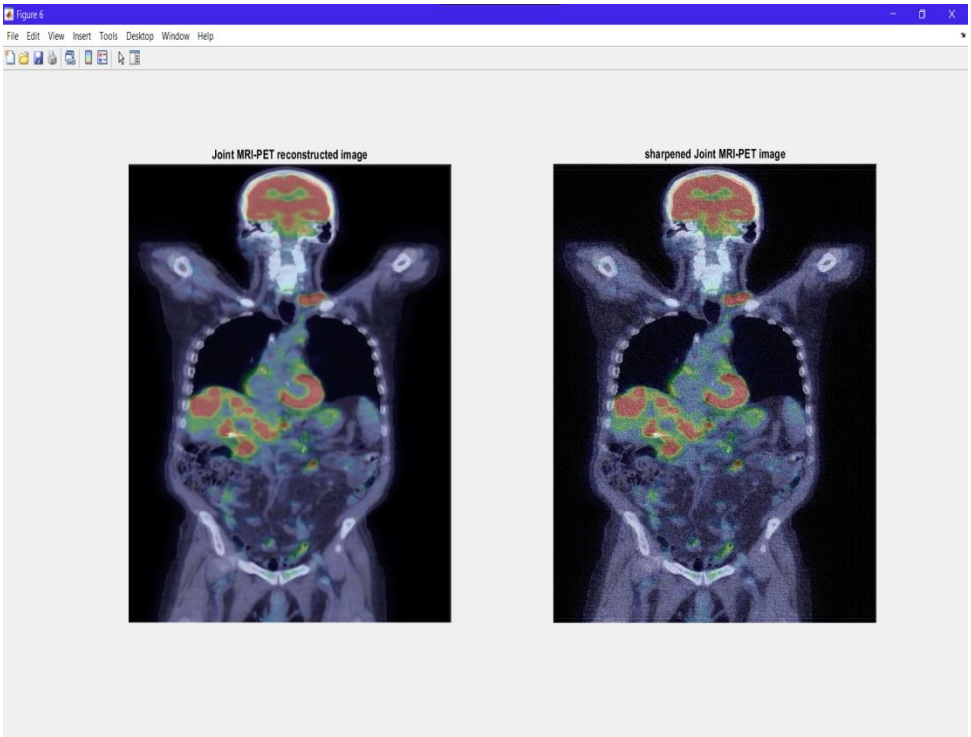


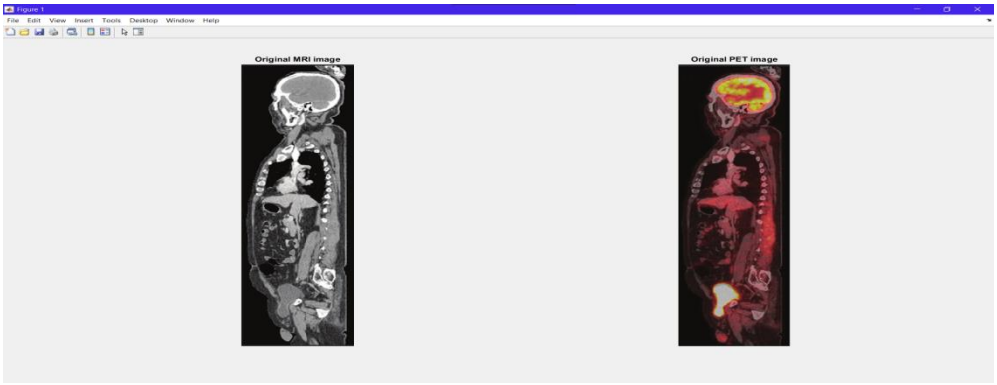
Fig. (j) PET image before sharpening and after sharpening using unsharp masking sharpening operator



(k)

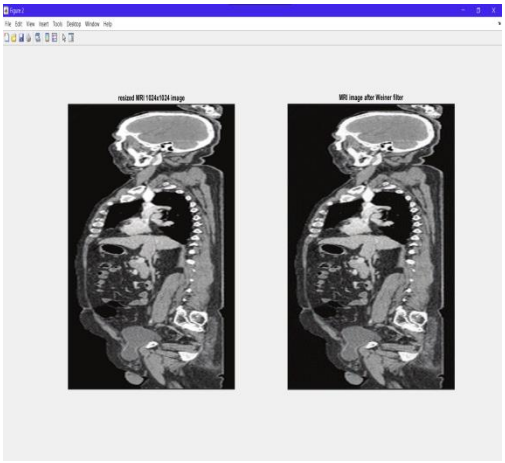
Fig. (k) Joint reconstructed MRI-PET image and sharpened joint MRI-PET image.

2) For images of patient suffering from brain cancer:

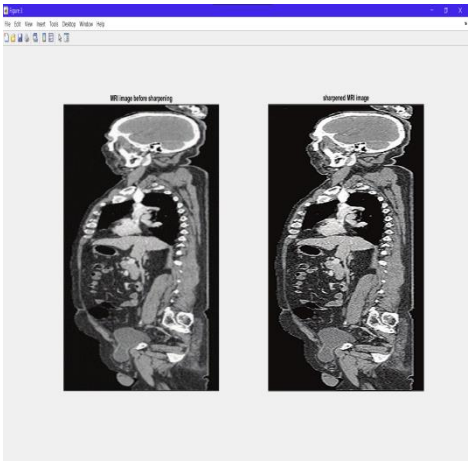


(l)

Fig. (l) Original MRI and PET images

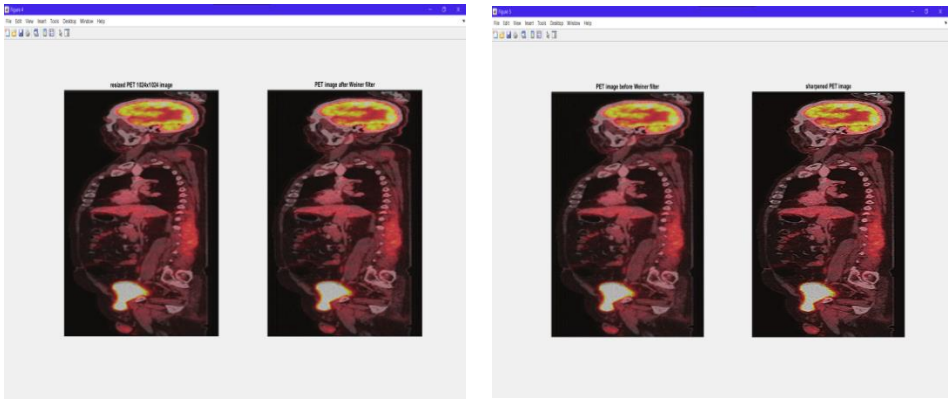


(m)



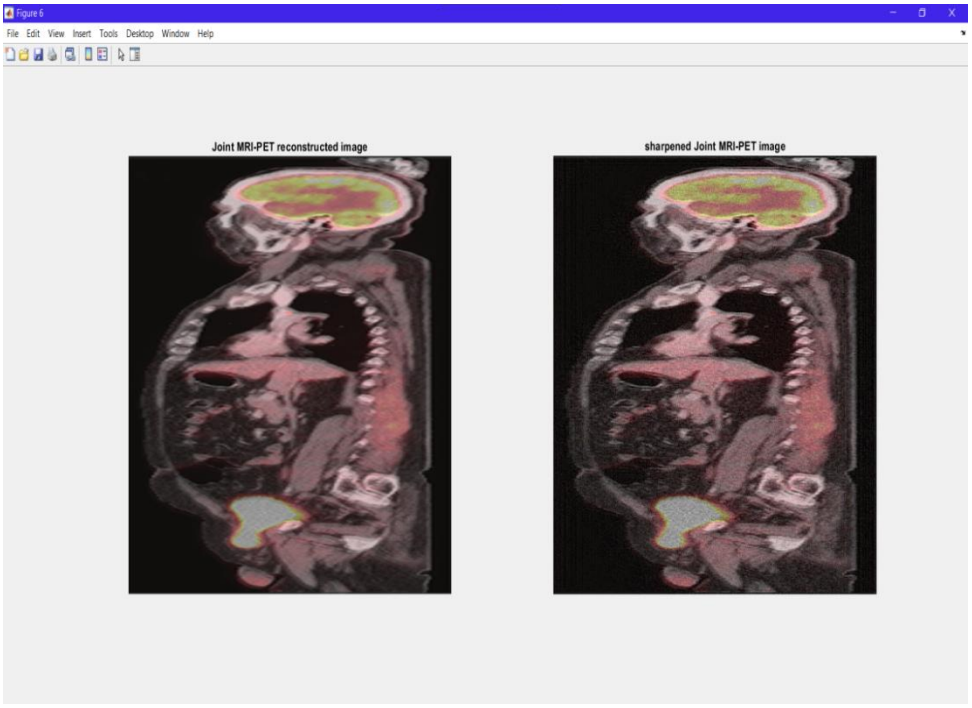
(n)

Fig. (m)MRI image resized to 1024x1024 pixel size and MRI image after filtering using wiener filter. Fig. (n) MRI image before sharpening and after sharpening using unsharp masking sharpening operator.



(o) (p)

Fig. (o) PET image resized to 1024x1024 pixel size and PET image after filtering using wiener filter. Fig. (p) PET image before sharpening and after sharpening using unsharp masking sharpening operator.



(q)

Fig. (q) Joint reconstructed MRI-PET image and sharpened joint MRI-PET image.

7. Conclusion

MRI scans enable us to get structural, anatomical image profile of body whereas the PET scans enable us to get the functional image profile. Joint MRI/PET enable concurrent procurement of both anatomical and functional imaging. By combining both these techniques,

so that both the anatomical and functional imaging can be observed in a single combined image, reduces the overall cost of scanning the patient and makes biomedical imaging very affordable and beneficent.

Fusing the reconstruction of PET and MRI can be advantageous to both techniques. The fused reconstructed image also enables us to view more precise, accurate, displays relatively fewer artifacts, reduced errors and has sharper edges than separate reconstructions. The joint MRI/PET reconstruction technique helps the doctors to effortlessly detect and identify the abnormality or disease which the patient is suffering from, in an affordable, cost effective and less time consuming way than individual imaging.

References

1. S. R. Cherry, "Multimodality in Vivo Imaging Systems: Twice the Power or Double the Trouble?" Annual Review of Biomedical Engineering, vol. 8, pp. 35–62, 2006.
2. D. W. Townsend, "Multimodality Imaging of Structure and Function," Physics in Medicine and Biology, vol. 53, no. 4, pp. R1–R39, 2008.
3. E. Haber and D. W. Oldenburg, "Joint Inversion: A Structural Approach," Inverse Problems, vol. 13, pp. 63–77, 1997.
4. L. A. Gallardo and M. A. Meju, "Characterization of Heterogeneous Near-Surface Materials by Joint 2D Inversion of DC Resistivity and Seismic Data," Geophysical Research Letters, vol. 30, no. 13, p. 1658, 2003.
5. B. Bilgic, V. K. Goyal, and E. Adalsteinsson, "Multi-Contrast Reconstruction with Bayesian Compressed Sensing," Magnetic Resonance in Medicine, vol. 66, no. 6, pp. 1601–15, 2011.
6. F. Knoll, T. Koesters, R. Otazo, K. T. Block, L. Feng, K. Vunckx, D. Faul, J. Nuyts, F. Boada, and D. K. Sodickson, "Joint Reconstruction of Simultaneously acquired MR-PET Data with Multi Sensor Compressed Sensing based on a Joint Sparsity Jonstraint," EJNMMI Physics, vol. 1, 2014.
7. M. J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B. F. Hutton, and S. R. Arridge, "Joint Reconstruction of PET-MRI by Exploiting Structural Similarity," Inverse Problems, 2014.
8. J. Matthias Ehrhardt; Kris Thielemans; Luis Pizarro; Pawel Markiewicz; David Atkinson; Sébastien Ourselin; Brian F. Hutton; Simon R. Arridge "Joint reconstruction of PET-MRI by parallel level sets". 2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) 8-15 Nov. 2014
9. J. Gaa, , Rummeny, E. J. , Seemann, M. D. "Whole-body Imaging with PET/MRI". Eur. J. Med. Res. 9, 309–312 (2004).
10. L. Blomqvist, Torkzad, M. R. "Whole-body Imaging with MRI or PET/CT: The Future for Single-modality Imaging in Oncology?" JAMA 290, 3248–3249 (2003).
11. B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: Color Filter Array Interpolation," IEEE Signal Processing Magazine, pp. 44–54, 2005.
12. M. J. Ehrhardt and S. R. Arridge, "Vector-Valued Image Processing by Parallel Level Sets," IEEE Transactions on Image Processing, vol. 23, no. 1, pp. 9–18, 2014.
13. M. Lustig, D. L. Donoho, J. M. Pauly, "Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging," Magnetic Resonance in Medicine, vol. 58, no. 6, pp. 1182–1195, 2007.

Integrating Efficient Machine Learning Studies with Fluoroscopy

Soumya Suresh, Sreerag M V, P Prarthana, Sivakumar R

Dept. Sensor and Biomedical Technology, Vellore Institute of Technology, Vellore, India
rsivakumar@vit.ac.in

Abstract. This paper reviews current trends in research on fluoroscopy. It traces the integration of Machine Learning into the fluoroscopic imaging technique for analysis and diagnosis. The general concept of fluoroscopy has always been a mode of interacting with X-rays to perform tasks live on the body. A video fluoroscopic record is essentially a video which is used for diagnosis of the target location being analyzed by the doctor. The generic method used is more resilient in using the data for live treatment of patients, as Fluoroscopy gives you a moving X-ray image of the body. The problem statement here is to make use of the video efficiently to view the multiple frames it consists, for further comparison and analysis to the reference data to diagnose the issue accordingly. Our research signifies that, by using fluoroscopic data and converting them into frames and using that data to run machine learning on it and to identify various problems with that data alone. Hence, Fluoroscopy could be used for more detailed disease analysis and treatments.

Keywords. Fluoroscopy, Machine learning, Reference Datasets, X-Ray frames, Imaging Diagnosis

1. Introduction

The aim is to develop a drastically different approach to processing generic fluoroscopic data. The generic method was more resilient in using the data for live treatment of patients, as Fluoroscopy gives you a moving X-ray image of the body. Then these data gets stored in video formats and are later used for reference. Even in the affected parts, X rays are preferred by the doctors for eye to eye verification, as it is easier for the human eye to understand still photos. Furthermore ,it becomes extremely hard for the later use of Fluoroscopic data for research and education as we will need more complex machine learning algorithms.

Our research tested this by performing analysis on normal and Pneumonic lung X-rays. This application can be extended to this model and hence Fluoroscopy could be used for detailed disease analysis and treatments. Different data sets were put through trials for the analysis. A shoulder bone Fluoroscopic video was used to initiate. Then later , the original test data set, that is the Pneumonia test data set for analysis was introduced .

Production of a detailed output for the level of similarity with the affected data set of the Pneumonia patient ,should be possible. This research focuses on obtaining the fluoroscopic data and making it directly usable by the doctor by converting it into multiple x ray images and finding out the frame with the highest percentage defect from all different angles, and running it through similar Fluoroscopic data of an infected patient and getting a similarity report. So that the presence of a defect can be directly diagnosed. Thereby the image with the highest resemblance to the defective image can be directly used by the doctor for diagnosis.

2. Defining Integration of Fluoroscopic frames as a medium of Diagnosis

Fluoroscopy has always been the live monitoring device to perform complex procedures with the vision of what's inside the body. But now that Fluoroscopy is split into multiple frames of X-rays, it can now be used for external reference by the doctor.

Multiple angles of the fluoroscopic data could be used to provide an advanced X-ray frame that can not only diagnose the problem from one specific angle but from multiple angles. With this Fluoroscopy can now be used in a variety of needs from the classic monitoring while the procedure is done, to using it for basic diagnosis of the problem as a Fluoroscopic frame would provide far more definition to the problem compared to the conventional X-ray method.

Conventional X-ray would still have to be used wherever it has been previously used, but in treatments that required Fluoroscopy, could now also have a detailed Fluoroscopy frame, that can elaborate and be used to diagnose the problem. It could even be used for further research.

3. Defining Integration of Fluoroscopic frames into Research

Fluoroscopy can now take a turn from being a medical instrument to also help in building research material. Saved Fluoroscopic video files would be extremely bulky as well as not so easy to handle when stored at extreme large quantities. Using these Fluoroscopic data to generate frames, and using the frames with the greatest precision to generate an advanced Fluoroscopic frame would enable easier storage for a medical document that can exactly pinpoint the problem and could be used for diagnosis. Now when all the Fluoroscopic data is passed through this algorithm to obtain the Fluoroscopic frames, these frames could be used as the primary reservoir for X-ray and Fluoroscopic research.

Patients from across the globe, and their medical Fluoroscopic data being stored and pushed through ML algorithms could facilitate some high quality research productivity, compared to the current stack of X-ray images that dwell with less data and inaccessible and extremely difficult to process data of Fluoroscopic videos.

This could serve as a breakthrough for the research in open ended medical procedures where there needs to be visibility into the body in order to diagnose. With higher quality and lesser storage spaces documents get stacked up and serve as the primary files for analysis. There is a higher probability of discovering newer strategies in this field.

4. Data storage complexities and Fluoroscopic frames

Storage of data is a primary concern whenever we deal with any kind of medical procedure. Very few devices have live monitoring facilities, and even fewer have video storage for diagnosis. Fluoroscopic data has been video driven since the beginning and because of this it is generally not used or preferred over X-rays for general diagnosis. Fluoroscopic frames would enable the information provided by the video to be juiced into a single frame using the frames that had the highest precision.

This would enable Fluoroscopy to now be used in general diagnostics as well as use the same data over and over physically by doctors, or stored and fed into an ML algorithm that could make the research in this field way advanced to it is now.

5. Prospects with other diseases

With the advancement in medical science there could be machine learning algorithms that could detect a lot of diseases solely based on their Fluoroscopic frames and not just constricted to Pneumonia or any particular disease. This application can be extended to this model and hence Fluoroscopy could be used for detailed disease analysis and treatments. Different data sets were put through trials for the analysis. A shoulder bone Fluoroscopic video was used to initiate. Then later , the original test data set, that is the Pneumonia test data set for analysis was introduced. If most disease diagnosis could be made possible with just the mere fluoroscopic frame, then a lot of the current medical practices of Fluoroscopy could be avoided and the diagnosis and treatment could be provided without much hassle.

6. Conclusion

This algorithm can be replicated for various other diseases,or open sourced on a platform that can detect diseases from videofluoroscopic records or x-ray datasets.Rather than just identifying the frames that are affected a detailed ,specific report on the percentage affected along with other parameters with input dataset can be generated as well. Fluoroscopic data sets can now be stored as frames for easier observation and manipulation. The videofluoroscopic data is being utilized as X-ray data frames and imported into various machine learning algorithms for identifying diseases.This also easens the medical storage of such data for manual analysis etc.

References

1. Mujtaba SF, Saghir T, Sial JA, Rizvi NH. Procedural determinants of fluoroscopy time in patients undergoing cardiac catheterization. Pak J Med Sci. 2019 Jan-Feb;35(1):166-171. doi: 10.12669/pjms.35.1.294. PMID: 30881417; PMCID: PMC6408651.
2. Yevich S, Odisio BC, Sheth R, Tselikas L, de Baère T, Deschamps F. Integrated CT-Fluoroscopy Equipment: Improving the Interventional Radiology Approach and Patient Experience for Treatment of Musculoskeletal Malignancies. Semin Intervent Radiol. 2018 Oct;35(4):229-237. doi: 10.1055/s-0038-1669962. Epub 2018 Nov 5. PMID: 30402005; PMCID: PMC6218250.
3. Yin Y, Wang Y, Wang Z, Qu W, Tian W, Chen S. Accuracy of fluoroscopic examination in the treatment of Bennett's fracture. BMC Musculoskelet Disord. 2021 Jan 4;22(1):3. doi: 10.1186/s12891-020-03867-1. PMID: 33397336; PMCID: PMC7783999. 7
4. Iaquinto JM, Kindig MW, Haynor DR, Vu Q, Pepin N, Tsai R, Sangeorzan BJ, Ledoux WR. Model-based tracking of the bones of the foot: A biplane fluoroscopy validation study. Comput Biol Med. 2018 Jan 1;92:118-127. doi: 10.1016/j.compbiomed.2017.11.006. Epub 2017 Nov 8.
5. Banks SA, Hodge WA. Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. IEEE Trans Biomed Eng. 1996; 43(6):638-49. doi: 10.1109/10.495283 . PMID: 8987268
6. Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G. Image analysis and machine learning for detecting malaria. Transl Res. 2018 Apr;194:36-55. doi: 10.1016/j.trsl.2017.12.004..
7. Berg, S., Kutra, D., Kroeger, T. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat Methods* 16, 1226–1232 (2019). https://doi.org/10.1038/s41592-019-0582-9
8. Marleen de Bruijne,Machine learning approaches in medical image analysis: From detection to diagnosis, Medical Image Analysis,Volume 33,2016,Pages 94-97,ISSN 1361-8415,https://doi.org/10.1016/j.media.2016.06.032.
9. Ge Wang, G. Schweiger, M. W. Vannier, "An iterative algorithm for X-ray CT fluoroscopy," IEEE Transactions on Medical Imaging, vol. 17, no. 5, pp. 853-856, Oct. 1998, doi: 10.1109/42.736058
10. Sakata, Y., Hirai, R., Kobuna, K., Tanizawa, A., Mori, S. (2020). A machine learning-based real-time tumor tracking system for fluoroscopic gating of lung radiotherapy. Physics in Medicine & Biology, 65(8), 085014. doi:10.1088/1361-6560/ab79c5

Applications of Deep Learning in Mammography

S. Abhiramalakshmi, Mayukh Mittal, Sivakumar R

Uschool of Electronics Engineering, Vellore Institute of Technology, Tamilnadu, India
abhiramalakshmi.s2019@vitstudent.ac.in, mayukh.mittal2019@vitstudent.ac.in, rsivakumar@vit.ac.in

Abstract. Over the last decades, Computer Aided Detection (CAD) structures have been enhanced to aid radiologists to examine screening mammograms, but blessings of present day CAD technology look contradictory, consequently they have to be advanced to be in the long run taken into consideration useful. Beginning around 2012, profound convolutional neural organizations (CNN) have been an unprecedented accomplishment in photo recognition, accomplishing human by and large performance. These procedures have extraordinarily outperformed the traditional methodologies, that are similar as by and by utilized CAD arrangements. Deep CNN-s have the capacity to upset scientific photograph investigation. We underwrite a CAD gadget principally dependent on one of the greatest achievement thing discovery systems, Faster R-CNN. The gadget identifies and groups dangerous or harmless injuries on a mammogram with no human intercession. The proposed approach units the country of the fine art class by and large execution on the overall population INbreast information base, AUC = 0.95. The strategy characterized here has done every second locale with inside the Digital Mammography DREAM Challenge with AUC = 0.85. When utilized as a finder, the gadget arrives at over the top awareness with a couple of phony enormous imprints reliable with the photograph at the INbreast dataset

Keywords. Deep learning, Neural Networks, Mammography, CAD Solutions, Breast Cancer

1. Introduction

Mammography alludes to the most common way of using low energy X-Rays to investigate the bosom and mammary tissues for diagnosis and screening. It is basically used to identify growths and some other irregularities in the bosom. A mammogram is a X-Ray picture of the breast. The primary point of mammography is to recognize breast cancer right on time by identifying discernable masses or miniature calcifications.

Miniature calcifications are little stores of calcium salts dissipated or bunched all through the mammary organs. These are too little to even think about being felt yet can be recognized through imaging. Presence of miniature calcifications can demonstrate the early indications of bosom malignant growth.

2. Benefits of Deep Learning in Mammography

These discoveries show that programmed deep learning techniques can be promptly prepared to accomplish high exactness on heterogeneous mammography stages, and hold gigantic guarantee for working on clinical devices to decrease false positive and false negative screening mammography results. The neural organizations are prepared and are profoundly precise in their discoveries and can be utilized for programmed location and grouping of cancers whether they are harmless or threatening.

3. Current models of neural networks in use for mammography

The current models employed in the detection and classification of micro calcifications are:

1. Precision and recall model:

In simple terms, precision is ratio of true positives to all positives (true and false). That means it helps understand as to what ratio of all positive cases detected were actually positive. It basically gives an idea of how likely a positive result would actually be true. E.g. If the precision score of the ML is 89% it means that if tested positive, it is 89% chance it is true.

Mathematically,

$$\text{Precision} = \frac{TruePositive(TP)}{Truepositive(TP)+FalsePositive(FP)}$$

Recall is the measure of correctly identifying true positives. Recall also gives a measure of how accurately our model is able to identify the relevant data. We refer to it as Sensitivity or True Positive Rate.

Mathematically,

$$\text{Recall} = \frac{TruePositive(TP)}{Truepositive(TP)+FalseNegative(FN)}$$

2. F-1 Score: F-1 Score is the harmonic mean of the precision and recall of the given data set. It gives us an idea about the balance between the precision and recall of the model used.

4. Model used and working of the code

Support Vector Machine (SVM) model:

Support Vector Machine or SVM is one of the most famous Supervised Learning calculations, which is utilized for Classification just as Regression issues. In any case, basically, it is utilized for Classification issues in Machine Learning.

The objective of the SVM calculation is to make the best line or choice limit that can isolate n-layered space into classes so we can undoubtedly put the new element in the right classification later on. This best choice limit is known as a hyperplane.

SVM picks the outrageous focuses/vectors that assistance in making the hyperplane. These outrageous cases are called as help vectors, and henceforth calculation is named as Support Vector Machine.

The reason we chose this method over others is due to following advantages in our code:

1. SVM works moderately well when a clear margin of separation between classes is present. Since in our dataset, there are only 2 distinct classes, the model is most efficient.
2. SVM is more efficient in high dimensional spaces.
3. SVM is potent in cases where the number of dimensions is more than the number of samples.
4. SVM is relatively memory efficient. Hence it can be utilised on devices with lower end specifications and still run as efficiently.

Code: The team used the Google Colab platform for the ideation and execution of the code. Due to lockdown restrictions this made it easy for the team members to track changes as well as work in tandem with each other.

Number of patients recorded: 569

Number of characteristics measured: 33

Importing data set from .csv file, giving information about shape, texture, smoothness etc. for the calcification.

Classification whether the given calcification is benign or not.

Benign: 357

Malignant: 212

Graphical analysis of tumor types and distributions

Heat map showing correlation of all data points

Splitting the dataset into 75% training and 25% testing

Applying various models for analysis and classification of data

5. Observations

```
#Get the new count of number of rows and columns
df.shape
```

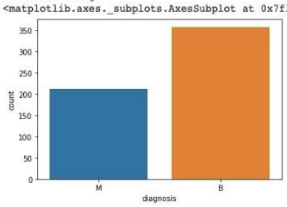
```
(569, 32)
```

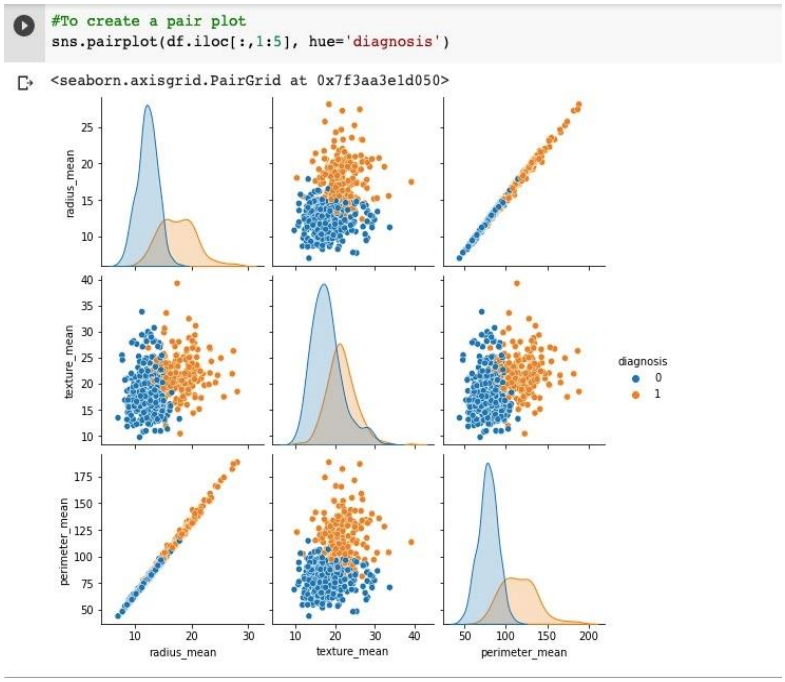
```
#Get a count of the number of Malignant (M) or Benign (B) cells
df['diagnosis'].value_counts()
```

```
B    357
M    212
Name: diagnosis, dtype: int64
```

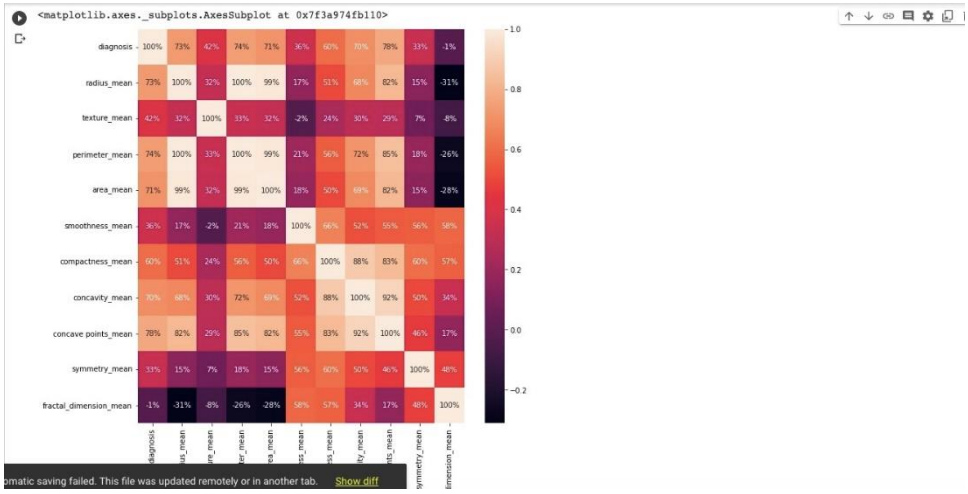
```
#Visualise the count
sns.countplot(df['diagnosis'], label='count')
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only
  FutureWarning
<matplotlib.axes._subplots.AxesSubplot at 0x7f3aa66ca2d0>
```





6. Inferences



```
#To show another way to get metrics of the models
from sklearn.metrics import classification_report
from sklearn.metrics import accuracy_score

for i in range(len(model)):
  print('Model', i)
  print(classification_report(Y_test, model[i].predict(X_test)))
  print(accuracy_score(Y_test, model[i].predict(X_test)))
  print()
```

```
Model 0
              precision    recall  f1-score   support

           0       0.97      0.96      0.96        90
           1       0.93      0.94      0.93        53

    accuracy                           0.95       143
   macro avg       0.95      0.95      0.95       143
weighted avg       0.95      0.95      0.95       143

0.951048951048951

Model 1
              precision    recall  f1-score   support

           0       0.98      0.92      0.95        90
           1       0.88      0.96      0.92        53

    accuracy                           0.94       143
   macro avg       0.93      0.94      0.93       143
weighted avg       0.94      0.94      0.94       143

0.9370629370629371
```

```
Model 2
              precision    recall  f1-score   support

           0       0.98      0.97      0.97        90
           1       0.94      0.96      0.95        53

    accuracy                           0.97       143
   macro avg       0.96      0.96      0.96       143
weighted avg       0.97      0.97      0.97       143

0.965034965034965

Model 3
              precision    recall  f1-score   support

           0       0.97      0.98      0.97        90
           1       0.96      0.94      0.95        53

    accuracy                           0.97       143
   macro avg       0.96      0.96      0.96       143
weighted avg       0.96      0.97      0.96       143

0.965034965034965
```

As seen in the above result table, for the various models tested our model tests on par or better with the current existing models. With more effort put into this and a larger data set we aim to make this system more efficient and reliable.

7. Conclusion

It was observed that the F1 score gave 1-2% better results compared to conventional models, other models like SVM model also gave similar or infinitesimally better results compared to the base model in use. This system although efficient cannot eliminate the involvement of doctors or other support staff but can definitely help in aiding them for better diagnoses and therefore efficient and quick treatment

References

1. Ribli, D., Horváth, A., Unger, Z. *et al.* Detecting and classifying lesions in mammograms with Deep Learning. *Sci Rep* **8,** 4165 (2018). https://doi.org/10.1038/s41598-018-22437-z
2. Hongyu Wang, Jun Feng, Qirong Bu, Feihong Liu, Min Zhang, Yu Ren, Yi Lv, "Breast Mass Detection in Digital Mammogram Based on Gestalt Psychology", *Journal of Healthcare Engineering*, vol. 2018, Article ID 4015613, 13 pages, 2018. https://doi.org/10.1155/2018/4015613
3. Adam Yala, Tal Schuster, Randy Miles, Regina Barzilay, Constance Lehman, " A Deep Learning Model to Triage Screening Mammograms: A Simulation Study" From the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Mass (A.Y., T.S., R.B.); and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, 55 Fruit St, WAC 240, Boston, Mass 02114-2698 (R.M., C.L.). June 18, 2019.
4. Alejandro Rodríguez-Ruiz, Elizabeth Krupinski, Jan-Jurre Mordang, Kathy Schilling, Sylvia H. Heywang-Köbrunner, Ioannis Sechopoulos, Ritse M. Mann, " Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System " From the Department of Radiology and Nuclear Medicine, Radboud University Medical Center, HB Nijmegen, Geert Grooteplein 10, 6525 GA, Nijmegen, the Netherlands (A.R.R., I.S., R.M.M.); Department of Radiology & Imaging Sciences, Emory University, Atlanta, Ga (E.K.); ScreenPoint Medical BV, Nijmegen, the Netherlands (J.J.M.); Lynn Women's Health & Wellness Institute, Boca Raton Regional Hospital, Boca Raton, Fla (K.S.); Referenzzentrum Mammographie Munich, Brustdiagnostik München and FFB, Munich, Germany (S.H.H.); and Dutch Expert Centre for Screening, Nijmegen, the Netherlands (I.S.). September 28, 2018.
5. Laurie R. Margolies, Rowena Yip, Esther Hwang, Rebecca H. Oudsema, Varun R. Subramaniam, Harvey Hecht, and Jagat Narula, " Breast Arterial Calcification in the Mammogram Report: The Patient Perspective " American Journal of Roentgenology 2019 212:1, 209-214
6. M., Samir & A., Alaa & E., Hassan & Gadallah, Mohammed. (2017). Breast Cancer Detection with Mammogram Segmentation: A Qualitative Study. International Journal of Advanced Computer Science and Applications. 8. 117 – 120. 10.14569/IJACSA.2017.081016.
7. Hariraj, Venketkumar & Khairunizam, Wan & Vijean, Vikneswaran & Ibrahim, Zunaidi & Abu bakar, Shahriman & Mohamad Razlan, Zuradzman & Thavasimuthu, Rajendran & Sathiyasheelan, R.. (2018). Fuzzy multi-layer SVM classification of breast cancer mammogram images. International Journal of Mechanical Engineering and Technology. 9. 1281-1299. 14. https://www.hindawi.com/journals/bmri/2017/3640901/
8. Daniela Bernardi, Petra Macaskill, Marco Pellegrini, Marvi Valentini, Carmine Fantò, Livio Ostillio, Paolina Tuttobene, Andrea Luparia, Nehmat Houssami, "Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study, The Lancet Oncology, Volume 17, Issue 8, 2016, Pages 1105-1113, ISSN 1470-2045, https://doi.org/10.1016/S1470-2045(16)30101-2. (https://www.sciencedirect.com/science/article/pii/S1470204516301012)
9. T. Kretz, K. -R. Müller, T. Schaeffter and C. Elster, "Mammography Image Quality Assurance Using Deep Learning," in *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 12, pp. 3317-3326, Dec. 2020, doi: 10.1109/TBME.2020.2983539.
10. M. Bustamante *et al*., "Algorithm for Processing Mammography: Detection of Microcalcifications," in *IEEE Latin America Transactions*, vol. 16, no. 9, pp. 2460-2466, September 2018, doi: 10.1109/TLA.2018.8789569.

Brain Tumour Detection and Segmentation Techniques

Desta Yakob, Andualem Atnafu, Sivakumar R

School of Electronics Engineering, Vellore Institute of Technology, Vellore, India.
destayakob.doda2018@vitstudent.ac.in
andualematnafu.mekonen2018@vitstudent.ac.in,rsivakumar@vit.ac.in

Abstract. A Brain Tumor is an inappropriate growth of tissues within the skull. The tumour starts growing inside the skull and interferes with the normal functioning of the brain. It needs to be detected at an early stage using MRI or CT scanned images when it is as small as possible because the tumour can result in cancer. Image processing is highly required in the medical Application of brain tumour detection after MRI. This paper aims to dig out the best and more accurate Brain Tumor detecting Technique from MRI images. First, we introduce Brain tumours and their class, next we will implement and analyze different Brain tumour detection and segmentation techniques. In this presentation Segmentation using anisotropic, SVM Based Classification, FCM and Level Set Algorithms will be seen in detail with comparison and we will select the best technique from them.

Keywords. Brain tumour, MRI images, Segmentation, Support vector machine(SVM).

1. Introduction

One of the global public health problems is cancer. According to the report by World Cancer Researcher, the primary cause of death is cancer. The usage of cell phones at a high rate is one reason for the incidence of brain tumours in the young generation. The brain tumour is a result of abnormal cell formation in the brain. The main types of tumours are malignant (fast-growing) and benign (slow-growing) tumours. Primary brain tumours (malignant) affect the surrounding tissues. Secondary brain tumours are spread to other parts of the body from the brain. According to the location and type of tissue used for detecting brain tumours are classified. Around 120 types of tumours are identified and classified by World Health Organization. Early detection of tumours probably can increase the life span of a person. In diagnosing brain tumours imaging plays an important role. Diagnosis of brain tumours can be done with CT (computer tomography) and MRI (Magnetic resonance image) scans.

Classification of brain Tumour Based on Origin of Tumour:

- Primary Tumour: the origin of tumour cell is the brain itself
- Secondary Tumour: known as a metastatic brain tumour, the origin of tumour cells is another part of the body where it finally reaches the brain (from lung, breast). It is the most common type of brain tumour.

Based on Malignancy:

Benign Tumour: Non-cancerous that are normal in size and growth they don't generally invade the other part of the body.

Malignant Tumour: Cancerous tumour cell. Abnormal in size and growth tend to invade neighbouring healthy cells.

Based on appearance: Tumour cells are graded during the period of treatment. The graded measurement gives information about the rate of growth of Tumours.

Table 1: Tumour types based on appearance and growth

Grade	Appearance	Growth Rate
1	Nearly normal	slow
2	Slightly abnormal	slow
3	Abnormal	Active growth
4	Most abnormal	Quick growth

Current trends in brain tumour detection and detection techniques:

Applied techniques for brain Tumour detection and classification

- Segmentation using anisotropic filtration method
- SVM Based Classification of BrainMRI Images using DWT and PCA Technique
- Modified FCM and Level Set Algorithms

2. Brain Tumour detection based on Segmentation using anisotropic filtration method

Anisotropic filtering (abbreviated AF). The anisotropic filter improves the image quality of textures on surfaces.

Proposed methodology

The proposed methodology consists of six stages i.e. pre-processing, filtration, edge detection, tumour outline, segmentation tumour detection. Preprocessing stage involves converting the original image into a grayscale image we are using the uigetfile () command to browse the "jpg" image from the computer and then we used the imread command to read the input image and in the filtering part we are removes the noise if present or crept in.

This is followed by edge detection using Sobel, Prewitt and Cannyalgorithms with image enhancement techniques. Next, segmentation is applied to display the tumour affected region in the MRI images. Finally, the image is clustered using the k-means algorithm. Here we have used MATLAB and an anisotropic filter for the development of the paper.

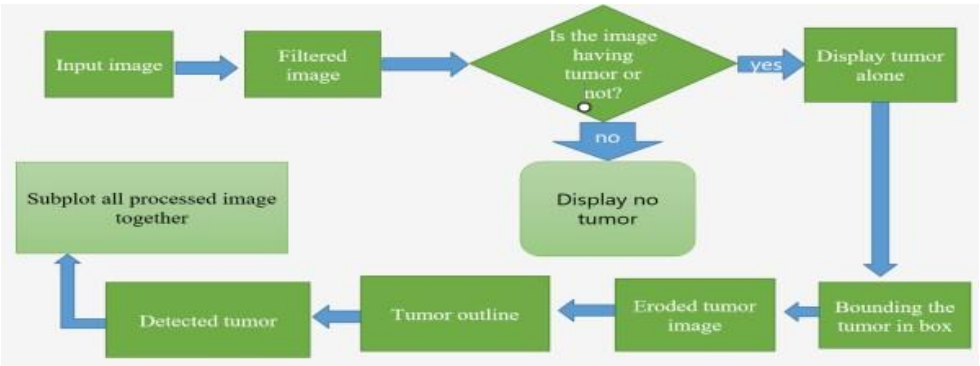


Fig. 1. Flow chart for anisotropic filtering methods

Anisotropic filtering Technique implementation using MATLAB:

A- Input as brain Image having a tumour

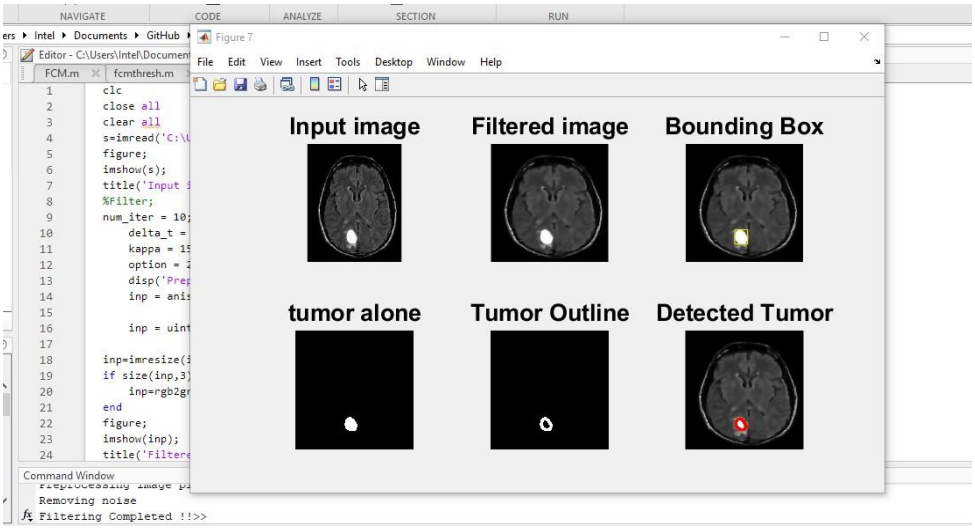


Fig. 4. Detected brain tumour output images using an anisotropic technique

B- Input as normal magnetic resonance brain image:

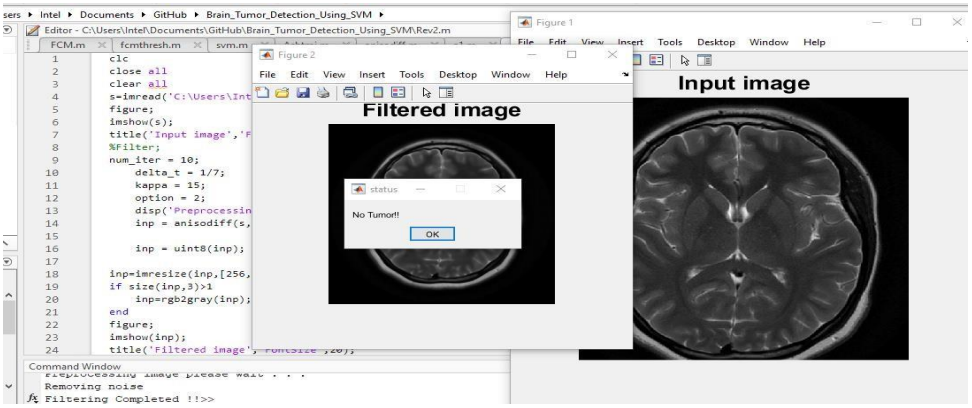


Fig. 3. No tumour output images using an anisotropic technique

Limitation:

In this technique, we are unable to identify the type of tumour. It takes much time to identify the tumour area since it's a little bit not faster. Sometimes the tumour area is not clear for filtration and edge detection Getting no tumour results since the accuracy is poor. Edge detection is also not accurate sometimes it is bound beyond the tumour area.

3. Support Vector Machine(SVM) based Classification of Brain MRI Images using DWT and PCA to Detect Different Types of Brain Tumour

Proposed methodology

The discrete wavelet transform(DWT) technique is used to extract the features in the MR Images. Over and above that the Reduction of the MR image is done by Principal Components Analysis (PCA) method. Once the features are completely extracted then the classification methods (SVM) is run to classify which type of Tumor is either benign or malignant. Input data set is provided for both

• Principal Components Analysis (PCA): It is used to reduce the data dimensions of the given input image. Such reduction is effective in quite a few cases: for image compression, data representation, calculation reduction necessary in subsequent processing.

• The discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translations obeying some defined rules. In other words, this transform decomposes the signal into a mutually orthogonal set of wavelets, which is the main difference from the continuous wavelet transform (CWT), or its implementation for the discrete-time series sometimes called discrete-time continuous wavelet transform (DT-CWT)

• Support Vector Machine (SVM): It is one of the most commonly used machine learning algorithms. It is mainly used in classification problems. As the flow chart indicated in fig.6, the algorithm takes an image as an input then the preprocessing of the image will be done by filters. After that feature extraction process will take place by using the Discrete Wavelet Transform(DWT). Finally, the classification process will be determined by a support vector machine(SVM).

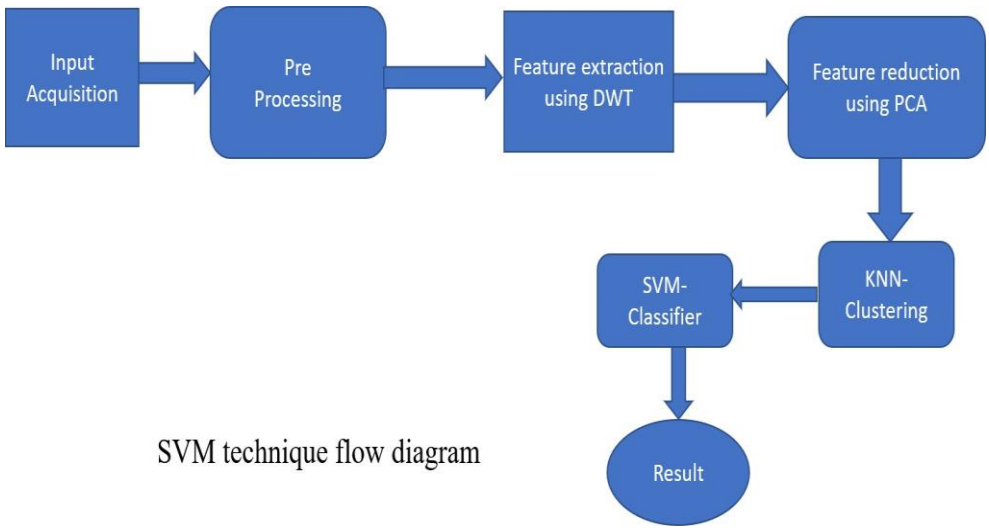


Fig. 4. SVM technique flow diagram

Implementation of SVM Based Classification using DWT and PCA:

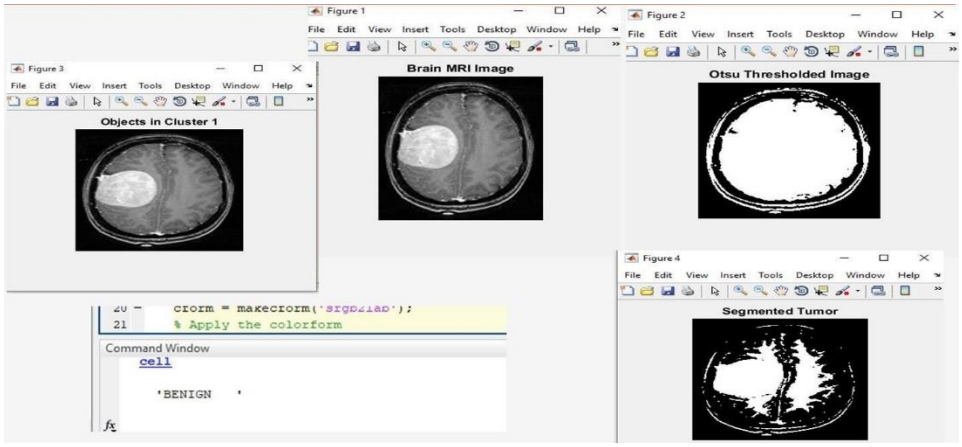


Fig. 5. Benign tumour class using SVM techniques

Case 2:

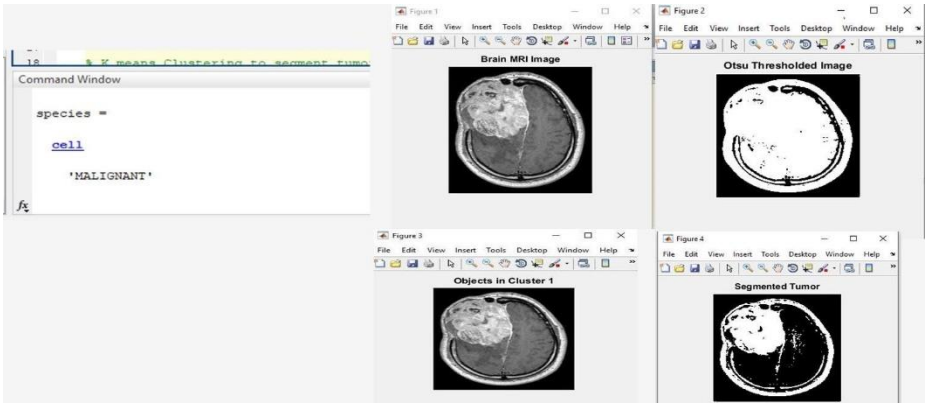


Fig. 6. Malignant tumour class using SVM techniques

4. Brain Tumour Detection by using a Modified FCM and Level Set Algorithms

Fuzzy c-means (FCM) is a data clustering technique in which a data set is grouped into N clusters with every data point in the dataset belonging to every cluster to a certain degree. For example, a data point that lies close to the centre of a cluster will have a high degree of membership in that cluster, and another data point that lies far away from the centre of a cluster will have a low degree of membership to that cluster.

FCM and Level set algorithms are used for MRI image segmentation purposes. To find the boundary of the target area and to simplify the MRI image without losing the important image features. FCM= Fuzzy C-means Algorithm: is used for Image clustering. It is a technique to obtaining more clusters/groups of images from the input image with different features. It gives membership values more than one different cluster. That makes it special from K-means clustering. Obtained by calculating the mean value from each data value.

Level Set Algorithm: used for fixing the boundary of tumour region, partitioning an Image into sub-regions with continuous boundaries.

FCM and Level set algorithms are used for MRI image segmentation purposes. To find the boundary of the target area and to simplify the MRI image without losing the important image features.

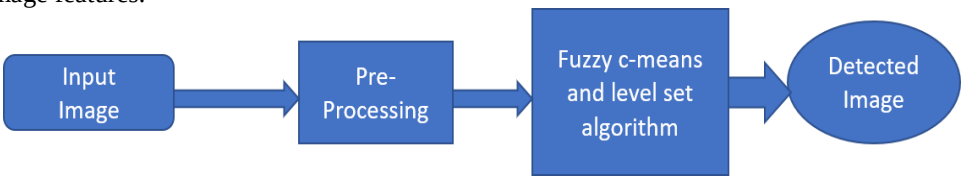


Fig. 7. Flow Chart of modified FCM and Level set algorithm method

The FCM function performs FCM clustering. The algorithm initially assumes the cluster centres; The location for each cluster will be given. Then, FCM assigns each statistics factor a random membership grade for every cluster. By iteratively updating the cluster facilities and the membership grades for every statistics point, FCM strikes the cluster facilities to the right area inside a statistics set and, for every information point, finds the diploma of membership in every cluster. This generation minimizes a goal feature that represents the distance from any given information factor to a cluster core weighted by way of the membership of that statistics factor in the cluster.

Fig. 8. FCM tumour detection sample diagram

Tumour detection using FCM

- Start from initial large curve to encircle given vision of interest
- The couture moves towards the centre
- When the couture faces sudden changes in intensity, the circle stops moving inside.
- Finally, region of interest (tumour area) will be detected

Limitations:

- Long iteration time if the image with The complete Iteration takes more than 2 minutes to get the image output.
- It can not exactly bound the region of Interest
- It can not differentiate the tumour type

MATLAB Implementation:

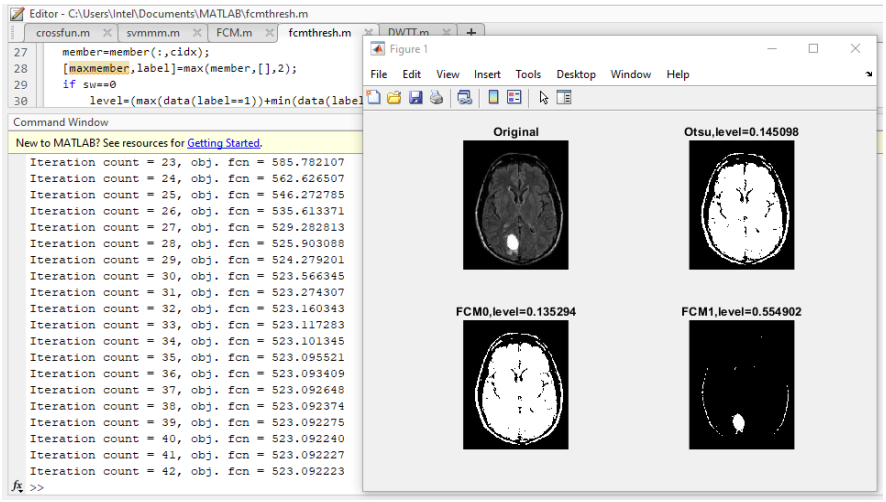


Fig. 9. Tumour class using FCM techniques

5. Conclusion

In this paper, from all the three techniques of brain tumour detection, we have identified the better techniques by comparing all of them. So SVM based along with DWT and PCA is the better technique than the remaining two techniques of brain tumour detection and classification. The SVM method works relatively well when there is a clear margin of separation between classes. And it is more effective in high dimensional spaces. SVM techniques are effective in detecting and classifying the type of brain tumour. SVM is relatively memory efficient and takes less time to display the result. Since using this technique we can able to detect

References

1. Introduction to the Basics of Magnetic Resonance Imaging by Klaus Mo¨ llenhoff, Ana- Maria Oros-Peusquens, and N. Jon Shah published in 2012
2. A Novel methodology for brain tumour detection based on two-stage segmentation of MRI images by Anjali Joshi, V. Charan, Sr. Member, IEEE Shanthi Prince Dept. of ECE, SRM University, year of publication 2015
3. Application of FCM Clustering on AM-FM to Detect MRI Disease Progression For Multiple Sclerosis by S.P. Washimkar Electronics & Telecommunication Engg. And S.D.Chede Electronics & Telecommunication Engg. Deptt, Rajiv Gandhi College of Engg.& Research Nagpur, year of publication 2015
4. Brain Tumor Detection Based on Segmentation using MATLAB by Animesh Hazra, Ankit Dey, Sujit Kumar Gupta, Md. Abid Ansari Assistant Professor, Dept. of CSE, Jalpaiguri Govt. Engg. College, Jalpaiguri, West Bengal, India, year of publication 2017
5. Detection of Brain Tumor in MRI Images, using Combination of Fuzzy C-Means and SVMby Parveen Dept. of computer science and engineering Govt. women engineering college Ajmer, India and Amritpal Singh Dept. of computer science and engineering Govt. women engineering college Ajmer, India year of publication 2015
6. MRI Image Distortion due to Magnetic Materials in Medical Implants by Dick W. Harberts Philips Innovation Services Electromagnetics Competence Center Eindhoven, The Netherlands year of publication 2017

7. Segmentation of Brain MRI Image Based on Clustering Algorithm by Siti Noraini Sulaiman, Noreliani Awang Non, Iza Sazanita Isa, Norhazimi Hamzah, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM), 13500 Permatang Pauh Penang, Malaysia year of publication 2014
8. Brain Tumor Detection Using Artificial Neural Networks Eltaher Mohamed Hussein 1, Dalia Mahmoud Adam Mahmoud 2 1 Biomedical Engineering Department, Sudan University year of publication 20012
9. Brain Tumor Detection Using SVM Classifier T. Sathies Kumar K. Rashmi Sreevidhya Ramadoss AP/ EIE EIE EIE Sri Sai Ram Engineering College Chennai, India Chennai, India Chennai, India, year of publication 2017
10. Brain Tumor Detection Using SVM Classifier T. Sathies Kumar K. Rashmi Sreevidhya Ramadoss AP/ EIE EIE EIE Sri Sai Ram Engineering College Chennai, India Chennai, India Chennai, India, year of publication 2017
11. Brain Tumor Detection by using a Modified FCM and Level Set Algorithms Bouc if BEDDAD Communication Technology Laboratory T AHAR Moulay University andKaddour HACHEMI.

Breast Cancer Classification using Convolution Neural Network

Mansi Narade, Nitikaa Samundeeswari, Priyadharshini, Sivakumar R

Dept. of Sensor and Biomedical Technology, Vellore Institute of Technology, Vellore, India
mansi.sanjeevkumar2019@vitstudent.ac.in, samundeeswari.rn2019@vitstudent.ac.in
priyadharshini.r2019@vitstudent.ac.in, rsivakumar@vit.ac.in

Abstract Breast Cancer is one of the most frequent disease in our beings around the world. In 2012, it represented around 12% of all new tumors in ladies and 25% of all diseases. Breast cancer happens when the cells in the breast run out of control. These cells typically structure growths that are regularly seen on x-beams or felt as protuberances. Growths become harmful (disease) or otherwise malignant when cells develop (attack) into encompassing tissues or spread (metastasis) to far off region of the body A benign tumor is a mass of cells (growth) that comes up short on the capacity to either attack adjoining tissue or metastasize (spread all through the body). We have built an algorithm based on Convolution neural network (CNN) a form of artificial neural network mainly used for image recognition and processing, mainly dedicated for processing pixel data to automatically identify whether a patient is suffering from breast cancer (malignant or benign) or not by looking at biopsy images [1]. We used Google Collab platform for the implementation. We derived the dataset to disengage the patches of the picture from training the network and finally giving the shot as an input to analyze the picture Our Training folder has 1000 pictures in every category while the validation folder has 250 pictures in every classification. The input is a training dataset that comprises of N image , each marked with one of 2 unique classes. Then, at that point, we utilize this training dataset to train a classifier to realize and learn what all of the classes resembles. Eventually, we assess the nature of the classifier by requesting that it predicts labels for set of images that we have never seen before. We will then, at that point, analyze the genuine marks of these images to the ones anticipated by the classifier. We first loaded all the libraries and dependencies and then transferred JPG images to array images to create train and test datasets. Then we created labels and shuffled the datasets followed up by splitting of data into train and test dataset. Then the images are displayed and data is generated. Finally, the model is built and trained. We can observe the results in performance metrics in a graphical format.

Keywords. Malignant, Benign, Convolution neural network, Google Collab platform, training datasets

1. Introduction

Breast cancer is one of the most common cancer in human beings worldwide. In 2012, it represented around 12% of all new tumors in ladies and 25% of all diseases. Breast cancer happens when the cells in the breast run out of control. These cells typically structure growths that are regularly seen on x-beams or felt as protuberances. Growths become harmful (disease) or otherwise malignant when cells develop (attack) into encompassing tissues or spread (metastasis) to far off region of the body. Here are some simple facts:

- Around 1 of every 8 ladies in the United States (around 12%) foster intrusive breast cancer sooner or later in their lives.

 In 2019, an expected 268,600 new instances of intrusive breast cancer will be analyzed in US ladies, and 62,930 new instances of in situ breast cancer will be analyzed.

 Roughly 85% of Breast cancer's occur in ladies who have no family background of Breast cancer. These are not brought about by inherited transformations, yet by hereditary changes that happen because of the maturing system and general life.

 At the point when first-degree family members (mother, sister, little girl) are determined to have breast cancer, the danger of the disease in ladies nearly duplicates. Under 15% of ladies who foster bosom malignant growth have a family with Breast cancer.

2. Definition of Convolution Neural Network (CNN)

A convolutional neural network (CNN) is a type of artificial neural network used for image recognition and processing, specifically designed for processing pixel data.

The CNN is a powerful image processing, artificial intelligence (AI) that uses deep learning to perform both generate and write tasks. Often, we use machine vision, which includes image and video recognition [2].

A neural network is a system that looks like hardware or software. With the function of neurons in the human brain. Traditional neural networks are not ideal for image processing, and images need to be split and fed at lower resolutions. CNN has placed "neurons" like the frontal lobe, the area that processes visual stimuli in humans and other animals.

The layers of neurons are arranged to cover the entire field of view, avoiding the problems of piecewise image processing in traditional neural networks.

CNN (Convolution Neural Network) is a type of neural network model which allows us to extract higher representations for the image content. Unlike the classical image recognition where you define the image features yourself, CNN takes the image's raw pixel data, trains the model, then extracts the features automatically for better classification. CNN algorithm has two main processes are convolution and sampling [3].

3. Methodology

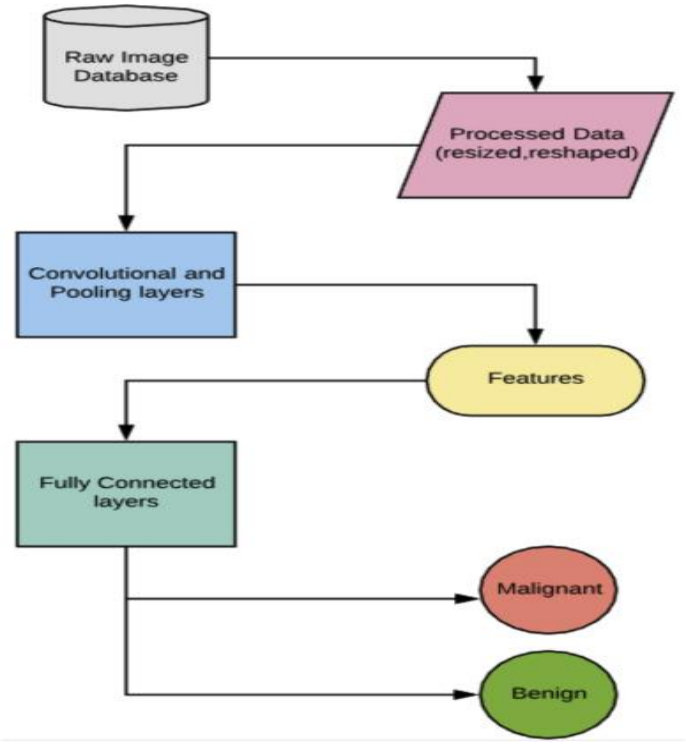
1. Initially to create a model to automatically find out whether a patient is suffering from breast cancer or not without looking at biopsy images. The algorithm had to be extremely precise because human being's lives is at stake.
2. We used Google Colab platform for the implementation.
3. Data is derived from dataset

Breast Cancer Histopathological Database (BreakHis) - Laboratório Visão Robótica e Imagem

4. Our Training folder has 1000 pictures in every category while the validation folder has 250 pictures in every classification.

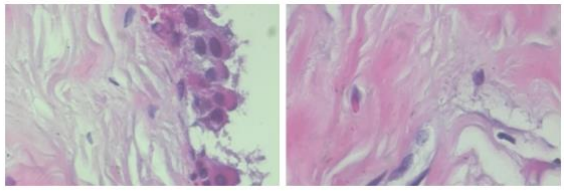
4. Algorithm and Flowchart

The input is a training dataset that comprises of N images , each marked with one of 2 unique classes. Then, at that point, we utilize this training dataset to train a classifier to realize and learn what all of the classes resembles. Eventually, we assess the nature of the classifier by requesting that it predicts labels for set of images that we have never seen before. We will then, at that point, analyze the genuine marks of these images to the ones anticipated by the classifier [4].

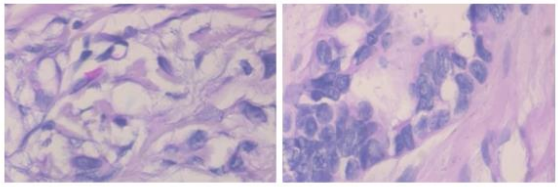


5. Output and Discussion

The training folder has 1000 images in each category while the validation folder has 250 images in each category.



Benign sample



Malignant sample

After loading the images to respective folders we created a NumPy array of zeroes for labeling benign images and similarly a NumPy array of ones for labeling malignant images. Shuffled the dataset and converted the labels into categorical format. Then we split the data-set into two sets — train and test sets with 80% and 20% images respectively [5].

We used a batch size value of 16. Batch size is one of the most important hyperparameters to tune in deep learning. We prefer to use a larger batch size to train my models as it allows computational speedups from the parallelism of GPUs. However, it is well known that too large of a batch size will lead to poor generalization. On the one extreme, using a batch equal to the entire dataset guarantees convergence to the global optima of the objective function.

However this is at the cost of slower convergence to that optima. On the other hand, using smaller batch sizes have been shown to have faster convergence to good results. This is intuitively explained by the fact that smaller batch sizes allow the model to start learning before having to see all the data. The downside of using a smaller batch size is that the model is not guaranteed to converge to the global optima. Therefore it is often advised that one starts at a small batch size reaping the benefits of faster training dynamics and steadily grows the batch size through training.

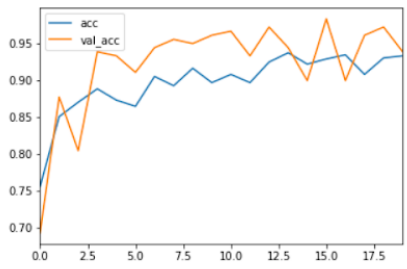
We used following CNN layers to create our model:

1. Global Average Pooling generates one feature map for each corresponding category of the classification i.e. Benign and Malignant. Feature map captures the result of applying filters to an input image.
2. Dropout layer to reduce the data by 50% to reduce the overfitting problem. Overfitting occurs when the model is too complex.
3. We used batch normalization and a dense layer with 2 neurons for 2 output classes which are benign and malignant and we used SoftMax as the activation function.
4. SoftMax transforms input values into values between 0 to 1 so that they can be interpreted as probabilities.
5. Adam as optimizer which maintains single learning rate and learning rate does not reduce during training which is helpful to reduce the losses.
6. Binary cross entropy as Loss function. Loss function compares the distance between current output and actual output. Binary cross entropy compares each of the predicted probabilities to the actual output which can be 0 or 1.

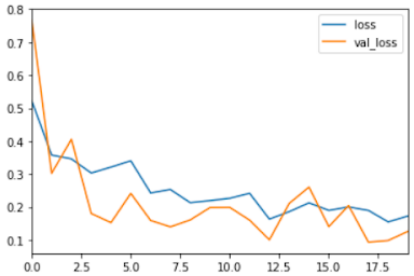
These layers will apply various filters and it will classify the image into Benign or Malignant [6].

5.1. Performance Metrics

The most common metric for evaluating model performance is Accuracy, Confusion Matrix , Precision Recall and F1 score , ROC AUC and Log loss. Accuracy is the simple ration between the number of correctly classified images to the total number of images. We trained the model for 20 epochs and as the number of epochs increases the accuracy increase and the loss decreases . The accuracy of this model is 98.3% with 2% loss of the data [4].



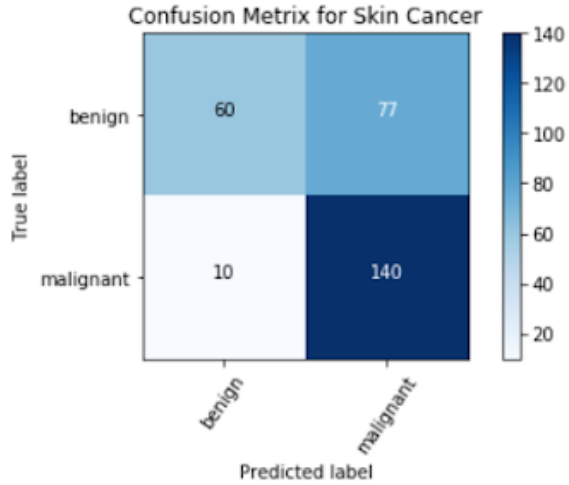
Accuracy



Loss

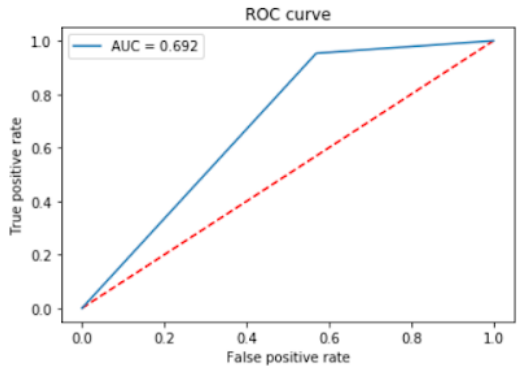
5.2. Confusion Matrix

Confusion matrix is used for analyzing any misclassifications in the model. Each row in the confusion matrix represents the instances of the predicted class. Each column in the matrix represents the instances of the actual class and the diagonals represent classes that have been correctly classified [7]. In this model 60 images of benign cancer and 140 images of malignant cancer of the training dataset have been correctly classified.

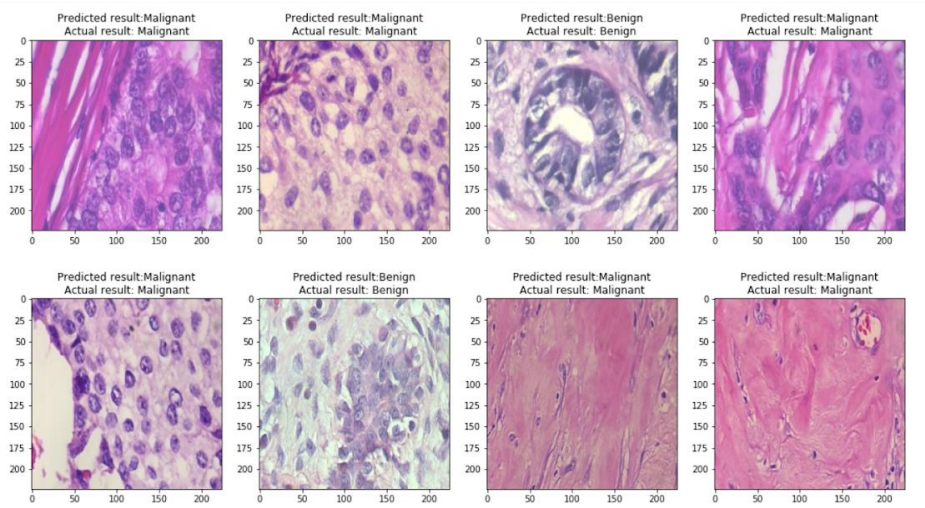


5.3. ROC Curves

ROC is Receiver operating characteristic curve. The 45 degree line is the random line, where the Area Under the Curve or AUC is 0.5 . The plot is between false positive rate versus the true positive rate [8]. Higher is the AUC better is the model. Highest a model can get a AUC is 1 . The ROC curve forms a Right angled triangle. ROC curves can identify the misclassifications and also can debug the model [9]. In this ROC curve bottom corner is closer to the random line, this implies that the model is misclassifying at y=0. Also at the top where y=1 shows that there is misclassification at y=1.



In the Prediction model we have used the TTA that is Test time augmentation on the training dataset [5]. This involves creating multiple augmented copies of each image in the test data set and model makes prediction for each of the image.The model was successfully able to predict the results of some of the biopsy images and it was compared to the actual results as show in the below figure.



7. Conclusion

Although this project is far from complete but it is remarkable to see the success of deep learning in such varied real-world problems. In this paper, we have demonstrated how to classify benign and malignant breast cancer from a collection of microscopic images using convolutional neural networks and transfer learning [10]. It classifies the type of cancer and easy magnification with Convolution Neural Network with a standard Accuracy of 93.45% with the train split of 0.2.

References

1. D. Bardou, K. Zhang and S. M. Ahmad, "Classification of breast cancer based on histology images using convolutional neural networks", IEEE Access, vol. 6, pp. 24680-24693, 2018.
2. N. Tajbakhsh et al., "Convolutional neural networks for medical image analysis: Full training or fine tuning?", IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1299-1312, May 2016.
3. Arevalo, J., Gonzalez, F., Ramos-Pollan, R., et al.: Representation learning for mammography mass lesion classification with convolutional neural networks. Comput. Methods Programs Biomed. 127, 248–257 (2016)
4. Aboutalib, S. S. et al. Deep Learning to Distinguish Recalled but Benign Mammography Images in Breast Cancer Screening. Clin. Cancer Res. (2018).
5. Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer P (2000) The Digital Database for Screening Mammography, in international workshop on digital mammography. In: Medical Physics
6. Jiamei Sun, Alexander Binder, "Comparison of deep learning architectures for H&E histopathology images", Big Data and Analytics (ICBDA) 2017 IEEE Conference on, pp. 43-48, 2017.
7. Dimou, I.N., Manikis, G.C., Zervakis, M.E., "Classifier fusion approaches for diagnostic cancer models," in 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 5334 –5337 (2006).
8. Hillis S. L., Obuchowski N. A., Schartz K. M., and Berbaum K. S., "A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette methods for receiver operating characteristic (ROC) data," *Stat. Med.* 24, 1579–1607 (2005).10.1002/sim.2024
9. Hillis S. L., "A comparison of denominator degrees of freedom methods for multiple observer ROC analysis," *Stat. Med.* 26, 596–619 (2007).10.1002/sim.2532
10. Sundarambal, B.; Mathana, J.M.; Subramanian, S.; Sandesh, H.; Omprakash, A. A. Detailed Investigation on Reduction of False Positive Rate in Breast Cancer Detection. IEEE Int. Conf. on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6-7 March 2020.

Bone Cancer/Tumor Detection from X-ray using Image Processing

Shubhanki Sengupta, Shishira P A, Dr. Mythili A

Vellore Institute of Technology, Vellore, India
shubhanki.sengupta2020@vitstudent.ac.in , shishira.pa2020@vitstudent.ac.in
mythili.asaithambi@vit.ac.in

Abstract. The bone cancer is the most threatening type of cancer and detection of bone cancer was the most challenging problem earlier. Bone tumor in primary stage often shows up with clueless indication which makes a delay in diagnosis and can be severe for the patient if not diagnosed on time. These indications are difficult to predict accurately. This paper aims in understanding different image processing methods used on bone X-ray images and analyze the results to come up with an appropriate algorithm. An algorithm in MATLAB has been proposed to foresee bone tumors/cancer based on the its dimensions and features from the x-ray image. Segmentation of the image is done using K-means clustering, which is most suited for bone tumor detection and compared with region growing and other edge-based segmentation like Prewitt, Sobel and Canny. Detection of the tumor is done using morphological operations like dilation and opening and extracting features like area, perimeter, eccentricity, mean, standard deviation from the tumor. Then the tumors are classified as healthy or cancerous. The algorithm can be further improved to analyze different types bone cancer and features to predict the condition.

Keywords. Bone tumor, MATLAB, Gaussian filtering, Segmentation, K-mean clustering, SVM.

1. Introduction

Bone is a rigid part of skeleton which provides support to the human body. The external portion of bone is made up of rigid tissue, collagen and calcium phosphate matrix. Cortical bone forms the outer part which covers trabecular bone inside and periosteum covers the outside of bone. The bone marrow and the medullary cavity forms hollow bones.

Tumor, a mass or lump of tissue is formed when cells divide abnormally and uncontrollably. The cancerous tissue can evade the healthier tissue, when the tumor grows. There can be malignant or benign type of tumor growth. Bone tumor in the initial stage comes up with unsuspecting indications, like any fracture, abnormal growth, swelling of soft tissues surrounding a bone etc. Only 1 percentage or less than that is the occurrence of primary type bone tumor according to estimation by National Cancer Institute (NCI). Oftentimes, diagnosis is done by X-ray imaging as the cancerous bone image will be distinctive from the healthier bone tissue and surrounding area.

1.1 Types of benign bone tumors

a) Osteochondromas: Osteochondroma type accounts for between 35-40% of all benign bone tumor. Bone and cartilage form this kind of tumor.
b) Non ossifying unicameral fibromas: This type is a single cyst in bone, which is mostly found in the leg.

c) Giant cell tumor: Giant cell tumor grow aggressively and are found as round in bone end and not shown in growth plate. This is a rare type of tumor.

d) Enchondromas: formed as cartilage-cyst which will grows in the bone marrow and usually found in hands, feet, arm and the thigh.

e) Fibrous-Dysplasia: This type of tumor can be caused because of mutation in gene which causes bone fracture vulnerable.

1.2 Types of malignant bone tumors

a) Osteosarcoma: this type of tumor formed in shoulder, hip and knee region. They grow rapidly and tends to spread to other parts of the body. The common sites for this type of tumor to spread are the growth plates (areas where the bones are most actively growing).

b) Ewing sarcoma family of tumors: ESFTs are seen in leg, human skull, arm, on ribs, on backbones and pelvis region. They begin from medullary cavities.

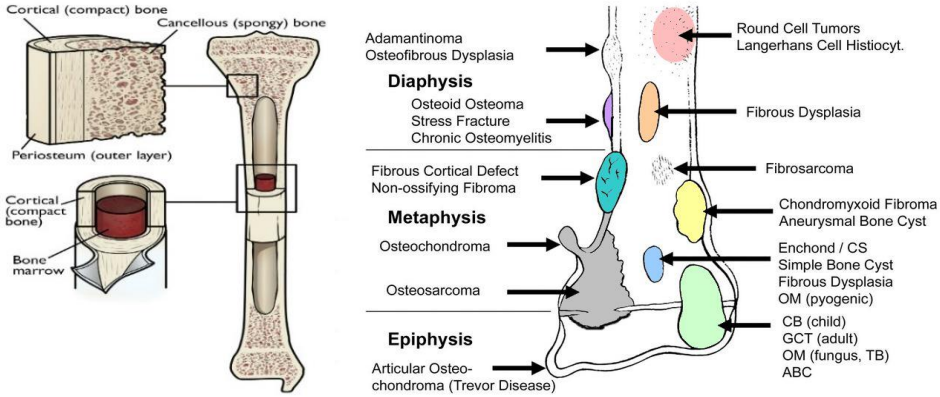


Fig 1. (a) Structure of normal bone (b) Different types of bone tumor (Source: epomedicine.com, orthoinfo.aaos.org)

2. Literature survey

2.1 Bone cancer diagnosis through imaging tests such as:

- X-rays: Diagnosis is primarily done by X-ray; imaging shows the tumors and the size of tumor.
- CT scans: They are taken to make more detailed pictures and imaging of bone and soft tissues.
- MRIs: Shows the exact extent of tumor detected, the marrow inside bone, soft tissues around the tumor, including the blood vessels and nerves.
- PET-bone scans: Radiotracer is injected to vein and PET scanner visualize the tumor region where radioactive glucose uptake is more. This help to show the spread of bone cancer to other bones and other parts of body, and also shows the how the cancer is responding to treatment during the treatment phase.
- Bone scan: To look inside the bone after the radiotracer is injected. It collects in bone tumor area and detected by special camera which help to determinate the stage of a bone sarcoma.

Image analysis of X-ray bone image is the cost-effective tool for bone cancer/tumor detection primarily, as absorption of radiation is different for the cancerous bone cells. Therefore, the cancerous bone x-ray image is seen as "ragged" form or as hollow regions/ holes [11]. Orthopedics practitioners usually suggests x-ray scans as a diagnosis tool.

Biomedical image processing is the process of enhancement and analyzing medical images, and are usually used in x-ray scan image, CT-images, MRI-scans, etc. methods of image segmentation are often used on the x-ray scans and MRI scan image for identifying the presence and type of tumor. Any abnormal growth of bone can either be malignant (cancerous) and benign(non-cancerous) can be decided based on its dimensions and other features, and form of bone tumor be also detected. As there are more bone cancer occurrence, x-ray image-based segmentation and detecting the type of tumor has gained interest in biomedical image analysis research area. Several researches have been done in this field for identifying type of bone tumor and there are a number of segmentation methods used for bone tumor detection. Also, different softwares and algorithms are studied for automated processing and providing precise result.

2.2 Bone tumor detection related research works

a) Prachi B.Tamgadge et. al used morphological operations such as dilation, erosion, opening etc. and segmentation based on zone, mean, entropy, standard deviation to determine tumor from X-ray images [1].
b) Dr. Abhilash Shukla et. al used region image segmentation method and k-mean clustering to detect tumor bone from X-Ray image [2].
c) Oishila B et.al developed algorithm that extract features from bone x-ray, using support vector machine method to differentiate healthier and cancer affected bone. Decision tree classification method is performed to analyze the pattern of bone destruction [3].
d) Torki Altameem detected tumor in his paper with 98% accuracy applying intuitionistic MATLAB fuzzy rank-correlation and processing derived features by applying the deep neural networks [4].
e) Bhagyashri Giradkar et. al used semi-supervised learning: ANN, SVM. They also used an unsupervised clustering approach, label each cluster with class, and optimize the misclassification error of the resulting clusters [5].
f) K. Sujatha et. al Performed MRI wavelet denoising, K- means algorithm, principal components of genetic algorithm (ie, GA) method, to identify if growth of bone tumor is malignant or beningn. [6].
g) Anil K. Bharodiya et. al has proposed binary blob-based detection og bone tumor (ie, BTDBB) along with gaussian filtering method to remove noise. These blob-based analysis and thresholding distinguished types of bone tumor from bone x-ray images. [7].
h) M. Vardhana et. al implemented edge detection algorithms using VLSI using MATLAB and Modelsim. Image classification is here performed using convolution neural network(CNN)[8].
i) Y. Rajesh Sk et.al in their paper preprocessed MRI image using a median filtering and thresholding & performed feature extraction and classified the images using the extracted features[9].

2.3 Project description

Diagnosis of bone tumor/cancer should be done without much delay and by preliminary diagnosis itself, by the biomedical image analysis method. The goal of the project

is to propose a method for the preliminary detection of the bone tumor. Developing a GUI, image processing algorithm in MATLAB helps in giving an alert to patient in the initial stage of cancer, so that treatment can be started before the situation gets worse. The work proposed here discusses a methodology developing an algorithm for bone tumor detection and thw whole process been divided into 3 sections: Input X-ray image and its pre-processing, then image segmentation and finally classification into healthier and tumor images.

- First step is the collection of cancerous and non-cancerous bone x-ray scan images.
- Creating Graphical User Interface in MATLAB.
- Preprocessing of the images.
- Segmentation of the images and detection of tumor.
- Classification of healthy and cancerous bones.
- Feature extraction (such as tumor dimension) from x-ray image after segmentation
- Method of SVM Support vector machine is used for classification of healthier bone image and tumor detected bone images.

3. Methodology

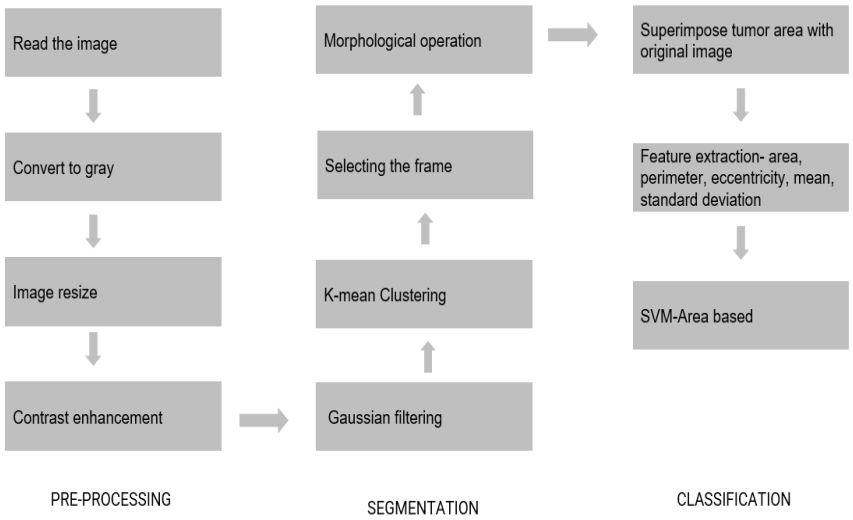


Fig 2. Block diagram of proposed method

Proposed algorithm: Segmentation and classification of bone tumor for given x-ray images.

Input provided : X-ray scan images- bone tumor affected and healthy bone scan image.

Expected output: Segmented bone tumor x-ray image.

Step_1: Input human X-Ray image.
Step_2: Performing image resize and converting the given image to gray-scale.
Step_3: Enhancing the image & removing noise by gaussian filter
Step_4: Segmentation & selection of tumor region from the enhanced image.
Step_5: Tumor features measurement/extraction.
Step_6: SVM-Tumor Classification
Step_7: Display Output.

3.1 Image pre-processing & filtering stage

Images that are collected from dataset is inputted, then they are resized and converted to grayscale. contrast enhancement is done because for segmentation it requires high contrast images. For noise removal in the x-ray image and enhancing the image for further processing, Gaussian filtering and median filtering method is performed. Filters help in salt and pepper type of noise and random noise/gaussian noise removal and also remove false-edges formed by the noise. This image enhancement will provide accurate output.

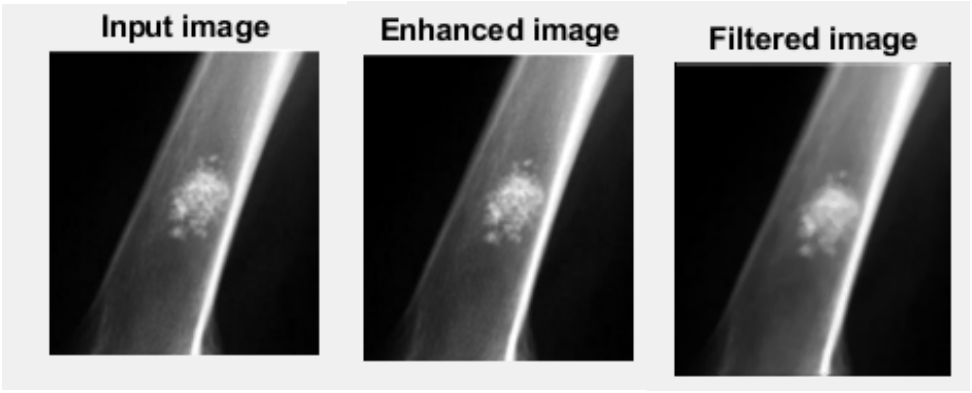


Fig 3. Demonstration of each steps in image pre-processing

3.2 Detection and segmentation stage

After the x-ray image enhancement, the next step will be segmenting only the tumor region from the image. Both edge-based segmentation and region-based segmentation methods are discussed here.

a) Edge detection: Comparing Prewitt, Sobel, Canny

Edge-based segmentation is used in biomedical image processing, which detects the brightness discontinuities in the given image and find the object boundaries. Prewitt, Sobel, Canny are the common methods in edge-based filters. Sobel filter is used to fing the edges, curve and tracing the boundaries in an image by using n*n convolution kernel operators. Horizontal and vertical kernels are used to define the edges.

The Prewitt filter is similar to that of sobel, which finds horizontal vertical edges with masks. For the pixels of image that are close the center of applied mask, prewitt does not give much emphasis.

Canny Edge detection algorithm is multistage, they are able to define both weak and strong edges in given image. It is used after gaussian filtering and then computing the gradient.

b) Region based: K-mean clustering, Region growth

K mean clustering in image processing is used to perform segmentation of an area of interest from the image background, and is an unsupervised learning algorithm. Therefore, it can be used in tumor detection application. They partition the image or group the objects into K-number of clusters by the object features. The minimizing can be done by choosing particular distance metric as per the image pixels and centroid clustered.

Region growth is another image segmentation method, which is based on image pixel. In this method, initially need to select a seed point and then its surrounding pixels are

iteratively compared, and further chose which pixels can be a part of region. It is been observed that the region-growth and k-mean method gives better result and are good for tumor detection application as compared with edge-based segmentation.

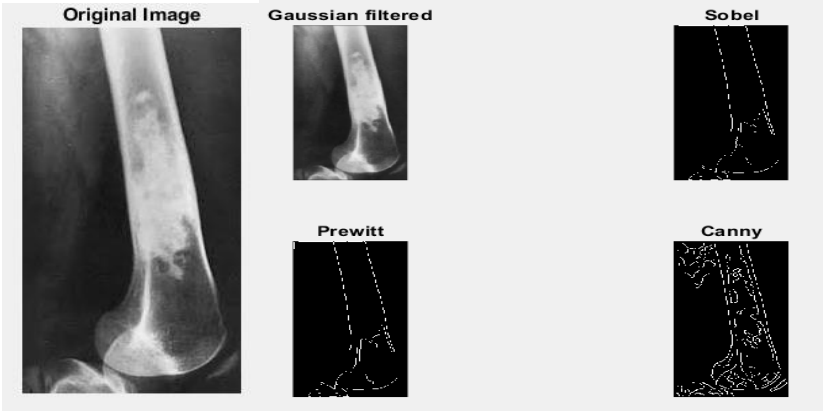


Fig. 4a. Demonstration of bone tumor segmentation technique (Prewitt, Sobel, Canny)

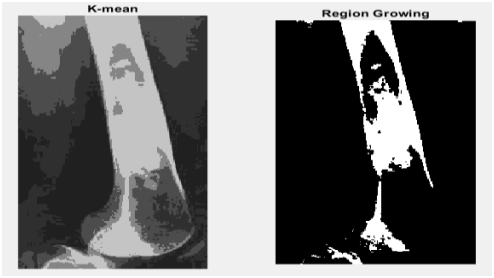


Fig. 4b. Demonstration of bone tumor segmentation technique (K-mean clustering, Region growth)

It is observed that region growth alone is not accurate for accurate segmentation. So, we extracted frames from the image and done morphological operations open and dilation. Image opening has been done to eliminate the objects which are smaller from given image, usually used in finding particular region or shape from the image. Dilation is performed to add or enlarge the boundaries of pixel and thereby holes in that region become smaller. The features such as area, perimeter, eccentricity, entropy, mean, standard deviation can be calculated.

3.3 Classification stage

The machine learning tool, Support Vector Machine (SVM) is used which is based on the idea of classification of large margin data. SVM classification has gained attention in image processing recently and are found to be used in many studies which provides good classification of images. Collected dataset of bone tumor X-ray images used for this work includes only 55 images of which 35 are tumor affected.

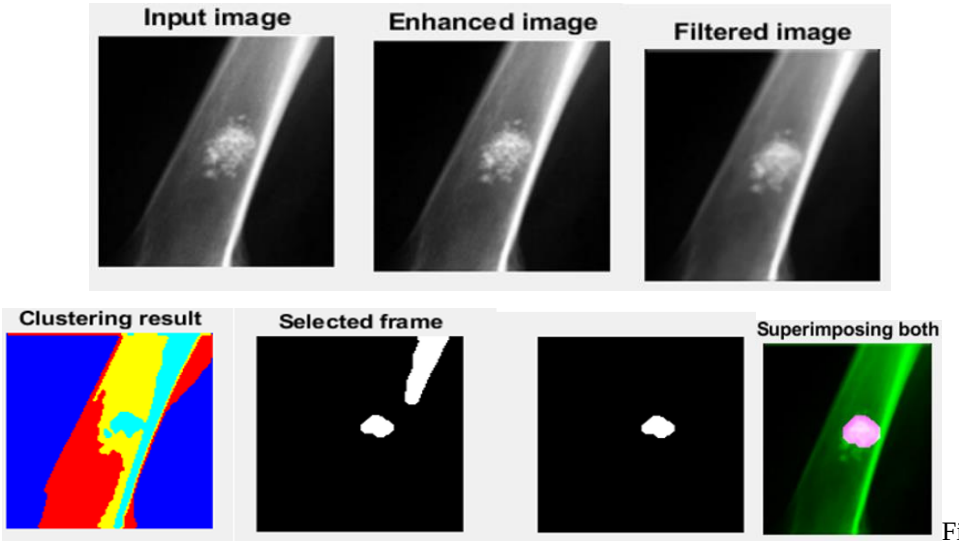


Fig. 5. Demonstration of bone tumor clustering and morphological operations

4. Result and Discussion

Bone tumor if not diagnosis in proper time it can be hazardous to patient. This paper has designed a digital way of image acquisition and processing techniques for preliminary assessment of bone tumor which required x-ray image filtering for image enhancement, segmentation and morphological operation for tumor identification, followed by feature extraction and classification by SVM method.

The GUI is made to import an X-ray image, converting it to grayscale, resizing the image, enhancement of the image and then, filtering of the image using Gaussian and median filter. Sobel, Prewitt, and Canny edge segmentation methods and region growing, k-mean clustering methods are discussed in this paper. This method can be used for interpretation of MRI scan image also. The result of each segmentation method are shown below and finally chose region growing and k- mean segmentation method in the algorithm as they are suited for bone-tumor detection.

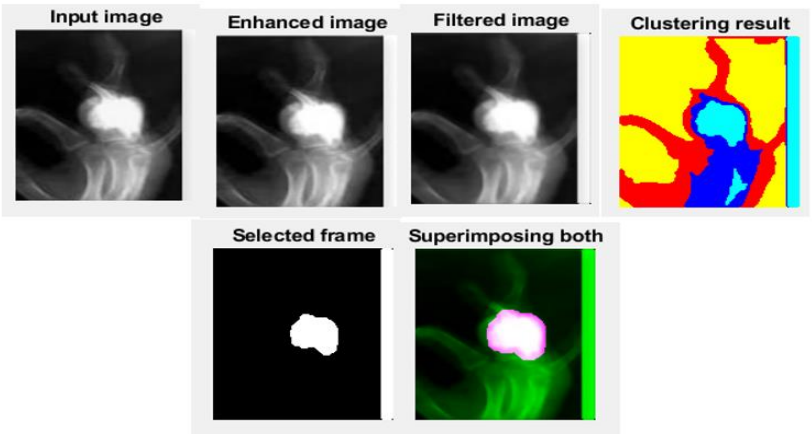


Fig. 6. Demonstration of all the image processing steps

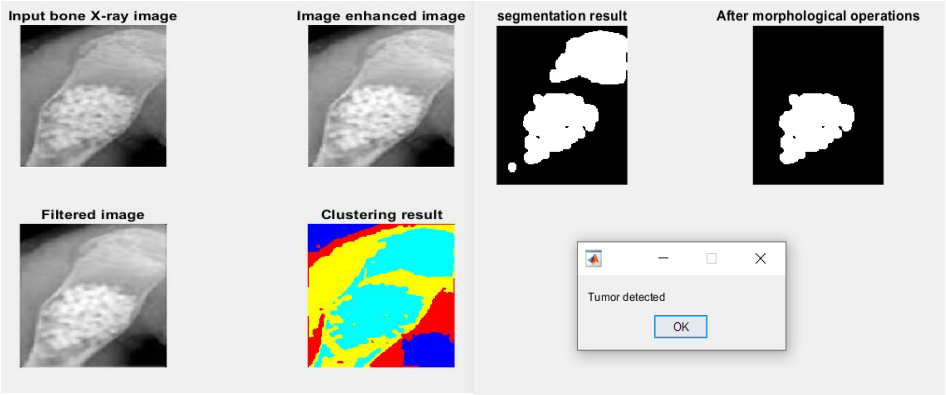


Fig 7. Tumor affected image shown after SVM

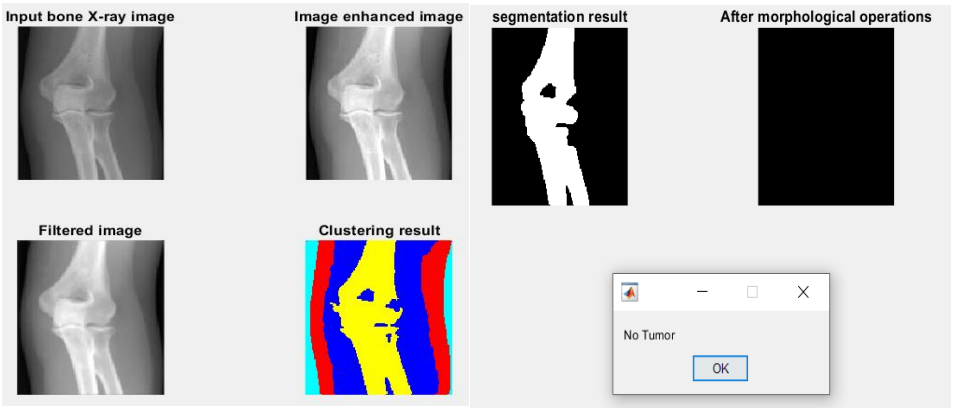


Fig. 8. Non-tumor image after SVM

Impact of the proposed solution

- Application of artificial intelligence in Bio-Medical image processing
- Easy and preliminary classification of bone tumor/cancer based on biomedical image analysis.
- Developing a GUI, image processing algorithm in MATLAB
- Assisting radiologist and doctors to identify tumor

5. Conclusion

There are 75 types of different cancers existing, according to the American Cancer Society. Bone cancer is one from them and can be again divided into 19 types. Ewing tumor type and osteosarcoma are more often formed tumor. In the initial stage symptoms of bone tumor/cancer cannot be easily predicted accurately. Therefore, there is a need of preliminary bone tumor/cancer prediction algorithm. In the inputted image if there is tumor present this will generate a message warning ‘tumor is detected’, and for a healthy bone image ‘no tumor’ message is given. K-mean clustering and gaussian filtering is performed on bone x-ray scan images here. The developed algorithm has observed as efficient and will be helpful in assisting

the radiologist and doctors to determine the presence of bone tumor, so there will be no delay in starting treatment.

Future aspect. There should be minimum of 10,000 image datasets to train the model providing an accurate result and auto detecting the bone tumor features using Deep Learning method with lesser computational time and usage cost. The features that are retrieved from the X-rays images may be classified into different tumor types.

References

1. Prachi B.Tamgadge, Dr. N.K Choudhari, Dr. D.M. Kate *"A Review Paper on Detection of Bone Tumor using Comparative Analysis of Segmentation Technique"* International Research Journal of Engineering and Technology (IRJET) Volume: 06 Issue: 02 Feb 2019
2. Abhilash Shukla,Dr. Atul Patel *"Bone Cancer Detection from X-Ray and MRI Images through Image Segmentation Techniques"* International Journal of Recent Technology and Engineering (IJRTE) volume-8, issue-6, march 2020.
3. Oishila Bandyopadhyay, Arindam Biswas, Bhargab B. Bhattacharya *"Bone-Cancer Assessment and Destruction Pattern Analysis in Long-Bone X-ray Image"* Journal of Digital Imaging (2019)
4. Torki Altameem *"Fuzzy rank correlation-based segmentation method and deep neural network for bone cancer identification"* Neural Computing and Applications.Springer Nature 2019.
5. Bhagyashri Giradkar, Nilesh Bodne *"Bone Tumor Detection using Classification in Deep Learning using Image Processing in MATLAB"* International Research Journal of Engineering and Technology (IRJET) Volume: 07 issue: 06 | June 2020
6. K. Sujatha, S.Jayalakshmi ,Sinthia. P, Malathi M, Ramkumar. K.S, Su-Qun Cao *"Screening and Identify the Bone Cancer/Tumor using Image Processing"* IEEE International Conference on Current Trends toward Converging Technologies, 2018.
7. Anil K. Bharodiya, Atul M. Gonsai *"An intelligent assistive algorithm for bone tumor detection from human X-Ray images based on binary Blob analysis",* International Journal of Information Technology (2020)
8. M. Vardhana a, N. Arunkumar b, Sunitha Lasrado, Enas Abdulhay d, Gustavo Ramirez-Gonzalez e "*Convolutional neural network for bio-medical image segmentation with hardware acceleration"* Cognitive Systems Research 50 (2018) 10–14
9. Y. Rajesh, Sk. Mastan Shareef, G. Ajay Kumar, K. Bhargav, *"Identification of bone cancer using image processing",* Journal of Engineering Science, Vol 12,Issue 7, July/2021 ISSN NO:0377-9254
10. Harish Babu Arunachalam , Rashika Mishra, Bogdan Armaselu, Ovidiu Daescu, Maria Martinez, Patrick Leavey, Dinesh Rakheja, Kevin Cederberg, Anita Sengupta, Molly Ni'suilleabhain, *" Computer aided image segmentation and classification for viable and non-viable tumor identification in osteosarcoma",* PMID: 27896975 DOI: 10.1142/9789813207813_0020
11. Bone Tumor: Types, Causes and symptoms, https://www.webmd.com/cancer/bone-tumors

A Portable Self-Diagnostic Device for Preliminary Monitoring of Lung Health

Shubhanki Sengupta, Shishira P A, Dr. Sasikumar K

Vellore Institute of Technology, Vellore, India
shubhanki.sengupta2020@vitstudent.ac.in shishira.pa2020@vitstudent.ac.in, sasikumar.k@vit.ac.in

Abstract. In the human body, the lung is the vital internal organ which is also vulnerable to various kind of diseases and infections. WHO estimated that 4 million people die due to lung disease per year. Lung abnormalities are frequently undiagnosed, which can delay disease management interventions. Lung abnormalities symptoms needs to be screened as early as possible to prevent progression serious lung disease. Conventionally the lung diseases are diagnosed by tests like x-ray of chest, CT imaging and laboratory tests like blood oxygen level and sputum tests, which are expensive and need experienced staffs to perform it. This demands a self-diagnosing device that check the lung condition independently at home or clinics in rural areas. This paper discusses designing of a portable device that is able to detect lung condition using body temperature, pulse rate and color of the fingernail as parameters. LM-35 temperature sensor, TCS 3200 color sensor and SEN-11574 pulse sensor are used to collect data. The sensor data are taken for classification of lung health. This non-invasive portable device is intended to perform real time monitoring of lung condition at an affordable cost.

Keywords. Self-diagnosis, LM-35, TCS-3200, SEN-11574, Lung health classification.

1. Introduction

The respiratory system depends on the most important organ, which is the lungs. They are responsible for inspiration and expiration mechanisms. The air we inhale contains oxygen, which reaches the lungs during respiration, and is moved to the blood stream. The oxygen is exchanged with carbon di-oxide in the cells of the body. This carbon di-oxide is then passed to the lungs and then comes out when exhaled. The lung and the respiratory system perform this vital process known as gas exchange.

Lung abnormalities must be detected as early as possible to prevent serious lung disease as this is the most important part of the respiratory system. According to India Today statistics Pneumonia which is also a symptom of the severe stage of covid-19, kills 1,678 persons every month In India on an average. Tuberculosis on the other hand kills 3,129 people per month.

Conventionally, CT Scan, chest X-ray, blood test, sputum test, oxygen level of blood helps in detecting lung diseases [3-5]. These tests are time consuming and needs assistance of doctor or nurses making the tests costly and clinic based.

Therefore, this demands a self-diagnosing device that can check the lung condition independently at home or clinics in rural areas or at home or clinic using several parameters of human body. Parameters like body temperature, pulse rate and finger nail color are considered in this paper for diagnosing lung condition.

1.1 Body temperature and lung condition

Body temperature has a vital role to play in detection of lung condition. Unhealthy lungs are affected by bacteria and to reduce the infection in body, the body temperature increases to kill the bacteria. So, people with unhealthy lungs often have higher body temperature than healthy people[1-2]. LM-35 sensor is used in this paper to get the body temperature due to its direct contact with human body[8].

1.2 Pulse rate and lung condition

The normal heart rate of a human being is 60 – 72 beats per minute. Whereas the pulse rate of people with unhealthy lungs is often more than 100 beats per minute. So, pulse rate also plays a key role in detecting unhealthy lung. SEN-11574 heart rate sensor is used in this work to get the pulse rate in bpm.

1.3 Fingernail color and lung condition

The nail color plays a significant role in identifying lung diseases. Oxygen level in our body has a significant effect on the nail color. Respiratory problems may occur before or after the nails start changing color and shape. But a healthy human being has a pink fingernail color. They are uniform in color and consistency and free of spots or discoloration. From the fingernail color a healthy lung person and a non-healthy person can be differentiated [6-7].

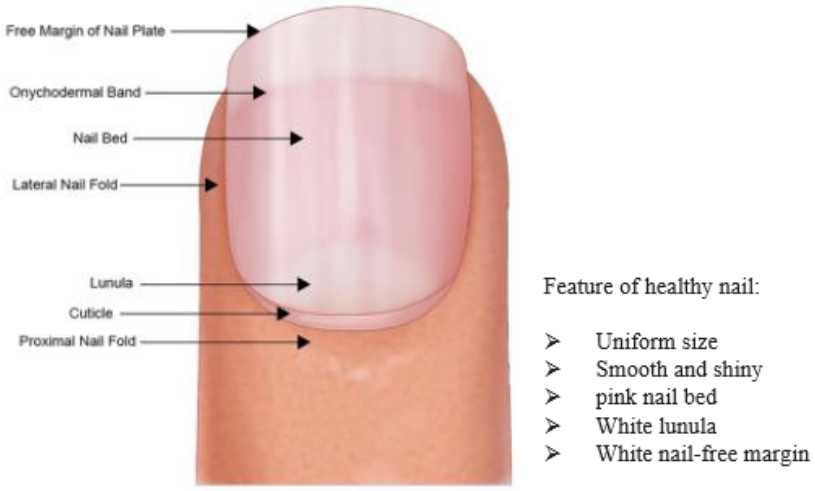


Fig 1. Pink color of healthy person nail
(Source: sciencedirect.com)

Lack of oxygen circulation in RBCs causes bluish color of fingernail. This condition is called cyanosis. Conditions which might be causing cyanosis includes:

- COPD (Emphysema)
- Acute respiratory distress syndrome
- Pneumonia
- Pulmonary embolism-Blood clot in lungs

Fig 2. Bluish/gray tint on fingernail when blood doesn't circulate enough oxygen throughout the body (Source: https://www.drthindhomeopathy.com/)

Tcs3200 color sensor module consists of 4 LED (Light Emitting Diode) and a photo-diode matrix for detecting the color of the object with more accuracy [8].

IoT in Healthcare. Internet of Things is an emerging trend where embedded devices can be internet connected. The device is communicable with human being and can store the data in cloud where the data can be processed and analyzed. IoT solutions are helpful for environmental monitoring and control, healthcare monitoring, industrial monitoring and automation of home. Thing Speak is the IoT platform used, allowing aggregation, visualization and analysis of live data from the stream.

The objective of this paper is to design a self-diagnostic device for identifying preliminarily the lung condition using the body temperature, pulse rate, fingernail color as features and identifying weather of person is healthy or is having pneumonia or hypoxia. Sensors used in this paper includes, LM35-temperature sensor, SEN 11574-pulse sensor, TCS3200- color sensor. Classification of lungs is based on the readings/data obtained from used sensors. Paper also objects to storing data in cloud with the help of Wi-Fi and sharing the data securely using Internet of Things.

2. Design and methodolocy

The flow of the device goes as follows:

2.1 Body temperature detection

LM35 temperature sensor was used for measuring the body temperature. LM35 has a vast range of -55 degrees centigrade to 150 degrees centigrade to be measured with high accuracy of 0.5 degree C.

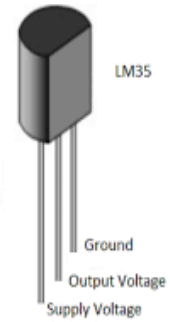


Fig 3. LM35- Temperature sensor

The LM-35 sensor gives the output as voltage and this is proportional to the temperature value measured. The LM35 sensor needs to be clamped between the thumb and index finger to get the body temperature. This method gives the accurate temperature of the body without interference of the environment temperature. 15 seconds is required to get the accurate result.

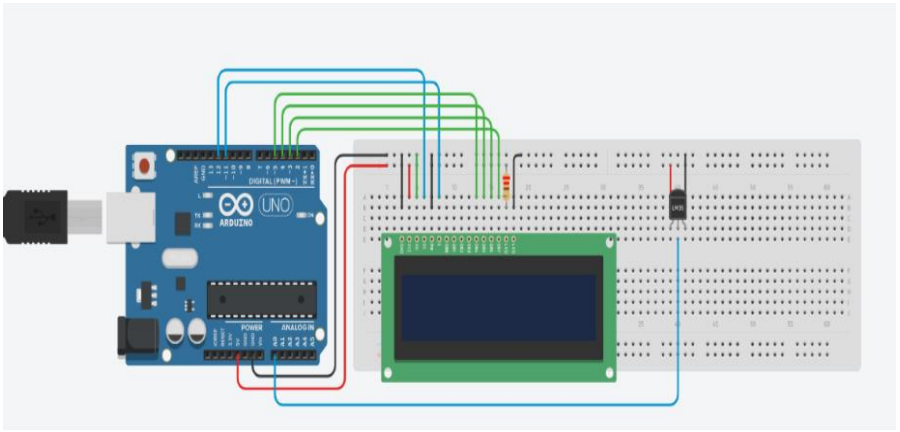


Fig 4. Simulation of LM 35 temperature sensor in Tinkercad

2.2 Pulse rate detection

Pulse rate sensor (SEN-11574) helps in detection of pulse rate with the help of blood flow through the veins. On the one side of the sensor is a LED light and on the other side there is the circuitry of the sensor. The sensor needs to be placed Just below the fingertip so that the light from the LED can fall directly on the veins of the fingertip. The veins have blood flow whenever the heart pumps. This can be detected by the pulse rate sensor and can be counted using Arduino IDE to get the pulse rate in beats per minute.

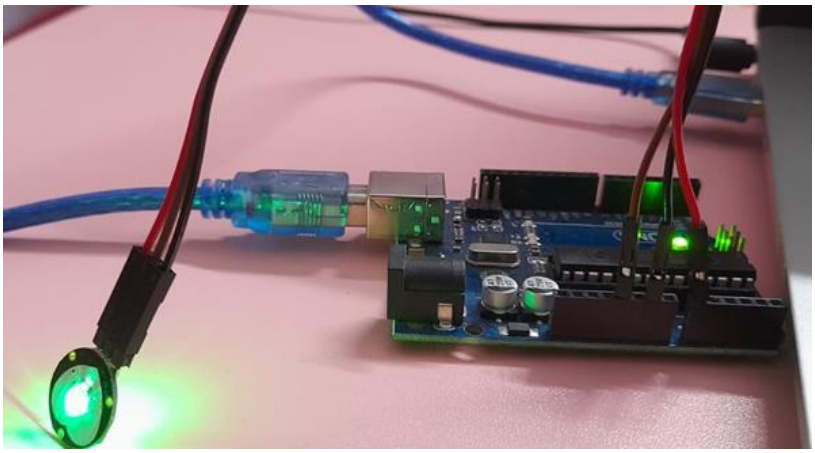


Fig 5. Circuit for SEN-11574 pulse sensor

2.3 Fingernail Color Detection

TCS-3200 consists of 4 LED and one photodiode matrix for detecting the color of the fingernail. This sensor gives the result in RGB values.

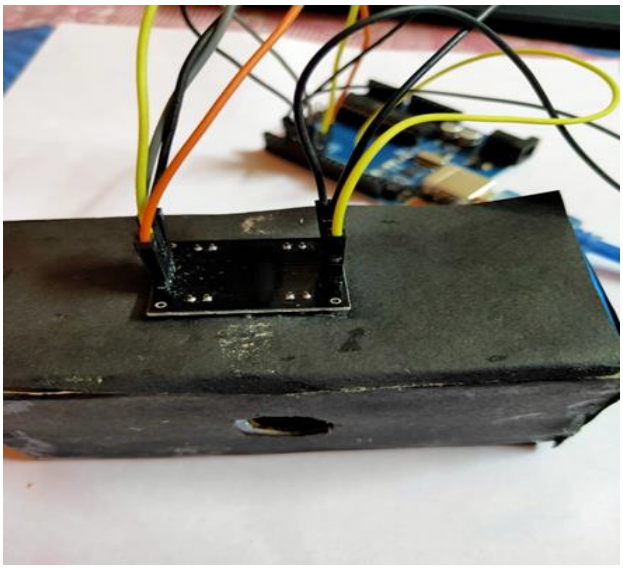


Fig 6. The color sensor is wired and attached to a black box with hole for placing finger for accurate RGB value of nail

2.4 3D model of the device

The device is black in color and portable. It has a slot or hole for placing the finger and it is battery operated. The tcs-3200 was placed above hole. The sensor was placed upside down inside the black box so that it can detect the color of the fingernail directly. The hardware box is made black in color to minimize the external light interference that may cause inaccuracy in the detection of fingernail color. The device also contains pulse sensor which is just below the

hole. It has the temperature sensor, LCD and push buttons outside the black box. The LM-35 was placed outside the black box as direct contact is required between the temperature sensor and human body for detecting the body temperature. The push button triggers the microcontroller for acquiring the sensor data.

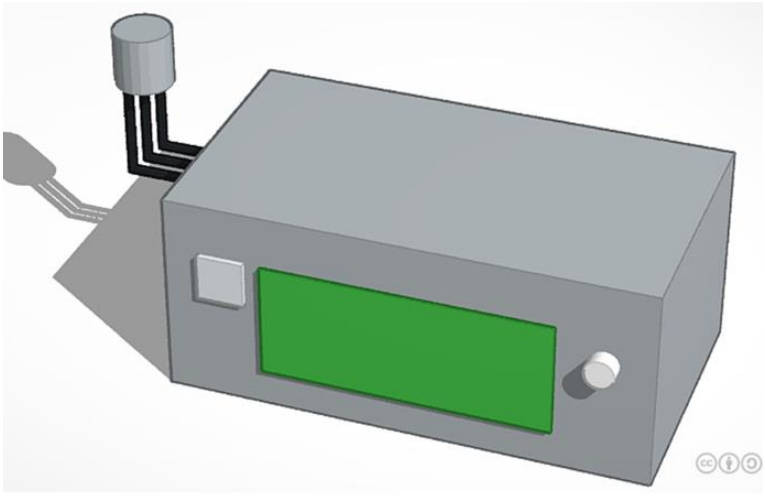


Fig 7. 3D modelling of the device in Tinker cad

2.5 Sharing the data in ThingSpeak

The ThingSpeak lets us store data in cloud and helps in development of IOT applications. To store the readings from sensors, project aims in sending the data from the sensors to ThingSpeak using the ESP8266. The first step is creating a channel on ThingSpeak, then connecting the ESP8266 to Wi-Fi network, then sending the data to the ThingSpeak IP address and API key.

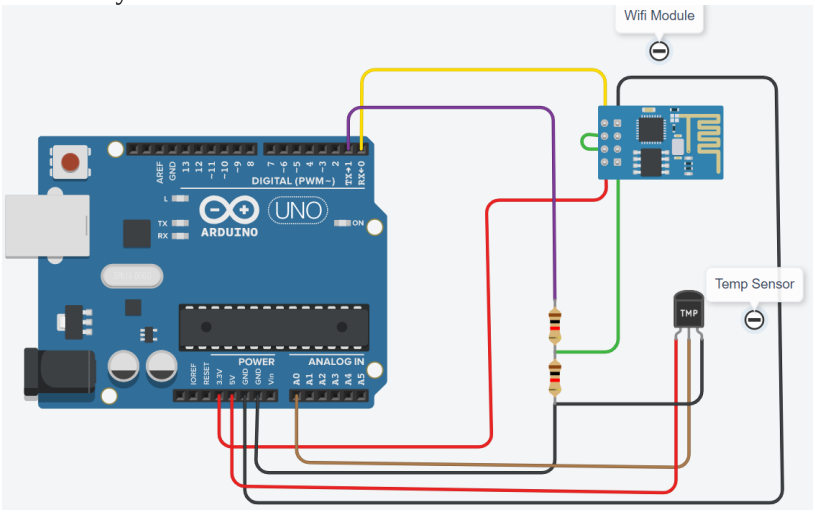


Fig 8. Simulation of Arduino circuit interfacing LM35 and ESP8266

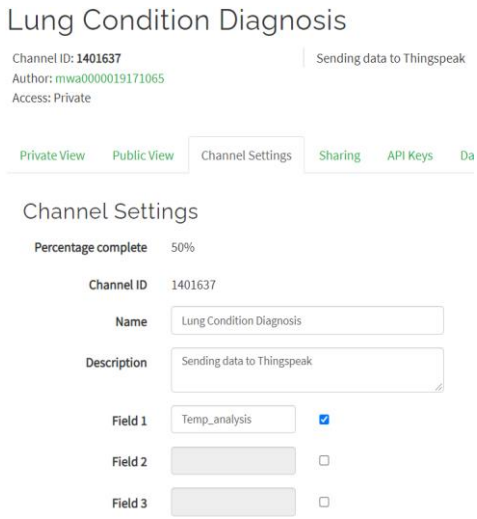


Fig. 9.1. Channel creation in ThingSpeak

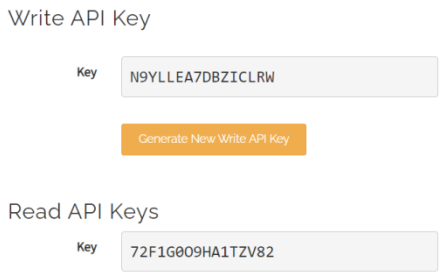


Fig. 9.2. API key of ThingSpeak

2.5 Starting with ThingSpeak

In theThingSpeak.com website "Get Started for Free" is chosen.

In the sign-up form the information entry is required and signing up is required according to the project details.

"New Channel" is to be chosen to create a channel for storing the information

the API key is to be copied and used in Arduino code.

Connect to Wi-Fi and can store sensor data in ThningSpeak platform and user can download all of this Channel's feeds in CSV format.

3. Result and discussion

3.1 LM35- Temperature sensor reading

The temperature sensor LM 35 will measure the body temperature every time and will display the temperature on the LCD. Buzzer is attached to the sensor so that whenever temperature goes high it gives indication.

Fig 10. Temperature sensor output in LCD

3.2 SEN11574- Pulse sensor reading

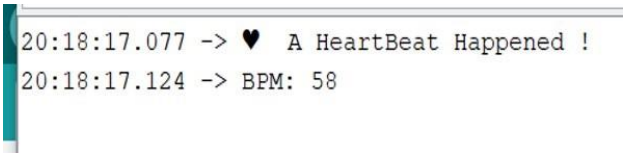


Fig 11. Pulse sensor reading of a healthy person in serial monitor

If the pulse sensor reading goes beyond 100, at steady condition, that indicates a diseased person. Here in the output we got bpm as 58 which is of a healthy person.

3.3 TCS3200- Color sensor reading

TCS3200 is first programmed to RGB values with random color strips. Then we trained it with different pink, yellow and blue shades fingernail color from healthy and affected persons, where the palette is obtained from Adobe CC software. Based on the blue color frequency value the classification of lung condition, whether it is healthy or TB/Pneumonia or hypoxia conditions is done. Fig 10 shows the classification results for each color strips.

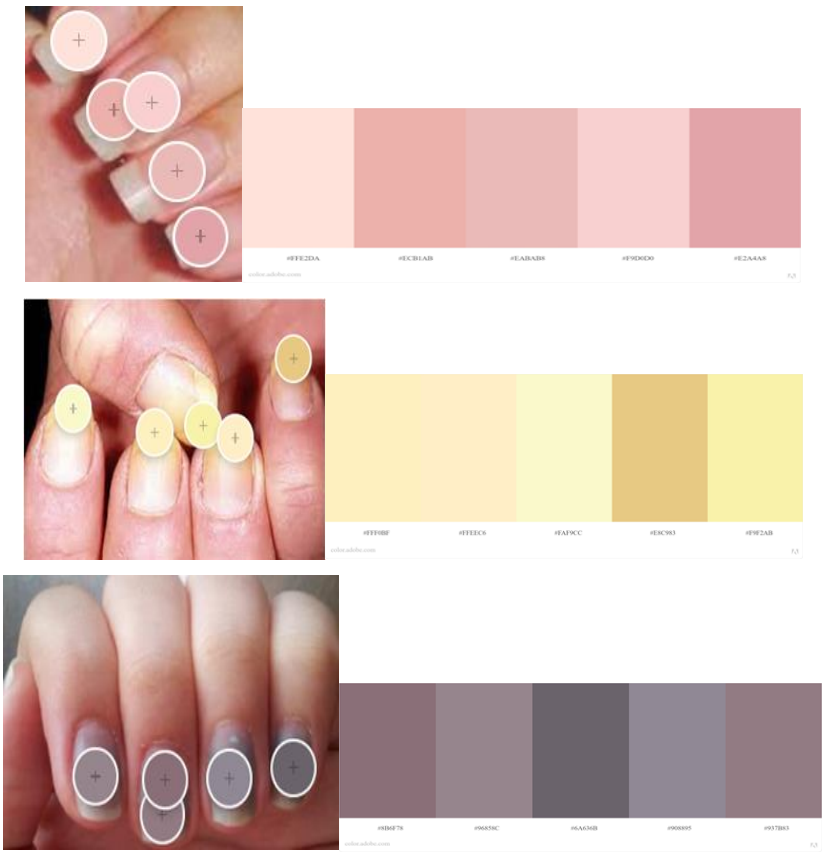


Fig. 12. Color palette from Adobe CC which is used to train the sensor

```
23:52:47.904 ->  R = 47 G = 65 B = 65
23:52:49.208 -> Healthy
23:52:49.208 ->  R = 47 G = 69 B = 69
23:52:50.475 -> Healthy
23:52:50.475 ->  R = 46 G = 68 B = 68
23:52:51.780 -> Healthy
23:52:51.780 ->  R = 46 G = 69 B = 69
23:52:53.072 -> Healthy
23:52:53.119 ->  R = 46 G = 69 B = 69
23:52:54.379 -> Healthy
```

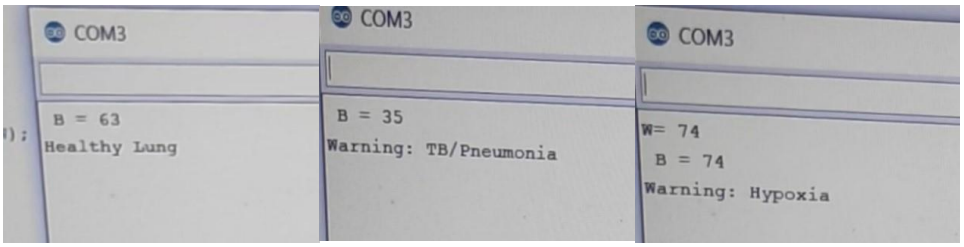


Fig 13. Color sensor reading of healthy person and pneumonia/hypoxia affected in serial monitor

3.4 Thingspeak

In this work a channel on ThingSpeak is created, and the ESP8266 is connected to an Wi-Fi network, and then the data is sent to the ThingSpeak IP address and API key.

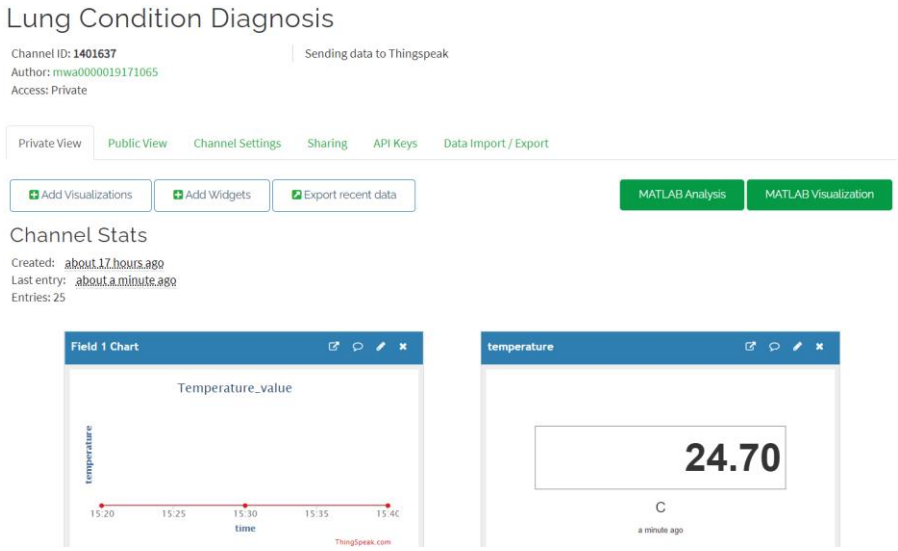


Fig. 14. Data streaming in ThingSpeak

4. Conclusion

This paper discusses designing a smart medical portable device that will help in diagnosing the lung condition at the earliest. The current lung tests are expensive and need experienced staffs to perform it. This demands a device check the lung condition independently

at home, instantly and in a cost-effective way. When any infection is present on lung, body tends to increase temperature as a first line of defense and pulse rate will be more than 100 beats per minute. The yellow and blue discoloration of nail is directly linked to pneumonia and hypoxia condition. The sensor data are taken for classification of lung health. The paper also shows the steps to storing the data in cloud using Wi-Fi and sharing the data securely in ThingSpeak platform. This work can further be extended to get the result instantly using machine learning and feeding the data. This device can show a path to diagnose a lung disease at the earliest and reduce mortality rate due to lung diseases.

Future aspects: Temperature, pulse rate and color sensor are the three-parameters used now. In future we can add other parameters that can be used to identify changes in lung health. Classification of lung condition can be done accurately from the result of sensors using different machine learning algorithms (e.g., Naive-Bayes). And also different classes can be detected like healthy-lung, hypoxia, pneumonia condition and other diseases.

References

1. A.J. Heuer and C.L. Scanlan, “Wilkin’s Clinical Assesment in Respiratory Care,” Elsevier, 2014.
2. N. Deepika, P. Vinupritha, and D. Kathirvelu, “Classification of lobar pneumonia by two different classifiers in lung CT images,” International Conference on Communication and Signal Processing, pp. 552-556, 2018.
3. R. Spilsbury, “Health and Disease: Investigating a TB Outbreak,” London: Raintree, 2014.
4. A. Rao, J. Ruiz, C. Bao, and S. Roy, “Tabla: an acoustic device designed for low cost pneumonia detection,” IEEE Healthcare Innovations and Point of Care Technologies, pp. 172-175, 2017.
5. J. Gibson, “Clinical Test of Respiratory Function,” Hodder Arnold, 2008.
6. A. Singal and R. Arora, “Nail as a window of systemic disease,” Indian Dermatology Online Journal, vol. 6, pp. 67-74, 2015.
7. Y. Fu and J. Liu, “System design for wearable blood oxygen saturation and pulse measurement device,” International Conference on Applied Human Factors and Ergonomics, pp. 1187-1194, 2015.
8. R. Maulana, D. Kurniawan, M. H. H. Ichsan,"Smart Devices for Self-Diagnosing of Lung Condition Based on Body Temperature and Fingernail Color",2020.
9. A. Rao, J. Ruiz, C. Bao, and S. Roy, “Tabla: an acoustic device designed for low cost pneumonia detection,” IEEE Healthcare Innovations and Point of Care Technologies, pp. 172-175, 2017.
10. J. Gibson, “Clinical Test of Respiratory Function,” Hodder Arnold, 2008.
11. M. Miller, “The Internet of Things: How Smart TV, Smart Cars, Smart Homes and Smart Cities are Changing the World,” Indiana: Pearson Education, 2015.

Smart Medical Clinic

Barath Ram S., Niket Narayan, Muthu Raja S.

School of Electronics Engineering, Vellore Institute of Technology, Vellore, India
barathram.s2018@vitstudent.ac.in, niket.narayan2018@vitstudent.ac.in, muthurajas@vit.ac.in

Abstract. Over the last 2 years, we have witnessed the debilitating impact of the SARS COVID 19 virus worldwide, millions were affected by the disease and many lost their lives. One common issue faced by the people was the inability to get symptoms diagnosed on time due to crowded hospitals or living in rural regions with limited healthcare facilities. Therefore, in order to ensure timely diagnosis and to ensure that healthcare facilities are available for rural and low-income households, we propose a low cost smart medical clinic to enable remote check-ups and monitoring. We developed a smart medical device using the design principles of an ATM machine along with the concepts of telemetry and telemedicine. A small sized room (similar to the size of an ATM center) is proposed to store the device which is equipped with sensors to measure vital parameters such as temperature, blood pressure, pulse, etc., and it also involves video conferencing with a general physician and therefore the patients can walk in one by one to get themselves checked and the data will be analyzed by the doctor remotely and if any abnormalities are found, they are assigned to the nearest health care facility with the help of an inbuilt GPS sensor in the device. Passive IR sensors are installed to detect the entry and exit of a patient, after which the entire room and the device will be sanitized using a disinfectant robot before the next patient enters.

Keywords. Healthcare, Telemedicine, Medical Device Development, Internet of Things, Microcontroller, Telemetry

1. Introduction

The current pandemic has clearly affected the everyday activities of people. It instilled a sense of fear in the minds of people who suffer from diabetes, common cold, body pain, etc., from going to hospital as the existing hospitals are flooded with COVID-19 patients. This is one such time where people are more scared of doctors and hospitals than the illness itself. “Fear of going to hospital “ is the normal mindset of people but can be dangerous as some people should consult a doctor for their problems rather than taking matters into their own hands by getting medications without consulting a doctor. How can we help people with regular check-ups without going to hospitals, in a safe and secured way? There is a massive rural urban area gap when it comes to availability of well-equipped health centers.

Almost 71% of the country is predominantly rural but the proportion of doctors and nurses in these rural areas are only 34% and 33%, respectively. India has a lack of infrastructure. Only 1 doctor is available per 1700 people in India and just 0.7 beds are there per 1000 people. You can possibly find a seat in local trains but not a complete bed in a hospital. Around 5 lakh doctors are required in India today. Just like the lack of doctors, the low no. of health workers is also very alarming. There are around 61 nurses and mid-wise per lakh population. If those with a medical qualification were considered, their number reduces to 6% per lakh population.

This idea is developed keeping in mind the current pandemic where people find it

difficult to go to hospitals for regular check-ups, in the fear of getting contaminated by the COVID-19 virus. A Smart Medical Clinic is like a normal health center but in the size of an ATM center with an automatic door and a smart medical device that is installed which has the necessary sensors connected to a microcontroller to measure the vital parameters such as temperature, pressure, heartbeat and pulse of a patient while a General Physician will be connected through a video call with the patient to whom the data from the sensors will be sent via a cloud. The doctor will be monitoring the data remotely and if anything abnormal is found in any of the patient's vital parameters, the patients will be advised to go for a further check up in the nearest tertiary healthcare center. The entire clinic and the device will be sanitized once a patient exits the clinic and before the next patient enters with the help of a sanitation robot.

The main components used for this system is microcontroller 8051, and the radio-frequency identification (RFID) which is incorporated to verify each patient using their Aadhaar card. This system can also be extended to rural areas where accessibility to hospitals or medical facilities is a major issue. The Global Positioning System (GPS) has been used in our project in order to allow accurate determination of geographical location of nearest health centers and Pathology centers which are then given to the patients so that they can easily go to that place for further testing and treatments if necessary. This is a zero physical contact model to ensure that all people get access to healthcare without the fear of coronavirus. This center and the device will be completely sanitized each time the patient leaves the clinic before the next patient comes inside.

2. Literature Survey

Detection of body temperature with infrared thermography: accuracy in detection off ever: This paper deals with the infrared thermography technology. For the application of IRT in screening for travelers with elevated body temperature at airports and border crossings, the forehead IRT temperature differed substantially from the core temperature, and the maximum lateral IRT temperature should be used. The reading should also be taken at a defined distance from the camera. Overall the sensitivity of IRT in detecting fever is low unless the cut-off temperature is low. When the risk of an epidemic is high and high sensitivity is required, a low cut-off temperature (≤35.5ºC) should be chosen.

Pulse oximetry: This paper discusses the Pulse oximetry. It is one of the most common monitoring modality in the health care setup. This paper describes the latest technological advances in the field of pulse oximetry, the accuracy and their limitations are critically examined. Finally, the existing data regarding the clinical applications and cost-effectiveness of pulse oximeters are discussed in this paper.

Arduino Based Door Automation System Using Ultrasonic Sensor and Servo Motor: This report proposed automatic door lock which works on the principle of breaking an infrared beam of light, sensed by a photodiode. It consists of transmitting infrared diodes and receiving photo-diodes. The system is to detect whether someone is coming in or not. The photodiodes are connected to comparators, which give a lower output when the beam is broken and high output when transmitting normally

RFID Based Aadhaar Card: SMART AADHAAR card with RFID Technology paper is mainly based on RFID technology. In this paper the technology which is issued is Radio Frequency Identification technology (RFID); There is an RFID card which is issued to each and every citizen with a unique identification number in it. This card is used as unique identification number in various aspects like e-passport, smart parking, hospital details and driving license which has been explained in this paper.

Telemedicine as the New Outpatient Clinic Gone Digital: Position Paper From the Pandemic Health System REsilience PROGRAM (REPROGRAM) International Consortium (part 2): This article talks about how technology has become a great enabler for doctors in patient continuity through remote consultation, monitoring, and patient education using videoconferencing in the COVID19 pandemic era. It says that limiting the exposure of patients, healthcare workers, and systems is critical in controlling the viral spread and telemedicine offers that opportunity to improve health systems delivery, access, and efficiency. This article critically examines the current telemedicine landscape and challenges in its adoption, toward remote/tele-delivery of care, across various medical specialties.

Access, utilization, perceived quality, and satisfaction with health services at Mohalla (Community) Clinics of Delhi, India: This article talks about Mohalla or Community clinic, Delhi, India. Studies show that more than half to two-third of beneficiaries at these clinics were women, elderly, poor, and with school education up to primary level. One-third to two-third of all beneficiaries had come to the government primary care facility for the first time. A majority who attended clinics lived within 10 min of walking distances. There was a high rate of satisfaction (around 90%) with overall services, doctor–patient interaction time and the people were willing to return for future health needs. Most beneficiaries received consultations, medicines, and diagnostics at no cost. A few challenges such as dispensing of medicines for shorter duration, lack of awareness about the exact location of the clinics, and services available among target beneficiaries, and the incomplete records maintenance and reporting system at facilities were identified.

3. Methodology

The main component of our project is the microcontroller. For the automatic door opening, we use a passive infrared sensor (PIR sensor). It is an electronic sensor that measures infrared (IR) light radiating from objects in its field of view. They are most often used in PIR-based motion detectors. A PIR sensor can detect changes in the amount of infrared radiation impinging upon it, which varies depending on the temperature and surface characteristics of the objects in front of the sensor. Once a patient enters the clinic, the PIR sensor will detect the entry and a red LED light gets turned on indicating the next patient outside to not enter. The next patient will be allowed to enter the smart medical clinic only after the current patient exits the clinic which will be detected by the PIR sensor and a green LED light gets turned on after the whole room is sanitized before the next patient enters.

LM35 is an integrated analog temperature sensor which is used to measure the temperature of the patients, whose electrical output is proportional to degree centigrade. LM35 Sensor does not require any external calibration or trimming to provide typical accuracies. The LM35's low output impedance, linear output, and precise inherent calibration make interfacing to readout or control circuitry especially easy.

Our project has assigned 4 unique RFID Keys to 4 people. When this RFID card is read by the RFID reader, it will display the RFID key in the LCD. If it matches with the ones we have configured in the code, it will display the information about the person. If the key is wrong, it will not show anything.

To convert the analog temperature signal to digital, ADC0804 is used. It is a 20- pin Single channel 8-bit ADC module. Meaning it can measure one ADC value from 0V to 5V and the precision when voltage reference (Vref –pin 9) is +5V is 19.53mV (Step size). That is for every increase of 19.53mV on the input side there will be an increase of 1 bit at the output side.

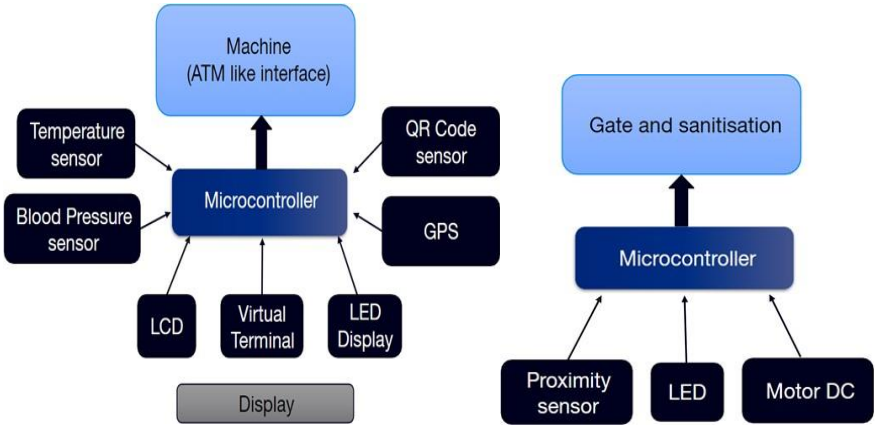


Fig. 1. Block Diagram of Smart Medical Clinic

To measure the blood pressure of the patients, an MPX4115 pressure sensor is used. The MPX4115 series is designed to sense absolute air pressure in an altimeter or barometer (BAP) applications. Motorola's BAP sensor integrated on chip, bipolar op amp circuitry and thin film resistor networks to provide a high level analog output signal and temperature compensation. The small form factor and high reliability of on chip integration makes the Motorola BAP sensor a logical and economical choice for application designers. The Global Positioning System (GPS) tells you where you are on Earth. The Global Positioning System, originally NAVSTAR GPS, is a satellite-based radio navigation system owned by the United States government and operated by the United States Space Force.

4. Working & Simulation

a) Automatic door opening with sanitization control: When a person exits the clinic, the proximity sensor detects and it sends the signal and opens the gate automatically by rotating the motors in clockwise direction, at the same time red LED is displayed, as the sanitization process starts. Once the sanitization process is over, the LED displays green color, which indicates the next patient is good to go as shown in figure 2. After which the motor moves in an anticlockwise direction and closes the door.

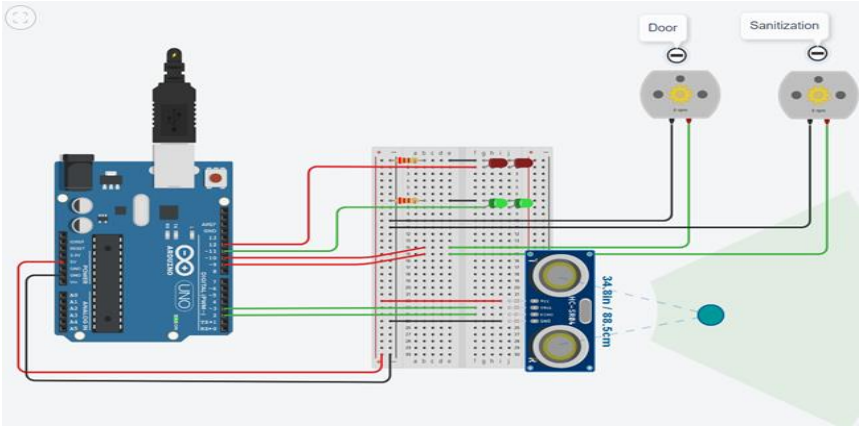


Fig. 2: Automatic door opening with sanitization control

b) Using the RFID scanner: The first step in accessing the machine is to enter the RFID key. We have made tags which are referring to a particular patient's /person's information, so when a person scans his or her ID proof though RFID sensor, if the tag is correct then the information about the person is displayed.

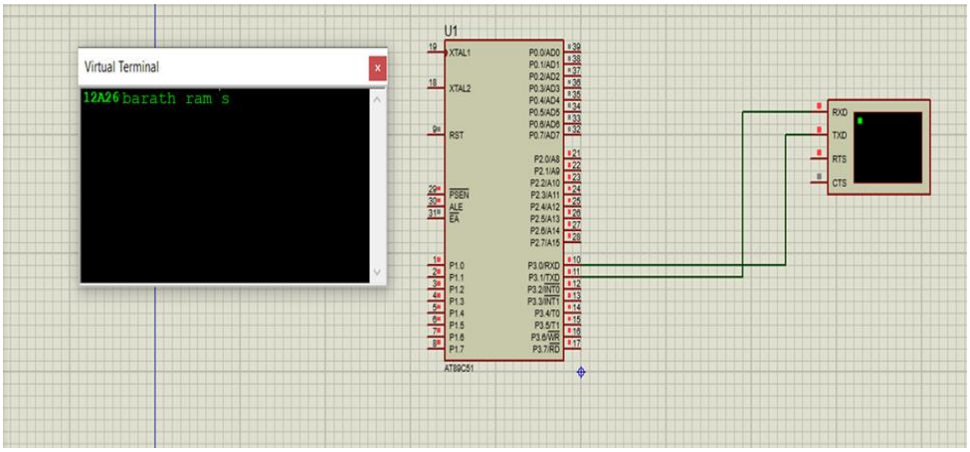


Fig. 3: Working of RFID Scanner

If the tag is wrong, no data is displayed. So as soon as the correct tag is detected it tells all the information about the patient which can be further used as shown in figure 3. So basically, it matches the bar code with the tag and if found, the information of the particular person is displayed.

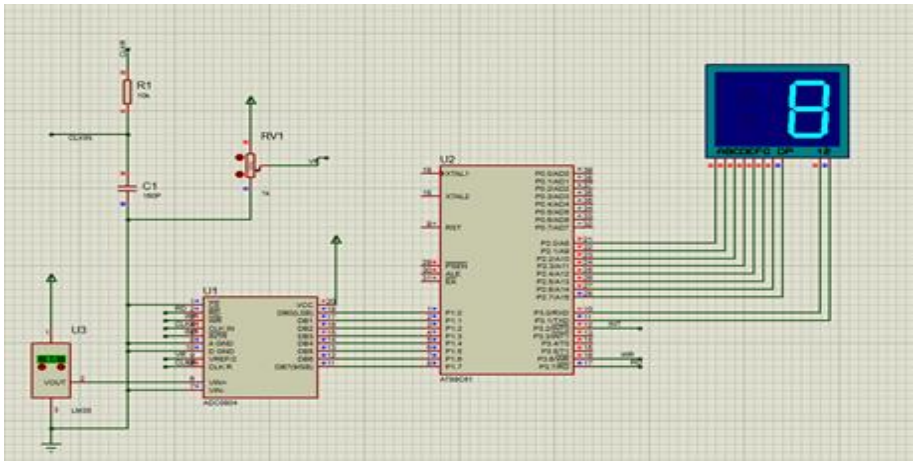


Fig. 4: Measuring the Temperature

c) Measuring Temperature: The patient's temperature is first measured using the LM35 sensor. It measures the value in Celsius scale. The output from the temperature sensor is in analog, to convert it to binary form, we use ADC0804, to convert analog signal to digital values. Port 1 will read the values from ADC when the conversion is completed by ADC, which is represented by the INTR pin when it goes from high to low and RD pin goes from

high to low. This conversion is done by the ADC when the WR pin goes from low to high. To Display the value sent from the ADC to the microcontroller, Port 2 of 8051 is connected to a 7-segment display, And the value is displayed on a 7-segment LED as shown in figure 4.

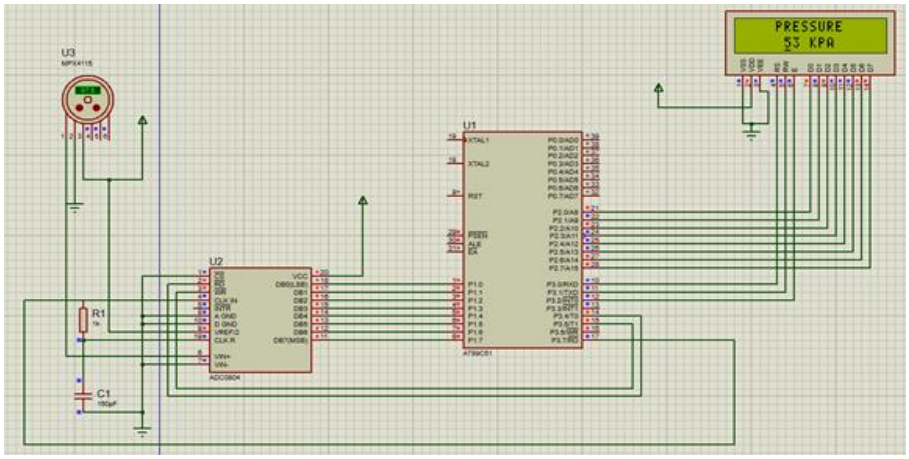


Fig. 5: Measuring the Blood Pressure

d) Measuring Blood Pressure: Patient's blood pressure is measured using a blood pressure sensor. Here we have used, MPX4115 pressure sensor to do the same. The output from this pressure sensor is analog, and we cannot send the analog values to the microcontroller, so we first convert it into binary using ADC0804, which does the conversion in the same fashion it did for the temperature sensor. The values are sent to port 1 of the microcontroller. And the value will be displayed on the LCD screen as shown in figure 5.

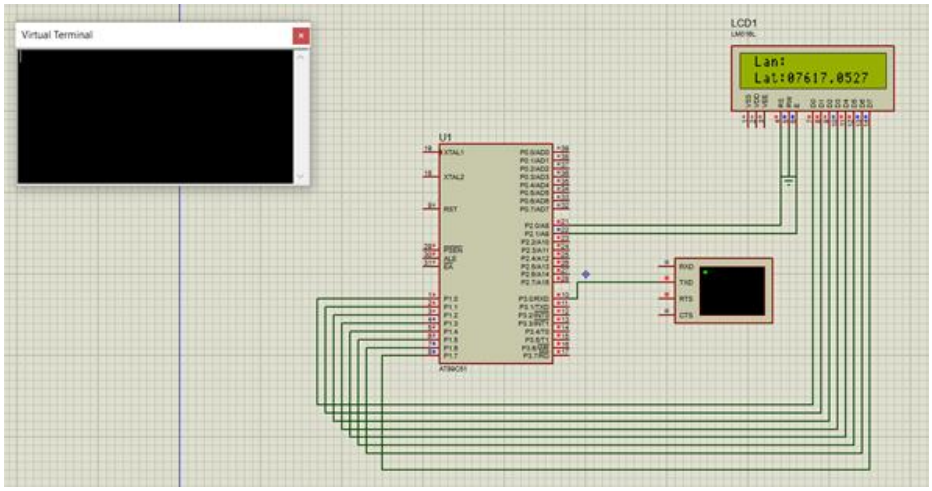


Fig. 6. Global Positioning System

e) GPS system: The GPS System will guide the patients by providing them with GPS coordinates of the hospital or Pathological labs which are nearest to the Smart medical clinic. The GPS module calculates the position by reading the signals that are transmitted by the

satellites. Each satellite transmits the message continuously which contains the time it was sent. GPS receivers measure the distance to each satellite based on the arrival time of each message. This information is used to calculate the position of the GPS receiver. The received raw data is Converted for the user as LATITUDE, LONGITUDE, ALTITUDE, SPEED and TIME.

The GPS receiver continuously transmits the data as per the NMEA standards using RS232 protocol. In this NMEA format, the LATITUDE and LONGITUDE values of location are available in GPRMC sentences. In this project LATITUDE and LONGITUDE values are extracted from NMEA format and displayed on LCD. We have to receive the data to the controller from the GPS module serially using UART protocol and now extract the latitude and longitude values from the received messages and display them on LCD.

If you want to check the latitude and longitude values of the location without any coding, then use Trimble Studio software. This software directly provides latitude, longitude, altitude, speed, time and date when you interface the GPS module as shown in figure 6. It also provides your location on Google maps for people with access to smart phones.

5. Results & Discussion

When a person comes near the door, the proximity sensor detects and it opens the gate automatically by rotating the motors, if the sanitization process is completed. There are LED and LCD displays outside the door, if the LED displays green color: it means the person is good to go. The LCD displays "Welcome to Smart Medical Clinic". After the door opens automatically using the PIR sensor, the person enters into the Smart Medical clinic, there will be a smart medical device which is an ATM-like machine with a similar user interface.

The first step to accessing the machine is by entering the RFID Tag. if the tag is correct then the information about the person is displayed. If the key is wrong, no data is displayed. The Patient's temperature is first measured using the LM35 sensor. It Measures the value in Celsius scale. And the value is displayed on a 7-segment LED. Next patient's blood pressure is measured using the blood pressure MPX4115 sensor and the value will be displayed on the LCD screen.

The doctors can advise the patient for some diagnostic tests and also may ask the patient to visit the hospital for further detailed check-up. So, the GPS System will guide the patients by providing them with GPS coordinates of the hospital or Pathological labs which are nearest to the Smart medical clinic.

After the patient consults the doctor via a video call, the patient leaves the place through the automatic opening gate. Once he leaves, the motor rotates, and the sanitization pumps get turned on and sanitize the complete place. During the process, the red LED turns on to indicate the next patient to wait before entering.

Advantages

1. Contactless, safe, innovative and effective way to diagnose patients, especially during this pandemic.
2. People from rural areas, where there are not many hospitals can consult doctors anytime using this
3. Usually in normal hospitals, patients have to wait hours before meeting a doctor. With this the waiting time for patients will be reduced.
4. While waiting in hospitals, there is always a chance to get infected from the people around us. Here, the spread of communicable disease can be reduced.
5. People don't have to travel far to consult some good and experienced doctors anymore.

Disadvantages

1. It cannot compensate for the direct interaction between a doctor and the patients.
2. Regular Maintenance Is Needed
3. Since this machine's functionality is similar to that of an ATM machine, people from rural areas could find it difficult to use, so they might need some help to use this machine.
4. The positioning of the GPS system cannot be fully accurate all the time.
5. It Might Be Expensive To Install And Maintain.

6. Conclusion & Future Scopes

This project presents a novel, innovative medical clinic, which addresses the issue of COVID-19. We designed it in a way to ensure safety for the patients, for which we have sanitization happening every time the person leaves. We have a lot of scope for expansion in the future, we can integrate many more sensors, also a compartment to store test samples of the patients, which will be collected later by hospital staff.

We can further extend where patients can choose to receive the prescription either via text or via image on their Aarogya Sethu App or get printed paper from the machine (for users who don't have a Smartphone, just like an ATM machine). Such medical clinics will benefit so many people and contain the spread of coronavirus and other contagious diseases. Good health care facilities will be easily provided in rural areas. These health care facilities will be in reach of everyone and will be easily accessible. With the help of a GPS module, a person new to a place also can easily locate the nearest health center and labs for testing. Rush in the hospital will be avoided and the waiting time of the patient will be reduced.

References

1. Orji, Everistus & Cv, Oleka & Nduanya, Ujunwa (2018) Arduino Based Door Automation System Using Ultrasonic Sensor and Servo Motor. The Journal of Scientific and Engineering Research. 05. 341-349.
2. Cheung BM, Chan LS, Lauder IJ, Kumana CR. Detection of body temperature with infrared thermography: accuracy in detection of fever. *Hong Kong Med J*. 2012;18 Suppl 3:31-34.
3. Jubran, Amal (2004) Pulse oximetry. Intensive care medicine. 30. 10.1007/s00134-004-2399-x.
4. Mrs. R. Angeline, Aditya Sharma, Ahitagnee Paul, Saurav Sanyal, "RFID Based Aadhaar Card". International Journal of Computer Trends and Technology (IJCTT) V58 (1):25-28, April 2018.
5. Tian, Shuo & Yang, Wenbo & Le Grange, Jehane Michael & Wang, Peng & Huang, Wei & Ye, Zhewei. (2019). Smart healthcare: making medical care more intelligent. Global Health Journal. 3. 10.1016/j.glohj.2019.07.001.
6. Bhaskar, Sonu & Bradley, Sian & Chattu, Vijay Kumar & Adisesh, Anil & Nurtazina, Alma & Kyrykbayeva, Saltanat & Sakhamuri, Sateesh & Moguilner, Sebastian & Pandya, Shawna & Schroeder, Starr & Banach, Maciej & Ray, Daniel. (2020). Telemedicine as the New Outpatient Clinic Gone Digital: Position Paper from the Pandemic Health System REsilience PROGRAM (REPROGRAM) International Consortium (Part 2). Frontiers in Public Health. 8. 10.3389/fpubh.2020.00410.
7. Lahariya C. Access, utilization, perceived quality, and satisfaction with health services at Mohalla (Community) Clinics of Delhi, India. *J Family Med Prim Care*. 2020;9(12):5872-5880. Published 2020 Dec 31. doi:10.4103/jfmpc.jfmpc_1574_20
8. Encardio rite, Temperature Sensors: Types, How It Works, & Applications, 2019
9. Naziyok, Tolga & Zeleke, Atinkut & Röhrig, Rainer. (2016). Contactless Patient Monitoring for General Wards: A Systematic Technology Review. Studies in health technology and informatics. 228. 707-711. 10.3233/978-1-61499-678-1-707.
10. Jamali, Mohammad. (2017). A Comparative Analysis upon Mobile Healthcare System and Services: A Case Study of Rural India.

Quality Management in University Education in the Context of Online Education

Alexandru-Lucian Manole, Cristian-Marian Barbu, Cătălin Deatcu

"ARTIFEX" University of Bucharest, Romania
amanole@artifex.org.ro, cbarbu@artifex.org.ro, cdeatcu@artifex.org.ro

Abstract. The paper reviews some of the authors' observations and preoccupations for quality management in the higher education system under the specific conditions posed by the online education. The internal quality management is a major component of the public responsibility of Romanian universities, as stated by the legal framework, and the internal quality assurance is an important process of the quality management. There are specific considerations to be set, monitored and reviewed with corrective actions, if necessary, according to the plan-do-check-act cycle, when online education is involved. In this paper, the authors propose some dimensions, and several indicators, that can be approached by the strategy and policies of the university in the aim to provide a proper quality in the online educational activities.

Keywords. Quality management, education, online, stakeholders, technology.

1. Introduction

Quality management in the academic system is a process whose organization, development and, most important, results and actual contribution to the progress of specific activities is of interest for both the university and stakeholders: students, financing authorities or organizations (founders, for private institutions), employers, graduates, quality assurance agencies, other national authorities, the public.

Online education, even if it is not a new phenomenon, has been brought forward as an alternative mean to ensure continuity of the education, at all levels of the system (from kindergarten to higher education), by the outbreak of the COVID-19 pandemics, whose occurrence and spread put a serious danger on the health of the population in the affected countries. One of the measures adopted by the national authorities was the restriction of on-site activities in the educational institutions. In these conditions, the option to migrate towards online activities become obvious and it was adopted as a solution in the face of the restrictions enforced in order to counter the dangers associated with the pandemics.

The process to migrate to a "more" online education was not easy, it involved allocation of financial and technical resources, dedication and professionalism of people involved, acceptance by the students.

The transition towards online education involves specific issues regarding the quality. Apart from the obligation and public responsibility to comply with external and internal quality standards, a proper approach towards quality can lead to an improved activity of the university, together with a better image among the students and a favorable context for the sustainable development of the institution. Thus, special attention must be given to quality aspects, as no educational activity can be developed outside the QA framework, and the benefits, even indirect and long-term ones, are not to be ignored.

2. Literature review

There are two main types of online education: synchronous (i.e. videoconference) and asynchronous processes (e-mail, course and homework/projects upload/download etc.). All the specific mechanisms associated with either synchronous and asynchronous education must be adopted and adapted according to the envisioned quality standards.

The study of Esfijani [1] on the quality of online education outlines resources, inputs and processes as specific factors. The quality of the resources themselves has a significant impact on the quality of education and, in the context of the online component, a great weight is associated with the technical resources, and on the technical skills of the staff who exploits these resources. Technical resources and trained staff involve expenses with acquisition of the respective equipment and services. Thus, a major preoccupation of the academic management is to skillfully allocate financial resources to cover these expenses.

Quality assurance in online education is not a new topic in the preoccupations of the researchers. The study of McGorry [2] builds a quality assessment model based on several factors, among which we outline technology, technical support, student satisfaction.

The quality assurance model proposed by Marciniak [3] outlines several components of an online educational program, including documentation and electronic resources, who are integrated in the quality assurance model proposed, and the authors emphasizes the desirable characteristics of these dimensions. Marciniak and Cáliz Rivera [4] have developed a system of indicators for quality assessment dedicated to didactic materials.

O'Keefe et al. [5] have published an elaborated handbook on organizing online courses, useful for both individuals and institutions, the study covers, most importantly, several aspects on the transition from classical education to online education. The availability of training on the use of technical resources is outlined, along with support for students. The complexity of effects of the pandemics on students' "surrounding environment" becomes more and more clear.

Xu and Xu [6] raise an important issue in the present quality assurance framework: there is a higher risk of individual and, therefore, accumulated dropout in the online environment. Furthermore, this study reflects the situation before the COVID-19 pandemics, the effects might be more adverse in this context. Significant dropout might jeopardize the financial capacity of the university, and it might have impact on the external quality evaluation process, where the maximum capacity of enrollment can be diminished, if the evaluators consider that the number of graduates, and even students in upper classes is far inferior against the number of students who have been accepted in the first year.

Undoubtedly, the continuous improvement of the technical resources, especially of the ones that come into direct contact with the beneficiaries, cannot be overlooked. One useful measure involves collecting opinions of the beneficiaries on what they like, what they do not like and what can be improved in order to offer them a more valuable experience. Chen et al. [7] have presented an analysis of the user satisfaction in the environment of online education platforms, their study covers China and reflects the context of the COVID-19 pandemics (its beginning, actually). One major conclusion of their study is that a significant factor in building the satisfaction of the user is the availability of the platform. Sun et al. [8] describe further effects of the COVID-19 pandemics on the Chinese higher education environment. The results of the survey presented by them is relevant for other researchers and professionals interested in collecting the opinions of the students on the online education. Butnaru et al. [9]

3. Aspects of quality management specific to online education

When thinking of quality assurance for online education, the construction of the criteria and indicators should take into consideration all the dimensions of quality: processes, resources, expected outputs, expectations of the stakeholders etc.

One institutional approach towards quality management in online education could consider the following framework:

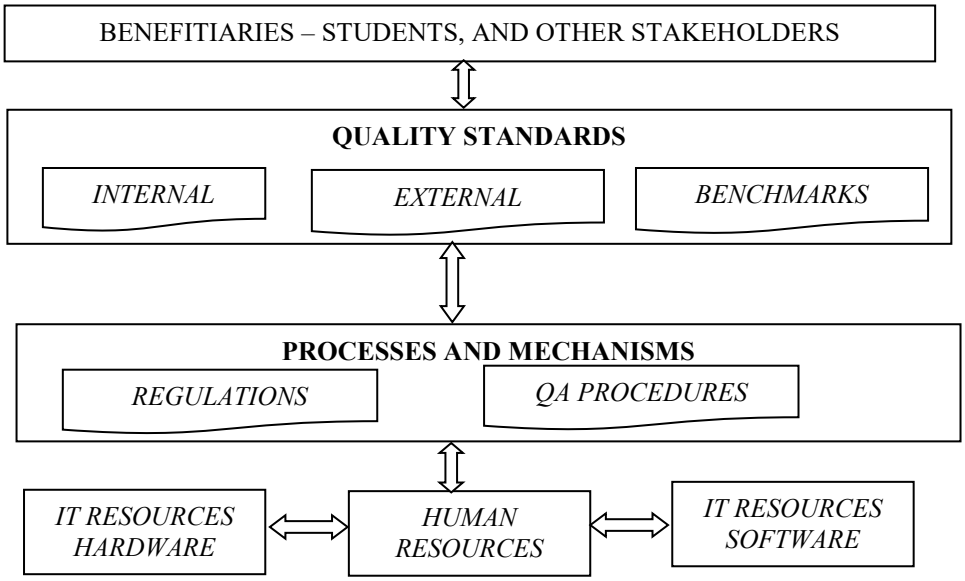


Fig. 1. An overlook on the quality assurance framework in an online education environment (Source: Authors' representation)

In this study, the aspects and indicators outlined by the authors refer to the IT infrastructure (hardware, software, network), the human resources (staff and students), institutional library and services of support for students, as dimensions of quality management and improvement.

A. The IT infrastructure. Resources and maintenance:

a) The number of fixed or portable computers/workstations dedicated for online educations (*NoC*). This number can be evaluated in accordance to the necessity for computers, in a similar way that the educational spaces (lecture or seminar halls) are assessed:
   a. is the number of posts (stations) enough to cover the schedule of classes without superpositions or interruptions?

$$NoC_{actual} =, <, >, NoC_{necessary} \quad (1)$$

where:
NoC_{actual} = actual number of computers
$NoC_{necessary}$ = necessary number of computers

b. number of posts (stations) per teacher or per student, in case of full online education:

$$NoCt = \frac{NoC_{actual}}{NoTh} \quad (2)$$

where:
$NoCt$ = number of posts (stations) per teacher
$NoTh$ = number of teachers

c. number of posts (stations) per student, especially important in case of hybrid online education, when students, or at least part of them, do come in campus to run specific activities:

$$NoCs = \frac{NoC_{actual}}{NoSt} \quad (3)$$

where:
$NoCs$ = number of posts (stations) per student
$NoSt$ = number of students

Corroborated with that: are there procedures and reserve equipment for contingencies?

d. What is the percentage of reserve equipment?

$$Wre = \frac{NoC_{actual} - NoC_{necessary}}{NoC_{actual}} * 100 \quad (4)$$

where:
Wre = percentage of reserve equipment / computers.

This percentage can also be computed considering (if any) the spare parts available in the storage (though this is uncommon), considered, each one in its category, as contributing to the functioning of one computer

e. What is the time necessary to deal with any technical issue that could have impact on the educational process? An average "lag" between the time of the occurrence of any issue and the time at which the issue is solved can be calculated as such:

$$Time\_lag = \frac{\sum(\text{ToIE} - \text{ToIS})}{\text{NoE}} \quad (5)$$

where:
ToIE: time at which the issue occurred;
ToIS: time at which the issue was solved, and functionality restored;
NoE: number of events.

b) The hardware/software configuration of the computers. This can be subjected to benchmarks against reference criteria such as:

a. The minimum / recommended configurations necessary to run applications dedicated to support the online educational processes, including videoconference platforms, educational software for certain topics (data management and analysis applications, programs designed to support management processes, decision support software, database management systems etc.) – are they compliant? Is the configuration far above the maximum recommended one (maximum, considering the ensemble of software running on an equipment)?;
b. Configurations used by other universities that, in support of their own educational processes, employ the same type of applications – benchmarking, if there is data available;
c. The quality of the educational software:
   - Does the function set of the software covers the entire or at least a significant part of the contents of each specific discipline?
     Percentage of topics from syllabuses covered by educational software functionalities:

$$ESC = \frac{\sum_{i=1}^{n} \mathrm{NoT}c}{\sum_{i=1}^{n} \mathrm{NoT}} * 100 \quad (6)$$

where:

ESC = educational software coverage;
$NoTc$= number of topics covered by educational software, per discipline;
NoT= total number of topics per discipline;
n = number of disciplines.

   - Are there official documentations available?
     Percentage of educational software covered by official documentation:

$$ESCD = \frac{\mathrm{No}SDc}{\mathrm{No}SDt} * 100 \quad (7)$$

where:

$ESCD$ = educational software coverage with documentation;
NoSDc = number of educational software covered with official documentation;
NoSDt = number of educational software - total;

This computation can be made by counting the individual software applications, or by counting the licenses available, when there are specific licenses acquired per user or per workstation

   - Are the members of the academic staff trained/certified to use the educational software? Is there specific documentation oriented towards the specific usage of software within a specific academic process, eventually authored by the academic staff of the university?
     Percentage of staff trained in use of the educational software:

$$ESTT = \frac{\mathrm{No}Tt}{\mathrm{No}T} * 100 \quad (8)$$

where:

$ESTT$ = percentage of teachers trained in educational software;
NoTt = number of teachers who have passed training;
NoT = number of teachers, total.

d. Is the software readily available for students?
   - If it is free/open source, how easy/difficult is for students to install on their own computers?
     The degree of difficulty can be assessed/monitored by applying specially designed questionnaires to the students. By processing the data in an adequate manner, the situation can be improved for the next cohort of students that use the same software.
   - Are there resources available regarding the installment of the software? Is there availability on behalf of the staff to help them in a synchronous environment?
     A similar indicator to *ESC* can be applied, by comparing the number of software for which installation resources are available against the total number of software.

e. If the software license involves a fee, does the university provides the software by acquiring itself the licenses for the use of students?
   The university can publish and make available to students and other stakeholders the amount of budgetary resources allocated to this type of operation.

f. Given the contemporary orientations on the academic ethics, use of plagiarism detection software is increasing, being made mandatory by internal and national regulations. Several issues are to be considered:
   - Do the students have direct access to the software (licenses per user)?
   - Is there a semi-automated form of communication with the operators of the anti-plagiarism software (e-mail, a module of the platform/portal with upload/download functionalities)?
   - Does the university offer free checks of documents against plagiarism?
   - If there is a fee, what is the relative size, compared with another relevant fee(s)?

c) The functions of the online software platform(s) adopted by the university:
   a. Is the platform a recognized solution or has it been developed in-situ?
   b. What is the capacity of the solution provider (years of experience, portfolio of products, list of customers)? Can the provider be characterized as reliable?
   c. What does the maintenance agreement stipulate on issues such as online, 24/7 support, costs of maintenance services compared to the costs of the platform itself, accepted downtime;
   d. Is the platform accessible outside the university campus?;
   e. What is the average time for offering proper response to technical incidents?
   f. What are the functionalities of the platform? Examples that can be pursued:
      - Electronic catalogue, with grades visible only to the teacher and the individual student for which the respective grades have been „assigned" by the teacher;
      - Support for statistical calculations on grades, credits (percentage of students who have promoted each session/semester, year etc. or scholar situation for each student), with rights assigned to registrars/dean's offices etc.;

- Capability to allow evaluation of teachers by the students and by their colleagues (collegial evaluation);
- Mechanisms that allow posting of course materials that the students are entitled to receive, in accordance with the Code of Rights and Obligations of the Student [10];
- Instruments that allow the upload of homework files, test files, projects by the students, according to the requirements of the teachers and the syllabus;
- Modules that implement a live chat solution, in which the students and teachers can interact within the schedule agreed by the teacher and then approved by the faculty/department management;
- Does the platform include interaction with the synchronous videoconference instrument used by the university? How this interaction runs (description of integrated processes should be documented as rigorously as possible)?;
- Implementation of the appropriate security and confidentiality processes, including compliance with the GDPR.

d) The capacity and characteristics of the network:
   a. Does the university have a broadband internet connection and public IP address?
   b. Is the performance of the connection in accordance to the needs of online synchronous activities, even if at "peak hours"?
      Indicators can be used to measure and monitor:
      - Average characteristics of the connection, against the number of teachers, respectively the number of students:
   c. Has the university its own internet domain and/or its own portal?
   d. Has the university its own site?
      - Is the site supported on a reliable platform, that provides integrity, security and confidentiality of information posted on the site?
      - Does the site include only static information, or it allows access to any dynamic function – interrogation of academic or financial situation?
      - Does the site include facilities for online payment – connection with an online payment processing company?
      - Are there institutional emails that can be used for official communications with students?
      - The students have the right/obligation to create institutional email accounts?

   Note: the same questions can be "answered" against the portal.

B. Human resource - training of the personnel and students. In this respect, the following aspects should be of interest:

a) The training regarding the own platform:
   - Has the solution provider included a documentation in the platform package?
   - Is the documentation updated whenever the platform receives a significant update in its main functionalities?
   - Are there alternative training materials created for the users of the platform – namely video or other form of tutorials, examples, case studies etc.

- Has the personnel received a formal training before the platform enters the "production stage"?
- Have the training programs concluded with an official, recognized certification issued by the provider?

b) Training regarding specifics of online education:
- Does the university include, in the organization chart, a structure that supports training in online education?
- Within this structure, is there appropriate personnel, certified in providing this kind of training?
- Do these persons have teaching and/or scientific degrees in the field of education sciences or similar fields, related to the specifics of training they provide?
- If no structure exists, are there teachers in the university that have the necessary competences to provide this kind of training to colleagues? Is the same kind of training provided to students? What is the number of students that have participated in training?

C. Institutional library – online resources:

a) Does the library offer access to online documentation?
b) Is the library connected to databases offering literature in the fields of interest for the university?
c) What are the characteristics of the license used to access the databases? Does it support unlimited (limited only by the number of students) access from outside the campus? Is the performance, as perceived by the users, limited in any way by the number of connections at a given time?
d) Is the library staff trained/certified in using the software supporting the online library?
e) Are all the courses in the curriculum provided with course materials stored in the online library?
f) Is access to library provided according to a proper security policy?

D. Support and counseling services for students. The study of O'Keefe et al. [3] outlined the major challenges that "change" the life of students. These issues must be properly dealt with, which increases the responsibility of the university towards the welfare of the students.

a) Do the students work? If so, they work from home or they can go at the office? Is the career counselling structure able to provide them online assistance, via individual or group videoconference session (or calls, chat/messenger-like applications);
b) Do the students feel the need to interact with their colleagues, even if physical meetings are not allowed? Then, does the university act in a pro-active way, offering its videoconference platform to allow informal meetings between students?
c) Is the videoconference a useful tool for tutorship and collegial tutorship? Are those forms of tutorship considered to be of help by the students?
d) How one can characterize the availability of teachers and management of the university to interact with students, even under the form of private online meetings?

4. Conclusion

Undoubtedly, online education is a constant of the times to come, regardless of the future evolution of the pandemic context. All universities have to invest time and other resources in order to offer a better, competitive answer to the challenges. The external quality assurance standards, which are conditioning the accreditation of the university/study program, can be adapted in the future in order to address the online education forms, as this experience, along with the resources and equipment/platforms acquired, together with the skills and competences of academic staff, can be capitalized in the pursuit for improved quality.

References

1. Esfijani A. (2018). Measuring Quality in Online Education: A Meta-synthesis, *American Journal of Distance Education*, 32:1, 57-73, DOI: 10.1080/08923647.2018.1417658
2. McGorry, S. Y. (2003). Measuring quality in online programs. *The Internet and Higher Education*, *6*(2), 159-177.
3. Marciniak, R. (2018). Quality Assurance for Online Higher Education Programmes: Design and Validation of an Integrative Assessment Model Applicable to Spanish Universities. *International Review of Research in Open and Distributed Learning*, *19*(2). https://doi.org/10.19173/irrodl.v19i2.3443
4. Marciniak, R. & Cáliz Rivera, C. (2021). A System of Indicators for the Quality Assessment of Didactic Materials in Online Education. *International Review of Research in Open and Distributed Learning*, *22*(1), 180–198. https://doi.org/10.19173/irrodl.v22i1.5069
5. O'Keefe, L., Rafferty, J., Gunder, A., Vignare, K. (2020, May 18). Delivering high-quality instruction online in response to COVID-19: Faculty playbook. Every Learner Everywhere. http://www.everylearnereverywhere.org/resources
6. Xu, D., & Xu, Y. (2019). The Promises and Limits of Online Higher Education: Understanding How Distance Education Affects Access, Cost, and Quality. *American Enterprise Institute*. https://files.eric.ed.gov/fulltext/ED596296.pdf
7. Chen T, Peng L, Yin X, Rong J, Yang J, Cong G. (2020). Analysis of User Satisfaction with Online Education Platforms in China during the COVID-19 Pandemic. *Healthcare*. 2020; 8(3):200. https://doi.org/10.3390/healthcare8030200
8. Sun, L., Tang, Y. & Zuo, W. Coronavirus pushes education online. *Nat. Mater.* 19, 687 (2020). https://doi.org/10.1038/s41563-020-0678-8
9. Butnaru GI, Niță V, Anichiti A, Brînză G. The Effectiveness of Online Education during Covid 19 Pandemic—A Comparative Analysis between the Perceptions of Academic Students and High School Students from Romania. *Sustainability*. 2021; 13(9):5311. https://doi.org/10.3390/su13095311
10. Romanian Ministry of Education (2012). Order no. 3666 from 30 March 2022 regarding the approval of the Code of Student's Rights and Obligations, Romanian Official Gazette no. 225 from 4 April 2012, retrieved from https://www.edu.ro/sites/default/files/_fi%C8%99iere/Invatamant-Superior/2016/studenti/OM_3666_2012_CODUL_DREPTURILOR_OBLIGATIILOR_STUDENTILOR%20.pdf

Interdisciplinary Opportunities in the Financial Planning of Romanian Holding

Ghenadie Ciobanu[1], Amalia Diaconu[2], Dinu Mihai[3], Victor Adrian Troaca[3]

[1] National Scientific Research Institute for Labour and Social Protection (INCSMPS), Bucharest, Romania
[2] ARTIFEX University, Bucharest, Romania
[3] Bucharest University of Economic Studies, Bucharest,Romania
gciobanu019@gmail.com

Abstract. In the conditions of developing the new knowledge economy, having as priorities intelligent growth, sustainable development and favourable inclusions, but also as a result of the Covid-19 pandemic crisis the role of economic agents, companies, enterprises and large companies is a priority. An important role in the forms of organizing the economic activity of the companies Belongs to the holding companies, which have been much discussed in Romania. Considering the fact that this form of organization exists in the economic activity in Romania, and after the pandemic crisis many companies go through crises of different order, we propose in this article to show the advantages of holding companies and to show the need for financial planning in the Activity of these economic agents.

Keywords. Holding, financial planning, fiscal policy, interdisciplinary, entropy, synergy, value.

1. Introduction

In the current conditions for a company to survive, it is not enough just to show initiative in business and get the primary capital, but it must also be able to calculate and predict how this capital will be used in each stage of activity. That is why planning is one of the most important elements of a successful business, because only with its help can be predicted and avoided possible internal and external risks.

Local entrepreneurs often have a careless attitude towards planning [1]. In Romania, over 99% of all enterprises are micro, small and medium enterprises. The widespread view is that the domestic SME sector does not focus on financial planning. Financial planning formulates ways and methods by which the company's financial objectives can be achieved.

However, taking into account the universal experience, we realize that precisely in such situations, of market fluctuation and intensification of competition, the importance of planning must increase.

Therefore, due to this business environment characterized by increasing changes, long-term planning is critical for both the survival and growth of an enterprise, a business.

Long-term (strategic) planning involves four steps in its completion: establishing the mission of the enterprise; defining the objectives generated by the mission of the enterprise; formulating the strategy for achieving the objectives; strategy implementation.

In the course of this process, resources have a central place: a strategy cannot be completed without the existence and consideration of the human, physical and financial resources of the enterprise. If the resources are not available, the strategy and, implicitly, the planning must be completed with the ways of obtaining them.

Consequently, a first and important role of financial management derives in direct connection with the strategic planning process, and consists in the construction of the long-term financial plan, from which the main tool of short-term financial planning, the operational budget, can be outlined.

More specifically, in developing long-term financial planning, financial management has the following important tasks: of quantitative identification of the capital necessary to reach the objectives of the enterprise; to identify in the structure of the capital necessary to achieve the objectives of the enterprise - to establish a target share of capital from external sources.

Regarding the operational budget built on the basis of the strategic financial plan, must fulfil 3 functions: basic landmark in the management of financial planning in the short term; necessary tool in completing the long-term financial plan (therefore it must always be linked to the latter); starting point in the construction of short-term financial budgets. In other words, the operational budget must be the link between the long-term financial horizons, the short-term financial horizons. The financial plan of an enterprise also sets the standards needed to measure the performance of the enterprise. Before the plans can be put into practice, the top management of the company must designate responsibilities and establish points of tolerance. These tolerances specify the minimum levels of performance (e.g., sales volume, rate of return on investment, etc.) and the maximum cost levels involved in each action of the plan. There is thus the necessary framework for future management corrective actions.

The purpose of the article is to investigate the essence and importance of financial planning of the investment project in Holding in the current conditions. Achieving this goal involved studying the following tasks: determining the concept of investment project and its classification; identification of the essence and operation of the Holding; studying and specifying the notion of financial planning; highlighting the basic features of the mechanism of elaboration of the investment project; the general characterization of the structure, the organization of the activity and the importance of the Holding for the Romanian economy; determining the influence of the Holding at international level. Important lack of serious financial projections in Entrepreneurial business, especially when it comes to cash flow projections [2].

2. Fiscal Provision in Romania – Support the Activity of Holding Companies

The activity of holding companies in Romania is not, regulated from a legislative point of view, which would detail the legal and fiscal regime, the general framework of the economic, financial and legal relations between the holding company and its subsidiaries and between its subsidiaries.

In order to be able to choose the legislation through which the holding companies of a group are established, in the fields: industrial, agricultural, trade, Ho-Re-Ca, or other activities, it depends on the operational approaches of infrastructure, language, time zone, administration costs, regulations, laws, etc., of the fiscal advantages associated with the regime. In a perfect jurisdiction. It is welcome not to be charged VAT or other taxes on sales or consumption, nor site taxes on various transactions between the founding companies and their subsidiaries. Romanian taxation allows several provisions in support of holding companies: tax exemptions for dividend income and for gains from the sale of shares or from the liquidation of companies.

Therefore, the respective provisions increase the attractiveness of the Romanian fiscal policy for investors, together with the measures of reduced single rate of taxation of profits and incomes of individuals, additional deductions for research and development activities,

maintaining the income tax exemption of persons creating programs on the computer, the possibility to carry over the internal fiscal credit etc.

From an operational point of view, the use of a holding company (company), within a group of companies, could have certain advantages: Consolidation of subsidiaries, for the implementation of an efficient and qualitative management for companies, for the provision of priority services.

Services for: financial-accounting, social responsibility reporting; flexible ways of acquiring or selling shares in subsidiaries; reinvestment of profits directly from the holding company; creating new opportunities to attract investors, both minorities and important investors; creating platforms for the acquisition of future businesses, creating efficient and viable partnerships, for the development of new businesses; Expanding and growing in new markets, developing new regions and new geographical areas. Development and implementation of innovation potential, which can be, capitalized and applied in the activity of new products and services. The respective advantages are valid for the groups of multinational companies, or with a regional presence, but at the same time, the group of local companies from Romania must implemented.

From the point of view of fiscal policy, the presence of a holding company in the activity of a group can have sufficient advantages such: reduction of capital gains taxation, consolidation of profits/losses at group level. The planning will be more efficient in more activities of the companies, but with priority the taxes and fees paid by the companies. Advantageous fiscal policy of repatriation of profit from subsidiaries; Implementation of the favourable provisions of the double taxation avoidance treaties, in situations when the holding company and the subsidiaries are in different regions and fiscal administrations, and collaboration agreements are signed between them. According to article 20, the Fiscal Code, starting with 2014(Law no. 123 of September 19, 2014) [3], the Romanian companies, which hold shares in other companies consider certain types of income as non-taxable when calculating the taxable profit and the profit tax: (a)" if the Romanian legal entity holds at least 25% for an uninterrupted period of 2 years, ending at the same time as the dividend payment date; b) differences in the value of equity securities, as a consequence of the establishment of reserves, benefits or issue premiums, in commercial companies (legal entities) in which securities are held in the form of shares; differences in the valuation of long-term financial investments; c) income derived from the cancellation of a company's obligations (legal person) (d) income obtained from: (1) cancellation of expenses that did not have deductions, (2) income from the reduction / cancellation of expenses, which did not have deductions, (3) income from the reduction / cancellation of provisions that did not receive deductions, (4) income from the recovery of non-deductible expenses.

According Decision no. 77 of February 5, 2014, Art. 20 [4], the following incomes are non-taxable when calculating the taxable profit) dividends that were received from a Romanian company (legal entity), or from a foreign company paying corporate tax or a tax similar to corporate tax, located in a third country, as defined in art. 20. Are non-taxable - income from the calculation of taxable profit, dividends received from the Romanian legal entity (Romanian company), or from a foreign legal entity (foreign company) paying profit tax or a tax similar to the tax on profit, located in a third country, as defined in art. 20 ^ 1 alin. (3), c), with which Romania has concluded an agreement to avoid double taxation. For an uninterrupted period of 1 year, at least 10% of the share capital of the legal entity that distributes dividends.

According Tax Code, Law no. 227/2015 applicable from 01.01.2016, Art. 23 [5] "Non-taxable income", when calculating the tax result, the following income is non-taxable: (a) dividends received from a Romanian legal entity; (b) dividends received from a foreign legal person paying corporate tax or a tax similar to corporate tax, situated in a third country, for an

uninterrupted period of one year, at least 10 % of the share capital of the legal entity distributing dividends; For Romanian or foreign investors, who want to develop a holding company in Romania, which holds shares in Romanian companies, it is important to know that the lack of control rules of foreign companies (CFC - Controlled Foreign Company). In Romania, there are no provisions on non-taxation (or non-withholding of interest income) or royalties between subsidiaries and holding companies. In Romania, there is only one regulation, regarding the payments of interest and royalties made by a Romanian company to a company from another EU member state. Another positive moment is that Romania has no provisions regarding the fiscal consolidation of the fiscal profit. Romanian holdings can manage equity portfolios. Under the conditions of the fiscal provisions that we have already discussed above, the activities recommended for Romanian holding companies are limited to owning or managing investment portfolios, nor to intra-group financing activities or to intellectual property rights management activities.

3. Bibliographic Study of Current Research in the Field of Holding Companies

Author Yenn-Ru Chen [6] examines the impact of corporate governance on the cash holding policies of firms with different investment opportunities. According to the author: It is difficult to determine the optimal level of cash holdings for "new economies" (companies in the computer, software, internet, telecommunications or networking industries), which require large amounts of capital for investments in business opportunities with high profitability potential. Unlike old economy, firms, for which investment opportunities are relatively limited, corporate governance in listed new economy firms can create shareholder protections that make investors willing to accept higher levels of corporate cash holdings. 1500 poor companies, evidence shows that CEO ownership and board independence affect cash holdings in listed and old companies with new economies differently.

Author Jagoda [7] presents results of empirical research on HR architecture by locating Decision -making and executive powers in 100 capital groups operating in Poland against the definitions of capital groups, their Polish origins, and the discussion focused on the main role played by a holding company that determines the type of capital group. The latter, in the perspective of model-oriented studies presented in the literature references, is the crucial determining element for the architecture of the HR function in an organizational structure of a group. The results of the research presented in the article, however, do not confirm the thesis presented, which seems to result mainly from the short history and specific nature of the capital groups established in Poland.

According to the authors Afshar-Jalili, Ilyadi and Esmaeilzadeh [8]: "In these modern times, knowledge management (KM) has attracted attention as an inseparable aspect of organizational leadership, especially in the public sectors". This paper aimed to detail the approach and methodology used in the Iranian National Water and Wastewater Engineering Company (NWW) to initiate, integrate and manage KM knowledge management in its 90 subsidiaries. The conclusions and further development plans proposed by the author directed towards the fact that this case study aims to establish how NWW wastewater has integrated knowledge management initiatives among 90 branches to initiate knowledge management in industry. This holding company has developed a roadmap for knowledge management, including the vision of knowledge management, the mission, the value propositions, as well as the strategies and activities of knowledge management. Finally, he used knowledge management software to facilitate knowledge in all W&W companies in 31 distinct provinces. However, these activities need to be improved. Therefore, the programming of the knowledge

management roadmap needs to be, reviewed based on the level of maturity of the knowledge management of the subsidiaries. In addition, some of the activities, such as knowledge assessment, should be, delegated to subsidiaries to speed up the process. The authors Robertson, Perkins and Taylor [9] contribute to the respective work on the current significant literature on neoliberalism and rural change and focus on multiple jobs and rural economic and social life in New Zealand: "The study of multiple jobs is a useful lens through which to examine aspects of economic and social change in rural areas. That research confirms that one way in which people have coped with the changing rural social and economic conditions associated with the neoliberal economic restructuring of the early 1980 is to adopt a multi employment strategy." The authors found ways to earn the desired income and have a satisfying job in a diversified rural economy, including agriculture, tourism and the service sector. This way of organizing their professional life and personal experiences and the financial resources derived from such jobs in turn affect the families and rural communities in which they live.

EU holdings, and access to EU directives, to the benefits of tax treaties, in the light of recent court decisions, (2020) Local contact, EY Global 13 Jan. 2020, the following conclusions are made, as clear as possible that the threshold for access to treaties (in the post-BEPS environment) is higher than in countries where treaties don't include, limitation on the granting of benefits. (2) The limitation of the Treaty or the Benefits Directive will arise as a result of comprehensive review of the purpose, effect and substance of the agreements. (3) The arrangements may be challenged, the actual and / or abusive beneficiary principles. (4) An income tax exemption should not be sufficient to deny the benefits of the tax treaty. (5) In cases where the benefits of the Treaties are economically significant, a strong defence against POEM, effective ownership and / or the principles of abusive purpose, either under EU law, based on the double taxation conventions, it will be annexed to the substantial economic basis underlying the legal relations between the entities in a corporate structure. (6) Holding companies have a wide range of forms, from simple legal intermediaries to companies headquartered worldwide or regionally, with significant integration with their branches. (7) Certain undertakings may require significant operational changes to ensure the benefits of the Treaty.

According to the authors Braga, Naftanaila, Avram [10]: "The concept of holding company, is characterized by the following aspects: company whose objective is to acquire other companies in order to have a portfolio of equity securities to ensure its controlling position; the holding company receives dividends from the controlled companies (subsidiaries) and participates in the financing of the allocation and/or guaranteeing of loans granted by subsidiaries and participation in the capital increase". According to the authors Bressan, Rammerstorfert, Weinmayer [11]: "The financial crisis has called into question the regulatory intervention on the financial markets. The arguments recommend that banking holding companies have clear preferences for double advantage, which does not have an unequivocal and explicit economic basis. "In this article, the authors analyse the effects of equity, debt and double leverage on the efficiency of banking holding companies. It shown that the efficiency of the banking company negatively affected by the financing of equity from the parent company to subsidiaries and this effect even more pronounced in the case of double advantage. The authors' conclusions indicate that further action needed by regulators to prevent inefficient double advantage financing, which can be, used to circumvent regulated capital requirements. Authors' conclusions in this paper, the authors obtained the efficiency of banking holding companies in the multidirectional efficiency analysis (MEA). This method allows the potential for improvement of each factor to be, taken into account by determining the overall level of efficiency.

According to the authors Di Salvatore and Avv Salvatore [12]: "The United Kingdom offers a favourable tax environment for the Italian entrepreneur. By using trusts and intermediate holding companies, an attractive and fiscally efficient structure can achieved." This article examines the possibilities. Italian entrepreneurial, demand for tax planning, techniques, through which Italian groups approach selected international markets, is growing rapidly. Choosing the location to operate their business worldwide is a priority. The elimination of certain capital gains taxes for UK holding companies, together with additional UK tax benefits, makes the UK an excellent choice. These other benefits include (i) the impressive number of tax treaties concluded; (ii) exemption for dividends received; and credited professional management. Approaching the international location. When dealing with cross-border transactions, a major concern is generally the effectiveness of the legal instruments to be, applied. Authors Jia, Li, Bian and Gan [13] using data from the 2014 China Family Panel Studies (CFPS) survey, investigated the effects of financial planning capacity and risk perception on choice. The study shows that households with a greater financial planning capacity are more likely to invest in financial markets and have a higher proportion of risky financial assets. Empirical results suggest that a higher level of risk perception leads to greater market share and ownership of risky assets. Compared to the insignificant effect of financial literacy, we find that financial planning capacity significantly affects household investment income, and high financial planning capacity tends to contribute to a positive return on investment.

The authors Jashim et al. [14] state: "Although corporate social responsibility is a well-researched concept, very few studies have studied organizational social responsibility in the face of internal (external) crises. This paper aims to examine the organizational social responses of different companies to COVID-19 as an external crisis, especially in terms of how organizations use their resources to address social issues and the real intention behind them, changing organizational factors that influence such behaviours. In the direction of design (methodology) approach, the authors explore the initiatives of seven organizations from three different industries in several countries and their actions following the COVID-19 pandemic.

4. Methodology of Synergies of Organizational Level of Financial Planning and Application of Value Management in the Holding Activity

a. Investment project - concept and classifications

The investment activity must be part of the coordinates of an economic and social development strategy, which are materialized through programs or projects. Programs have a more inaccurate identification, with a wider scope and, most of the time, for their implementation they need to be detailed in one or more projects. In this context, the investment project can be defined as "an optimal set of investment actions based on a sectoral, global and coherent planning, based on which, a defined combination of human resources, materials, etc. causes a determined economic, social development" [15]. The notion of investment (investissement - French, investment - English; inversion - Spanish; investion - German) is synonymous, from a linguistic point of view, with; allocation, placement, endowment [16] and by extension we reach the efforts made now in the hope of future rewards.

In the literature there is a diversity of opinions on the concept of investment, which may lead to conceptual confusion and perverse practical implications. Thus, we find that "... is a capital flow intended to change the existing stock which is, together with the labour factor in particular, one of the main factors of production." [17]. In a reference paper we find that the investment is "the operation which, for the enterprise, is a transformation of financial resources

into one or more elements which are used continuously for a relatively long time, in a sustainable manner, in order to achieve the social objective" [18]. The notion of investment, in a broad sense, is synonymous with: allocation, placement, endowment [19], and in a narrower sense (financial-accounting) is an expense made to obtain material goods with high value and long use. Within the strategies, development and policies of companies, the investment strategy is important for all activities of companies depend on investment, depend on the investment policy adopted [20]. The investment process is: "Resource insurance, made in the hope of obtaining long-term benefits in the future" [21] or "action, which consumes money resources, other resources, in the aspiration that in the future will be collected sums of money or other benefits will be obtained". The investment project is the concrete motivation of some present expenses in the hope of some future advantages. Its specific features that distinguish it, not only from current activities, but also from other types of projects are: scale, finality and structure. We summarize three more important criteria for classifying investment projects: Economic life (economic time associated with investment projects); the risk; dependence on other projects.

b. **The essence and operation of the holding company**

Holding is a modern form of business organization. A holding company is a private or public company that holds a significant part of the shares of other companies (enterprises) – producers or sellers of goods, in order to ensure control of their activities. Holding, has the quality of party, when it holds 10-50 % of the shares of other companies, these becoming "subsidiaries" only in the situation when the holding company owns over 50% of their shares.

In practical activity, the holding company fulfils two functions: a) ensures the majority of votes in the general meetings of shareholders and in the boards of directors of the companies in which it holds shares, having a decisive role in the control of the activity. In this sense, in order to ensure a great cohesion in the management process, the unitary coordination of the activity of several companies, the shareholders of some of them can be constituted in a holding company that will become the main shareholder of other companies. For example, a joint stock company can be set up in two ways: (1) through the direct contribution of the various shareholders who will each hold a certain number of shares; (2) through the direct contribution of a holding company that will hold the majority of the shares, as well as through the contribution of other shareholders; b) ensures the multiplication of the capital by association starting from limited financial means, in order to ensure the control of the activity of one or more commercial companies. In developed countries there are different forms of association on the principle of holding: the concern, in Germany; the group, in France and Japan; the conglomerate, in most developed countries, which includes companies from different sectors of activity in order to prevent financial risk. At present, holding companies have a wide distribution in industrially developed countries, having financial attributions in the elaboration of strategies, in the coordination of the activity of the subsidiaries, in the allocation of financial resources through functional compartments. In some holding companies there are divisions (departments) that depend on a central economic authority through different types of links, from financial participation (it can also be a minority), to the situation that they have common administrators. In the case of holding companies, there is a certain limitation of the decision-making powers of the component companies, given that, although they have legal personality, they do not set their own selling price. The parent company elaborates the overall strategy, approves the objectives, negotiates long-term loans, establishes the financial and accounting rules that apply in all branches, coordinates the personnel policy, etc.

c. Synergy as a characteristic of holding in the budget system

Elaboration of the concept of synergy and its reflection in the holding [22]. The strategy of creating and developing integrated structures of the economy, focused on increasing the value of the business. Synergy is an important strategic factor in managing the value of an integrated business. Scientific discussions take into account the issue of synergy mainly in the justification and implementation of strategic decisions related to mergers and acquisitions, where the real effect depends significantly on the level of costs for mergers and acquisitions. The problems of synergy, focuses on the activity of integrated structures, functioning is accompanied by synergistic potential, have not been sufficiently studied. The degree of implementation and, consequently, the magnitude of the synergistic effect of the existing integrated structures, is not only of integration nature, but also of the Effectiveness of the management. Main types and classifications of synergy: (1) Mobilize managerial efforts and create the impetus needed to begin the process of change. (2) Transfer of strategy to operational level; (3) Build the strategic alignment of the organization; (4) Transforming strategy into daily work; (5) Strategic Management as a continuous process. Adherence to the above principles results in the creation of intra-corporate synergies in which business units and ancillary services combined not only to achieve economies of scale, but also to reduce costs, both explicitly and implicitly, often without monetary value (time for approval, creation of additional management reports, etc.). Strategic alignment creates a new type of value - corporate - benefit from the synergy of the enterprise - by coordinating the efforts of individual enterprises.

Creating synergies at the organizational level by organizing financial planning. The issue of synergies between departments still too little studied to create value at the organizational level. However, in order to create this synergistic effect, it is necessary to introduce in the holding's business strategic management proposals, which describe how a separate unit proceeds from the interests of shareholders and customer needs, improving internal processes by coordinating efforts of personnel. Information systems and organizational culture. Thus, the intra-corporate synergistic effect achieved in four areas: 1. Intra-corporate financial synergy created due to the ability to distribute financial resources between operational units. In this case, separate business units create value for shareholders through operations in the national capital market, which is much more productive and efficient than each separate organization would operate. In addition, combining different business units under a single brand helps to promote common values or strategic directions. The optimization of financial processes achieved through the allocation and efficient management of resources. 2. The synergy of intra-corporate customers achieved with a common customer base in the process of cross selling a wide range of products and services, from various branches of business. 3. The synergy of internal business processes achieved through savings in the sharing of systems, capacities and staff, as well as through the integration of value creation processes associated with the value chain. 4. Intra-corporate synergy of learning and growth arises from the use of common skills to increase human, informational and organizational capital.

d. Model of the mechanism for obtaining value in the holding management activity

The environment of enterprises is not only an economic, industrial, agricultural and commercial one, it is a social environment, located in the natural environment, so it is both biological and biosocial, with all the laws of natural sciences and the environment. As a biosocial system, the enterprise includes dynamically and functionally: labour force, fixed capital, working capital, information as intangible assets and the natural environment. In a

broader sense, the enterprise system is in mutual inter conditioning with the social system and the natural environment. The enterprise also approached as a cybernetic system, able to function within the national and world economic system. Specific to cyber behaviour is the mechanism of dynamic self-regulation through the feedback system. Feedback, as a reaction of the economic environment, social and natural, allows centres with a decision-making role in the enterprise system to issue decisions to regulate their economic, financial, social or ecological behaviour. In the market economy, the main source of feedback is the market reaction through the mechanism of supply and demand. The role of information at the microeconomic level is constantly growing, this contributing to increasing efficiency by increasing the degree of order, organization and performance. The direct consequence is the reduction of entropy in all phases of economic activity, by organizing its qualitatively superior oriented towards maximizing the economic, social and conservation of the natural environment. In order to point out even more clearly the role of information in the economic process of value production, the author Bran in [23] defines in a remarkable way the essence of value incorporated in goods and services as follows:

Product = Information (s) implemented in the Substance (s) using Energy (s). Service = Energy. (s) Subordinated to an Information (s) using the material support (s) of the system. Along with energy (s), information (s) constitute a fundamental component of material reality (s) having a well-defined role and which cannot replaced by the other basic elements. Information, as a specific form of manifestation of energy in nature, has specific characteristics that decisively influence its behaviour in the structure of the enterprise.

The mechanism of obtaining value

The enterprise system includes - Man, Objects of Work, Means of Work, Information and Environment. Theory of value-low entropy. At the base of value we put the physical form of conservation of low entropy (a certain degree of organization) in our reference system- the natural potential existing in the three forms of existence of matter: substance, energy, information [24], page 139:

$$P_n + P_s + P_{e_s} + P_{e_{dr}} = P_{e_r} + P_{e_{m,s}} + D_R + D_{n_r} \quad (1)$$

P_n - Potential (low or negative entropy) attracted from the direct natural environment (solar energy, geothermal energy, minerals, water and wind energy, etc.) or indirectly (through plants, animals): P_s - Potentially attracted from the human society system in the form of social services, education, culture, protection, defence, transport arrangements, banking services, etc.: P_{e_s} - Economic potential in the form of services from the consumption-type economic processes carried out in the periods prior to the production process in t1; $P_{e_{dr}}$ - Economic potential from recoverable waste. P_{e_r} - Economic potential re-conserved following two economic processes (consumption and production) through which the conserved potential passed in the product obtained from the production process in t_1; we consider this level the real level of the value of the product obtained in t_1. $P_{e_{m,s}}$ - Economic potential ceded to the environment and society to feed the processes that take place in the respective systems (this potential can be considered as a loss for the current value circuit - in the period t1-tn); D_{n_r}- Waste that includes recoverable economic potential in future economic processes (potential$P_{e_{dr}}$), if economic and social processes are organized in accordance with environmental requirements, based on the principles of sustainable development;. D_{n_r}- Waste that includes economic potential that can no longer be recovered in future processes (entropic losses); this waste pollutes the environment, society and the economy, negatively influencing the systems engaged in the mechanism of obtaining value.

In the opinion of Boos [24]: "The use of a systematic approach to achieve this goal allowed us to draw the following conclusions. (1) The key indicators of enterprise development as a system is entropy. At present, there are already a number of studies that aim to study the relationship between the level of entropy of an enterprise and other indicators of enterprise performance. (2) The definition of entropy proposed in this article allows us to take into account the experience of researchers who have studied the concept of entropy in different fields of science, which may extend the possibilities of analysing the topic under consideration. (3) The classification developed in this study allows us to distinguish four types of entropy: the entropy of unpredictability and the entropy of disorder; the entropy of the connection structure and the entropy of the flow quantities.

5. Conclusion

In this paper, the situation regarding the need for financial planning of the investment project within the holding company analysed. Based on the analysis performed and the objectives set, some conclusions reached. Once the company has a strategy, it needs a plan, in particular a strategic plan, which is a set of company actions designed to implement this strategy. The investment opportunities that allow the company to implement this strategy constitute the company's investment strategy. Financial managers must ensure that the company has sufficient funds to meet the needs of the operating activity as well as the investment. This requires a financial strategy of the company. In this case, planning is necessary.

The integration of enterprises under the management of a single company in the face of increased competition and the deficit of funds for development allows the consolidation of financial resources, the implementation of a single economic policy, centralized strategic management and the expansion of spheres of influence.

Development of a new type of management of holding companies in Romania, in the conditions of building the knowledge economy, of ensuring the intelligent growth, sustainable development and favourable to social inclusion, with a support of the digitalization of the managerial and functional activity of the holding companies, we need to develop a strategy for the development, design and implementation of knowledge management, which would include the mission, value management, knowledge management strategies and activities.

For the development of a performance management both within the activity of an economic agent, as well as within a group of economic agents that form a holding group will allow us to build a design of the management by value of the activity of a holding company. Which included various elements for building a holding company and applying modern financial planning models with a high-performance digital support.

It is necessary to know the potential supply sources of economic processes. Necessary to follow the chain of successive processes, which will finalize the value of a product by recognizing by the systems engaged in obtaining the value.

The need for the integral assimilation of the action of the general laws of nature. Correlation of the TVE model with new economic models and theories - the value model based on entropy will include the other microeconomic and microfinance models with a vector of social and economic lines of force towards a real level of economic, social, ecological and intellectual potential.

References

1. Nagy, Csaba (2014) The importance of financial planning of romanian SMEs in the XXI century. Analele Universitatii'Eftimie Murgu'Resita. Fascicola II. Studii Economice.

2. Ghenea, Marius (2011), Antreprenoriat: drumul de la idei către oportunităţi şi succes în afaceri. Universul Juridic.
3. Law no. 123 of September 19, 2014 for the amendment of Law no. 571/2003 regarding the Fiscal Code, M.Of. nr. 687/19 sep. 2014.
4. Decision no. 77 of February 5, 2014 for the modification and completion of the Methodological Norms for the application of Law no. 571/2003 on the Fiscal Code, approved by Government Decision no. 44/2004.
5. Tax Code, Law no. 227/2015 applicable from 01.01.2016.
6. Chen, Y.-Ru (2008) Corporate governance and cash holdings: Listed new economy versus old economy firms. Corporate Governance: An International Review 16.5, pp.430-442.
7. Jagoda A. (2014) The Architecture of HR Function in Capital Groups – the Report of Empirical Research, Economics & Sociology, Vol. 7, No 1, 2014, pp. 61-71.
8. Afshar-Jalili Y., H. Ilyadi, and R. Esmaeilzadeh (2019) Employing knowledge management roadmap for integrating KM initiatives in a holding company: a case study of NWW. Journal of Knowledge Management, Hyderabad, vol. 17, issue 3, pp. 7-18.
9. Robertson N. J., H. C. Perkins, N. Taylor (2007) Multiple job holding: interpreting labour market change and economic diversification in rural communities. Sociologia Ruralis, 48.4, pp. 331-350.
10. Braga V.-F., C. A. Naftanaila, L. G. Avram (2016) Situation of Regulations for Holding Companies in Romania. Annals of Spiru Haret University, Economic Series, vol. 7.2, pp. 89-100.
11. Bressan S., M. Rammerstorfert, K. Weinmayer (2021) Internal capital markets and bank holding company efficiency. Review of Financial Economics, vol. 39(2), pp. 163-177.
12. Di Salvatore, Avv Salvatore (2002) Opportunities for Italian Tax Planning: Using a UK Holding Company and Trust." Trusts & Trustees, vol. 8, issue 7, p. 20.
13. Jia D., R. Li, S. Bian, C. Gan, (2021). Financial planning ability, risk perception and household portfolio choice. Emerging markets finance and trade, 57(8), ISSN: 1558-0938, pp. 2153-2175.
14. Ahmed J. U. et al. (2021) Corporate social responsibility in the wake of COVID-19: multiple cases of social responsibility as an organizational value. Society and Business Review. ISSN: 1746-5680.
15. Stoian, M. (2002) Investment management. ASE, Bucureşti, p.37.
16. Romanian Explanatory Dictionary (in Romanian: Dictionarul explicativ al limbii romane), Romanian Academy, Univers Enciclopédie Publishing House, Bucharest, Romania (first edition 1984, second edition 1996).
17. Bancel F., R. Alban 1995 Le choix d'investissement: methods traditionnelles, flexibilite et analyse strategique. Economica.
18. Dictionaire de management de projet. AFNOR, 1996, pag 236.
19. Dictionary of Synonyms/ in Romanian: Dicţionar de sinonime, Editura Albatros, 1972, pag. 215.
20. Vasile I. (1999) Financial management of the enterprise, Didactic and Pedagogical Publishing House Gestiunea financiară a întreprinderii, Editura Didactică şi pedagogică, Bucureşti.
21. Bierman H., S. Smidt, 1991 27th National Congress of the Order of Chartered Accountants and Diplomatic Accountants and of the French Association for Economic and Technical Cybernetics, Paris, September 1991 (Al 27-lea Congres Naţional al ordinului experţilor contabili si al contabililor diplomaţi şi ai Asociaţiei Franceze pentru Cibernetică Economică şi Tehnica, Paris, septembrie 1991)
22. Makarova V. A. (2012) Topical issues of the organization of financial management in holding structures. SPb: Publishing house of the National Research University "Higher School of Economics. (СПб: Изд-во НИУ «ВШЭ).
23. Bran P. (2003) Economics of Value, ASE Publishing House, Bucharest.
24. Boos V. (2011) Using the concept of entropy to describe the influence of intellectual capital on the financial results of an enterprise, Economics and Finance, Vestnik YurGu, nr.28.

Republic of Moldova Evolutions and Opportunities for European Integration

Ghenadie Ciobanu, Cristina Lincaru

National Scientific Research Institute for Labour and Social Protection (INCSMPS), Bucharest, Romania
gciobanu019@gmail.com, cristina.lincaru@yahoo.de

Abstract. "Under the purpose of sketch, the evolution and tendencies of relationship between European Union and Republic of Moldova, in terms of agreement and programs post 1989 we perform a qualitative analysis of the main programmatically documents coupled with literature exploration. The main findings of our analysis are that the European Union and Republic of Moldova relationship has difficulties, but regardless the huge background transformation, improved continuously. On the background of globalisation, increases the vulnerability to crises to any state, regardless its dimension. Peace stability, democracy and prosperity for all Europeans citizens, the European Union's goal, is present in the complex relationship developed with Republic of Moldova. In the last three decades, Europe build a complex political relationship framework, structured in a multifaceted architecture, where Moldova plays different roles: Cooperation and Association Beneficiary of the bilateral political framework with European Union; Beneficiary in European Council Joint Programs, Eastern Partner in European Neighbourhood Policy with the role of Associate in the Association Agreement including a Deep and Comprehensive Free Trade Area, partner in an Association Agenda focused on economic integration and political association; Partner in Mobility Partnership focused on legal migration, development and fighting against illegal migration, beneficiary of Visa liberalisation and visa facilitation to support labour mobility. Now, in 2020, the New Joint Programming document and the New Association Agenda (2021-2027) confirm the continuous European Integration of Republic of Moldova, in a deepen manner. Originality of this work is the structured synthesis of the main policy framework Documents, Agreements and Programs that shapes the European Union and Republic of Moldova relationship, since its independence. The main conclusion is that the further the successful evolution of relationship between European Union and republic of Moldova, demand its statistics as the main tool that aims at raising awareness on its positive impact."

Keywords. European integration, Neighbourhood policy framework, Cross-Border programs, Local labour markets, Indicators system

1. Introduction

We live in a global world. Castell's paradox says, "European integration is, at the same time, a reaction to the process of globalisation and its most advanced expression" [1]. As a response to global crises, EU establish its Integration goals in the Treaty of Rome on 25 March 1957. Post Second World War, Monnet, Schuman and Adenauer understands to restore peace on Europe continent in a sustainable manner, "through the integration of economies and the development of common policies" [2]. The anticipated outcome is the peace stability, democracy and prosperity for all Europeans citizens.

Post 1989 relations between EU and Republic of Moldova (RM) grow deep, deep, and increasing in complexity. EEAS [3] points that EU is one of the most influential external factors in determining RM's progress providing financial and technical support, "but also because of its involvement in attempt settlement of Transnistrian conflict" [4].

Montesano, Van der Togt and Zweers [5] in the Clingendael Report argue that "given the window of opportunity currently, offered by Russia, Moldova also offers a chance for the EU to test Russia's willingness to cooperate in a pragmatic way in the so-called shared neighbourhood". Europe build a complex political relationship framework, structured in a multifaceted architecture, where Moldova plays different roles:

a) Cooperation and Association Beneficiary of the bilateral political framework with EU

b) Beneficiary in European Council Joint Programs with Moldova joint programmes are focused to improve democracy, the rule of law and human rights.

c) Eastern Partner in European Neighbourhood Policy (ENP), focused on free trade agreements, visa liberalisation and EU-funded projects.

d) Associate in the Association Agreement including a Deep and Comprehensive Free Trade Area (DCFTA) and an Association Agenda focused on economic integration and political association.

e) Partner in Mobility Partnership focused on legal migration, development and fighting against illegal migration.

Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur? (But who can rightfully criticize him who wants to be in that pleasure which results in no discomfort, or he who avoids that pain in which no pleasure is produced?)

We state the Research Question (RQ): what is the evolution and tendencies of relationship between Moldova and EU in terms of agreement and programs post 1989?

Methodology: *Qualitative analysis* of the main programmatically documents coupled with literature exploration for the RQ.

2. Moldova's Relations with the European Union – Agreements and Programs

RM relationship with EU progressed slowly but firm from strangers in first decade 1990 to 2000, then new neighbours in the second decade in the years 2000-2010 and in the last decade 2010-2020 the decade of Partners [6].

Mamaliga [7] identifies four stages in European Integration of RM in EU following the 27th August, the Independence Day of RM. The first stage was during 1989-1994, marked by the signing of the Agreement between the European Economic Community and the European Atomic Energy Community and the Union of Soviet Socialist Republics, signed on 18 December 1989. This agreement was focused on trade and commercial and economic cooperation. This period was market by changing the socialist system administrative control principles with the market economy and international law principles. The second stage (1994 – 2011/2014) is the bilateral Partnership and Cooperation Agreement (PCA), signed on 24th November 1994 [8]. The third stage (2014-2020/24) is initiated since the 27th June 2014, when was signed in Brussels the Association Agreement (AA). (National Bank of RM, 2018) add as second basic document of EU integration of RM the Association Agenda, agreed between EU Moldova on 26 June 2014. Finally, according to [7], the fourth stage starts" with becoming a Member State of the European Union after 2020".

a) The bilateral political framework of the relationship between European Union and Republic of Moldova

Partnership and Cooperation Agreement (PCA) 1994-1998 [9] states the legal framework of the bilateral relations between the RM and the European Community (Community). PCA signed on 28 November 1994 entered into force on 1 July 1998 for an initial period of 10 years, subject of tacit extension. PCA covers the fields: political, commercial, economic, legal, cultural and scientific [10]. PCA widen its objective in a large and complex spectrum that endorse: the political framework, trade and investment and harmonious economic relations, to set the basis for legislative, economic, social, financial, and cultural cooperation, to support the democracy and the transition into a market economy" [11]. This framework creates the opportunities for EU business and Investments in RM on economic integration pathway.

*Another distinct plan is the free movement of labour***,** liberty that starts the process of labour market integration. The 23 Article of PCA state the labour conditions, for Moldovan nationals legally employed in the territory of a Member State. Moldovan nationals beneficiate, since 1998, of the same working conditions, remuneration or dismissal, without any discrimination based on nationality, as compared to its own nationals**.** The RM social security coordinate with EU Member States in the field of Insurance, for the purpose of pensions in respect of old age, invalidity, death and for medical care. The same provision are adopted for the nationals of the Member state legally employed in RM [11].

The third important step the evolution of EU - RM is the Association Agreements (EU - RM AA) active since 1st of July 2016, initiated since 2014, which replace the previous PCA. This a crucial point in relaunching EU – RM relationships. Moldova, in 206, on the background of 2016 corruption persistence, profound social and economic problems, prolonged conflict in Transnistria, raise in discussion the success of EU's EaP [5, page 4]. As a response to these negative tendencies, European Council urges the RM's reform through Association Agreement implementation. EU - RM AA state the *political association and economic integration, in condition of democracy and to political, economic and institutional stability in the RM. The specific character of this relationship marked, EU's objective to "promote, preserve and strengthen peace and stability in the regional and international dimensions, including through joining efforts to eliminate sources of tension, enhancing border security, promoting cross-border cooperation and good neighbourly relations"* [12]. EU - RM AA covers *freedom, security, justice mobility areas. RM supported to develop its economic potential via international cooperation, also through the approximation of its legislation to that of the EU. Deep and Comprehensive Free Trade Area (DCFTA)* [13], ratified and active since July 2016, spot an important shift of EU from the focus on ENP towards trade policy. DCFTA, according to Koeth [13] assure the access *"into the Single Market, although with some limitations: limited access to EU agricultural markets, limited access to EU funds and no say in setting the common rules. Whereas economic benefits and incomes from new sources (VAT or income tax)". DCFTA is an important economic document, but rather political according to* Koeth [13], that express RM, an other countries' aspirations to become a member of the EU. Mamaliga [7] points that" RM goes to another area of influence, then CIS, on political, economic, social, cultural and scientific aspects". Association Agenda establish priorities for joint work of EU and RM first for the period 2014-2016, then updated for the period 2017-2019. National Bank of Moldova (NBM) 2018 through the National Action Plan for the Implementation of the Moldova – EU Association Agreement for 2014-2016 (NAPAA), updated by the 2017-2019 Plan, assumes as priority the implementation of EU banking and financial standards. NATIONAL ROADMAP for the Integration of the RM into the European Research Area (ERA) for 2019-2021, approved by Government Decision No 1081 of 8th

November 2018, make RM the first EaP Country that showed this intention, associated with EU Framework Programme for Research and Innovation Horizon 2020 (2014-2020) [14].

b) European Council Joint Programs with RM

The RM joined the Council of Europe on 13 July 1995. Upon accession, it accepted the statutory obligations and entered into a number of specific commitments as laid out in the PACE Opinion 188 (1995) on RM's application for membership of the Council of Europe [15]. The first Council of Europe Action Plan focused on support democratic reforms in the RM 2013-2016: functioning Democratic Institutions and good governance, Justice Sector Reform and Independence of the Judiciary, Protection and Promotion of Human rights, Freedom and Pluralism of Media, Local democracy and Decentralisation. The second Council of Europe Action Plan 2017-2020 (CoE, 2017), adopted in 2017, focused on *multiple sectors and policy areas "constitutional reform and elections, the fight against corruption, support to the justice sector reform, assistance to national human rights institutions, combatting ill-treatment and impunity, combatting trafficking in human beings, freedom and pluralism of media, education for democratic citizenship, local democracy and decentralisation as well as Confidence Building Measures across the Dnistru river*".

RM's strategical framework, National Development Strategy "RM 2020", the National Public Administrating Reform Strategy 2016-2020, the National Decentralisation Strategy and other sectorial strategies aligned with European standards. CoE country-specific programmes in the RM implemented alongside a number of regional joint programmes with the EU, in the framework of the Partnership for Good Governance (PGG). PGG actions geared towards enhancing *judicial reform, supporting free and fair elections, combatting corruption as well as cybercrime*. The Action Plan consists of a total of 17 country-specific and regional/multilateral areas of intervention with an overall financial volume of almost 18 million Euros. Again, the EU-RM Action Plan, prove to be a mixed tool, that provide both economic and political cooperation, with focus on security at EU border. This security plan promote the "deeping the political cooperation in area of foreign policy and resolution of the conflict in Transnistria conflict". The economic dimension of the plan, next to Economic growth and poverty Reduction strategy paper (EG-PRSP) promote economic growth and poverty reduction [16].

c) European Neighbourhood Policy (ENP) – bilateral and multilateral cooperation

RM, as an EU neighbouring country, beneficiate of a special relationship based on the Treaty of EU' foreign policy. (Article 8, Treaty of EU) These relationships include "cooperation, association and partnership to further strengthen economic and political ties". This special relationship is designed in 2003 by the ENP, launched with the Communication 'Wider Europe', (COM(2003) 104 final). ENP is a large framework that cover EU relations with 16 countries. RM is included in the ENP in the EASTERN group since 2009, next to: Armenia, Azerbaijan, Belarus, Georgia, and Ukraine. The other group, since 2008, is the South – Mediterranean and East which include: Algeria, Morocco, Egypt, Israel, Jordan, Lebanon, Libya, the Palestinian Autonomy, Syria, Tunisia. Russia and Belarus refused to join ENP, even if these countries are in the EU neighbourhood. EU and Russia relationship built in the framework of "the Partnership and Cooperation Agreement (PCA), signed in June 1994 and in force since December 1997. Russia and the European Union agreed at the St. Petersburg Summit in May 2003, to create four Common Spaces: economic; freedom, Security and Justice (mainly Visa Facilitation and the Readmission Agreements); Space on External Security and Research, Education, Culture. RM, through ENP, increases its "the opportunity of closer

economic integration with the EU and the prospect of increased access to the EU's Internal Market. The plan was for integration to be progressive, through the implementation of challenging political, economic and institutional reforms, and a commitment to common values." [17]. The economic integration target implies sector cooperation across a broad range of areas (including energy, transport, agriculture and rural development, justice and home affairs, customs, taxation, environment, disaster management, research and innovation, education, youth, culture, health, etc.). Under ENP's umbrella, RM is partner with EU on the following priority areas: trade, connectivity, security, migration and mobility, engagement with youth, creating opportunities for women.

Among the tools which RM beneficiates under the NEP are: *Visa liberalisation on the 28th April 2014, and visa facilitation* to support mobility, the Association Agreements and the Deep and Comprehensive Free Trade Areas (AAs/DCFTAs) [17].

Starting November 2013, Ukraine, an EaP country member, become the stage of a series of dramatic events. Following the Euromaidan movement (November 2013 – February 2014), "Russia illegally annexed Crimea in March 2014, and the eastern part of Ukraine plunged into an armed conflict spurred by Russia-backed separatists" [18]. Naumescu [19] points that on the Crimea's crises from 2014 was launched the second Cold War between West and Russia. Following the failure of Minsk Trilateral (2015 between OSCE, Russia and Ukraine), Damen [18] signals that the EU initiate the economic sanctions to Russia, until Minsk agreement signed and process still active. These events increased the tensions in RM diminish the progresses of EU-RM programs implementation.

ENP include RM as partner in three main programs: Eastern Partnership (EaP), Neighbourhood Wide (NW) and Cross Border Cooperation (CBC) (Figure 1):

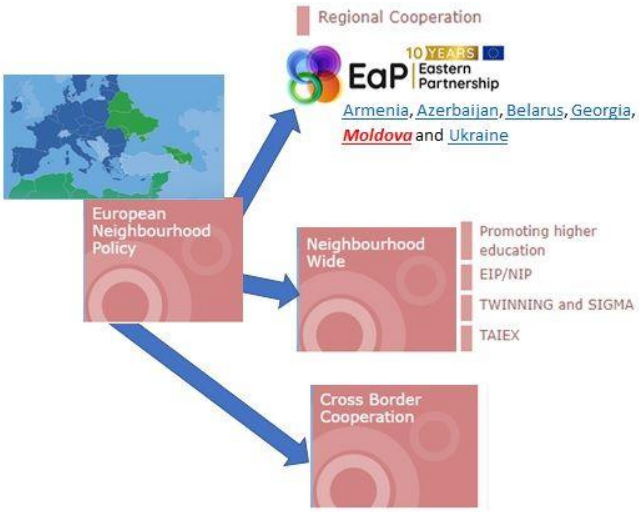


Fig. 1. Main Programs of the EU Neighbourhood Policy, that address RM
(Source: synthesis made by authors, from the site: https://ec.europa.eu/neighbourhood-enlargement/neighbourhood/european-neighbourhood-policy_en.)

- *Eastern Partnership (EaP)*

RM, signed with EU in Prague, in May 2009 a joint declaration that established the Eastern Partnership (EaP) (Council of European Union, 2009). The main goal of EaP is to "create the necessary conditions to accelerate political association and further economic

integration between the European Union and interested partner countries". EaP develops in parallel with the bilateral cooperation between EU and member states. In this framework, RM beneficiate of the EU support to develop political and socio-economic reforms that increases the stability, security and prosperity of entire region and even entire continent. The EU RM Action Plan is a political document laying out the strategic objectives of the cooperation between RM and the EU. It covers a timeframe of three years [16]. The recent Civil Society Forum analysis from the EaP Platform, in regard to Moldovan Government Action Plan for 2020-2023 (from 7th of February 2020, in short GAP), specify that "*one policy objective refers to the bilateral relations with neighbours i.e. Romania and Ukraine (3.10). Majority of the policy actions are planned by the end of 2020 or 2021 and are less proactive in seeking of settlement of bilateral issues and strengthening bilateral cooperation*" [20]. The same document, states regarding the Regional development and spatial planning points on "infrastructure projects such as water supply and sanitation, roads, waste management and local services". At policy level, the GAP "envisages the approval of a new *Regional Development Strategy for 2021-2025,* a new approach towards regional development, focusing on economic development, approval of the urban development program, amendments to the Law on regional development, adjustment of the work of the regional development councils investment projects, particularly in infrastructure and delivery of services, which will be implemented from the National Regional Development Fund with the involvement of the Regional Development Agencies [20].

Georgia, Ukraine and Moldova, signed a joint statement, to create the future EaP format of the EU plus the three countries. This EaP format should make the sectoral integration in transport, energy and other fields, assuring the free movement of goods, services, capital and the labour force [21].

"Foreign Ministers of Georgia, Ukraine and Moldova, David Zalkaliani, Vadym Prystaiko and Aureliu Ciocoi, have addressed the European Commission over the importance of "relevant financial support" for the Eastern Partnership (EaP) in the next EU Multiannual Financial Framework for 2021-2027." [22].

The EaP policy beyond 2020 is the subject of the EU Joint Communication, adopted on 18 March 2020, under objective to 'Reinforcing Resilience - an EaP that delivers for all'. All the EaP priorities chapters are subordinated to resilience and well-tuned with EU' new cycle Strategical objective, respectively in: economy, institutions, law and security; environmental and climate; digital transformation and societies. The new set of deliverables and policy measures that will shape the EaP policy for the future is set to be adopted at the June 2020 EaP Summit. EaP will remain an inclusive project, to further advance its countries', including RM, in the "*legislative harmonisation and economic integration* with the EU" [23].

- *Funding Sources for RM's development. EU is the first development partner for RM, beneficiary of the highest financial assistance per person among the EaP countries [24]*

*Neighbourhood and Partnership Instrument (ENPI) is the first funding source for ENP in the period 2007-2013***.** Scope of the ENPI is to support political, economic and social reform, across all sectors of common interest, at regional and local development, and regional integration though the RM's participation in Community programmes and agencies. The budget for the ENPI amounts to 11181 million for the period 2007-2013, of which 95 % is for national and multi-country programmes and 5 % for cross-border cooperation programmes. Other donors were the European Investment Bank (EIB), Member States and financial intermediaries (the international and regional organisations) sharply increased from €40 million in 2007 to €131 million in 2014 (Regulation (EC) No 1638/2006) [25].

BEI and BERD offered in a 50% un-refundable loans RM 2 billion euros during 2010-2020 trough programs of technical assistance, microfinance assistance, direct budgetary support, etc. (eu4moldova.md).

European Neighbourhood Instrument (ENI), replace the ENPI. EU allocate grants for ENI over EUR 15.4 billion for 2014-2020, from which for the EaP there are 741-906 million Euro. In addition to the ENI, RM is eligible for funding under the EU thematic programmes: the European Instrument for Democracy and Human Rights, Instrument contributing to Stability and Peace, Civil Society Organisations and Local Authorities, Human Development and Migration & Asylum. The main challenge has been the incomplete use of EU assistance provided through budget support programs and macro-financial assistance. The 131 million euros under budget support programs allocated in 2014 were suspended in 2015 by the European Commission, following a "banking fraud scandal" [18]. In 2014 the pro EU government coalition declared intention that RM to become EU member by 2020, but discredit the process by its corruption scandal. This was the biggest corruption scandal, under the oligarchic regime of Plahotniuc, in the country's history, according to Shupac [26] equivalent of the 12% from GDP disappeared from three Moldavian banks. The lack of progress in investigating the fraud in the banking system caused this first decision. The funding process prove to be difficult. In 2018, following the cancellation of the local elections in Chisinau as a democratic slippage, European Commission suspended for the second time the first tranche under the EU's € 100 million Macro-Financial Assistance Program.

Damen [18] fact that: "*after an agreement was struck at the end of 2016 on a programme between RM and the International Monetary Fund (IMF), the main aim of which was to stabilise the banking sector, the EU resumed its disbursement of budget support. Bilateral assistance to RM under the European Neighbourhood Instrument (ENI) ranged from EUR 335 million to EUR 410 million during the 2014-2017 period. The new EU multiannual programme for 2017-2020 (EUR 284 million to EUR 348 million) was adopted in September 2017 and focuses on the following priority sectors: economic development and market opportunities; strengthening institutions and good governance, including the rule of law and security; connectivity, energy efficiency, environment and climate change; and mobility and people-to-people contacts*".

The Pro-European support of Moldovan citizens rise up again to 59% in 2019 after the 2014 corruption shock. This was the result of the 2019 consensus between the EU, Russia and the US, followed by the Moldova Parliament's Declaration on the Recognition of the Captive Character of the Republic of Moldova and Plahotniuc's flee from the country [26].

An important tuning point of the EU-RM's relationship recovery was in the autumn of 2019. On 18th of October 2019, Johannes Hahn, Commissioner for ENP and Enlargement Negociations Hahn announced a new assistance package. The new EU budget "support rule of law, energy and public finance reforms This new payment follows the budget support payments made in July 2019 (€14.54 million), the recently announced first instalment of Macro-Financial Assistance (€30 million), and the EU-RM Association Council, last 30 September 2019." Council Conclusions of 26 February 2018 "recognises the important steps the government has taken in addressing the problem of significant politization of state institutions, corruption and the lack of independence in the judiciary and prosecution services. Positive developments have taken place in the reforms in the electoral system. Re-established relations with the International Monetary Fund (IMF). On 20 September, the IMF completed the fourth and fifth reviews of its programme with RM" [27].

On the Covid pandemic background, in May 2020 EU identified and reallocated over 87 million euros for immediate necessities of RM. EU also increases the Omnibus micro-financial emergency assistance from 70 million euros in 2017 with 100 million euros in 2021,

under some conditions. (EU Commission - '*Team Europe' strategy*, 22April 2020) EU joined with BEI and BERD increased with 963 million euros, coupled with technical assistance funded through Civil Protection European Mechanism CPEM. CPEM is focused on improve prevention, preparedness and response to disasters. Under this Mechanism, Romanian Government provided approximatively 3.5 million euros in a consistent support package in efforts to combat the COVID-19 pandemic [28].

- *Neighbourhood Wide (NW)* is the second main instrument of ENP. NW includes Superior education promotion (Erasmus + and Tempus), eTwinning Plus and SIGMA, and TAIEX

Erasmus+ and Tempus funds higher education from 2014. The ENP countries beneficiate by three key actions: Learning mobility of individuals **(**mobility for master up to 12 months), Cooperation for innovation and the exchange of good practices **(**capacity building, institutional reform and related student and staff mobility) and Support for policy reform **(**policy support and policy dialogue). *Tempus* and the *Erasmus Mundus Programme Overview* support to higher education in 2007-2013. *eTwinning Plus* Programme promotes Cooperation between Schools for the countries (Azerbaijan, Armenia, Georgia, Moldova, Tunisia and Ukraine).

Twinning is a European Union instrument for institutional cooperation between Public Administrations of EU Member States (MS) and of beneficiary countries (BC). It was originally designed in 1998 to help candidate countries of the time to acquire the necessary skills and experience to adopt, implement and enforce EU legislation. Twinning aims to provide support for the transposition, implementation and enforcement of the EU legislation (the *Unionacquis*).

SIGMA - Support for Improvement in Governance and Management is a joint European Commission and OECD initiative, principally financed by the EU. It focuses on strengthening public management in areas such as administrative reform, public procurement, public sector ethics, anti-corruption, and external and internal financial control.

TAIEX - Technical Assistance and Information Exchange – sharing experience since 1996, contribute effectively to EU legislation implantation in RM and support to justice reform in RM. TAIEX addresses mainly to public administration representatives (civil servants working in public administrations at national level; the judiciary and law enforcement authorities; and civil servants working in parliaments and legislative councils, as well as representatives of social partners, trade unions and employers' associations) and not on civil society, private citizens or individual companies. TAIEX main actions are workshops, expert missions and study visits. Workshop on Justice reform, held on 4th and 5th of September 2019, bring together experts from the European Commission and EU Member States in Chisinau to share peer-to-peer expertise on judicial accountability, independence, integrity and performance evaluation, in the presence of international partners.

- *Cross Border Cooperation TESIM –ENI CBC* is the second main instrument of ENP. Cross Border Cooperation (CBC) is a key element of the EU policy towards its neighbours

It supports sustainable development along the EU's external borders, helps reducing differences in living standards and addressing common challenges across these borders. (Figure 2) It first recognised as such in the European Neighbourhood and Partnership Instrument (ENPI) regulation for the period 2007-2013. This confirmed for the period 2014-2020 in the European Neighbourhood Instrument (ENI) regulation, adopted in March 2014.

ENPI CBC, generate 7.000 applications, involving over 37.000 organisations and leading to the approval of nearly 1.000 projects [29].

CBC promotes cooperation between EU countries and neighbourhood countries sharing a land border or sea crossing. Funding can also be provided for a programme between several EU and neighbourhood countries which, for example, are part of the same sea basin. A new Technical Assistance project, i.e. "Technical Support to the Implementation and Management of ENI CBC programmes (TESIM)", replaces the previous *Regional and Capacity Building Instrument (RCBI)* and *ENPI Interact projects*. The objective of this project is to continue to give support to the CBC programmes on both programme and project level focusing on improving the capacity of Partner Countries to participate in the programmes." (ec.europa.eu)

TESIM Program Romania – RM, for the period 2014-2020, funded by EU with EUR 81 million through the Joint Operational Programme Romania - RM, under the ENI, approved by commission in 2015. Romania - Republic of Moldova ENI CBC programme has the General Objective (GO) "to enhance the economic development and to improve the quality of life of the people in the Programme area through joint investments in education, economic development, culture, cross border infrastructure and health while ensuring the safety and security of the citizens in the two countries [30]. This GO implemented in four Thematic Objectives (TO): to Support to education, research, technological development and innovation; Promotion of local culture and preservation of historical heritage; Improvement of accessibility to the regions, development of transport and communication networks and systems; and Common challenges in the field of safety and security. This programme cover full Moldova's territory and Botosani, Galați, Iași and Vaslui Romanian counties. (Figure 2) among the 484 ENI CB projects contracted by 10 January 2020 there are only 4 between Romania and Moldova.

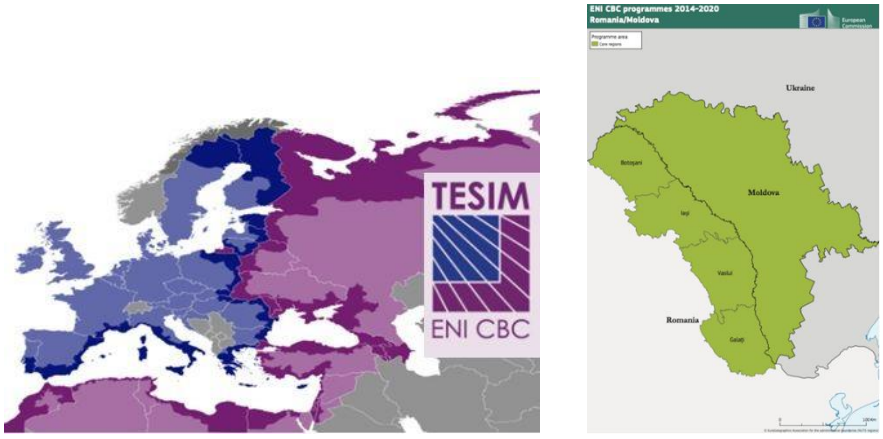


Fig. 2. TESIM Programmes and detail with Romania – Republic of Moldova ENI CBC
Source: https://interreg.eu/programme/eni-cbc-romania-republic-of-moldova/

The Platform Open Moldova [31] the Romanian Ministry of Public Works, Development and Administration announces on February 7th 2020, the launch of a new Joint bilateral program, Romania-Republic of Moldova for the period 2021-2027. The EU neighbourhood include RM in the objectives: smarter (unemployment, education, health and social inclusion) and more connected (sustainable, smart and intermodal infrastructure and digital connectivity). It is included also in the specific objectives' of INTERREG 1 better cooperation governance (institutional capacity, civil society, minorities) and INTERREG 2's

safer and more secure border [30]. The CBC ENI program reports the Large Infrastructure projects, approved by EU Commission approved (on November 14, 2018) and granted on 28 April, the following 2020: Communication Infrastructure; Regional Cooperation for Preventing and Combating Cross-border Crimes between Romania and RM; A safer Romanian – Moldavian cross border area infrastructure through the improvement of the operating infrastructure of the Mobile Emergency Service for Resuscitation and Extrication (SMURD); Rehabilitation and modernization of customs offices from the border of Romania and RM (customs offices Albița – Leușeni, Sculeni – Sculeni and Giurgiulesti – Giurgiulești) [32].

d) EU – RM Mobility Partnership, visa facilitation and readmission agreements

RM as eastern neighbourhood EU country beneficiate of the Mobility Partnership (MP) cooperation framework. This framework includes negotiations of visa facilitation and readmissions agreements between EU and RM. The migratory and transit routes from RM, as an origin country, are a top priority to EU.

The EU external and asylum policy is treated by the he Global Approach to Migration and Mobility (GAMM) established since 2005. GAMM set six objectives: Preventing and fighting illegal migration; Readmission, voluntary return and sustainable reintegration; Legal migration with a special emphasis on labour migration; Integration of legally residing migrants; Migration, mobility and development; and Strengthening capacities in the area of asylum and international protection. EU launched a regional policy, a *targeted migration dialogue promoting migration partnerships* through the Prague Process Joint Declaration in April 2009. The Prague Process Action Plan 2012-2016, adopted during the 2nd Ministerial Conference in 2011, set out for each GAMM objective, a thematic area for cooperation. In 2017 the Strategic Group transformed the single state leadership of the Prague Process into collective leadership (Czech Republic, Hungary, Lithuania, Poland). (icmpd.org) The current stage of the Prague Process is shaped by the Bratislava Ministerial Declaration, endorsed at the 3rd Prague Process Ministerial Conference in Bratislava in September 2016. (pragueprocess.eu)

The EU and Republic of RM signed a *Joint Declaration on Mobility Partnership in June 2008.* The Mobility Partnership have the purpose of facilitating legal migration including circular and temporary migration. The Joint Declaration on Mobility have decided that based on the ”labour market and socio-economic situation, establishing cooperation on migration and development, and preventing and combating illegal immigration and trafficking in and smuggling of human beings, as well as promoting an effective readmission and return policy, while respecting human rights and the relevant international instruments for the protection of refugees and taking into account the situation of individual migrants and the socioeconomic development of the Partner countries” [33].

The EU-RM Visa Facilitation Agreement, facilitate the issuance of visas to the citizens of the Republic of RM for an intended stay of no more than 90 days per period of 180 days, starting with 1st of January 2007 [34].

In the same year, 2007 EU-RM signed the Readmission [35]. RM readmit, all national citizens that,” or who no longer, fulfil the conditions in force for entry to, presence in, or residence on, the territory of the Requesting Member State”.

An amended version of the Visa Facilitation Agreement with the RM entered into force in July 2013 [36]. It changes in the Agreement title the word ‘Community’ by the word ‘Union’. In April 2014, RM became the first country in the Eastern Partnership to benefit from a visa-free regime.

According to Prague Process Platform (pragueprocess.eu), in 2018 RM from the 3.5 mil persons total population (Stat MD and Eurostat) there are 140 thousand immigrants (UN

Immigrant Stock) in 2017 and according to STAT MD are 3.7 thousand persons. The emigration is estimated as a stock by UN at 973.6 thousand persons in 2017 and at 2111 persons by the STAT Md in the same year. Also, UN announce in 2017 401 refugees and 81 asylum seekers.

e) Other partnership Black Sea Synergy and Energy Community

RM face a great threat to economy competitiveness as an effect of an entirely dependent energy sector on the import of natural gas from the Russian Federation through Ukraine. Perju and Crudu (2019) [37] found that "*the Moldovan Association Agreement with the EU offers an opportunity to modernize the energy sector by implementing the third package for the electricity and gas market, which involves the separation of electricity, gas and natural gas production, transport and distribution activities, as well as free contracting of suppliers by consumers and infrastructure integration into the EU energy market*." Authors iterates some important EU- RM energy projects: *Iasi-Ungheni-Chișinău interconnection project Phase II, since the 19th of February 2019 have the objective of consolidation of energy security and integration into the energy market of the European Union* [38].

Gas Pipeline Iași-Chișinău project is 90% finished on 29th of April. This gas pipeline, built by romaine Transgaz company," will link RM to the European energy system via Romania and reduce its energy dependence on Russia, according to the Moldovan Economy and Infrastructure Minister, Serghei Răileanu [39].

RM Energy and Biomass develop during the 2015-2018 (Phase 2), with an EU budget of over 10,5 billion USD, ENPI program. It was implemented by UNDP and aims to contribute to a more secure, competitive and sustainable energy production in the Republic of RM from biomass sources, the most readily available renewable energy source in the country. The Phase I, during the period 2011-2014, funded by EU, co-founded and implemented by UNDP," lays the basis for the establishment of functional markets for biomass technologies while increasing the use of renewable energy sources, in particular for heating public buildings and households in rural areas. New jobs and income are created and secured through the establishment of value-added chains at the local and regional level through the supply of biomass fuel and technologies." [40].

*Moldovan Sustainable Energy Financing Facility in 2009***,** offer" a credit line of 42 million Euro combined with a 5-20% grant component was provided for on lending to Moldovan companies through EBRD's partner banks. MoSEFF provides also technical assistance to the projects through Fichtner - a leading German engineering and consulting company*"* [41]**.**

INOGATE is one of the longest running energy technical assistance programmes funded by the EU. It started in 1996 and RM works through the EaP framework [42]. RM is partner in the Programmes:

- New INOGATE Technical Secretariat and integrated programme in support of the Baku Initiative and the Eastern Partnership energy objectives 2012-2016. Were built Sustainable Energy Information Centres and events (workshops). Also was provided assistance to Moldelectrica in defining and developing its TSO planning function (CWP.02.MD);
- Supporting Participation of Eastern European and Central Asian Cities in the 'Covenant of Mayors' (INOGATE-related project) in the period (20/09/2011 -20/07/2015).

3. Main Results of the EU – RM Agreements and Programs

The EU - RM relationship has difficulties, but regardless the huge background transformation improves continuously. Korosteleva [43] arguments the ENP and later the EaP inherited limits in defining the nature of partnership until 2014, that "does not automatically offer the prospect of membership to former Soviet republics apart to Baltic States". Danii and Mascauteanu [4] point that "not offering a membership perspective to Moldova, while using conditionality to facilitate reform, the EU policies - both ENP and EaP EU policies caused hesitation and even rejective attitude among the population and elite of RM". Morari [44] emphasize that EU has a major importance role "in the realization of the European integration process of the RM". As an effect of the 2016 reform, EaP's new instruments for RM: Association Agreement, DCFTA and Visa liberalization [45] prove to have visible effects, overpassing the declarative role (declarations, visits and official meetings).

The new geo-politics changes dramatically the background of big forces. Naumescu [19] signal the "westernisation" process of the Eastern Europe, area that become the new Central Europe. Before 1989 Eastern Europe covers the countries beyond the Iron Curtain. The author concludes, "the tragedy of Eastern Europe, in its old and new version, is chiefly represented by the same paradoxical cleavage between high geopolitical significance and the economic weakness of the region". The entrance of East-Central Europe in EaP does not get out from the Russian market and trade at the 2015 moment. Also, Naumescu points that energy dependency on Russian gas keep a major vulnerability towards EU integration of RM. Until 2015 "the economic aid from the West is slow and insufficient" [19]. Perju and Crudu [37] conclude that relationship between RM and EU are intensifying in the period 2007-2018,"and this, fortunately, affect positively the economy of the Republic of RM. The authors' remarks are as follows [37]:

- positive evolutions: AA and the DCFTA prove to be vital for RM's development, with visible positive impact on the economic growth of the country, improving the quality of life of Moldovan citizens, "through the creation of jobs, the modernization of public transport, the improvement of infrastructure, the provision of mobility studies and exchange programs for students and teachers, the provision of efficient energy supply etc.";

- further efforts for RM to: "work hard on its reform agenda, ensure the transparency of elections, fight against the high level of corruption, as well as to recover the misappropriated funds and bring all responsible persons to justice, irrespective of any political affiliations.".

The literature is scarce in regard the impact and effects of EU-RM relationship. Lately, more and more research results increased the transparency and put in light results and progresses. The platform eu4moldova.md/en points that "the RM is the largest recipient of EU aid per capita in the European Neighbourhood. The EU's assistance has provided tangible improvements to Moldovan citizens lives: … EU renovated 700 kilometres of roads, and provided residents of Chişinău and Bălți with modern trolleybuses. 15,700 people now have safe drinking water thanks to new water pipes built with EU support." (eu4moldova.md/en). Vîțu [46] claimed in 2019 that EU is the main economic partner for RM. Also the independent team of researchers [47] concluded that: "After five years of implementation of the Association Agreement, today the sectoral cooperation between the European Union and the Republic of Moldova is further strengthened, in particular in the fields of research, innovation and education. The benefits also enhanced by the visa-free regime with the EU, launched on April 28, 2014. Cooperation in the field of foreign policy and security been deepened. However, the functioning of democratic institutions, the strengthening of the rule of law and the independence of justice have not undergone clear changes, which has severely affected the

quality of the EU-RM dialogue, implicitly the full use of European support. Despite these constraints, due to the DCFTA, the RM has come significantly closer economically to the European Union. The EU has become the country's main economic partner. Today, about 70% of Moldovan exports directed to the European Union market, and imports represent almost 50%. The net impact of exports to the EU estimated at over 367 million euros, contributing to the creation of over 15,000 jobs and helping increase the budget revenues by 5% and by 320 million euros in investments in the private sector. However, the potential offered by the AA and by the DCFTA is yet be harnessed. The key to progress in this regard remains the existence of a clear and consistent political will in implementing all the provisions of the AA. And this means that the results of the transformations must contribute to the sustainable development of the country, bring more visible benefits to the citizens and help the continuous integration of the RM with the European Union".

The Head of the EU Delegation to RM Peter Mihalko claims "we all want to be more strategic for the next five years, especially when the programming of a new financial instrument for development and international cooperation for the period 2021-2027 is being discussed at EU level. Moreover, we will start updating the joint programming document for the RM at EU level and work on the new common agenda next year." [46].

Vitu [46] points that "corruption is still perceived as the main problem of society". According to Transparency International the Corruption Perceptions Index (CPI, higher the best) decreases from its best performance of 36 in 2012 to minimum of 30 in 2016, while, its level in 2018 reached the value of 33, a proof of a slightly improvement in terms of corruption persistence [48].

4. Conclusion

The EU - RM relationship is the result of over three decades of continuous construction. The complexity of the agreements and programs shapes an architecture growing in arborescent manner. The branches are the many dimensions and plans in which the policy framework develops. Variate degrees of integration driven by economic and trade relationship, extended to education, labour mobility, R&D, culture, security, energy, transport etc. RM integrate in global market through EU regional market, fact that make visible the full transition towards market economy, value changes and EU standard adoption in critical sectors (Finance and Banks, Law, Energy, Higher Education and RDI, trade, transports and with some progresses in public administration).

EU, under its ENP plays a major and unique role in Moldova's conflict management**.** Transnistria since 1990 its declared unilaterally independence from Moldova generate a persistent conflict. Damen [18] references that "*the EU participates as an observer in the 5+2 negotiation process on the settlement of the Transnistrian conflict*", and over the political tensions with Gagauzia. The recently conflict from Ukraine, the war in Donbas accentuated the two-block relationship deterioration. Vardanean [49] remarks the high polarisation of RM society, more divided between geopolitical preferences, missing the role of a bridge between Russia and the West. The cumulated result of these evolutions „reinforces the perception of uncertainty about the country's future". These consequences are dramatic over the population. The conflict areas are under the poverty risks, with low investments rates and active population exodus.

The EU-RM evolve from bilateral relationship through Cooperation and Association to Neighbourhood policy developing also multilateral relationships. Całus and Kosienkowski [6] points that" EU has finally acknowledged that RM can hardly be considered a 'success story' of the EaP." EU membership becomes more desirable with the 2019-2020 progresses

increasing visibility. Both EU and RM learn from the latest evolutions. For the New Association Agenda (2021-2027), emerge as new priorities, according to Shupac [50] "the domestic modernization rather than a geopolitical one", increasing the demand for "more actions for a more inclusive society". According to CSF-EaP platform [20] the high fragmentation of the country demand a new regional development structures managed by proper institutions. The complex architecture of policies and programs, in the specific case of RM, generates mixed tools that provide both economic and political cooperation, coupled with focus on security at EU border. This makes very difficult the functions of policy development, monitoring and evaluation effective implementation. These functions demand new institutions i.e. the creation of "an Agency, as a separate body responsible for the implementation of the National Regional Development Fund, while the Ministry of Agriculture, Regional Development and Environment preserve its functions of policy development, monitoring and evaluation" [20]. This demand identified also for the EU need to adopt the objective criteria measurement of governments' successes and failures [26].

The policy framework built, starting to function. In a global world, the only way of functioning is efficient and effective, increasing the capacity to impact evaluate and objective results measurement. All these aspects silhouette the statistical indicators systems of EU-RM relationship. Joint Communication on the EaP policy beyond 2020, (focused on 'Reinforcing Resilience') proposes the strengthening of the EU strategic communication in partner countries and aims at raising awareness on the positive impact of EU policies. Also, the new set of deliverables and policy measures that will shape the EaP policy for the future is set to be adopted at the June 2020 EaP Summit. In short, the successful evolution of EU - RM demand its statistics. The European Council in June 23 and 24, 2022 the decided to grant candidate country status to Ukraine and the Republic of Moldova [50]. Therefore, we hope for a not easy stage of connecting to all the requirements of the European Union, which requires new studies, new analyses with new prospects for economic and social development.

Acknowledgements

This work was supported by a grant from the Romanian Ministry of Research and Innovation - Ministerul Cercetarii si Inovarii din Romania, The Government of Romania, the National Research and Development Plan – Programme NUCLEU, 2019-2022, InovSoc program of the INCSMPS – the National Labor Research Institute of Romania, project no: Perspective funcționale a piețelor locale ale muncii în România, în contextul economiei inteligente și innovative / [Functional perspectives of local labor markets in Romania, in the context of smart and innovative economy] PN 19130101, coordinator Dr. Speranța Pîrciog.

References

1. Castells, M. (2000) End of Millennium, second edition, Oxford: Blackwell, p. 24.
2. European Voice (2004) EU's big chance to extend peace, stability, democracy and prosperity [WWW Document]. POLITICO. https://www.politico.eu/article/eus-big-chance-to-extend-peace-stability-democracy-and-prosperity/ (accessed 7.8.20).
3. EEAS (2016) The Republic of Moldova and the EU. EEAS (European External Action Service) -EC Commission. https://eeas.europa.eu/delegations/moldova/1538/republic-moldova-and-eu_en
4. Danii O., M. Mascauteanu (2014) Moldova Under the European neighbourhood Policy: ~Falling Between Stools". In Eastern Partnership: A New Opportunity for the Neighbours? https://books.google.com/books/about/Eastern_Partnership_A_New_Opportunity_fo.html?id=yzrJBQAAQBAJ
5. Montesano F. S., T. Van der Togt, W. Zweers (2016).The Europeanisation of Moldova: Is the EU on the Right Track? Clingendael Report, Nedderlans Insitute of International Relations, p. 32.

https://www.clingendael.org/sites/default/files/pdfs/Clingendael%20Report%20The%20Europeanisation%20of%20Moldova%20-%20July%202016.pdf
6. Całus K., M. Kosienkowski (2018) Relations between Moldova and the European Union. In P. Flenley & M. Mannin (Eds.), The European Union and its eastern neighbourhood. Manchester University Press. https://doi.org/10.7765/9781526109118.00017
7. Mamaliga I. (2014) The Evolution of the Integration Process as Effect of the Treaties Signed between Republic of Moldova and European Union. Acta Universitatis Danubius. Administratio, 6(2) Article 2. http://journals.univ-danubius.ro/index.php/administratio/article/view/2541
8. Partnership and Cooperation Agreement (PCA), signed on 24th November 1994.
9. Cooperation Agreement (PCA) 1994-1998.
10. Euratom (1998) 98/401/EC, ECSC, Euratom: Council and Commission Decision of 28 May 1998 on the conclusion of the Partnership and Cooperation Agreement between the European Communities and their Member States, of the one part, and the Republic of Moldova, of the other part, Pub. L. No. 31998D0401, 181 OJ L (1998).
11. Partnership and Cooperation Agreement between the European Communities and their Member States and the Republic of Moldova, no. L 181/3 (1998). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:1998:181:TOC
12. Association Agreement (2014) Association Agreement between the European Union and the European Atomic Energy Community and their Member States, of the one part, and the Republic of Moldova, of the other part, (2014). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.260.01.0004.01.ENG&toc=OJ:L:2014:260:TOCp. 4.
13. Koeth W. (2017) The 'Deep and Comprehensive Free Trade Agreements': An Appropriate Response by the EU to the Challenges in its Neighbourhood? 8.
14. GD 1081. (2018). National Roadmap for the Integration of the Republic of Moldova into the European Research Area for 2019-2021. https://mecc.gov.md/sites/default/files/national_roadmap_for_the_integration_of_the_republic_of_moldova_into_the_european_research_area_for_2019-2021_and_the_action_plan_for_its_implementation.pdf
15. CM-Public ODGProg/Inf(2013)17-final. (2013). Action Plan to support democratic reforms in the Republic of Moldova 2013-2016 of the Council of Europe (CM-Public ODGProg/Inf(2013)17-final).
16. IBP (2013) Moldova Country Study Guide Volume 1 Strategic Information and Developments. International Bussiness Publicatioons, USA.
17. European Commission (2015) Towards a new European Neighbourhood Policy, Joint Consultation Paper. https://ec.europa.eu/neighbourhood-enlargement/sites/near/files/neighbourhood/consultation/consultation.pdf
18. Damen M. (2019 November). Three Eastern Partnership neighbours: Ukraine, Moldova and Belarus Fact Sheets on the European Union/European Parliament. https://www.europarl.europa.eu/factsheets/en/sheet/171/three-eastern-partnership-neighbours-ukraine-moldova-and-belarus
19. Naumescu V. (2015) Introduction: The New 'Eastern Europe' and Cold War II. In The European Union's Eastern Neighbourhood Today: Politics, Dynamics: Cambridge Scholars Publishing. https://books.google.ro/books?id=7fumBgAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
20. CSF -EaP platform (2020). Analysis of the Moldovan Government Action Plan for 2020-2023. https://eap-csf.eu/wp-content/uploads/Analysis-GAP_2020.pdf
21. Agenda.ge. (2019, December 5). Georgia-Ukraine-Moldova deal, requesting EU+3 format for sectoral integration, new benefits. Agenda.Ge. https://agenda.ge/en/news/2020/484
22. Agenda.ge. (2020, February 16). Georgia, Ukraine, Moldova address EU over future development of Eastern Partnership. Agenda.Ge, https://agenda.ge/en/news/2020/484
23. EC Join 7, final. (2020). Eastern Partnership policy beyond 2020. Reinforcing Resilience—An Eastern Partnership that delivers for all. Euroepan Commission. https://eeas.europa.eu/sites/eeas/files/1_en_act_part1_v6.pdf
24. Vîțu V. (2020, May 5) Cum ajută Uniunea Europeană R.Moldova? RFI România: Actualitate, informații, știri în direct. https://www.rfi.ro/politica-120856-cum-uniunea-europeana-ajuta-rmoldova
25. European Commission (2006) Regulation (EC) No 1638/2006.

26. Shupac C. (2020, May 24) Eastern Partnership – Moldova> Republic of Moldova: Captured State vs. Success Story. Polis180. https://polis180.org/regionale-schwerpunkte/perspektive-ost/eastern-partnership-moldova/
27. European Commission. (2019, October 18). New EU budget support to Republic of Moldova [Text]. European Commission - European Commission. https://ec.europa.eu/commission/presscorner/detail/en/IP_19_6117
28. Aurescu, B. (2020). Declarația ministrului afacerilor externe Bogdan Aurescu după întrevederea cu Prim-ministrul Republicii Moldova, Ion Chicu / Ministry of Foreign Affairs. https://www.mae.ro/node/52366
29. TESIM ENI CBC. (2020). About ENI CBC. Tesim ENI CBC. https://tesim-enicbc.eu/about-eni-cbc/
30. JOP RO MD (2015) Joint Operational Programme Romania – Republic of Moldova 2014-2020, approved 12.2015. Managing Authority Ministry of Regional Development and Public Administration Romania. http://www.ro-ua-md.net/wp-content/uploads/2016/01/0-Romania-Republic-of-Moldova-JOP-approved3.pdf
31. Deschide.Md. (2020, February 7) Un nou program bilateral România-Republica Moldova 2021-2027, finanțat de Uniunea Europeană. Deschide.MD, https://deschide.md/ro/stiri/economic/60845/Un-nou-program-bilateral-România-Republica-Moldova-2021-2027-finanțat-de-Uniunea-Europeană.htm
32. Aron O. (2020) RO-MD - Large infrastructure projects. https://ro-md.net/en/about-the-programme/large-infrastructure-projects.
33. Council of European Union (June 2008) Joint Declaration on Mobility Partnership between the EU and Republic of RM. https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/policies/international-affairs/global-approach-to-migration/specific-tools/docs/mobility_partnership_republic_of_moldova_en.pdf
34. OJ L 334/169 (2007). AGREEMENT between the European Community and the Republic of Moldova on the facilitation of the issuance of visas,. Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:22007A1219(11)&from=EN
35. OJ L 334/149 (2007). Agreement between the European Community and the Republic of Moldova on the readmission of persons residing without authorisation. Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A22007A1219%2810%29
36. OJ L 168/3 (2013). AGREEMENT between the European Union and the Republic of Moldova amending the Agreement between the European Community and the Republic of Moldova on the facilitation of the issuance of visas. Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:22013A0620(01)&qid=1499690341710&from=EN
37. Perju D., R. Crudu (2019) The Evolution of Economic Relations Between the Republic of Moldova and the European Union in the Period of 2007-2018. CSIE Working Papers, vol.12, pp. 67–83. https://ideas.repec.org/a/aem/wpaper/y2019i12p67-83.html
38. Energy Industry Review. (2019, February 19). Iasi-Ungheni-Chisinau interconnection project Phase II has started. Energy Industry Review. https://energyindustryreview.com/oil-gas/iasi-ungheni-chisinau-interconnection-project-phase-ii-has-started/
39. Necsutu, M. (2020, April 29). Gas Pipeline Linking Moldova to Romania Nears Completion. Balkan Insight. https://balkaninsight.com/2020/04/29/gas-pipeline-linking-moldova-to-romania-nears-completion/
40. UNDP. (2019). [Closed] Energy and Biomass (Phase 2) | UNDP in Moldova. UNDP. https://www.md.undp.org/content/moldova/en/home/projects/moldova-energy-and-biomass-project2.html
41. MoSEFF (2009). Moldovan Sustainable Energy Financing Facility-Stakeholders-ETIP-B-SABS 2. https://www.etipbioenergy.eu/databases/stakeholders-db/513-moldovan-sustainable-energy-financing-facility
42. INOGATE. (2016). INOGATE a regional cooperation programme between the EU and 11 Partner Countires in Eastern Europe, Caucas and Central Asia. http://www.inogate.org/pages/1?lang=en
43. Korosteleva (2014) Eastern Partnership: A New Opportunity for the Neighbours? Routledge.
44. Morari, C. (2016). EU Role in the Republic of Moldova European Integration within Eastern Partnership. 16. http://www.ceswp.uaic.ro/articles/CESWP2016_VIII3_MOR.pdf

45. Association Agreement, DCFTA and Visa liberalization
46. Vițu V. (2019, October 31) Uniunea Europeană, principalul partener economic al R.Moldova. RFI România: Actualitate, informații, ştiri în direct. https://www.rfi.ro/politica-115566-uniunea-europeana-principalul-partener-economic-al-rmoldova
47. Lupușor A., Pîntea, D., Ghileți, E., Groza, I., Rusu, I., Ghileț, S., & Gumene, V. (2019) Shadow Report. Association Agreement Republic of Moldova—European Union. Five Years of Implementation (2014-2019). Progress, Constraints, Priorities. „Konrad Adenauer" (KAS), institutul pentru Politici și Reforme europene IPRE .Md, Expert-Grup Centru analitic Independent. https://www.expert-grup.org/media/k2/attachments/Shadow_Report_5_years_AA_en.pdf
48. Transparency International, 2020 Results—2018—CPI. Transparency.Org. https://www.transparency.org/en/cpi/2018/results
49. Vardanean E. (2018) Moldova between Russia and the West. 28. https://www.ifri.org/sites/default/files/atoms/files/rnv_110_vardanean_moldova_russia_and_the_west_2018.pdf
50. European Council meeting (23 and 24 June 2022) Bruxelles, 24 iunie 2022 (OR. en) https://www.consilium.europa.eu/media/57467/2022-06-2324-euco-conclusions-ro.pdf

Intelligent Search of Cross-Media Content

Stanislav Miroshnyk, Igor Grebennik

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
stanislav.miroshnyk@nure.ua, igor.grebennik@nure.ua

Abstract. The paper reviews current trends in intelligent search by media content used to develop media content dissemination systems, like music or movie streaming services. The symbiosis of the data mining and the semantic annotations is proposed to make a new intelligent search method. Experimental research was made to compare the developed method with exist vector method. Finally, promising directions for future research are outlined.

Keywords. Derivation of new knowledge, implicit semantic relations, media content, semantic annotation.

1. Introduction

There are currently many services that contain images, graphics, videos, and other types of content. However, with the significant growth that continues each year, the problem of finding the necessary content has arisen, as it is distributed among billions of services. Therefore, the development and research of content search methods are actual for now.

The main problem now is that the active phase of development of most of these services, and because of the study of relevant methods, ended many years ago and to form recommendations and search, these services do not use all the opportunities that have emerged during the development of artificial intelligence in recent years.

The purpose of this report is to research current methods used to search among cross-media content based on similarities and propose a solution to the class of problems of the intelligent search by digital cross-media content in various subject areas because of developing the intelligent search algorithm that more accurately offer content, focusing on user desires by identifying implicit semantic links between different resources.

2. Definition of Content

Content is certain information that is usually created by man [1].

There are several main types of content:

a) video content is any content format that contains video features. Typical forms of this type of content are vlogs, cinematographic works, animated GIFs, video broadcasts, etc.;

b) textual content - is information that is formed in textual forms, such as scripts, books, posts on social networks, etc.;

c) audio content is any material or information consumed due to listening. This type includes songs, music, audiobooks, podcasts, voice messages, etc .;

d) game content - includes game locations, objects, characters, etc .;

e) images. These include paintings, images, infographics, presentations, and more.

The content is present in a person's life almost always. One of the first examples of content can be called rock paintings, which are images. People learned to depict specific events and demonstrate them to others.

Over time, songs began to appear. Finally, people began to write to communicate with each other through sounds and pass on specific knowledge through writing - and this could already be considered the first profound example of textual content.

In addition, humanity began to paint pictures, put on plays. Over time, the camera was invented, and it was possible to take photographs. Later, the ability to record moving footage began to appear, and the world saw the first example of video content.

In the 20th century, the first digital computers appeared in the world [2], with their text and then audio-visual content. Thus, they laid the foundation for developing digital content, which is the basis of current work.

3. Current Trends in the Intelligent Search by Media Content

Many applications currently use methods to search for cross-media content intelligently. The most common include:

a) applications for the distribution of video and audio content, which include, among other things, recommendation systems.

b) various resources dealing with the distribution of textual and visual content, such as blogs, various news resources, etc.

c) applications dealing with the distribution and sale of images and other content.

d) software that focuses on finding image recognition used to identify people (such as FaceID), find music, and so on.

The first known legal music distribution service was iTunes. It is a media player used to organize and play audiovisual content. The client can purchase music, series, movies, and more as part of this service.

In this application, for the first time, a kind of recommendation system was developed, which, with some modifications, is still used today, namely - Genius. The main idea of this system is to analyze the current media library of users and build recommendations based on its comparison with the media libraries of other users. This algorithm has several exciting features.

First, the current algorithm is quite well optimized because the content information is replaced rather slowly. Hence, there is no need for continuous content analysis with each request from the user.

Second, to perform the search for similarities between objects, Apple uses information search algorithms based on the vector space model [3], thus comparing similarities between the current content with its other units, calculating the angle between them, where the vectors d_1 and d_2 are documents or examples of content that are compared with each other, and q is the song for which the search is performed [4].

Third, the factors for which the weights are provided during the comparison process are identified. To do this, the algorithm calculates the frequency of repetition of content from the user on a specific set of factors, such as genre, rhythm, artist, etc., for music and then uses this information to compare it with other users' media libraries.

Most IT giants now have their services for distributing music, videos, or searching using media content. Consider music distribution services. Such services among music distributors include Apple Music, Spotify, Youtube Music, and Netflix. After such services appeared and became widespread, musicians' incomes began to grow for the first time in a long time.

Each of these data giants of the IT industry has specific aspects related to the recommendation system, and each of them has its positive and negative nuances. Consider a few examples.

The first of them will be the YouTube Music service. At one time, he promoted the idea that in most cases, users like new content, but over time the service began to rely more on popular music, offering content that has already been listened to many times as opposed to trying to find some new content that could potentially interest users.

A negative feature of this service is also the principle of building referral systems, which initially generate a specific "baseline" for the user, thus solving the cold start problem. This level changes minimally to indicate the personalization of the service [5].

The second example is Apple Music, and as described earlier, this service uses a vector model to search for similar content for later use in building music collections.

A third example is Spotify, which also uses a vector model [3], offering users works by specific artists with similarities. To perform the search, this resource adds a factor that affects the formation of the result. Namely, the algorithm analyzes the presence of common listeners. As a result, several groups, for example, may have 20,000 common listeners out of a total of 50,000. As a result, there is a high probability that the system may recommend another 30,000 people to the group [6].

In addition to the examples discussed above, we can list a vast number of applications and services that perform analysis and search for media content, determine the degree of similarity and on its basis try to produce results that best meet the criteria set at the beginning.

After analyzing the above, we can understand the general principle of modern systems and conclude that modern application that has the task of searching and determining the similarity of digital content usually turn to the vector model in the formation of recommendations used for these tasks for over 15 years.

4. An Intelligent Method of Search by Media Content

As mentioned earlier, the previously developed method is based on data mining methods and semantic annotation.

Data mining is the process of identifying patterns and other valuable information based on large data sets using machine learning methods, statistics, and databases. The main idea of data mining is to use optimization, genetic algorithms, etc. The results of this analysis can be further used to solve problems of object classification, forecasting of certain events, and modeling [7].

Data mining models are used to solve several problems, which include [8]:

a) tasks related to the analysis of available information and the formation of expected results based on it, better known as forecasting tasks. The results of such tasks can be used to forecast the level of sales.

b) grouping tasks, such as solving the problem of grouping commercial offers on a specific topic for users.

c) the task of finding sequences that are used, including for predicting the actions of users in applications.

d) tasks of risk calculation. These include, for example, calculating the probability that a customer will repay a debt.

Data mining methods are relevant in developing an intelligent method for searching cross-media content. There is a need to form and select the most relevant search results.

Artificial neural networks were implemented to develop a new search method.

Artificial neural networks (ARNs) are computing systems inspired by conventional biological neural networks that make up the brains of animals and humans. Thus, they are built based on natural biological networks and consist of neurons, a very simplified adaptation of biological neurons.

There are several types of neurons.

Specimens of the first type receive signals from outside the network, then perform their processing and transmit them using signals to neurons in the hidden layer.

In the second step, the generated list of attributes of the object will be transferred to the hidden layer of the ANN. This layer can consist of several levels. For example, during the formation of ANN at synapses, the value of their weight w is defined by means. Thus, it is possible to characterize the level of importance of each of the signals given for calculation in neurons.

There is another type of neuron, namely the original neurons. These neurons process the results obtained by calculating the hidden layer and determining the resulting value.

The calculation of input values in neurons is performed by formula (1).

$$H_{i\ input} = (v_1 \times w_1) + (v_2 \times w_2) + \cdots + (v_n \times w_n) \quad (1)$$

The calculation of the initial value of the neuron is calculated by form (2), which can be modified depending on the needs of the system.

$$H_{i\ output} = f(H_{i\ input}) \quad (2)$$

Resource annotations, in most cases, are a set of metadata related to a given resource, i.e., a collection of values through which it is described. Thus, each resource has its own set of values that explains it, refers to specific groups, allows you to identify and perform other necessary actions based on this information.

In general, there are several main types of annotations [9]:

a) formal. This type of annotations has a formed structure, which is further subject to software processing and is described using specific languages.

b) informal. This type of annotation cannot be processed programmatically because it does not have a well-defined, clear structure that could be processed, and its formation is usually done using spoken language. With the current development of machine learning technologies, when using many resources, it is possible to process it, but this will require a large number of resources.

c) ontological. This type is based on a semantic model that contains a set of critical concepts and forms the correspondence of the object to software.

The essence of semantic annotation is to perform a formal description of each instance of the object under study, for which a plurality of tuples is used. Thus, the formation of metadata acts as an assistant in classifying different types of content because this type of annotation makes it possible to perform a description of a unit of content using only a set of tuples.

The tuples used in the description execution must correspond to the following structure (3) [10].

$$< s, p, o > \quad (3)$$

where s is the subject identifier, p is the predicate identifier, o - object identifier.

To compare the formed tuples, we determine the criterion of similarity of songs based on calculating the semantic distance using Hemming's length, calculated by the formula (4). This method calculates the total number of items according to which you can distinguish existing concepts.

$$L = \frac{\sum_{i=0}^{n} \frac{W_i}{maxW} \ (C_{i1}|C_{i2})}{N \sum_{i=0}^{n} \frac{W_i}{maxW}} \quad (4)$$

where

W_i – weighting factor of the category within the domain;
maxW – weight of the category with the maximum weight within the domain;
C_{i1} and C_{i2} – are the i-th characters in the tuple;

N – the number of concepts in the train.

The developed algorithm is divided into two parts.

The first part uses the calculation of the similarity between objects to detect implicit semantic relationships. Its implementation and use occur during the download of a new content unit, after making changes to existing ones, and is periodically performed for the entire content based on a specific schedule. Therefore, this part of the algorithm should be performed in the following sequence:

a) to perform the calculation of similarity coefficients, one copy of media content is selected, based on metadata about which a set of tuples is formed.

b) form a set of tuples with other instances, in which the value of the subject in at least one of the existing tuples coincides with the value of the issue in at least one tuple of the current instance.

c) perform the calculation of the semantic distance between the tuples of the current instance of cross-media content and other instances formed for it, using the method of finding the Hamming length for current calculations (4).

d) we obtain the result during which the formation of the relationship between the instances is based on the coefficient of similarity, which is additionally stored for its further use.

For the second part of the algorithm, consider it in the context of using the first part results to build a recommendation system. We use user data to analyze and search for results, namely the history of viewing or using content units. In this way, we will be able to obtain a sufficient amount of metadata from the content to form specific user preferences and, on their basis, to search and generate the resulting selections for recommendation systems.

The algorithm for developing a system of recommendations will look like this:

a) select the units of content from the database that the user used.

b) using the obtained sample to form a set with similar copies of media content, using the previously calculated similarity factor.

c) get from the database data on user activity, analyzing which we build a list of references for each attribute of content elements. To do this, calculate the value of attributes (4) and calculate the value of the weighting factor (5).

$$w = \frac{n_p - n_n}{2 \times N} + 0{,}5 \tag{5}$$

where n_p – repeatability of the attribute in the content with a positive rating;

n_n – repeatability of the attribute in the content with a negative rating;

N – repeatability of the attribute with any value.

d) perform the formation of formulas for calculating the values in the neurons of the network (1) (2) (6).

$$f(H_{i\ input}) = \sum_{i=0}^{n} \frac{H_{i\ input} \times w_i}{N} \tag{6}$$

where $H_{i\ input}$ – input data of the neuron;

w_i – weighting factor;

n – number of input signals;

N – number of neurons in the layer.

e) calculate the values in the neurons and transmit the results to the original neuron, which determines whether the received copy of the content or set of copies in the interests of the user.

e) for the obtained results, if necessary, we generate separate selections using previously calculated similarity coefficients.

f) display the result.

The new intelligent method was used to compare it with the vector method. A software tool for experimental research has been developed.

To test the method through a console application, it is necessary to obtain a training sample, which will contain comprehensive information about each unit of content, and a pre-formed result sample, which should be used to check the accuracy of the search.

During the preparation of this stage, two input samples were formed:

a) to check the algorithm's accuracy, the sample contains a small number of records, namely data on 96 songs that have varying degrees of similarity.

b) to generate information about the algorithm's running time formed a large sample of data with 60192 songs, containing 1444608 lines with data about the song data. Songs in the sample can be repeated because the current software implementation does not reject duplicates and needs to be refined if used in an actual project.

First, let's analyse the software testing results of algorithms used for intelligent search when finding links between units of content.

A sample with data on 96 songs was involved during the search quality testing.

During the testing, the song Billie Eilish - Bad Gay was chosen as the basis for the search. In preparation for the study, a resulting sample of 8 songs was formed, which included Billie Eilish - "burry a friend", "I love you", "COPYCAT", "No time to die", Imagine Dragons - "Natural", Twenty One Pilots - “Chlorine”, Maroon 5 - “Girls Like You”, Lady Gaga & Bradley Cooper - “Shallow”.

A sample of 10 songs with the highest coefficient of similarity with the current song was obtained (Table 1). We can note that the final sample includes all eight songs presented in the initial condition. Thus, in a small sample, the search accuracy was 100%.

Table 1. The results of the experimental study

Musical composition	Measure of similarity
Billie Eilish – burry a friend	0.68528330449827
Billie Eilish – I love you	0.40549307958478
Billie Eilish - COPYCAT	0.19869160899654
Billie Eilish – No time to die	0.19869160899654
Twenty One Pilots – Chlorine	0.10137326989619
Imagine Dragons – Natural	0.10137326989619
Drake – God’s Plan	0.10137326989619
Maroon 5 - Girls Like You	0.10137326989619
Drake – In My Feelings	0.10137326989619
Lady Gaga & Bradley Cooper – “Shallow”	0.10137326989619

It is worth noting that when using a search using semantic annotation, the search accuracy also depends on the completeness of the data on the units of media content so that the search accuracy may vary depending on the volume and accuracy of the input data. Still, in general, its value will exceed 95%.

The next step was to test the load on the algorithm, which involved a sample of 60,192 copies of media content (a total of 1,444,608 attributes). The search execution time was checked during the testing, and the result is 0.51869201660156s.

The correct comparison will be the Flat-CPU method, as it uses minimal indexing compared to other methods and is only partially optimized. Namely, its principle is to discard half of the most dissimilar records in advance.

The time to find implicit semantic links in the developed method is directly proportional to the number of attributes in the sample, then calculate it for 3 million records as follows - 0.51869201660156 * 3000000/1444608 = 1.07716145128.

For comparison, the Flat-CPU algorithm for 3 million records is executed in 9.1 seconds.

As a result, the developed algorithm should be performed for this sample in 1.077s, which is a significant improvement over the Flat-CPU method. Furthermore, in further research and improved algorithms in the case of clustering methods to discard the most different content groups and use indexing, search results can be significantly improved and compete with the fastest algorithms that use a vector model.

5. Conclusion

The review clearly shows a number of theoretical studies. During the analysis of data mining methods, based on data, on their advantages and disadvantages, the choice of the method of neural network implementation for the implementation of part of the algorithm was made.

The next step described the possibility of using semantic annotation to solve the problem of organizing metadata sets, choosing an ontological annotation model to formalize data about content instances for further processing, and finding Hamming distance that distinguishes implicit semantic relationships.

The study describes the new intelligent method of searching for cross-media content and describes the possibilities of its adaptation to work when used in referral or search engines.

A description of experimental research is carried out, during which it is determined that this topic is a promising area for future research.

References

1. Segura A. The Top 12 Types of Social Media Content to Create. Mailchimp, 2018. (https://mailchimp.com/resources/top-12-types-of-social-media-content-to-create/).
2. The Modern History of Computing // Stanford Encyclopedia of Philosophy, 2017, (https://plato.stanford.edu/entries/computing-history/).
3. Steinberg E. Semantic vector search: the new frontier in product discovery / Eugene Steinberg // Grid Dynamics, 2020, (https://blog.griddynamics.com/semantic-vector-search-the-new-frontier-in-product-discovery/).
4. Mimsarchive C. How iTunes Genius Really Works. MIT Technology Review, 2010, (https://www.technologyreview.com/2010/06/02/91325/how-itunes-genius-really-works/).
5. Madrigal A. C. How YouTube's Algorithm Really Works. The Atlantic, 2018, (https://www.theatlantic.com/technology/archive/2018/11/how-youtubes-algorithm-really-works/575212/).
6. How "Fans Also Like" Works. Spotify for artists, 2019, (https://artists.spotify.com/blog/how-fans-also-like-works).
7. Sytnik V., Krasniuk M. Datamining. Kyiv: Kyiv National Economic University, 2007. 376 p.
8. Olson D. L. Data mining in business services. Service Business, 2007, №3, pp. 181–193.
9. What are semantic annotations? / E. Oren, et. al. National University of Ireland, Galway: Digital Enterprise Research Institute. 2006. №1. P. 15.
10. Golovianko M, Plisko D. Building a distributed ontology system based on technology Peer-To-Peer (P2P). Digest of scientific works of Kharkiv University of the Air Force. 2010. №3. P. 131–133.

Information and Communications Technologies in the Preservation of Historical and Cultural Heritage

Yuliia Trach, Maryna Tolmach, Olena Chaikovska, Tetiana Gumeniuk

Kyiv National University of Culture and Arts, Kyiv 01601, Ukraine
lena@knukim.edu.ua

Abstract. The main problem of digitization of historical and cultural heritage sites in Ukraine is the lack of scientifically based developments on the rules and standards for creating a digital resource to ensure accessibility, active and efficient use, compatibility and data exchange at the local and international levels, saving and updating created resource. The need to develop a strategy for building a future digital collection is indicated. Emphasis is placed on the need to recognize the existence of digital culture objects, digital cultural heritage objects, to provide universal access in cyberspace to digital culture objects at the state level. The state of digitalization of cultural heritage in Ukraine is analyzed, the description of projects on digitization of cultural heritage objects is given: the project "Pocket country" (provides to digitize 40 architectural heritage objects, realized 3D-tours of prominent cultural objects of Lviv, Chortkiv, the project "Tartakiv & Tartak", etc.); project "Feel Ukraine to the touch" (3D-models of architectural structures for the blind); project "Trakhtemyriv-DIGITAL" (3D-model of the landscape of the Trakhtemir Peninsula with fixation of known and discovered monuments of archeology and nature); digital infrastructural tourist service "Castles 360: Transcarpathia" (Transcarpathian castles in the digital world 4.0) and many others..

Keywords. Historical and cultural heritage, information and communications technologies, digitization, national interests.

1. Introduction

The analysis of existing strategic documents on historical and cultural heritage in Ukraine gives quite reasonable grounds to state that the Ukrainian state, on the one hand, declares the need to protect cultural heritage, but on the other, the state represented by state bodies, local governments in many cases plays a negative role in this process, because in their inaction, cultural heritage not only does not receive the necessary protection but gets its objects destructed. Under such conditions, in a society where legal nihilism prevails, economic rather than cultural problems come to the fore, the desire to obtain material benefits in any way prevails, despite the threat of destruction of objects that must be sacred to every nation. Such blatant disrespect for Ukraine's cultural heritage in the pursuit of greater profits undoubtedly poses a significant internal [1] threat to Ukraine's national interests, especially given that Ukraine loses more than 100 cultural heritage sites each year [2].

The analysis of international acts and Ukrainian legislation, as well as global trends in the protection of world cultural heritage give grounds to distinguish the principles of protection of historical and cultural (tangible and intangible) heritage: the generality of cultural heritage; neutrality of cultural heritage; accounting of cultural heritage (creation of a single register); promotion of cultural heritage; harmless access to cultural heritage; informing the public; reproduction or restoration (restoration) of cultural heritage; preservation of cultural heritage (conservation); control over the protection of cultural heritage [3]. However, the development

of international relations towards finding new ways and methods of conservation of world heritage significant, for example - *the systematic digitization of heritage sites included in the UNESCO World Cultural and Natural Heritage and the List of Intangible Heritage, to visually accessible to all interested persons should be a prerequisite for the promotion of new principles of their protection and use*. The reason for this statement is the fact that today the popularization of cultural heritage sites in the form of transfer of traditional art forms into their technologically accessible modern virtual analog is a very effective direction in the long run. The development of information technology and computer networks has reached a level where they can make a significant contribution to the preservation and dissemination of digital culture objects, which over time can be transformed into cultural heritage sites as digital culture objects. The latter undoubtedly has great potential for present and future generations, as they enable the preservation of cultural heritage, even in digital form, provide access to digital cultural objects, crossing existing borders, and promote cultural heritage through a copy in the form of a digital object. Combined, this is sure to be an effective means of attracting society to cultural content. That is why it is necessary not only to recognize the existence of objects of digital culture, objects of digital cultural heritage, but also to ensure universal access in cyberspace to objects of digital culture at the state level. Moreover, the digitization of cultural heritage sites makes it possible to ensure unrestricted access to its sites, improve the quality of education and tourism, and so on.

However, unfortunately, in Ukraine, despite the general development of the IT-sphere and strong creative ideas, the dialogue between culture and technology is just beginning. This is indicated not only by Ukrainian and foreign experts, whose activities are to some extent related to these areas, but also by the "embryonic" state of digitization of cultural heritage sites in the country. Thus, the American expert in the field of cultural heritage and museums Linda Norris and the head of the department dealing with museums, libraries, and archives in the Ministry of Culture of Lithuania, Vaiva Lankelienė, among the main challenges and problems in the cultural heritage sector in Ukraine note: a) the slow process of transformation of society and overcoming the post-Soviet syndrome; b) outdated legislation governing the cultural heritage sector; c) virtually no strategic planning that would ensure the development of the sector; d) lack of digital archives of cultural heritage; e) limited funding for the sector, including the restoration and preservation of architecture and museum collections; f) lack of skills and understanding of how to receive support outside public funding [4]. Among the recommendations, experts point to the need to ensure the dissemination of information on the Internet to a wide audience, which includes the digitization of cultural heritage sites. However, there are some problems with the latter recommendation, as well as with others: despite all the advantages of using information and communication technologies, Ukrainian legislation restricts their use. For example on cloud services, it is possible to store the state information and personal data, but thus it is forbidden to publish information on owners of objects by numerous regulations [5]. Unfortunately, the situation has not changed for the better even with the adoption of the Law of Ukraine "On Personal Data Protection" (2011).

In the field of preservation and management of historical and cultural heritage opportunities of information and communication technologies are quite wide: from the creation of information systems, the use of various security systems in the preservation of cultural heritage, control and monitoring of architectural monuments and other cultural values, creation and maintenance of cultural values, web-technologies - to the use of visualization technologies and virtual reality technologies used to reproduce lost or damaged cultural heritage sites, etc. In particular, cloud technologies can promote the national cultural heritage: digitized books, paintings, ornaments, old films, and other information can be hosted on a remote server and made available to anyone in the world. This will increase the tourist

attractiveness of the country and the state will preserve cultural heritage. Not only the state but also creative business could be interested in this because such content can be used creatively.

Creating virtual tours of cultural heritage sites is one of the most promising areas of application of IT achievements.

Today, this format is already an integral part of the content of many museums, theaters, hotels, and entertainment complexes in the world. In Ukraine, virtual tours are very popular, drawing attention not only to the history and natural environment of the country but also to specific cultural heritage sites. These are virtual tours: "Open-air Museums of Ukraine"[6], "The Carpathians in 3D" [7], 3D-tour "Chocolate House [8] and other similar projects.

Regarding the digitization of museum collections, until recently museums in Ukraine used the regulatory framework of the mid-80s of the twentieth century, focused exclusively on paper ways of thinking and information processing. As a result, today more than 95% of the information in museums is stored on analogy media, moreover, there is still no list of cultural heritage objects to be digitized. Only recently has the Government of Ukraine identified the formation of cultural heritage registers as a whole as a priority in cultural policy. Another problem with the digitization of museum collections is compliance with standards for the electronic description of museum objects, which will allow museums that generate information to freely exchange data and transfer it to the state register. The main requirements for the information system of the museum are clearly spelled out in the Law of Ukraine "Procedure for an accounting of museum objects in electronic form" (2016) [9], in particular, the requirements for the software of the museum information system: "export-compatibility with the ontological model CIDOC Conceptual Reference Model (www.cidoc-crm.org) and XML-scheme for the exchange of data on museum objects codeUA".

But the main problem of digitization of historical and cultural heritage sites in Ukraine is the lack of scientifically sound developments on the rules and standards for creating a digital resource to ensure accessibility, active and efficient use, compatibility and data exchange at the local and international levels, preservation of the created resource and its actualization. Therefore, it is appropriate to point out the need to develop a strategy for the future construction of digital collections, which should take into account the following aspects:

- planning of the digitization process: determination of the purpose of the project, target audience, and rules of selection of objects for digitization; assessment of available assets for the project implementation (eg, condition of the object and other characteristics); determination of necessary resources for the project implementation (necessary technologies, specialists, etc.); work planning according to available resources;
- coordination of legal issues: determination of rules of access to resources under current Ukrainian and international legislation;
- compliance with generally accepted standards and management principles: definition of metadata standards for cultural heritage sites; definition of image storage formats; organization of the system of description of cultural heritage objects, accounting and management of digital resources; choosing the system architecture and method of resource integration;
- support for access and preservation of the created digital resource.

Thus, the state of legal, institutional and organizational support for the preservation of historical and cultural heritage sites in Ukraine urgently requires the development of a clear national policy and strategy for the preservation of digital heritage. The European model for the development, promotion, and implementation of strategies for digitization, access, and preservation of digital heritage can become a model for the development of an appropriate strategy in Ukraine.

2. Ukrainian Experience in the Reconstruction of Historical and Cultural Heritage Sites

Although the dialogue between culture and technology in Ukraine is just beginning, there is already some experience in the application of information and communication technologies for the preservation of cultural heritage sites.

The Pocket Country project was launched in 2020. Pocket Country is the largest 3D digitization of objects of the material cultural and architectural heritage of Ukraine with the support of the Ukrainian Cultural Foundation. The executor of the project is the Lviv company "SKEIRON" - one of the few in Ukraine, which has set ambitious goals: monitoring the state of preservation of the architectural and the cultural heritage of Ukraine; drawing attention to the problems of preserving the architectural and cultural heritage of Ukraine; digitization of architectural monuments to carry out high-quality repair and restoration works; popularization of new methods of working with architectural heritage; creation of a modern (virtual) archive of 3D models of monuments of Ukraine for further use by architects, sculptors, and restorers to restore the original state of objects [10].

Among the completed and presented to the public projects "SKEIRON" - 3D tours of such prominent cultural sites as the Church of Saints Olga and Elizabeth (Lviv, built in 1911), Church of St. Mark (Varyazh village, Lviv region, built in 1693), Villa of Emil Dunivsky (Lviv, built in 1898), Church of St. Joseph (Pidhirtsi village, Lviv region, built in 1766), Monument of St. Yuri Zmieborets (Lviv, opened in 1999) [11]. This is the result of the promotional project *"Pocket City: Lviv"*, implemented at the initiative of Skeiron and Gwara Media, a project that not only introduces technologies and their capabilities but also indicates the need to digitize architectural monuments. As of 2019, 24 buildings were scanned in Lviv: the Potocki Palace, the Opera House, the Boim Chapel, Mykhailo Hrushevsky's Villa, and others. An application “Lviv. Pocket City ", the main function of which is to interact with architectural objects interactively was created for this project. The presentation of this project took place at the AR-exhibition, which lasted for a month and allowed everyone to see images of architectural monuments in AR mode with the help of a special application. The app also works with 3D postcards and 3D books that have appeared in libraries and bookstores. Part of the proceeds from the sale of printed materials was used to digitize abandoned architectural monuments, and the digital models created were provided to specialists to develop restoration and reconstruction projects for buildings that had already lost hope.

The Skeiron team has implemented a number of other projects: "Pocket City: Chortkiv", "Tartakiv & Tartak", as well as the social project "Feel Ukraine to the Touch: Lviv".

"Pocket City: Chortkiv" is a project aimed at preserving the architectural heritage in a modern way and a clear presentation of the opportunities that the digitization of monuments offers today. The project was implemented by Skeiron together with the Center Initiative NGO, with a grant from Re_HERIT [12] and the city council for work with cultural heritage. Eight monuments representing Chortkiv’s multiculturalism were selected for the project: the Clock Tower and shopping malls; Chortkiv Fortress; Church of the Ascension of the Lord; Church of the Assumption of the Blessed Virgin Mary; Church of St. Stanislaus; New Chortkov Synagogue; Chortkiv City Hall; Monastery of the Sisters of the Order of Mercy. Selected sights are depicted on postcards, where you can use the smartphone app to view them in 3D and listen to their history. The project was presented in June 2020 [13].

As part of the *Tartak & Tartak project*, the SKEIRON team measured and calculated the smallest details of the Tartak Palace, which in the future will allow solving the problem of restoring the architectural monument by carrying out design restoration drawings.

The project *"Feel Ukraine to the Touch"* is based on the idea of helping blind people "see" the world around them, which arose in cooperation with the Resource Center of Lviv Polytechnic National University with the IT company ELEKS. The ELEKS team together with the Skeiron and Polyspase studio teams created 3D models of architectural structures for the blind. Among the digitized objects: Lviv State Circus, House of Organ and Chamber Music, monuments to Adam Mickiewicz and Ivan Franko, Latin Cathedral, Church of the Ascension of Christ and Powder Tower. The Skeiron team digitized the objects [14].

It was these SKEIRON projects that prompted them to continue their work to draw attention to the destruction of architectural monuments in various regions of Ukraine and to launch a large-scale digitization campaign to capture their current state. The next project of SKEIRON *"Pocket Country"* envisages to cover all 24 regions and the Autonomous Republic of Crimea of our country from East to West and from North to South, and to digitize 40 objects of architectural heritage. Crimea and Donbas are Ukraine, so to create a comprehensive interactive map of Ukraine's landmarks, it is planned to develop a model of the Swallow's Nest (Yalta) and Donetsk Airport with the help of photos and videos from the network. All digitized objects will be available in the format of 3D models and augmented reality on iOS and Android in the free application Pocket City AR with an audio guide in Ukrainian and English. Pocket City AR is an application for fans of intellectual and interactive tourism and unexpected discoveries of cultural heritage, where digital objects and information about their uniqueness will be collected. Those interested will be able to virtually walk through the tourist business cards and landmarks of the regions of Ukraine in an augmented reality format. Pocket City AR will be created exclusively to promote architectural monuments as tourist attractions at the national level. 3D-scanning will allow capturing the current state of monuments, and the use of AR-technologies - to develop souvenirs to promote the Ukrainian architectural heritage within the country and abroad. Those interested will be able to virtually "walk" with tourist business cards and landmarks of the regions of Ukraine in the format of augmented reality.

One of the completed and available for viewing objects is the State Historical and Cultural Reserve "Trakhtemyriv" (Kaniv, Cherkasy region). For the first time in Ukraine, the technology of laser scanning of the Earth was widely used to study the historical and cultural heritage and preserve archeological monuments in Cherkasy region. The purpose of the project "Trakhtemyriv-Digital" - to create a 3D-model of the landscape of the Trakhtemyr Peninsula with fixation of known and discovered monuments of archeology and nature, to create a scientific description and catalog of all discovered monuments, to disseminate for specialists the technique of landscape photography objects. The project uses LIDAR technology (Light Identification Detection and Ranging), which is one of the world's newest digital technologies for the study of archaeological heritage, examples of its application are still isolated, but mostly bring sensational results [15].

The project will also allow a deeper and more detailed study of the condition, location, volume of monuments of different eras, possibly to find unknown objects (ancient roads, mounds, remains of settlements), thus, it will provide a better understanding of the course of historical and cultural processes on the territory of the Trakhtemyr Peninsula in the all-Ukrainian and all-European context. It is worth noting that the innovative experience gained by the reserve in the use of modern methods for monument protection is widespread in the relevant expert community, in particular, among fellow scientists from foreign research centers.

In 2019, Kropyvnytskyi hosted a presentation of the final of the project "Digitalization of cultural heritage monuments", according to which 30 plates with QR-codes were installed on the city's cultural heritage monuments. The signs describe the sights in two languages -

English and Ukrainian, and QR-codes - in six languages. There is also an NFS tag with which you can access the audio guide [16].

In Transcarpathia, the first in Ukraine digital infrastructure tourist service "Castles Transcarpathia" is being developed. The authors of the idea are the founders and organizers of the first cultural and technological festival DRONEVAR (Uzhhorod, 2017) TM "Creative Industries of the Territory" and NGO "Cultural Platform of TranscarpathiaThe idea of the project is to transfer the castles of Transcarpathia to the digital world 4.0 and open access to the cultural and historical heritage of "Silver Land" to the world, using immersive technologies in a new format to tell the history of the region. To do this, the organizers plan to develop a cross-media platform for iOS and Android systems, which will be an interactive map of Transcarpathian attractions, restore and supplement local cultural sites, as well as modernize their presentation for future generations. Technologies used include Dronetech, VR, AR, Live video, Compositing (creating a holistic image by combining two or more layers of film or video footage, as well as computer objects and animations. Widely used in modern computer technology to create visual effects in movies, television production, advertising, when you need to perfectly simulate reality or artistic effect), 3D animations, Interactivity, Photogrammetry, Equirectangular, 360 ° projection, 3D sound.

At present, a three-dimensional structure of Serednyansky Castle has been created, more precisely a dungeon (or tower of the last defense), the only fortress of the Order of the Templars in Ukraine, in the village of Seredne, Uzhhorod district, Zakarpattia region. Now only the ruins of this fortress remain on the territory of this castle, so the designers recreated the appearance of a medieval marvel of architecture: the first stone of the castle was laid in 1146 at the expense of Froy de Barmagne, chaplain of Transcarpathia in the XII century and the prototype was the ancient Roman border watchtowers on the Rhine and the Danube.

In addition to the three-dimensional constructions of the castles of the region, according to the project, the originators plan to create an interactive audio-AR tour of the object; VR-tour, available anywhere in the world with the Internet and VR-headset; VR-AR lessons for schoolchildren; as well as 3D-reconstruction of historical spaces. This is a large-scale project, the materials of which will be available in Ukrainian, English, Hungarian, Slovak, Czech, and Polish. VR materials will show the historical and geographical facts of epochs and figures, as well as artifacts that were directly related to cultural heritage sites. School VR-AR materials will be guided directly by the teacher, who during the thematic lesson will be able to control all the students' glasses from the teacher's device. At present, software solutions for the production of prefabricated models of locks on a 3D printer have already been created based on the complex Dronetech solutions on the methods of scanning architectural objects. The ultimate goal of prototyping is to create puzzle toys and game designs for the educational and developmental component of public presentation of cultural heritage sites of Ukraine inside the country and abroad [17].

The Pixelated Realities team scans the buildings and sculptures of Odesa creates a database with information about the local cultural heritage. Pixelated Realities is a non-profit public organization from Odesa that develops digital tools available to people for the preservation and restoration of cultural heritage, as well as for the design and renovation of public space. The purpose of Pixelated Realities is to establish a dialogue between government, business, and the public on cultural heritage and public space, as well as to promote digital methods of solving these problems.

Using 3D modeling, the Pixelated Realities team makes virtual city tours, as well as augmented reality applications for smartphones and computers. In this way, Pixelated Realities tells people about heritage sites and the problems associated with them. In addition, they show how to use modern technology to preserve and continue to work with the cultural heritage of

the city. Among the scanned objects are Duke Square, statues of the founders, and prominent figures of Odesa, Dox's dacha, and Passage

In order to create a database of digitized monuments, the team is implementing the ARCH F6 project. These materials will make it possible to control and significantly improve the quality of restoration and reconstruction of historic buildings in the city. The project has a Google map, which can be marked with the object to be digitized. Everyone who cares can join in preserving the cultural heritage of their city. Pixelated Realities' portfolio includes digitized objects not only in Odesa. For example, Pixelated Realities is working to create an *Interactive Digitized Museum on Post Square in Kyiv*, the site of ancient wooden houses and fortresses in ancient Kyiv. The team is digitizing all the archeological finds in the area and developing a VR tour that can be downloaded from Google Play. The list of digitized objects is constantly updated both in Odesa and in other cities of Ukraine.

The Dialog Art Council together with the Lviv Art Gallery is implementing the *Pinzel.AR* project, which involves the creation of three-dimensional high-precision models of Pinzel sculptures in Ukraine and abroad, and is developing a website and applications for smartphones. Thus, the project promotes new approaches to non-formal education using modern 3D modeling technologies and augmented reality. The idea of the project is to make sense of Pinzel's legacy through IT technology. The project to digitize the legacy of the Baroque sculptor Johann Georgy Pinzel with augmented reality, which allows you to watch 3D models of the sculptor's work in different modes, contains information about individual sculptures and their features, allows you to learn more about the artist. The Lviv Museum of Sacred Baroque Sculpture in Lviv is the first to present its collection in digital format in Ukraine [14].

AERO 3D has been digitizing cultural heritage for more than three years. During this time, they managed to digitize the objects of museums, architectural monuments, and territories. With the help of Matterport technology (ed. technology that allows you to create an interactive 3D-model of the room), the team digitized and created a virtual *3D-tour of the Municipal Museum named after O.B. Bleschunov* in Odesa. The technology of creating virtual tours allows you to conveniently view a virtual tour from your phone using the application. Among the projects, you can also find a virtual tour of the *National Museum of the Holodomor Victims Memorial* or view digitized monuments of Kyiv, including the *Independence Monument,* which is included in the TOP 10 digitized models of cultural heritage according to Sketchfab (ed. platform for publishing, distribution, sale, and purchase of 3D, VR, and AR content) [14].

In March 2019, a joint project of the Ministry of Culture and Google Ukraine "Authentic Ukraine: Intangible Cultural Heritage" [18] was presented. Over the past two years, as part of the Authentic Ukraine campaign, the Ministry of Culture of Ukraine and Google Ukraine have been working to create unique web resources that introduce users to the cultural heritage of different regions of Ukraine. During this time, 8 outstanding monuments of the Carpathian region, wooden churches on the UNESCO World Heritage List, 7 open-air museums and opera houses of Ukraine were digitized. The resource provides coverage of areas and elements in which the intangible cultural heritage is manifested: oral traditions and forms of expression; Performing arts; customs, rites, celebrations; knowledge and practices concerning nature and the universe; traditional crafts. The resource contains a brief description of each area in which intangible cultural heritage is manifested, information on elements included in the National List of Intangible Cultural Heritage of Ukraine, a small test for mastering the material and a list of additional literature.

By means of the newest information and geodetic technologies the State historical and cultural reserve "Tustan" has started realization of the project "Tustan_virtualna"

("Tustan_virtual"). Tustan is a medieval rock fortress-hail and customs of the IX-XVI centuries, an ancient Russian rock defense complex, the remains of which are located in the Ukrainian Carpathians. It is a landmark of archeology of national importance. The project "Tustan_virtualna" was launched primarily as a research tool to study a unique type of medieval wooden construction and its graphic reconstruction. Today, virtual and augmented realities are the main means of non-destructive visualization of the lost architecture of Tustan. Also, the 3D model of rocks and fortresses is the basis for planning emergency, conservation and restoration work, museumification of the monument. Tustan_virtual will also become an information platform for intelligent real-time monument management [19]

Thus, many projects on the digitization of cultural heritage have been implemented in Ukraine. Some of them are funded by the Ukrainian Cultural Foundation - a state institution established, in particular, to promote the development of national culture and art in the country, wide access of citizens to national cultural heritage, support cultural diversity, and integration of Ukrainian culture into the world cultural space [20]. One of the UCF projects is *"Digital Museum Collections of the Kyiv Fortress"* - a large-scale development of a comprehensive modular software and technology solution for electronic accounting and management of digitized collections in the National Historical and Architectural Museum "Kyiv Fortress" as well as the creation of a web resource where 3D tours and 3D models of museum exhibits - more than 13 thousand digitized museum objects – are available [21].

The project is aimed at the introduction and development of modern information technologies and digitalization in the museum: electronic accounting in the funds in accordance with the requirements of the "Procedure for accounting museum objects in electronic form" (Order of the Ministry of Culture of Ukraine № 784 from 09.09.2016), digitization and digital collections museum funds with presentation to their users through the web resource of the museum. The project will also create technological and methodological principles for further export of information about museum collections to a single electronic information resource at the Ministry of Culture of Ukraine and integration of digital collections into national and global web resources, including the cross-European web resource of digitized heritage "EUROPEANA". The project will be implemented on 1000 museum objects, using the available software "DC-Visu Digital Collection Management Systems". The result of such a large-scale project was software and technological solutions for electronic accounting and management of digital collections in the museum, in particular: digital collections of the Kyiv Fortress; 3D tour and 3D models and videos 3600; management and accounting system [22].

The project *"Fortress without Barriers"* was implemented in open-air exhibitions on the territory of the museum and in the building "Oblique Caponier", as well as partially - on the online resource of the institution. The exhibits and artifacts from the museum's collection have been adapted for the perception of the visually impaired. 3D-models for tactile perception were also created, the tour was adapted and guides to the territory and exposition of the Kyiv Fortress were published in Braille. To test the results, a guided tour was conducted [23].

The program "Museum Funds of Ukraine" for 10 users has already been developed in Ukraine. "Museum Funds of Ukraine" is a configuration created on the UA-Budget platform. Its purpose is to record museum objects in Ukrainian museums. This accounting includes registration, admission for temporary / permanent storage, distribution by funds, transfer of museum objects to exhibitions, restoration, etc., classification of objects according to all the rules of museum business, obtaining various regulatory and analytical reports [28].

3. Conclusion

The considered projects are only the beginning of the presentation of the Ukrainian historical and cultural heritage in the modern universe of digital format. At present, conscious Ukrainians plan to launch a project to unlock the potential of Ukrainian culture and science through the innovative platform-marketplace *"Ukrainian National Intelligence"*, based on which they are going to implement the principle of "Culture in a smartphone". This is the main task of the newly created Office of Cultural Innovations, which includes representatives of such public organizations as "Ukrainian National Intelligence"; Humanitarian Development Fund of Ukraine; Institute of Digital Transformation; Arma Ukraine Audiovisual Authors' Society; The Owners group of companies; a group of companies The Owners, European Profiles SA Ukraine; IPS company. The main goal of the Office of Cultural Innovations is to use the unique digital platform "Ukrainian National Intelligence" to help Ukrainian culture and Ukrainian art become accessible and open to Ukraine and the world with the help of digital technologies. After all, it will be a strong stimulus for cultural tourism, for the creation of horizontal international ties. The founders of the platform are convinced that unique Ukrainian artifacts should become part of the world cultural space and serve as a powerful foundation for domestic and world science. The first step will be the full digitization of national cultural heritage - outstanding works of Ukrainian artists stored in museums, archives, and various scientific and educational institutions. The Office plans to launch a pilot project in the near future to connect the first key Ukrainian museums to the digital platform. According to the plans of the Office of Cultural Innovations, 2020 was to be the year of total digitization of Ukrainian art, which is stored in museums, cultural, educational, and scientific institutions of our country. However, due to the coronavirus pandemic, these plans have not yet been implemented [29].

References

1. Trach, Yuliia, Tolmach, Maryna, Chaikovska, Olena, Gumeniuk, Tetiana. (2020). Problems of Cultural Heritage Preservation in the Context of the Armed Conflict Growth. 10.1007/978-3-030-48939-7_4.
2. Oliynyk O. Social program for the preservation of the cultural heritage of Ukraine.URL: https://blog.liga.net/user/ooliynyk/article/15000
3. Shestova S. M. Principles of World Cultural Heritage Protection and Global Challenges. URL: https://cyberleninka.ru/article/n/printsipy-ohrany-vsemirnogo-kulturnogo-naslediya-i-globalnye-vyzovy
4. Norris L., Lankeline V. Report on the development of the cultural heritage sector in Ukraine. URL: https://www.culturepartnership.eu/upload/editor/2017/Subsector_report_Ua_ukrainian_Fin.pdf
5. Prykhodko L. Digitalization of cultural heritage objects according to the regulatory documents of the european society. URL: http://au.archives.gov.ua/index.php/au/article/ view/29
6. Open-air museums of Ukraine. URL: https://museums.authenticukraine.com.ua/ua/
7. The Carpathians in 3D. URL: https://xn--3-7sbar0a5ags5f.com/uk/primenenie.html
8. Chocolate House Art Center in 3D. https://my.matterport.com/show/?m=ZfDK3wJC2Ue
9. Law of Ukraine "PROCEDURE for the accounting of museum objects in electronic form", dated November 15, 2016, for № 1478/29608. URL: https://zakon.rada.gov.ua/laws/show/z1478-16#Text
10. Pocket City Air [Official website].URL: https://pocketcityar.com/about//
11. SKEIRON [Official website].URL: https://skeiron.com.ua/
12. Re_HERIT – a project of rethinking, actualization, and development of the cultural heritage of Uman and Lviv together with other cities of Ukraine.

13. Chortkiv in 3D: eight sights of the city that are worth seeing.URL: https://chortkiv.city/read/misto/84732/chortkiv-u-3d-visim-viznachnih-pamyatok-mista-yaki-varto-pereglyanuti
14. Rudy V. How IT helps to preserve and restore architectural monuments. URL: https://euprostir.org.ua/stories/141806
15. Pocket Country project launches in Ukraine. URL: https://api.visicom.ua/uk/posts/poctkercitymap230720
16. The project "Digitalization of cultural heritage monuments" was presented in Kropyvnytskyi. URL: https://dozor.kr.ua/post/proekt-didzhitalizatsiya-pamyatok-kulturnoi-spadschini-prezentuvali-u-kropivnitskomu-foto-11489.html
17. Digitization of cultural heritage. URL: https://cultpz.org/otsyfruvannya-kulturnoyi-spadshhyny/
18. Authentic Ukraine. URL: https://authenticukraine.com.ua/
19. Tustan. Fortress city. URL: https://old.tustan.ua/tustan_3d/
20. Ukrainian Cultural Foundation. URL: https://ucf.in.ua
21. A virtual visit to the expositions of the Kyiv fortress. URL: https://my.matterport.com/show/?m=nBY4tFp9GuG
22. Digital museum collections of the Kyiv fortress: introduction of electronic accounting and creation of a web resource. URL: http://dig-content.com.ua/
23. "Fortress without Barriers". URL: https://ucf.in.ua/archive/5f08732ff299ad7a72663d73
24. The program "Museum Funds of Ukraine". URL: https://soft.integra.ua/uk/product/programa-muzejni-fondy-ukrajiny-na-10-korystuvachiv/
25. The Office of Cultural Innovations has opened in Ukraine. URL: https://www.prostir.ua/?news=v-ukrajini-pochav-pratsyuvaty-ofis-kulturnyh-innovatsij-2
26. Digital museum collections of the Kyiv fortress: introduction of electronic accounting and creation of a web resource. URL: http://dig-content.com.ua/
27. "Fortress without Barriers". URL: https://ucf.in.ua/archive/5f08732ff299ad7a72663d73
28. The program "Museum Funds of Ukraine".URL: https://soft.integra.ua/uk/product/programa-muzejni-fondy-ukrajiny-na-10-korystuvachiv/
29. The Office of Cultural Innovations has opened in Ukraine. URL: https://www.prostir.ua/?news=v-ukrajini-pochav-pratsyuvaty-ofis-kulturnyh-innovatsij-2

Cyber-attacks Scenarios Analysis

Serghei Ohrimenco[1], Valeriu Cernei[1], Eduard Ryzhkov[2]

[1] ASEM, Chisinau, Moldova
[2] DSU, Dnipro, Ukraine
osa@ase.md, valeriu.cernei@bsd.md, revord924@ukr.net

Abstract: The digital transformation, in all its aspects, is accompanied by a rapid growth of cyber threats and attacks on various critical infrastructure facilities. Some of cyberattacks target personal data, operations management processes, financial reporting results, etc., having as a final objective monetization of the results. Accordingly, there is a need for a complex and comprehensive analysis of information, which presents key aspects of cyber-attacks. The current paper presents the results of a study aimed to identify and analyze possible scenarios, stages and facets of cyberattacks, highlighting potential motives, goals, objects, means of attack, specific actions, as well as the final potential results of cyber-attacks.

Keywords: Cybersecurity, Attacks, Shadow Digital Economic, Cybercrime, Advanced Persistent Threat

1. Introduction

Cybercrime is a successful business. Reports of world technology and information security leaders confirm this fact. We present some of the analysis results, which are presented further. In 2016, Cybersecurity Ventures predicted that cybercrime would cost the world $6 trillion annually by 2021, up from $3 trillion in 2015.

If it were measured as a country, then cybercrime — which is predicted to inflict damages totaling $6 trillion USD globally in 2021 — would be the world's third-largest economy after the U.S. and China.

Cybersecurity Ventures expects global cybercrime costs to grow by 15 percent per year over the next five years, reaching $10.5 trillion USD annually by 2025, up from $3 trillion USD in 2015. This represents the greatest transfer of economic wealth in history, risks the incentives for innovation and investment, is exponentially larger than the damage inflicted from natural disasters in a year, and will be more profitable than the global trade of all major illegal drugs combined [1].

"In 2021, cyber criminals adapted their attack strategy to exploit vaccination mandates, elections and the shift to hybrid working, to target organizations' supply chains and networks to achieve maximum disruption," said Maya Horowitz, VP Research at Check Point Software. "The sophistication and scale of cyber-attacks will continue to break records and we can expect a huge increase in the number of ransomware and mobile attacks. Looking ahead, organizations should remain aware of the risks and ensure that they have the appropriate solutions in place to prevent, without disrupting the normal business flow, the majority of attacks including the most advanced ones. To stay ahead of threats, organizations must be proactive and leave no part of their attack surface unprotected or unmonitored, or they risk becoming the next victim of sophisticated, targeted attacks" [2].

Should we look into the future, global cyber-security predictions for 2022 [3], we state that the cyber space becomes a field for misinformation campaigns, and technology related threats, focused on mobile devises continue to grow. Accordingly:

1. Fake news 2.0 and the return of misinformation campaigns.
2. Supply chain cyber-attacks continue to grow, and governments will address the challenge.
3. The cyber 'cold war' intensifies.
4. Data breaches are larger scale and more costly.

Technology cyber-security predictions for 2022:

1. Mobile malware attacks increase as more people use mobile wallets and payment platforms.
2. Cryptocurrency becomes a focal point for cyberattacks globally.
3. Attackers leverage vulnerabilities in micro-services to launch large-scale attacks.
4. Attackers weaponize deep fake technology.
5. Penetration tools continue to grow.

2. Literature Survey

Among the many specialized sources, we consider important to mention the "classic" work of Adam Shostack "Threat Modeling. Designing for Security" [4]. The work contains is focused on modeling threats for the Windows operating system. This is one of the works, which provides an approach to threats description and modeling. A list of key threats is proposed and their analysis is carried out with the allocation of successive stages: Mitigating threats; Eliminating threats: Transferring threats: Accepting the risk.

A number of interesting works have been presented in the recent years. Among them are such as "Threat Modeling. A Practical Guide for Development Teams" [5] (overview of threat modeling, the future of threat modeling methodologies, automation, and Agile development methodologies and etc.) "An Introduction to Cyber Modeling and Simulation" [6] (An Introduction to Assessment and Maturity Frameworks (including DevOps automation, Technical and Operational Scenarios), "Mental Modeling Approach. Risk Management Application Case Studies" [7] (Overview of Mental Modeling Research Methodology, Key Benefits of Mental Modeling, Mental Modeling Core Technique, Key Steps in the Mental Modeling Process), Foundations of Multi -Paradigm Modeling for Cyber-Physical Systems (Multi-Paradigm Modeling for Cyber-Physical Systems, Unifying Framework for Modeling of Physical Systems, Petri Nets [8]) and others.

3. Cyberattacks key characteristics

Cyberattacks are evolving from the perspective of targets, how they affect organizations, and the methods used. Improving cybersecurity protection can unlock economic value by reducing the cost of cybercrime and opening up new revenue opportunities. By understanding where they can gain value in their cybersecurity efforts, leaders can minimize the consequences and even prevent future attacks. By prioritizing technologies that improve cybersecurity protection, organizations can reduce the consequences of cybercrime and unlock future economic value as higher levels of trust encourage more business from customers.

Improving cybersecurity protection can decrease the cost of cybercrime and open up new revenue opportunities.

According to Bissell et al. (2019) [9], the three steps to unlocking the value in cybersecurity are as follows:

1. Priorities protecting people-based attacks: countering internal threats is still one of the biggest challenges with a rise in phishing and ransomware attacks, as well as malicious insiders.

2. Invest to minimize information loss and business disruption: already the most expensive consequence of cyberattacks, this is a growing concern with new privacy regulations such as GDPR and CCPA.

3. Target technologies that reduce rising costs: use automation, advanced analytics, and security intelligence to manage the rising cost of discovering attacks, which is the largest component of spending.

4. Targeted cyberattacks analysis.

The history of attacks on information systems dates back several decades. It started with the spread of computer viruses. Over time, cyberattacks have changed significantly in terms of the tools used, penetration methods, etc., and currently represents a serious weapon of attack against governments, commercial organizations, as well individuals.

The Table 1 below presents a classification of cyber security threats.

Table 1. Classification of Cyber Security Threats

Target application	**Findings and Contributions**
Wireless sensor networks	Categorized the security threats into three levels are: data security level (anonymity and freshness), access security level (accessibility, authorization and authentication) and network security level
Information systems	Established a hybrid model for classifying the security threats for information systems. They classified the security threats into three types: human threats, technological threats and environmental threats
Smart grid	Classified the security threats of smart grid into technical and non-technical resource threats. Technical threats was categorized into three types of threats are infrastructure threats, technical operational threats and system data management threats. While non-technical threats were classified into environmental threats and governmental threats
Wireless sensor networks	Classified the attacks that could occur in all layers from application layer to physical layer. For example, at the application layer level, a malicious attack can be added along the communication link to generate fake messages and data in order to attack the ongoing communication and increase the data collision. The attack in transport layer happen through sending unlimited connection request in order to minimize the node's energy and exhaust its resources and this lead to denial of service. Other attack can be occurred in a network layer in several forms such as spoofing, sinkhole, flooding and replay attack in order to create and send fake messages or causing congestion in the network. Jamming attack at the Data link layer can cause loss of signals and data and destroy the channel and

	increased interference. At the physical layer level, the attacker can allow unauthorized nodes to access to the network and damage it
Cloud computing	Categorized the security attacks and threats on cloud computing into four levels: authentication Attacks, side Channel Attacks, cloud Malware injection attack and Denial of Service (DoS) attacks
Mobile edge computing	Classified the security threats of mobile edge computing five assets are: (1) Network infrastructure threats such as man in the middle and denial of service attack, (2) Edge data center threats such as physical damage, privacy leakage, privilege escalation and service manipulation, (3) Virtualization infrastructure threats such as denial of service, misuse of resources, privacy leakage and privilege escalation, (4) core infrastructures threats such as privacy leakage, service manipulation, rogue infrastructure and (5) User devices such as injection of information and service manipulation
Blockchain technology threats	Divided the security threats for blockchain technology into five categories are: (1) Double spending threats, (2) Mining/Pool threats, (3) Wallet threats, (4) Network threats such as DDoS attack, and (5) Smart contracts threats
Cyber security threats	Classified the common cyber security threats by using mapping study, which include phishing, denial of service (DoS), injection attack, man-in- the-middle attacks, session hijacking, SQL injection attack and malware
Classification of RFID attacks	Classified threats associated with Radio Frequency Identification systems. They distinguished attacks in the physical layer, network transport layer, application layer, strategic layer, and multilayer
Social engineering malware	Social engineering malware is both pervasive and persistent. Emphasized the importance for organizations to develop a shared social responsibility to combat social engineering malware and not solely on technical solutions. Social engineering malware proliferation through a variety of infiltration channels such as e-mail, social software, websites, and portable media
Social engineering semantic attacks	Introduced a structured baseline for classifying semantic attacks by breaking down into components and identifying countermeasures

Sourse: Yassine Maleh, Youssef Baddi, Mamoun Alazab, Loai Tawalbeh, Imed Romdhani (2021). Artificial Intelligence and Blockchain for Future Cybersecurity Applications. ISBN 978-3-030-74575-2, https://doi.org/10.1007/978-3-030-74575-2 [10]

Majority of security experts and analysts split targeted cyberattacks (Advanced Persistent Threat (APT) into a separate class. A feature of targeted attacks (APT) is that attackers are interested in a specific government organization, company or individual.

Targeted attacks are usually well planned and include several stages – starting with passive information gathering and analysis, to systems penetration and evidence destruction. Because of a targeted attack, attackers gain access to the victim's infrastructure and remain undetected for months or even years. During all this time, they have access to corporate information.

According to the well-known consulting firm A.T. Kearney, the main targets of such attacks are the following [11]:

- Companies' Head Offices. Often the equipment is not adequately protected from physical damage (for example, by cleaning or maintenance personnel).

- Research & Development. Usually this department requires the highest level of protection. Companies' knowledge is generated and concentrated there. However, it is often no better protected than other departments.

- Data centers. They provide secure environments for hosting private clouds. The problem is to ensure secure functioning of numerous servers, as well as applications running on these servers.

- Supply Chain. With the increasing use of networked solutions in dealing with suppliers, there are risks associated with the fact that relatively small supplier companies are usually less protected.

- Cloud computing. Basically, providers are responsible for security of cloud environments. In addition, this is a potential problem. Cloud services providers protect only in the case of SaaS services. All other models (IaaS, PaaS) remain to be the responsibility of the service beneficiary. Another point is the fact that the level of data protection depends on the specific adopted law in each country.

- Production systems. Many legacy specialized systems (e.g. SCADA) are increasingly being connected in networks. These are difficult to monitor and control. Attacks by intruders in this case can lead to production losses or even to collapse of a company or even an entire industry (e.g. Critical Infrastructure).

- Databases provide secure storage of important information. Main worries are that hackers can use administrators as "tools" to break into databases.

- Office networks. Offices, systems and PCs are interconnected, which provides many opportunities for network penetration.

- Sales. Leakage of marketing plans, pricing and customer information may lead to reputation loses as well as loss of competitive advantages.

- Mobile devices. If unsecured, sensitive data from a mobile phone can be easily stolen by hackers. One of the basic principle in using mobile devices is to avoid using personal devices for fulfilling job duties.

- Online stores. Use of stolen credit cards data and/or personal data are main tools applied by hackers.

- Phone calls. By exploiting people's willingness to help each other, attackers can use phone calls as a way to easily get the information they need, including financial data, passwords, personal information, etc.

We further present an overall attacks classification which is focused on mobile platforms [10] (Table 2).

Table 2. Attacks Classification which is Focused on Mobile Platforms

Cyber security attack	Description
Access attack	Allow unauthorized users access to the network or devices such as smart phones with no right to access
Reconnaissance attack	Attack allows an attacker to capturing, discovering and mapping of system vulnerabilities such as scanning traffic network, network ports and IP address information
Physical attack	This type of attack aims to tamper with hardware devices, for example some technologies such as IOT devices operate in outdoor environments may highly susceptible to physical attacks
Denial-of- service (DoS) attack	Denial-of-service (DoS) attack allow the attacker to make the network or device services unavailable to its intended users due to several reasons such as limited computation resources and low memory capabilities. This make mobile platforms are vulnerable to DOS attack
Attack on privacy	Attack on privacy through using remote access methods and malware to spy or stole sensitive information of users or organizations. Privacy protection in mobile devices has become increasingly challenging due to share large amount of information between mobile devices
Password- based attack	Attackers in two ways can do password- based attack: (1) brute force attack by using cracking tools to guess the correct password in order to access valid password, (2) dictionary attack depends on trying several letters and numbers to guess user passwords
Supervisory Control and Data Acquisition attack	Supervisory Control and Data Acquisition attack using malware such as Trojan to take control of the system. Mobile applications are vulnerable to many cyber-attacks like Trojan virus
Spoofing attack	Spoofing attack is based on obtaining the IP address of the devices to attack the users through enabling attackers to access users' confidential data and use it for malicious purposes
Botnet attack	Botnet attack is based on a collection of Internet-connected devices that have been breached and ceded to a malicious device known as botnet controller. The botnet controller able to direct malicious activities in order to damage the network or exploit users and data for materialistic gain
Sybil attack	Sybil attack is a threat in which attacker attempt to obtain identity of honest user, pretend as a distinct user, and then attempt to create relationships with honest users. If the attacker is successful in compromising one of the honest users, he will gain unauthorized privileges that help in the attacking process

From the perspective of its life cycle, a cyberattack can be split into 4 Stages [12,13]:

1. Preparation - the main objective of the first phase is to find the target, collect enough detailed private information and analyze it. Based on the information analysis, identify weaknesses in the infrastructure, build an attack strategy, select available hacking tools or develop exploits. Typically, the attack strategy will be thoroughly tested in order to ensure non-detection measures.

2. Penetration - the active phase of a targeted attack. Social engineering techniques and zero-day vulnerabilities are applied to initially compromise the target and conduct internal diagnostics. Additional malicious code or special settings are being applied when the ownership of the compromised host (server/workstation) is done.

3. Distribution – the objective is to extend hacker's control as much as possible by inspecting the IT environment and adding new unauthorized code through control centers.

4. Achieving the goal is the key phase of a targeted attack. Depending on the final objectives and chosen strategy, it can be theft of classified information; deliberate change of classified information; manipulation of the company's business processes and so on.

A mandatory activity is to, at every stage, to hide traces and evidence of a targeted attack. It often happens that cybercriminals create "Points of Return", allowing them to return in the future.

There are sources, [14] which split the life cycle of an attack into 6 stages, including:

1. The cyber-criminal, or threat actor, gains entry through an email, network, file, or application vulnerability and inserts malware into an organization's network. The network is considered compromised, but not breached.

2. The advanced malware probes for additional network access and vulnerabilities or communicates with command-and-control (CnC) servers to receive additional instructions and/or malicious code.

3. The malware typically establishes additional points of compromise to ensure that the cyber-attack can continue if one point is closed.

4. Once a threat actor determines that they have established reliable network access, they gather target data, such as account names and passwords. Even though passwords are often encrypted, encryption can be cracked. Once that happens, the threat actor can identify and access data.

5. The malware collects data on a staging server, then exfiltrates the data off the network and under the full control of the threat actor. At this point, the network is considered breached.

6. Evidence of the APT attack is removed, but the network remains compromised. The cybercriminal can return at any time to continue the data breach.

5. Cybercrime Attackers Taxonomy

There are many hacker categories; these categories include different terminology that create controversy over the computer attacker terms. Many of hackers' activities cannot be considered as illegal. We have to distinguish hackers who commit crimes and cybercriminals. The difference rests upon their attitudes and the motives [22-24].

The proposed taxonomy is based on technical experience, behavior, motivation and the level of moral development and is split into seven categories:

- Script Kiddies (SK) - the least skilled and youngest members using the tools created by elite hackers, who run precompiled software to harm individual users, systems and networks.
- Cyber-panks (CP) – these people have disrespect for Governments. They disregard social and moral norms. Recognition offered by peers and the society represent key drivers.

- Haktivists (H) – the name "hacktivist" sounds more respectful for themselves than calling criminals. These actors tend to justify their destructive behavior, including defacing websites, by labeling them as civil activists and attributing to their actions political and moral correctness.
- Thiefs (T) – this group targets information systems for financial gain. They will be focused on illegal collecting of credit card and bank account numbers that can be used for immediate personal gain. This group can accurately be called "light" criminals, as their actions are usually not sophisticated.
- Virus Writers (VW) – this category includes both, technically trained and beginner actors. This category combines four subcategories, namely: teenager, college student, adult and former virus writer. Viruses still form a very profitable segment of the crime software market.
- Cyber terrorist (CT) – members of this group may be part of military or paramilitary forces of a nation. They may be seen as "soldiers" or "freedom fighters" in the new cyberspace. Their activities are associated with the commission of terrorist acts in the cyberspace.
- Professionals (P) / Elite – They have the knowledge and skills of the highest level. This is the most elite of the cybercriminal groups. This status can be gained by a particularly famous exploit, hack or longevity on the scene. Representatives of this group may be involved in sophisticated scams or corporate espionage. They will sell confidential information and intellectual property for the highest price. Only limited information is known about this underground group as they use strict anonymity to hide their activities. The criminal activity is associated with day-to-day work.

In [15], the model includes eight main categories: 1. Novice (NV), 2. Cyber-punks (CP), 3. Internals (IN), 4. Petty Thieves (PT), 5. Virus Writers (VW), 6. Old guard hackers (OG), 7. Professional criminals (PC), 8. Information Warriors (IW).

Taking into account the motivation of criminals, cybercrime can be conditionally divided into the following categories:

- cyber fraud with a final objective to steal funds;
- cyber fraud a final objective to steal information (for own use or for subsequent sale);
- information systems disruption in order to gain control (for deliberate damage for a fee or to damage competitors);
- Other crimes.

6. A new segment – cloud computing

One of the new targets for cybercriminals is the cloud computing (CC) environments. The cloud computing technology in used not only for optimizing investments, better accessibility, less risks, increased collaboration, etc., but also hackers, spammers, and scammers use criminal software (Crimesoftware-as-a-Service) to access resources, fulfill their computing needs, organize a set of attacks on a competitor's resources, or develop new malicious technologies [16-18].

Hackers exploit benefits of cloud technologies as well. They do not need to invest heavily in building their own computer networks of information resources with high bandwidth. All these services are provided through cloud computing at a low cost. By developing and using criminal software as a commercial product (Software-as-a-Service) to gain access to resources, they are able to organize sets of attacks on the resources of a competitor or develop new malicious technologies.

Using the cloud computing environments facilitates realization of attacks as distributed denial of service (DDOS), password decryption (password decryption), hash cracking, organize activities related to shadow economy (e.g. drugs/arms/malicious software trafic) etc. It should be noted that there are automated technologies aloowing to plan and conduct DDoS attacks, Port Scanning, Flooding Attack DNS attacks, Sniffer attacks, Prefix Hijacking, IP Address reuse, IP Spoofing, Fragmentation Attack, Deep Packet Inspection, Active and Passive Eavesdropping, and so on.

Gaining access to the cloud-computing environment by simply passing the registration process enables easy distribution of criminal products and services. As it was mentioned earlier, this environment is also used as a commercial platform and distribution of malicious products (e.g. viruses, trojan horses, keyloggers, ransomware, botnets, etc.). Subscribed users get access to functionalities allowing distribution of protected archive files, they get guarantees for the performance of certain applications, get acquainted with the reviews of other users and so on.

7. Conclusion

The threats forms and composition is constantly changing and a constant increase in the number of negative cyber impact is expected.

One of the main threats is the COVID-19 pandemic, which has revealed many problems related to remote user access and "home" work. First, these are information security risks - connecting personal computing devices to information networks. Once the COVID-19 [19-21] problem spread all over the world, phishing activities grew significantly. Remote access devices, not prepared and configured to face the new reality, have become themselves additional threats. The goals of cybercriminals have changed - if earlier financial organizations were considered as main targets of cyberattacks, during the pandemic there was a shift. Hackers understood that by compromising end users, they can easier target government organizations, industrial enterprises, the energetic sector, medical institutions and so on.

References

1. Steve Morgan (2021). Cybercrime To Cost The World $10.5 Trillion Annually By 2025. https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/
2. Check Point Software 2022 Cyber-security Predictions also anticipates an increase in supply chain attacks in the new year. https://blog.checkpoint.com/2021/10/26/deepfakes-cryptocurrency-and-mobile-wallets-cybercriminals-find-new-opportunities-in-2022/
3. Cybercrime predictions for 2022: Deepfakes, cryptocurrencies and misinformation. https://www.weforum.org/agenda/2021/11/2022-cybercrime-predictions-checkpoint/
4. Shostack A. (2014). Threat Modeling/ Designing for Security. John Wiley & Sons. ISBN: 978-1-118-81005-7
5. Izar Tarandach, Matthew J. Coles (2021). Threat Modeling. A Practical Guide for Development Teams. O'Reilly Media .ISBN 9781492056553
6. Jerry M. Couretas (2019). An Introduction to Cyber Modeling and Simulation. John Wiley & Sons. ISBN 9781119420835
7. Matthew D. Wood, Sarah Thorne, Daniel Kovacs, Gordon Butte, Igor Linkov (2017). Mental Modeling Approach. Risk Management Application Case Studies. Springer Science+Business Media. ISBN 978-1-4939-6616-5 DOI 10.1007/978-1-4939-6616-5
8. Paulo Carreira, Vasco Amaral, Hans Vangheluwe (2020). Foundations of Multi-Paradigm Modelling for Cyber-Physical Systems. ISBN 978-3-030-43946-0 https://doi.org/10.1007/978-3-030-43946-0
9. Kelly Bissell, Larry Ponemon (2019). The Cost of Cybercrime. Ninth Annual Cost of Cybercrime Study Unlocking the Value of Improved Cybersecurity Protection. https://www.accenture.com/_acnmedia/pdf-96/accenture-2019-cost-of-cybercrime-study-final.pdf

10. Yassine Maleh, Youssef Baddi, Mamoun Alazab, Loai Tawalbeh, Imed Romdhani (2021). Artificial Intelligence and Blockchain for Future Cybersecurity Applications. ISBN 978-3-030-74575-2 https://doi.org/10.1007/978-3-030-74575-2
11. Advanced persistent threat (APT): Targeted or targeted attacks, ttps://www.tadviser.ru/index.php/Статья: APT_Таргетированные_или_целевые_атаки
12. Advanced persistent threat (APT). https://www.imperva.com/learn/application-security/apt-advanced-persistent-threat/
13. Advanced Persistent Threat (APT) Attacks. https://www.cynet.com/network-attacks/advanced-persistent-threat-apt-attacks/
14. Anatomy of Advanced Persistent Threats. https://www.fireeye.com/current-threats/anatomy-of-a-cyber-attack.html
15. Marcus K. Rogers (2005). The Development of a Meaningful Hacker Taxonomy: A Two Dimensional Approach. CERIAS Tech Report 2005-43
16. Malte Ackermann (2021). Mobility-as-a-Service: The Convergence of Automotive and Mobility Industries. ISBN 9783030755898
17. Peter Elger, Eoin Shanaghy (2020). AI as a Service: Serverless machine learning with AWS. Manning Publications. 9781617296154
18. Enrique Castro-Leon, Robert Harmon (auth.) (2016). Cloud as a Service: Understanding the Service Innovation Ecosystem. Apress. ISBN 9781484201046
19. Adrian T. H. Kuah, Roberto Dillon (2021). Digital Transformation in a Post-COVID World. Sustainable Innovation, Disruption, and Change. CRC Press. DOI: 10.1201/9781003148715
20. Levi West (2020). The Coronavirus Cybersecurity Survival Guide. Top Tips to Protect You from A Cyber Attack.
21. Robert Slade (2021). Cybersecurity Lessons from CoVID-19. CRC Press. ISBN: 978-1-003-13667-5
22. Ben Buchanan (2020). The Hacker and the State. Cyber Attacks and the New Normal of Geopolitics. Harvard University Press. ISBN 9780674246010
23. Clement Guitton (2017). Inside the Enemy's Computer Identifying Cyber Attackers. Oxford University Press. ISBN: 9780190699994
24. Akashdeep Bhardwaj, Varun Sapra (2021). Security Incidents & Response Against Cyber Attacks. ISBN 978-3-030-69174-5 https://doi.org/10.1007/978-3-030-69174-5

Mapping Sustainability Risk on Upstream Sugarcane Supply Chain Network

Nina Aini Mahbubah [1], Abdul Muid [2]

[1] Industrial Engineering Dept, Universitas Muhammadiyah Gresik. Gresik, Indonesia
[2] Industrial Engineering Dept, Universitas Negeri Malang, Malang, Indonesia
n.mahbubah@umg.ac.id, Abdul.muid.ft@um.ac.id

Abstract. An upstream supply chain commonly causes supply chain disruptions in both the manufacturing and service industries. As one of the prominent national industries, the sugarcane industry in Indonesia has been facing such disruption. Upstream risks in the sugar cane supply chain need to be evaluated to avoid disrupting the national sugar industry supply chain. This study aims to map network configuration and the potential sustainability risks and causes through the upstream sugarcane supply chain. The Supply Chain Operation Reference is used to classify business activities in the upstream sugarcane supply chain. Stakeholders of the upstream sugarcane supply chain have participated as respondents for this research. This study has begun by conducting the Delphi interview method to validate sustainability risk events and agents on the supply chain. In addition, supply chain configuration has been validated to visualize a comprehensive view of upstream sugarcane supply networks. The results of this study found that potential factors and causes of sustainability risk. This study also shows that environmental risk factors are the most prominent factor contributing to the impact of economic losses compared to social risk factors and economic risk itself.

Keywords. Upstream, sugarcane, risk, supply chain, sustainability

1. Introduction

Sugarcane is the leading agricultural product in Indonesia. A tropical climate is one of the main factors for growth. Sugarcane is the essential raw material for sugar production. Sugar Factory has been established in Indonesia since the sixteenth century. Indonesia is one of the top ten sugarcane producers in the world. Indonesia is known as a world-class sugarcane producer. However, the national productivity is ranked ninth among the tenth biggest sugarcane producers worldwide. In addition, the national sugarcane productivity is 65.604 kilograms/hectare or only a level higher than Pakistan's 64,322 kilograms/hectare in the tenth position. Guatemala is a country with the highest productivity globally at 107.673 kg/hectare. India follows them at 80.105 kg/h, China at 77.388 kg/h, Australia, Brazil, and Mexico in fifth, sixth, and seventh place with average productivity of 74 thousand kg/h. With 71 thousand kg/h, Thailand and Colombia are in eighth place [1].

Meanwhile, national consumption has increased in line with population growth. However, the increase in national sugar consumption is not comparable to the national sugar production. Moreover, the national average sugar production of 2 million tons can only meet at least half of the total national sugar consumption with an average of 5 million tons [1].

The sustainability of the sugar industry in Indonesia is essential to improve food independence. The low sugar production is caused, among other things, by the shrinking area of agricultural land on the island of Java. Another contributing factor is the centrality of farmers and sugar factories only in the Java archipelago. There are 43 sugar factories in East

Java, Central Java, and West Java. While outside Java, there are only the Sumatra Islands and Sulawesi Island. Many government policies have been established to increase the productivity of sugar production or develop sugarcane farming outside Java.

The business sustainability trend in the world is experiencing a shift in business orientation. The business orientation, which was only based on economic factors, then considered environmental and social factors as considerations for business sustainability in the 1980s. The United Nations launched the 1987 Brundtland Report through the World Commission on Environment and Development. Business practices do consider economic factors and consider people and the planet [2]. Sustainability practices initially implemented in developed countries have subsequently developed in developing countries such as Indonesia [2]. This research aims to identify configuration of upstream sugarcane supply networks, to identify risk event and risk agent on upstream sugarcane supply network, and to map sustainability risk on upstream sugarcane supply network. This paper begins with an overview of the national sugar industry, followed by a literature review of sustainability and supply chain networks. Sampling and procedure are part of the methodology described in the third section, continued with the results and discussion, and finally, the conclusion.

2. Literature Review

a. Sugarcane Cultivation

Sugarcane cultivation is carried out in four stages. The first preparation is based on climate, weather, and sugarcane cultivation land [3]. The second is seed preparation. The third is planting, harvesting, and shipping sugar cane to an ethanol factory. Agroclimatic, the nature of the climate, is related to the distribution of rainfall. Sugarcane plants need water during the vegetative period. The ideal monthly rainfall of 200 mm/month. It is planted in the lowlands with an optimal temperature of around 24-300C. Wind speed is not more than 10 km/hour to avoid falling tall plants and reduce sugar content/random. Land conditions for growth with a pH between 5.7-7. The land management system that is generally used is Reynoso. The principle of the system makes sewers for the disposal and storage of water. The first step is to turn the soil by plowing to lose, and air can enter adequately. After that, make a ditch or ditch to smooth water disposal [3].

The second part is seed preparation and planting. The government has provided seeds, and many sugar factories have provided them for farmers. After the sugarcane seeds are ready for planting, we need to determine the planting time. Sugarcane planting time is usually in the rainy season, October - December. Sugarcane preservation includes seed selection. Sugarcane seedlings aged one week after planting must be replaced with new seeds if they do not grow or die. The second embroidery can be done four weeks after the second embroidery. Sugarcane plants require sufficient water for 4-5 months of growth. Watering is done every three days until the plants are two weeks old. When sugarcane plants reach 2-4 weeks, watering is done twice a week. When the plant is 4-6 weeks old, watering is done once a week. Finally, at the age of 6-16 weeks, watering is once a month [3].

b. Sustainability Risk of Supply Chain Network

Profit is the main thing that is considered by business people worldwide. Business sustainability depends on the profits generated so that a service and manufacturing business can survive in the global market [4]. The sustainability practices have been implemented not only based on profit-oriented but also considering all industry players' economic and social factors. The implementation of sustainability initiated by developed countries has developed well and was adopted by business people in developing countries in the past three decades

[5-7]. Agriculture as business backbones in most developing countries is also starting to realize the importance of business practices by paying attention to the environment and social aspects of all stakeholders [8, 9]. The green agriculture approach as an approach in managing agricultural business risks is carried out in developed countries and is also implemented in developing countries [5]. Agroindustry in developing countries setting risk mitigation scenarios by considering environmental and social factors in generating business is one example of implementing business sustainability [5, 6, 8].

Mapping the textile industry network in Sri Lanka shows a complete picture of business entities and examples of business practices that focus on triple bottom line sustainability [10]. The mapping of entities and supply chain networks is the right step to map business risks. Risk along the supply chain flow is seen from one factor and can be seen from the entire business supply chain. Determining the right strategy to manage business risk based on a sustainable approach is the right approach in determining the sustainability of the business.

3. Materials and Methods

a. Sampling

This research uses a descriptive qualitative approach. Research respondents are representatives of eleven stakeholders. These stakeholders include rental landowners, fertilizer distributors, sugar cane seed distributors, farmers. In addition, representatives from farmer associations, workers, truck owners, intermediaries were selling ready-to-cut sugarcane, and sugar producers were considered stakeholders. The determination of stakeholders in this study is based on the results of a field survey.

Furthermore, research participants consisted of two farmers, four workers, a founder of a farmer association, a log sugarcane middle man. There were also two truckload drivers, a truck owner, a landowner, a seed agent, a fertilizer distributor, two sugarcane cutters, a foreman. The selected respondents in this study can be categorized as experts in their respective fields because the selected respondents have experience in their field for more than five years.

b. Procedure

The three stages of the research are described as follows. The first stage is to collect data from the previous literature regarding risk events and risk agents for similar studies carried out in the sugar industry [11, 12]. The data collection results were used as a questionnaire design regarding the plan, source, make, delivery return. Questionnaire design refers to the Supply Chain Operation Reference Based Approach in Defining Risk. In addition, fieldwork was conducted to get an accurate picture in mapping the supply chain configuration and network in the upstream sugar industry.

The results of the questionnaire design in the first stage were then used in the second stage. At this stage, in-depth interviews were conducted with selected participants using the Delphi method. The Delphi method is a systematic way to get joint decisions from a group consisting of experts from various disciplines [12]. These experts were never found with each other. The advantage of this method is that it provides a broader perspective, usually for long-term planning, including technological changes. The downside of the Delphi method is that it takes longer and does not result in one agreement but a range of answers. The weakness of the Delphi method is minimized through respondents verifying the risk events and agents identified from previous research. Furthermore, the respondents also added the factors that cause and risk events according to their expertise in their respective fields.

The result of the second stage was the mapping of business activity processes, risk events, and risk agents based on the Supply Chain Operation Reference. The Supply Chain Network configuration results from fieldwork and in-depth interviews. It was based on the configuration of previous research [10]. Once the network configuration and risk identification were mapped, the last step was to classify risk events and risk agents based on the criteria for sustainability business practices from previous research [2, 5, 7].

4. Result and Discussion

a. Upstream Sugarcane Supply Chain Network

The configuration of the upstream supply chain of the sugar industry shows in Figure 1. The description of the upstream supply chain configuration is the result of direct observation and brainstorming with sugarcane farmers.

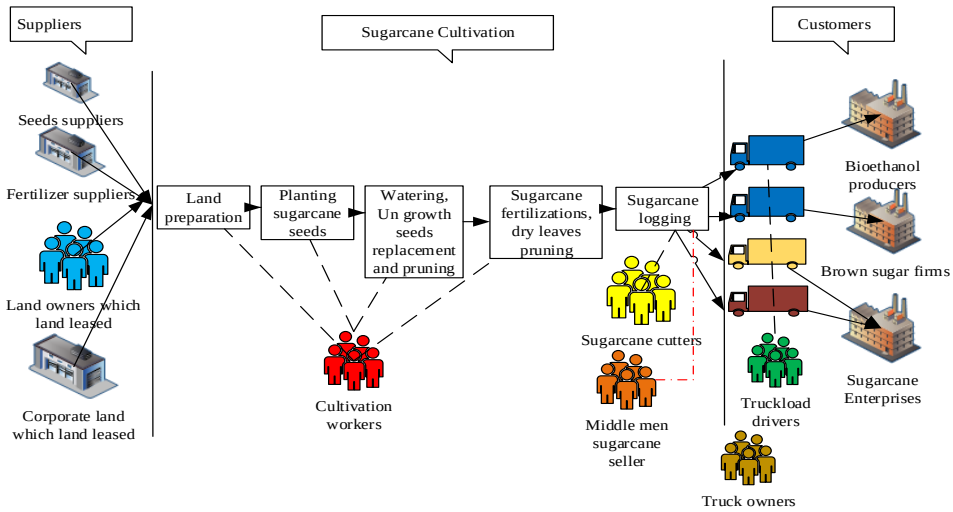


Fig. 1. Configuration of Upstream Sugarcane Supply Chain Network

Figure 1 illustrates the stakeholders involved in the upstream sugar industry. There are three classifications of supply networks in the upstream supply chain of the sugar industry, namely suppliers, sugarcane cultivation, and customers. Moreover, there are five supplier network entities, namely seeds, fertilizers, private landowners, which land leased, and corporate land which land leased. Furthermore, the first network consists of landowners, workers, supervisors, sugarcane loggers, and intermediaries of sugarcane sellers in the sugarcane cultivation process. The five stages of the cultivation process are land preparation, planting seeds, watering and pruning, fertilizing and watering, and finally, sugarcane logging. The last network entities consist of customers, truck owners, and truckload drivers who distribute sugarcane logs to sugar factories, brown sugar factories, and bioethanol factories.

4.2. Business Process Activity on Upstream Sugarcane Supply Chain Network

Mapping of business process activities refers to the Supply Chain Operation Reference (SCOR) approach. Five activities, namely plan, make, source, delivery, and return and business activities at each stage of SCOR, can be seen in table 1. The business activity was conducted in-depth interviews with sugarcane farmers, workers, and a supervisor.

Table 1. Business Process Activity

No	Activity	Business Process Activity
1.	Plan	– Forecast financial
		– Forecast land
		– Forecast seed and fertilizer
		– Forecast distribution
		– Forecast sugarcane periodic treatment
		– Forecast human resources
2.	Make	– Seed planting
		– Land treatment
		– Cultivation treatment
3.	Source	– Logging and harvesting
		– Worker's procurement
		– Logged sugarcane procurement
		– Seeds and fertilizer procurement
4.	Delivery	– Land rent procurement
		– Delivery harvested sugarcane
		– Logged loading to truck
5.	Return	– Return rejected harvested sugarcane
		– Return rejected seeds
		– Return rejected fertilizer
		– Return unsuitable land

Table 1 contains information about six forecast activities in plan activities, three business activity criteria in the make category, and four business activity processes in the source criteria. Three business activities in the delivery stage and four business activities in the return category result from mapping business activities on the upstream sugarcane supply network.

b. Risk Event and Risk Agent

The results of identifying potential risks and risks of make and plan business activities based on the Supply Chain Operation Reference in the upstream supply chain of the sugar industry are shown in Table 2.

Table 2. Risk event of Plan and Make Activity

Plan Activity	
1. Sudden change in the sugarcane area plan	6. The mismatch between the supply chain and financial planning
2. Mistakes in the plan to deliver the cut sugar cane	7. Unsuitable sugar production
3. Less land for sugarcane	8. Unable to fulfill all requests sugar
4. The significant error of forecasting	9. Significant increase in demand for sugar
5. Sudden change in plans production	10. Cultivation plan schedule delay

Make Activity	
1. Sugarcane yields lost during transportation and queuing	24. Receiving seeds that do not match the character of the land
2. Decrease in sugarcane yield from previous years	25. Insufficient seed requirement
3. Available seeds cannot be used	26. Mismatch of productivity and cultivation targets
4. Delay in logging implementation	27. Errors in fertilization and irrigation time
5. Inefficient cultivation process	28. Inaccurate raw materials and product prices
6. The arrangement process in the truck does not match the specifications and requirements of the sugar factory.	29. Sugarcane price fluctuation
7. Seasonal factor	30. The high rental price of sugarcane cultivation land
8. The quality of level sweetness does not match the existing standards	31. The high rental price for sugarcane transportation mode
9. Available seeds cannot be used	32. Working Accident during cultivation process
10. Delay in logging implementation	33. Working accident during logging and loading to truck
11. Inefficient cultivation process	34. Lack of personal protective equipment for workers
12. Harvest time delay	35. Lack of training for workers
13. Decrease in sugarcane quality during cultivation	36. Lack of safety training for workers
14. Planting time does not match the milling time	37. Worker's salary is not based on regional minimum wage
15. Sugarcane yields lost during transportation and queuing	38. No health and safety insurance for workers
16. Decrease in sugarcane yield from previous years	39. Irrigation difficulties
17. Low operating income from venture capital	40. Planting time does not match the milling time
18. Late payment of sugarcane by the factory	41. Lack of supervision during cultivation
19. Limited resources to harvest sugarcane on time	42. Sugarcane farmers do not understand the PH system
20. Limited availability of land for sugarcane cultivation	43. An error occurred in the amount of creation Delivery Order
21. Low skill and knowledge for better cultivation process	44. The flood of sugarcane cultivation land
22. High labour cost	45. The high rental price of sugarcane cultivation land
23. Fertilizer does not increase yield and productivity	

Table 2 shows the mapping of risks and potential risks event in the plan and makes business activities in the activity plan identify at least ten risks that disrupt the supply chain. Furthermore, in the making business activity stage, it is known that there are forty-five risks and potential risk events. The results of the risk mapping were obtained from stakeholders upstream of the sugar industry supply chain.

Moreover, table 3 shows the risks in the source, delivery, and return business activity processes. The mapping of risk and potential risk events based on business activities is obtained in Table 1.

Table 3. Risk Event of Source, Delivery, and return Activity

Source Activity	
1. Delay in receiving sugarcane seeds 2. Delayed loan disbursement by the bank 3. Delay in payment of seeds and fertilizers from farmers 4. sugarcane land conditions are not suitable 5. late receipt of sugarcane seeds 6. Weather factors (long drought and heavy rainfall rain) 7. Sugar land fires 8. Quality of seeds that are easily attacked by pests	9. The need for the number of seeds is not met requirements 10. lack of water availability 11. There is a sugar cane planting area that lack of slash and transport arrangements 12. An error occurred in the amount of creation Delivery Order 13. Late payments by sugar factories 14. Not evaluating supplier performance 15. Violation of supplier contract agreement.
Delivery Activity	
1. Lack of truck drivers on that day 2. Less and no transportation available 3. Traffic accident during delivery 4. Queue during weighing process more than 24 hours	5. Transportation disruption (trucks and lorries damaged or overturned) 6. The yield of sugarcane produced is low due to long queue during weighting process .
Return Activity	
1. Return of Defective Products to Supplier 2. Late arrival of the replacement product from suppliers 3. Returning sugarcane that is not sweet, clean and fresh to farmers 4. Returning seeds, fertilizers and land to supplier 5. Rejected Product Returns from sugar factory	6. Delay in the process of returning to the managing industry 7. Handling Returns from User industry 8. Delay in the return process from the user industry 9. Delay in delivery of risk products to industrial users.

Table 3 shows the mapping of risk and risk events in source, delivery, and return business activities. The mapping of risk events shows that there are fifteen risk factors for the source activity, six risk events for the delivery activity, and nine risk events for the return activity. The risk agent mapping is illustrated in Table 4. The risk agent shows the risk factors that have been identified in Table 2 and Table 3.

Table 4. Risk Agent

1. Low supervision of workers and plantations	12. Transport truck has an accident on the road
2. Manual cultivation and harvesting makes harvesting inefficient	13. Lack of supervisory supervision of workers
3. Lack of support and subsidies for production facilities to support farmers' welfare	14. Workers' negligence when cutting log and loading to truck
4. Floods, droughts, and other natural disasters	15. Product arrival from a supplier outside schedule
5. Weak bargaining position due to small land ownership	16. Price fluctuation
6. Limited availability of labor for planting and cutting which must be recruited from other areas	17. Lack of monitoring of worker attendance
7. Poor irrigation management	18. Damaged sugarcane log due to being on truck for too long
8. High rainfall affects the effectiveness of harvesting, which affects the amount of sugarcane transported	19. Drivers do not have a driver license
9. Low income from cultivation	20. Rare truck maintenance
10. The low performance of sugarcane harvesters during the harvest season	21. Truckload that exceeds the capacity
11. Bad weather	22. Workers commit a fraud
	23. No driver standby when needed
	24. Human error
	25. Improper fertilization technique and timing
	26. Forecasting that does not pay attention to external factors
	27. There is damage to the unit/transporting machine

It can be seen in Table 4 that twenty-seven causes of risk were identified that resulted in the occurrence and of risk events. The risk caused is based on events were obtained in the plan, make, source, delivery, and return business activities.

c. Discussion

Based on the configuration and mapping of risks and causes of risk on Indonesia's upstream sugarcane supply network, three points are obtained that could be further explored. The first is the configuration of the supply chain network. In the results of previous research, the upstream sugar industry supply chain has not comprehensively described how the maps and networks involved are in the upstream [11, 12]. he results of mapping the upstream sugar industry supply chain in this study describe three networks and eleven stakeholder entities involved along the upstream supply chain. The results of identifying potential risks and causes of risks in the upstream supply chain of the national sugar industry can provide a more in-depth picture of the risks and causes of risk based on the point of view of all stakeholders involved.

Furthermore, from the results of the identification of sustainability, it can be seen that economic and social factors are the most significant factors that cause the sustainability of the national sugarcane industry. Economic sustainability is considered the highest risk factor in upstream supply instability. The second rank is social sustainability risk, and the last is environmental sustainability risk.

5. Conclusion and Future Work

Upstream Sugarcane Supply Chain Networks consist of suppliers, cultivation processes, factory delivery. Moreover, upstream sugarcane stakeholders consist of farmers, cultivation workers, landowners, intermediaries to sell sugarcane, drivers, truck owners, sugarcane cutters. The farmers' association, supervisors, fertilizer distributors, seeds distributors have represented these stakeholders. In addition, Supply Chain Operation Reference Business activities in the sugarcane supply chain network consisting of six plan activities, and three make activities, four source activities, three delivery activities, and four return activities. Potential Risk Event identified Based on Supply Chain Operation Reference as follows. There are ten risks of plan activity, forty-five risks of make activity, seven risks of delay activity, nine risks of return activity. Sustainability risks on upstream sugarcane supply chain are related to the health and safety of workers, workers' rights, working conditions, the environmental impact of the cultivation process, and sugarcane delivery to factories.

Quantitative-based research to determine the resilience strategy of the upstream sugar industry supply chain is being carried out to continue the results of this research. This research is considered in the initial stage so that some weaknesses can be used as further research. The weakness in this study is that it does not involve consumer stakeholders, namely sugar factory entities, bioethanol factories, and brown sugar factories. The involvement of these stakeholders will further strengthen the risk description and configuration of supply chains and networks in the upstream sugarcane industry. The subsequent weakness is that the research results are only qualitative data in identifying risks. In the next stage, a quantitative approach is needed, for example, using the House of Risk and Multi-Criteria Decision Making Methods to determine mitigation strategies in controlling business risk [13]. Quantitative-based research to determine the resilience strategy of the upstream sugar industry supply chain is being carried out to continue the results of this research.

References

1. S. Pertanian, "Proyeksi Konsumsi Gula Indonesia," Jakarta, 2017. [Online]. Available: http://epublikasi.setjen.pertanian.go.id/epublikasi/outlook/2017/Outlook Tebu 2017/files/assets/basic-html/page60.html.
2. N. A. Mahbubah and A. Muid, "Sustainable supply chain management: Transferring from developed nations to developing countries," in *Proceedings - 2016 6th International Annual Engineering Seminar, InAES 2016*, 2017, pp. 239–244, doi: 10.1109/INAES.2016.7821941.
3. G. Windiastika, "Good Agriculture Practice Tanaman Tebu," *Dinas Pertanian Kabupaten Pasuruan*, Pasuruan, p. 1, Sep. 2019.
4. G. Behzadi, M. J. O'Sullivan, T. L. Olsen, F. Scrimgeour, and A. Zhang, "Robust and resilient strategies for managing supply disruptions in an agribusiness supply chain," *Int. J. Prod. Econ.*, vol. 191, pp. 207–220, 2017, doi: 10.1016/j.ijpe.2017.06.018.
5. D. G. Mogale, S. K. Kumar, and M. K. Tiwari, "Green food supply chain design considering risk and post-harvest losses: a case study," *Ann. Oper. Res.*, vol. 295, no. 1, pp. 257–284, 2020, doi: 10.1007/s10479-020-03664-y.
6. A. Braimoh *et al.*, "Increasing Agricultural Resilience through Better Risk Management in Zambia," 2018. doi: 10.1596/29779.
7. E. Doyle, Z. Friedman, and M. Johnson, "Sustainability Measurement of the American Sugarcane Industry," 2018.
8. K. Sartorius and J. Kirsten, "A framework to facilitate institutional arrangements for smallholder supply in developing countries: An agribusiness perspective," *Food Policy*, vol. 32, no. 5–6, pp. 640–655, 2007, doi: 10.1016/j.foodpol.2007.03.001.
9. Y. L. Everingham, R. C. Muchow, R. C. Stone, N. G. Inman-Bamber, A. Singels, and C. N. Bezuidenhout, "Enhanced risk management and decision-making capability across the sugarcane

industry value chain based on seasonal climate forecasts," *Agric. Syst.*, vol. 74, no. 3, pp. 459–477, 2002, doi: 10.1016/S0308-521X(02)00050-1.

10. B. L. MacCarthy and P. G. S. A. Jayarathne, "Supply network structures in the international clothing industry: Differences across retailer types," *Int. J. Oper. Prod. Manag.*, vol. 33, no. 7, pp. 858–886, 2013, doi: 10.1108/IJOPM-12-2011-0478.
11. M. Asrol, Marimin, Machfud, M. Yani, and E. Taira, "Risk management for improving supply chain performance of sugarcane agroindustry," *Ind. Eng. Manag. Syst.*, vol. 20, no. 1, pp. 9–26, 2021, doi: 10.7232/iems.2021.20.1.9.
12. E. Rizqiah and P. D. Karningsih, "Identifikasi Risiko Supply Chain dengan Mempertimbangkan Kepentingan Stakeholder Pada Indsutri Gula," *Prozima*, vol. 1, no. 2, pp. 71–81, 2017, doi: 10.21070/prozima.v1i2.1287.
13. Nafi'ah and N. A. Mahbubah, "Managing Risk on A Pharmacy Enterprise Supply Chain Using House of Risk Approach," *JKIE J. Knowl. Ind. Eng.*, vol. 8, no. 3, pp. 167–182, 2021, doi: https://doi.org/10.35891/jkie.v8i3.2752.

Challenges of Deterrence Theory in a Cyber-Environment

Nuran Mahmudov

Ministry of Defence of the Republic of Azerbaijan, National Defence University,
Military Scientific Research Institute, Baku, Azerbaijan
mahmudovnuran@outlook.com

Abstract. The purpose of this article is to discuss significant differences between nuclear, conventional and cyber deterrence. Because current deterrence theory is not valid for cyber deterrence theory. Traditional deterrence theory 70 years old, is not fit for cyberspace. Moreover, there have emerged significant issues and gaps in applying traditional deterrence theory to cyber. The concept of cyber deterrence is discussed on a global scale. The idea, which has been under discussion for almost two decades. However, the cyber environment is unlike the real environment. For this reason, the idea of cyber deterrence remains a controversial concept. The article provides a comprehensive analysis of deterrence theory and its application to cyber. It will the focus on determining the limitation/challenges of cyber deterrence.

Keywords. Cyber, deterrence, cyberspase, cyber attack, cyber environment.

1. Introduction

The concept of cyber deterrence will be discussed on a global scale. The idea, which has been under discussion for almost two decades, is often regarded as far from reality in connection with the specific nature of computer issues, in particular, the anonymity of the attackers or the use of surrogates (both people and servers). Nevertheless, many authors write about this concept. Inevitably, its comparison with the theory of nuclear and conventional deterrence, which prevailed in strategic thinking for almost seventy years. There is a significant difference between nuclear, conventional and cyber deterrence. In the digital age, in the case of cyber-attacks, days or weeks may pass, until it is generally known that it has occurred. There may be a misunderstanding of the purpose of cyber-attacks, whereas the designation of a nuclear strike is obvious. Nuclear weapons are currently being imposed by states alone, and cyber-attacks can be the result of actions by a lone hacker whose intentions and location can be carefully disguised. All of these and other important factors, the idea of or cyber deterrence remains, as before, a controversial concept. As a starting point, article has provided analysis of deterrence theory in cyberspace. The content, significance of documents were analysed, comments of contemporaries of these documents considered, as well as criticism from foreign authors. Moreover, the article elaborates cyber deterrence in International Relations. Furthermore, the work describes the challenges of deterrence theory in a cyber-environment. Moreover, the article also will explore and identify challenges of three essential elements of deterrence theory “capability, credibility, and communication" in cyber deterrence.

2. Deterrence Theory and Cyberspace

Information systems and infrastructures have become an integral part of everyday life as a result of the fact that the public and private institutions and organizations. Especially

communication, finance, energy, and security, operate on information systems, with the help of information systems and communication technologies developing every day. After the land, sea, air and space environments, as the 5th Operation Area (5th zone of action) the importance of the emerging cyber environment has increased even more. Because electronic and information systems provide opportunities for more efficient processing, transmission, protection and use of information in a short period of time. Especially in the economic, political and military powers of the states. In the cyber environment, malicious movements towards information systems and the onset and ongoing increases of attacks reveal the concepts of "Cyber Security or Cyber Defense." Moreover, the desire to harm or adversely affect the opponent's information systems reveals the concepts of "Cyber Attack." [1]. States have begun to develop strategies and policies in relation to the concepts of cyber defense and cyber-attack and to implement them effectively. Following these concepts, “Cyber War” has occurred. The information systems and infrastructures possessed in the cyber environment, and the ability to use them effectively are defined as “Cyber Power,” which is necessary for the initiation and maintenance of cyber warfare. Cyber Power means, in short, “dominance of cyberspace.” While, cyber is a key enabler of economic, military and political forces of power and their effectiveness. In particular, cyber power alone can be used effectively. Scientist Martin C. Libicki, a renowned scientist in cyber crime investigations of the United States, defines deterrence in general terms as "to create disincentives for starting or carrying out further hostile action," describing cyber deterrence as an act of aggression or punishment (retaliation threat) in the cyber environment.” It is alleged that the effect of the victim in the scope of deterrence of cyber can be accepted as the deterrence provided before diplomatic and economic sanctions after nuclear and conventional deterrence. According to Libicki this table has drawn as square and it shows that damages and result of these deterrents are same. As is known, damage and result of nuclear and conventional deterrence can immediately recognize by everyone. However, damage and result of cyber deterrence cannot be identified immediately. In addition, it will be clear to show their utilization as a triangle table form. However, cyber deterrence is more used than nuclear and conventional deterrence. The balance between distributive power and usage are vice versa for each other.

"Deterrent with Cyber Power," which has been using information and communication technologies in recent years, is the integration of the military and diplomatic influences in the cyber environment. In other words, using cyber-force, it is possible to refrain the states or non-state actors' aggression action or directed to a specific behavior [2]. The connection between cyber power and cyber warfare and results of cyber-attacks are researched by scholars for using cyber to deter state or non-state actors.

3. Cyber Deterrence

“The development of the internet has made our world a smaller space and opened the way for communication and interaction across borders.” Thanks to the Internet, the world is becoming a small village, and everywhere in the world is at the same distance for everyone. However, the lack of boundaries on the internet has also caused the cyber area to become a fertile breeding ground for attackers and cybercriminals. For this reason, countries have tried to establish new legislation in order to make a legal regulation of an efficient cyber deterrent system. However, this legal structure that is attempted to be created has to be out of the understanding of the traditional law. Because the activities in the cyberspace are not only specific to the states but have the ability to reach a wide range, and the new possibilities that cyberspace brings are different from many weapons that can contain elements of violence. For example, there are few nuclear states, “more than 140 nations have cyber weapons, and more than 30 countries have established cyber units within their military units” [3]. Cyber deterrence

is seen as a useful fact when viewed in the context of game theories that are effective in the discipline of International Relations. Cyber deterrence is seen as a valuable fact if we consider it in the context of game theories that are effective in the regulation of International Relations. However, in a speech in 2014, Eric Rosenbach, who served in the US Department of Defense, emphasized that those specializing in International Relations are more concerned with nuclear deterrence than with cyber deterrence. In his speech, Rosenbach also talked about the three stages of the study and research on deterrence, using the first of these as "legal deterrence," which was once again used to prevent criminals from once again committing the same criminal acts. The second, "nuclear deterrence," which restricts US and Soviet Union movements against each other in the Cold War era, and the latest, "cyber deterrence," which is used today. Rosenbach also stated that nuclear deterrence is easier than cyber deterrence, reminding that the use of the concept of attack in cyberspace is a different form of use from the nuclear side. Discussion of the official figures of the state shows that cyber deterrence will participate essential role for states strategy in the future.

Many states describe cyber space as the fifth battlefield." Therefore, the states have started to implement various strategies and implement them in order to be able to protect their strengths and to maximize their strength in deterrence. As an extension of competition in cyberspace, primary actors of the international arena are increasing their capacity in both cyber defense and cyber- attack, adding to the effects of cyber deterrence, and making enormous investments in their cyber military. In addition to this, the states have made many alliances to determine the sides and to provide cyber defense and intelligence. For example, former US President Barack Obama has increased the budget by 35% to $19 billion in 2016 to ensure security in this area [4]. Mike McConnell, the Director of National Intelligence of the United States, pointed out to the Washington Post on February 28, 2010, "How to win the Cyber War we are losing." He observed that, “the world has returned to the 1950s and that the methods used for the proliferation of nuclear weapons no longer have to be used to deal with cyber threats.” The same elements that dominate the debate over the link between deterrence and cyber warfare are dominant. Most of the studies on deterrence strategy and cyber warfare is based on the American point of view. The investigations examine the possibilities of successfully implementing the deterrence strategy to prevent cyber-attacks or analyze the US's use of cyber warfare to deter other threats confronting the United States. However, developed countries such as Israel, Japan and India should not be ignored. The fact that the United States is so far ahead is the reason why cyber is a vulnerable country against an attack. In comparison with countries such as Russia and China, military and civil infrastructure are dependent on cyberspace. Therefore, it is believed that the US is the most advanced state in both the defensive and offensive areas of cyber warfare.

4. General View of Challenges of Deterrence Theory in a Cyber-Environment

The deterrence of cyberspace is negatively related to deterrence strategies applied throughout history. The existence of these problems and rapid changes in cyberspace are questioning cyber deterrence as a whole. Today, as the use of cyberspace is becoming more widespread, it is a big debate that states cannot use cyberspace as part of its deterrence strategies. The results of this discussion do not give answers like "yes" or "no." Beyond that, it leads to different questions such as "against what actors/actors," "how," "over which control mechanisms." Deterrence is not an effective strategy for the cyber environment due to the difficulties of identifying the source of the attack and a large number of state and non-state actors in the cyber environment [5]. In his article titled "Applicability of Traditional

Deterrence Concepts and Theory to the Cyber Realm," Patrick Morgan pointed to the existence of items that are important in classical deterrence but are problematic at the point of providing cyber deterrence. These elements that Morgan stands on are elements of identity and motivation. It is obvious that the actor who was involved in the attack during the implementation of nuclear deterrence during the Cold War and how motivated it is with mobilization. While it is not possible to determine the identity of the attacker, it is hard to know what the actual reaction to being given to the attacker, whose motivation is uncertain [6]. In his work entitled "Cyber Warfare and Deterrence: Trends and Challenges in Research," Amir Lupovici questioned whether states could achieve deterrence in cyberspace. According to Lupovici, in order for deterrence to be successful, the capacity (defensive capacity) and the reality of the threat must be successfully transmitted, as well as the potential rival [7]. The "3 Cs" are essential references to best understand Deterrence theory. Capability, Credibility, and Communication. For any succeeded, deterrence must all these three elements [8].

5. Challenges of Capability

Capability refers to whether the actor, who produces the deterrent threat is capable of performing the threat. Capability is holding sufficient military force, to be able to accomplish probable military retaliatory threats. The percentage and character of force depends on the enemy and what benefits are being threatened. If these three primitive capacitors will examine the state of cyberspace, the important problem is that the information of state capacity is transferred to the enemy [7]. One of the most effective elements in deterrence is the performance of the attack capacity. Disclosure of the capacity to ensure the deterrence of the state is very important in the formation of policies that other states are applying [9]. For example, in the context of classical military power, states perform parades or regular exercises on national days to demonstrate their capacities. In cyberspace, it is highly doubtful how and to whom the right message about capacity will be given. There is a difference between the capacity to be shown against individuals and groups and the capacity to be shown at the intergovernmental level. Above all, cyberspace attacks are made using software vulnerabilities found in the system. The impulse of the developed capacitor also means that it is open to the vulnerabilities at the same time. When it is closed off, the developed capacity becomes meaningless despite all the investment and effort made [10]. Nye finds three similarities between cyberspace and land size after making similar comparisons with sea and air sizes in his work "Cyber Power." These similarities are "number of actors," "accessibility to the area" and "camouflage possibilities." Unlike the sea, air and space dimensions, large powers will not be able to exercise absolute dominance in the fifth dimension that cyberspace creates, similar to the land size. Because the state with the most network is the most vulnerable state for a cyber attack. For this reason, the weight of the great powers supporting their own dominance in other dimensions is becoming the greatest weakness in cyberspace [11]. For this reason, the great states, which is dominance in other dimensions is becoming the greatest weakness in cyberspace. The reason of this vulnerabilities is an open network. States cannot control all network system. The Internet gives more opportunity for a non-state actor to use these vulnerabilities.

Barriers to show and use capability: The ability to provide cyber field power should not be perceived only as a condition of states' capacities and skills. Because some obstacles come out of the nature of cyberspace and from those reasons, cyber deterrence getting difficult in cyberspace. Above all, the power to neutralize or abolish an enemy in cyberspace, or to use counter-force strategies effectively is limited. Therefore, the cyber field deterrence brings about difficulties due to the problems of attributing an attack source. Three factors seem to be

a barrier to a state's increasing cyber-field power or deterring aggression. These are asymmetry, attribution, and power. It will be more descriptive and understandable to assess these obstacles and challenges that arise from the attack of Russia on Estonia in 2007, which is an example of a cyber-force demonstration [12].

Attribution problem: Anonymity is a significant obstacle for the defender. The unknownness of the elements of identity and motivation underlined by Patrick Morgan is also an obstacle to the correctness of the message to be given [13]. In other words, it is very difficult to cite someone to find out who attacks cyberspace or to determine the intention of the attacker [14]. However, the nature of the attack and the reaction to it are at an important stage, so even nation states with intelligence capacity and resources cannot be directly authorized to take an immediate attack on them [15]. When looking at the example of Estonia, it is obvious that the attack is still carried out by Russia. Moreover, the anonymity factor plays an essential role here.

Asymmetry: It can be described as the struggle of the weaker side against the stronger side with different tactics or random methods, benefiting the weaker side. When we look at the context of our example; according to experts, even if Russia has carried out a cyber-attack against Estonia, this area may leave no space for Estonia to challenge. In this case, even if Estonia wanted to carry out an attack on Russia, Russia might have abandoned the environment in which it could carry out this attack for Estonia. This reveals the asymmetry of the cyber field [16].

Power: The Estonian attacks show how super-competent actors were on the internet. Although Estonia was confident that this attack was carried out by Russia, it was only an Estonian conviction of Russian origin. This shows how only one person (the individual) can have power over the cyber field by carrying out an attack on the state capacity [17]. This creates major problems both regarding the shaking of the capability of the country and in the deterrence of the opponents. Because it is so difficult to deter a state or compete with a state, it seems impossible to deal with "individuals" [9].

Having the US and European countries behind Estonia succeeded in achieving the least damage from the Russian attack. However, cyber-attacks are not limited to small countries that can only spread to restricted areas such as Estonia, or they do not give temporary damage in this way. On the contrary, the great powers are more vulnerable to cyber-attacks. Therefore, the attacks of the great powers may result in much more damage. Asia and Europe seem to be highly vulnerable to damage that can occur in their networks, while they are reliant on information networks. Therefore, the increase in the tools used for cyber-attack increases the technical capacity of cyberspace for malicious actors, which causes the cyber-area vulnerabilities to increase.

6. Challenges of Credibility

Credibility is perhaps the most complicated element in deterrence theory. The deterrent should be obliged to use force without any doubt, but it is even more important that the attacker should not doubt that the dangerous threats would be implemented [7].

The dominant element that distinguishes the concept of cyber deterrence from deterrence used in a classical/traditional sense is the identification of the source of the attack. The anonymity that cyberspace brings to the nature of the virtual world is an obstacle to the right source of retaliation. If a reprisal is to be made, first of all, it is necessary to know who was attacked. If deterrence is to be commissioned before it comes to retaliation in the first instance, others should be aware that the deterrent state will know precisely who is attacking

it. Otherwise, the state will be attacked by the wrong state or person, and this will cause new enemies to be born [18].

According to Morgan, it is important in classical deterrence, and there are also problems in trying to provide cyber deterrence [19]. One of these elements is the period in which the message is transmitted to the attacker. It is very clear who the opponent is and what motivation it is when nuclear deterrence is achieved in the Cold War. However, in cyberspace, it may not be possible to identify the attacker. In this case, the problem of transmitting the message to the unauthorized attacker arises [20]. Another important issue as much as the identity of the attacker is the motivation of the attacker. It is difficult to know what the right message should be given to an attacker whose motivation is unknown. Another important challenge Morgan has underlined is credibility. The threat to be successful in deterrence must be credible. The threat must be convincing, as far as the rival state is concerned, that there will be a threat [21]. In cyberspace, it is not possible for the actor who attacks to be identifiable, even if the identification and motivation of the actor are determined. Above all, the harmfulness of the cyber-criminal against the determined assailant remains unknown. The belief that cyber response will not harm the attacker can be caused by a bigger attack.

Cyber deterrence is still in its infancy, as nuclear deterrence was in 1950. Sometimes internal threats are intertwined with international threats, as they happen in "multi-polar" environments, and as the effects of cyber weapons can be instantaneous, they can create a much more insecure environment than other types of deterrence [22]. For example, the operation of Russia against Georgia in 2008 revealed this problem of cyber deterrence. These are problems of “Scalability and Temporality.” This, in turn, makes it highly complicated to foretell the attack indicators and warnings of the attack scale. During the attacks on our example of Georgia, hackers had collapsed government websites. However, this did not cause a long-term deterioration by causing slight damage. They have left their secret and time-sensitive viruses on government systems and have created havoc in Georgian networks after they have been through the intervention. Another problem, temporality, is the instant nature of cyber-attacks. Defenders give early warnings against attack. However, it is unclear when, what, and for what purpose these digital signals are. “The physical signals provide most or all of this information” [9].

Furthermore, the nature and motivation of such attacks cannot be predicted. Every operation can have unexpected results. This leads to the existence of many actors who are not deterred by cyberspace. Threat perceptions, play a dominant position both in the evolution of an effective national strategy and for the application of deterrence theory. It is important to understand the strategic responses of a country or an organization in the form of threat assessment. For this reason, the nature of the threat, the technical tools used and the potential target are crucial to the effectiveness of the attack against a cyber-attack.

Another important problem that needs to be addressed is the "realization of defense," one of the two elements of deterrence's functionality. Two types of measures, "attack" and "defense," are used interdependently in deterrence [23]. The fact that the attack and defense are single-centered also increases the deterrence and enables the state to use its resources effectively. However, the cyberspace defense is carried out locally by institutions, which are regularly attacked in the area where defense is needed rather than being centralized. This is a weakness in the defense as well as a holistic approach. Another point of this is after the nuclear bombardment of Nagasaki on August 9, 1945, no force in the world used the nuclear bomb in the war. Despite the limited data obtained, strategists have wasted their resources on hypothetical scenarios in the absence of adequate experimental experience. In the cyberspace, the attackers had more knowledge on the result of the data coming from the major attacks in the last decade. The data obtained at this point for cyber demolition distinguishes the field

from nuclear experience. Despite this separation and all the evidence at hand, there has not yet been a Cyber War. The DDOS attack in Estonia and Georgia or industrial sabotage with STUXNET to Iran gives only clues as to what can happen in the future. In spite of all these clues, it is not known how damages that cannot be counted in a cyber-war, or how the distinction of moderation and criminality that Lupovici emphasizes on the frame of war law is unknown [24].

7. Challenges of Communication

The final C is communication. The state must report its own threats. This action is not prohibited if the state does not know what actions are prohibited. Therefore, it is important for states to know what measures they should take, and know what happens when they act. Thus, good communication lines are very important [7]. For example, during the Cold War, this "hotline" between the Soviet Union and the United States was well implemented. One of the goals of deterrence is to avoid armed conflict, and relationships are crucial to building boundaries and revealing expectations.

The last principle is that the process of communicating the threat of the actor to challenge deterrence in cyberspace is quite difficult if we examine the transmission of the threat in a successful manner. At this point, two areas need to be clarified:

1. "Identification of the attacker,"
2. "Determination of the capacity of the aggressor" [25].

Detection of an attacker in a conventional or nuclear attack is possible within the physical pursuit or intelligence possibilities. However, the detection of the attack in the cyber environment has complicated. Due to the nature of the cyberspace is not like land environment. At this point, the difficulty of detecting the attacker is not only because the attacker is anonymous in cyberspace. It is also influential that the actor can be any one of a wide range of individuals, not only states but also non-state actors. In this direction, the states are trying to increase the control structures in the cyberspace. The difficulty of identifying the attacker, and even in some cases it is impossible, making it difficult to determine whether the attack is cybercrime, the use of cyberspace for terrorist purposes, or whether it is a cyber-attack. This detection problem of the attack is also difficult to determine the measurement of the countermeasure. For example, if the attack comes from an individual rather than from the state level, and if it is an economic purpose cybercrime, the punishment of the criminal should be proportional in the legal system. However, the hypothesis of the same attack coming from another state reveals the possibility of a conflict.

The identification of the person, who handles, as well as the committed crime, becomes important when providing the measure that must be provided. Although the anonymity provided by cyberspace after the attack of the cyber will be inflated to some extent using intelligence facilities, it will not be possible to detect it exactly. The implementation of the rivals' identity strategy has key prescriptions for determining the tactics in the process. In the cyber-attack on Estonia in 2007 and the STUXNET attack in Iran in 2011 was the problem of detecting the attacker [26]. However, in cyberspace, not only nation-states but also individuals and groups gathered around a certain purpose have the capacity to attack. In such a case, even if the source of the attack is narrowed down to a certain state country, this determination is not enough to answer the question whether it comes from the state or from the different actors within that state. This difficulty also makes it difficult to predict the aggressor's capabilities. Deterrence is, first of all, based on the willingness of the attacker to know the capacity of the defender and to avoid the threat of retaliation, which makes the situation ambiguous.

Another difficulty in communicating the message is that unlike the Cold War, there are new problems brought about by the increase in the number of actors. During the Cold War, a limited number of actors and limited post-attack detectable capacity have made the practice of these actors a deterrent to each other. The proliferation of actors who could threaten the existence of the state with the possibilities of cyberspace, the difficulties in determining the identity and capacity of these actors led to the nullification of the practices practiced in the Cold War and before. This rise in the number and form of threats in cyberspace also reveals the question of how to convey the message to be transmitted, as well as to whom the message should be transmitted in order to provide deterrence [7].

Thus, it seems very difficult to solve the problems described above in the short run. Although it is inevitable for cyberspace to be able to provide solutions for the changes that will be made on this dimension created by human-based space, it will also lead to the emergence of new problems for the same reasons. States have to provide deterrence. For this reason, studies and living developments in this area will lead to much more divergent views. Apparently, the word "deterrence" is a different context to our traditional understanding of deterrence theory. The main problem of applying 3C principles in cyberspace is the difficulty of matching the deterrence theory in this direction. However, development of cyberspace and new strategies can give new opportunity to understand and use deterrence theory in the cyber environment.

8. Conclusion

The end of the Cold War does not mean the end of the need for deterrence. Continuous and new attacks and threats will emerge. It is an important strategy for self-defense. Reality and rationality are influencing international actors as a result of their reflections. Unique and innovative tools must be developed to achieve national and international goals. Different deterrence methods should be developed such as cyber deterrence.

Furthermore, cyber deterrence can be carried out by any person anywhere in the world without the need for a base, headquarters or a soldier to hide invisibility and hacking outside of such deterrence and reprisals when approaching the issue. With the emergence of cyber deterrence, states have been struggling to gain advantage for centuries in physical areas (land, sea, and air) have now begun to carry the agreement and attack areas to a far different area, cyberspace. As a result, "the future growth and development of human civilization are linked to the growth and development of cyberspace." Moreover, it can be said that more public should be aware of this fact.

As mentioned above, in general, it is thought that cyber deterrence cannot be provided effectively. There are no problems with cyber attackers, and they are listed as untouchable and encourage more. However, an effective and reliable deterrent strategy developed by strategists can be successful in deterring the attackers at a certain point in their attacks. In addition, declaring the developed strategy in such a way that the attackers can be aware of the attackers in the case of actual warfare.

As a result, although new areas of work are being undertaken in each area, it is important to make new work in this area. Dealing with the different aspects of this subject, to write theses on this subject, to develop deterrent products and technologies, to make necessary investments in this subject. Moreover, most important to create and develop in cyber environment deterrent forces are essential. It is evaluated that it will make a significant contribution to the protection and safety of cyber.

References

1. Andress and Winterfeld, Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners, 69, 256.
2. Martin C. Libicki, Cyber deterrence and Cyberwar (The USA: RAND Corporation, 2009), 28.
3. Iasiello, Emilio, "Is Cyber Deterrence an Illusory Course of Action?" Journal of Strategic Security Vol 7, No. 1 (2013): 54.
4. The Third Way, "The President's 2017 Defense Budget," thethirdway.org, https://www.thirdway.org/report/the- presidents-2017-defense-budget (Accessed 10.03.2022).
5. Richard L. Kugler, "Deterrence of Cyber Attacks" in Cyber power and National Security. Franklin D. Kramer, Stuart H. Starr, and Larry K. Wentz, (The USA: National Defense University Press and Potomac Books, 2009), 309.
6. Patrick M. Morgan, "Applicability of Traditional Deterrence Concepts and Theory to the Cyber Realm" in National Research Council, Proceedings of a Workshop on Deterring Cyber-attacks: Informing Strategies and Developing Options for U.S. Policy (The USA: The National Academies Press, 2010), 70.
7. Amir Lupovici, "Cyber Warfare and Deterrence: Trends and Challenges in Research" 56.
8. Kenneth Geers, "The challenge of cyber-attack deterrence" Computer Law & Security Review, Vol. 26, No 3, (2010): 299
9. Will Goodman, "Cyber Deterrence Tougher in Theory than in Practice?" Strategic Studies Quarterly, vol. 4, no. 3, 116.
10. Eric Talbot Jensen, "Cyber Deterrence, Emory Int'l L. Rev. Vol 26, No 1 (2012): 781.
11. Joseph Samuel Nye, Cyber Power (The USA: Harvard College, 2010), 4. 92 Joseph Samuel Nye, Cyber Power (The USA: Harvard College, 2010), 3.
12. Guzman, G. (2017). Cyber power - the Great Equalizer: Estonian Cyber power Development. https://www.iapss.org/shop/budapest/uploads/2512_guzman_g_cyberpower_the_great_equalizer_03_2017_politikon_ela .pdf (Accessed 25.02.2022).
13. Patrick M. Morgan, "Applicability of Traditional Deterrence Concepts and Theory to the Cyber Realm," 70.
14. David E. Graham, "Cyber Threats and the Law of War" Journal of National Security Law and Policy Vol. 4, No. 87 (2010): 92.
15. M.N. Schmitt, Cyber Operations and the Jus Ad Bellum Revisited. Villanova Law Vol. 56, no. 3 [2011], 577. 569-605.
16. David Betz. (2012). "Cyber power in Strategic Affairs: Neither Unthinkable nor Blessed." The Journal of Strategic Studies. Vol 35, No 5 (2012): 695.
17. Matthew Reinbold, "Super Empowerment, Networked Tribes and the End to Business as We Know It." youtube.com, https://www.youtube.com/watch?v=LmieSvccWfg (Assessed 27.02.2022)
18. Brandon Valeriano, Ryan C. Maness, Cyber War versus Cyber Realities: Cyber Conflict in the International System. (The UK: Oxford University Press, 2015), 60-61.
19. Tim Stevens, "A Cyberwar of Ideas? Deterrence and Norms in Cyberspace," Contemporary Security Policy, Vol 33, No 1, (2012): 152.
20. Eric Talbot Jensen, "Cyber Deterrence, Emory Int'l L. Rev. Vol 26, No 1 (2012): 780.
21. John Stone, "Conventional Deterrence and the Challenge of Credibility," Contemporary Security Policy, Vol 33, No 1, (2012): 119.
22. George Perkovich, and Ariel E. Levite, Understanding Cyber Conflict: Fourteen Analogies (The US: Georgetown University Press, 2017), 56.
23. Kim Holmes, "Basing Deterrence on Strategic Defense," Heritage.org, https://www.heritage.org/defense/report/basing-deterrence-strategic-defense (Assessed 27.02.2022).
24. Andress and Winterfeld, Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners.
25. Jan-Frederik Kremer, and Benedikt Müller, Cyberspace and International Relations: Theory, Prospects and Challenges (Berlin: Springer, 2014), 166.
26. Karsten Friis, and Jens Ringsmose, Conflict in Cyber Space: Theoretical, strategic and legal perspectives (The UK: Routledge, 2016), 180.

Model and Method of Integral Risk of Emergencies Management

Svitlana Bordiuzhenko, Olena Liashevska, Oleh Stelmakh, Oleksandr Danilin, Serhiy Shevchenko

National University of Civil Defence of Ukraine, Kharkiv, Ukraine
kravtsiv1992@gmail.com

Abstract. The work is devoted to the development of a mathematical model and method for managing the risks of emergencies (dangerous events). The peculiarities of the developed mathematical model are investigated in the work. The objective function of the problem is linear, the constraints are linear, nonlinear and discrete. The objective function is the dependence of the integral risk on the main factors that characterize the response process of rescue units. The distribution of rescue units is carried out taking into account: the normalized response time to emergencies (dangerous events); belonging of high-risk objects and potentially dangerous objects to the areas of departure of rescue units (taking into account other integral risks of emergencies or dangerous events); placement of operational and rescue units in the united territorial communities (certain settlements). The paper also develops structural and logical schemes for the implementation of a mathematical model of distribution of rescue units according to the level of integral risk of emergencies (dangerous events) and taking into account its features, namely: existing areas of departure of rescue units and limited resources to create these units. These structural and logical schemes allow for various types of distribution of rescue units, taking into account the integral risk of emergencies or dangerous events).

Keywords: Emergency, Dangerous event, Integral risk, Mathematical model, Method management, Rescue unit.

1. Introduction

Emergencies in any country are always a danger to humans. The negative consequences of emergencies are both financial losses and human injuries and casualties. According to Article 3 of the Constitution of Ukraine, a person, his life, health and safety are recognized in Ukraine as the highest social value. One of the links in ensuring human security in the country belongs to the State Emergency Service of Ukraine. The purpose of the service is not only to protect people from emergencies (dangerous events), but also to prevent such events. Dealing with the consequences of emergencies is a resource-intensive business, which is why it is appropriate today to reduce the integral risk of emergencies (dangerous events)

2. Recent Research and Publications Analysis

In report [1] of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction determinate that the disaster risk is The potential loss of life, injury, or destroyed or damaged assets which could occur to a system, society or a community in a specific period of time, determined probabilistically as a function of hazard, exposure, vulnerability and capacity.

In work [2] is indicated that disaster management evolves into disaster risk management, there is greater concern and urgency for understanding the "how", "who" and "with what" of disaster risk reduction and resilience. This work is already aimed at combating the consequences of emergency situations. The manual states, there is no blueprint or pathway to follow. But there are many good lessons that guide what a strategy should look like, who should be involved, what mechanisms might be used and how it is possible to advance in implementation.

In Ukraine, such scientists as R. Kravchenko [3–5], V. Fedorchak [6–8], A. Bilousov [9] and others deal with the risk of emergency. However, these works do not contain models or methods for managing the integral risks of emergencies (dangerous events).

Paper objective: In this paper, it is necessary to develop the model and method risk of emergencies management.

3. Presentation of the Main Research Material

3.1. General model of integral risk of emergencies (dangerous events) management

In [10, 11] as a result of the correlation-regression analysis it was found that the specified integral risk of an emergency (dangerous event) depends on such factors as N_{events} – the number of emergencies (dangerous events) recorded in the region; $M_{victims}$ – the number of deaths due to emergencies (dangerous events) in the region; τ_{arrive} – time of departure of rescue units to the place of occurrence of an emergency situation (dangerous event); τ_{loc} – time of localization of emergencies (dangerous events); τ_{liq} – time of liquidation of emergencies (dangerous events). The main parameter that rescue units should focus on is the standard time of arrival of units to the scene of an emergency. The normative time of arrival of rescue units to the place of call [12–14] should not exceed: on the territory of cities – 10 minutes; in settlements outside the city – 20 minutes.

In [15] it was proved that the time of rescue units to the place of emergency (dangerous event) depends on the coverage of the area by the service areas of rescue units. This coefficient is calculated using the following expression

$$k_{\text{cover}} = \frac{S\left(\bigcup_{q=1}^{N_q} P_q\right)}{S(S_0)}, \tag{1}$$

where N_q – the number of existing RUs; P_q – area of departure q-th unit ; S_0 – given territory; $S(.)$ – area calculation function.

Thus, the following problem arises. Let a certain administrative-territorial unit S_0 be given in the form of a polygon in the global coordinate system. The region S_0 has discrete elements V_k, $k = 1,\text{K}, N_k$, which are settlements. Let $G_l \subset V_k$, $l = 1,\text{K}, L$, $L < N_k$, – united territorial communities (settlements), in which it is permissible to create rescue units in accordance with settlements in which (next to which) there are potentially dangerous objects and / or objects of the increased danger, we will designate through $S_d \subset G_l$, $d = 1,\text{K}, D$, $D < L$ (the specified objects can become factors of realization of technogenic risks).

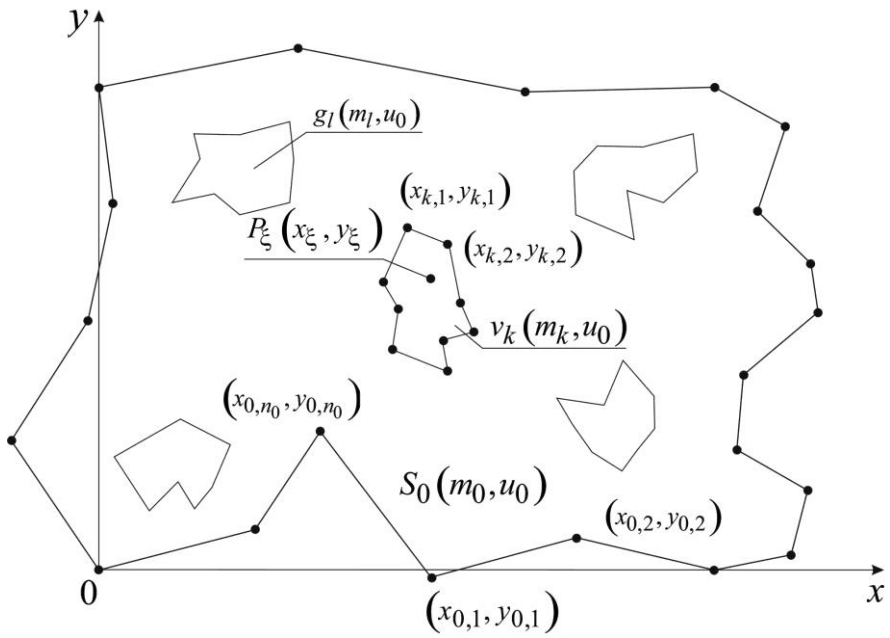


Fig. 1. The region $S_0(m_0,u_0)$ with distrete elements $v_k(m_k,u_0)$, $k=1,\text{K},N_k$, and $g_l(m_l,u_0)$, $l=1,\text{K},N_l$.

It is necessary to cover the area S_0 with the exit areas of rescue units P_i, $i=1,\text{K},N$ with the exit areas of rescue units

– minimum area of crossing of exit areas of rescue units;

– belonging of departure areas of rescue units of area S_0;

– affiliation of settlements V_k, $k=1,\text{K},N_k$, as well as settlements S_d, $d=1,\text{K},D$ areas of departure of rescue units (taking into account other integral risks of emergencies and hazards of man-made nature);

– the arrival time of rescue units to the most remote point of the departure area P_i, $i=1,\text{K},N$ should not exceed the specified T^* (in cities – 10 minutes; in settlements outside the city – 20 minutes);

– placement of rescue units is carried out in the united territorial communities (settlements) G_l, $l=1,\text{K},L$;

– minimum number of rescue units P_i, $i=1,\text{K},N$.

The mathematical model of distribution of rescue units according to the level of risk of emergencies (dangerous events) of man-caused nature is as follows

$$\min_{u\in W} R\left(\bar{\tau}_{arrive},\bar{\tau}_{loc},\bar{\tau}_{liq},N_{events}\right);\ u=\{m_i;v_i\};\ i=1,\text{K},N, \tag{2}$$

where W :

$$\omega\left(m_i, m_j, v_i, v_j\right) \to \min \tag{3}$$
$$i = 1,\text{K}\ , N;\ \ j = i+1,\text{K}\ , N;$$

$$\omega\left(m_i, m_{cS_0}, v_i, v_{cS_0}\right) \to \min, \tag{4}$$
$$i = 1,\text{K}\ , N;\ \ S_0 \bigcup cS_0 = R^2;$$

$$V_k \in \{P_i\};\ k \in \{1,\text{K}\ , N_k\};\ i = 1,\text{K}\ , N; \tag{5}$$

$$S_d \in \{P_i\};\ d = 1,\text{K}\ , D;\ i \in \{1,\text{K}\ , N\}; \tag{6}$$

$$\bar{\tau}_{arrive}\left(P_i\right) \le T^*;\ i = 1,\text{K}\ , N; \tag{7}$$

$$\bar{\tau}_{arrive} = f\left(k_{cover}\right); \tag{8}$$

$$u = \{m_i; v_i\} \in \{G_l\};\ G_l \in \{P_i\};\ i = 1,\text{K}\ , N;\ l = 1,\text{K}\ , L; \tag{9}$$

$$N \to \min. \tag{10}$$

In the mathematical model (2)÷(10) expression (2) is the objective function of the problem, with m_i – the coordinates of the vertices of polygons P_i, $i = 1,\text{K}\ , N$ in the local coordinate system, v_i – the parameters of the location of objects P_i (the position of the local coordinate system of the i -th object in the global coordinate system); expression (3) – is a condition of the minimum of mutual intersection of objects P_i and P_j, where is a $\omega(.)$ – ω-function, which is the area of intersection of polygons P_i and P_j; expression (4) – the condition of the minimum intersection of objects P_i with the addition of the region S_0 to the Euclidean space R^2; expression (5) – condition of belonging of settlements V_k, $k = 1,\text{K}\ , N_k$ areas of departure of rescue units P_i; expression (6) – condition of belonging of settlements S_d, $d = 1,\text{K}\ , D$ areas of departure of rescue units P_i; expression (7) – a condition regarding the arrive time of rescue units to the place of call; expression (8) – the relationship between the average arrive time of rescue units and the coverage ratio of the region S_0; expression (9) – the condition of placement of rescue units in the united territorial communities (settlements) G_l, $l = 1,\text{K}\ , L$; expression (10) – is a condition of the minimum number of rescue units.

Consider the features of the mathematical model (2) ÷ (10):

1. The objective function of the problem is linear, the constraints are linear, nonlinear and discrete.

2. In the case of taking into account the locations of existing rescue units, the following restriction must be added to the mathematical model.

$$\omega\left(m_i, m_q, v_i, v_q\right) \to \min; \tag{12}$$
$$i = 1,\text{K}\ , N;\ \ q = 1,\text{K}\ , N_q,$$

where N_q – the number of existing rescue units.

3. If the problem of minimizing the consequences of emergencies by dividing the rescue units by the level of integral risk of emergencies (dangerous events) is solved taking into account the limited resources, then in the mathematical model instead of expression (10) it is necessary to use the following restriction

$$Q_{res}(N) \le Q_{res}^{*}, \tag{12}$$

where $Q_{res}(N)$ – resources needed to create N rescue units;

Q_{res}^{*} – resources allocated for the creation of rescue units.

4. The type of rescue units is determined depending on the objects of protection located in the service area.

3.2. Method of integral risk of emergencies (dangerous events) management

Since the problem of emergency risk management (dangerous events) was reduced to the problem of optimal coverage of a given area by areas of rescue units (the problem of optimal coverage of convex polygons of a given area with discrete elements), a method related to the class was developed to solve it. combinatorial optimization methods. The essence of this method is that a complete search of discrete elements G_l, $l = 1,\text{K}, L$, taking into account the constraints (3)÷(9), also takes into account the condition of maximizing the population in the service areas of rescue units. The solution of the problem is completed provided that the whole set is covered G_l, $l = 1,\text{K}, L$. The solution that provides the minimum of the objective function (2) is chosen.

If the problem is solved taking into account the restriction (11), it determines the number of rescue units, the principles of which are given above.

It should be noted that the method of optimal coverage of convex polygons of a given area with discrete elements consists of the following methods

Method 1. The restriction is taken into account that rescue units must be created in the settlements in which (near which) the objects are located S_d, $d = 1,\text{K}, D$.

In this case, the service areas of rescue units will have the form shown in Fig. 2, and the boundaries of these areas (dashed lines) are determined based on the restriction (7).

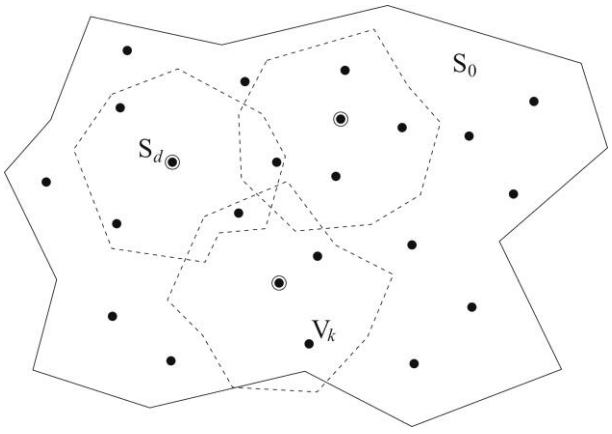


Fig. 2. Areas of service of rescue units which are located in settlements with objects S_d, $d = 1, \text{K}, D$

To determine the location of the next rescue unit (Fig. 3, dashed line), you must perform the following steps

– to carry out search of admissible locations of rescue units V_k, $k = 1,\text{K}, N_k$ (places in which rescue units are already located, are removed from the corresponding set);

– to form a set of locations of rescue units that provide coverage of settlements that are not part of the service areas of existing rescue units, with the largest total population;

– when determining the location of the rescue unit, priority is given to the case when the service area will belong to the object S_d, $d = 1,\text{K}, D$.

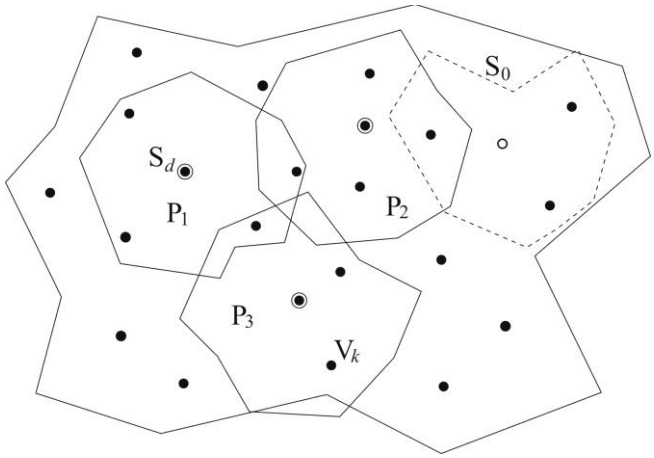


Fig. 3. Areas of service of rescue units which are located in settlements with objects S_d, $d = 1,\text{K}, D$.

An example of coating a given area using method 1 is shown in Fig. 4. The solution of the problem is completed under the condition that the restrictions (3) ÷ (9) are satisfied.

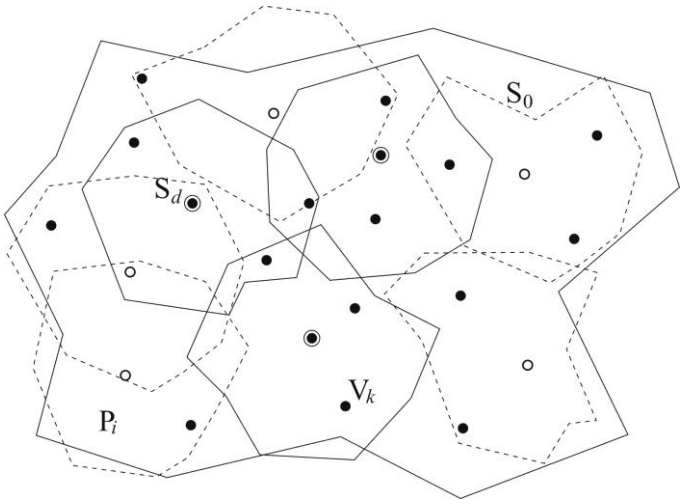


Fig. 4. Coating a given area using method 1.

Method 2. This method is that one rescue unit must service the maximum number of objects S_d, $d = 1,\text{K},D$.

Then, first of all, it is necessary to build an area $\bigcap_{d=1}^{D} P_d^{'}$ (Fig. 5), where objects $P_d^{'}$ are defined by restriction (7).

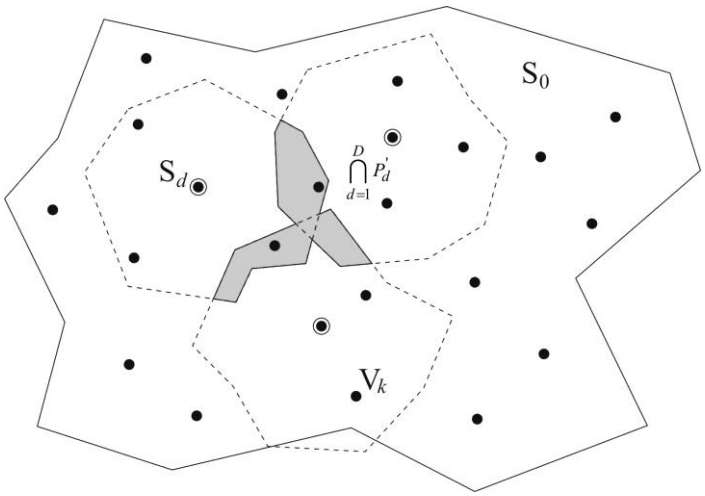


Fig. 5. Definition of the area $\bigcap_{d=1}^{D} P_d^{'}$.

The next step is to search for objects V_k, $k = 1,\text{K},N_k$ (Fig. 6), that belong to the area $\bigcap_{d=1}^{D} P_d^{'}$. Priority for the deployment of the rescue unit is given to the case when

– the largest number of objects is covered S_d, $d = 1,\text{K},D$;

– the settlements with the largest total population are covered.

The next location of the rescue unit is selected based on expression (6) and actions corresponding to method 1.

After fulfilling restriction (6), the principles of deployment of rescue units coincide with method 1.

It should be noted, if $\bigcap_{d=1}^{D} P_d^{'} = \varnothing$, then method 2 will be reduced to method 1.

In Fig. 7 shows an example of coverage of a given area by service areas of rescue units using method 2.

Thus, the problem of optimal coverage by nonconvex polygons of a given region with discrete elements belongs to the class of combinatorial optimization problems.

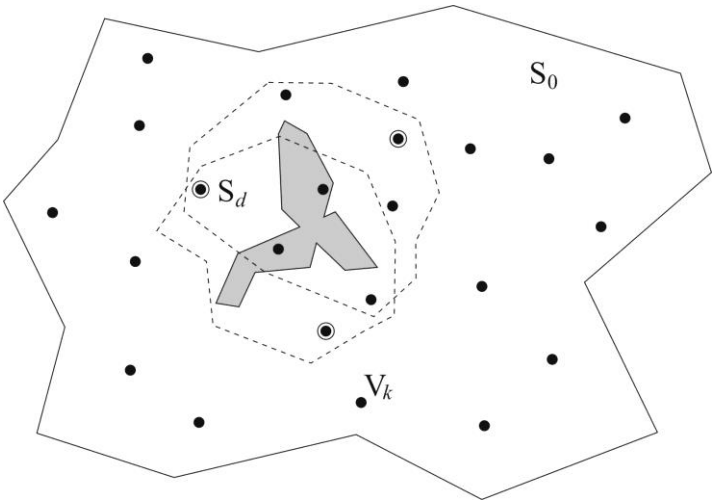


Fig. 6. Select a valid rescue unit location.

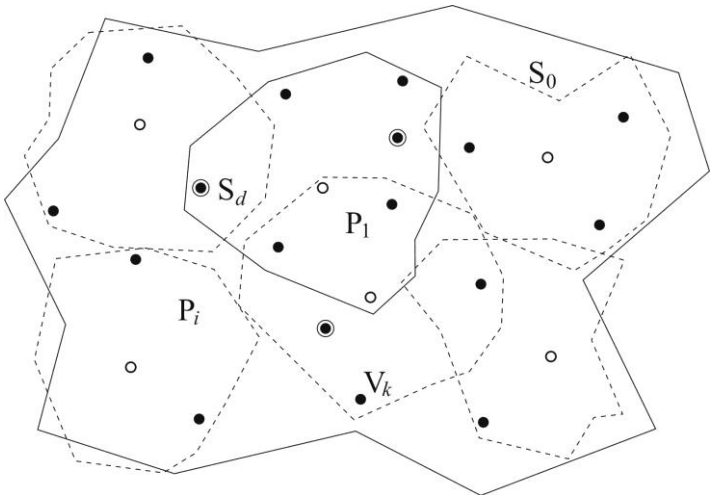


Fig. 7. Coating a given area using method 2.

Taking into account the locations of the existing rescue units P_q, $q = 1,..., N_q$, from the set of points V_k, $k = 1,\text{K}, N_k$, are removed those in which these units are already located. Then one of the above methods is used to solve the problem of optimal coverage.

If problem (2)÷(9) or (2)÷(10) is solved taking into account restriction (11), then the condition $V_k \in \{P_i\}$; $i = 1,\text{K}, N$; $k = 1,\text{K}, N_k$, is removed from the model, while determining the number of security centers N_{res}, that correspond to available resources Q^*_{res}. Next, the coverage of a given area using the developed method.

4. Conclusions

The paper develops a mathematical model for risk management of emergencies (dangerous events) and explores its features. The objective function of the problem is linear, the constraints are linear, nonlinear and discrete. The objective function is the dependence of the integral risk on the main factors that characterize the response process of rescue units. The distribution of rescue units is carried out taking into account: the normalized response time to emergencies (dangerous events); belonging of high-risk objects and potentially dangerous objects to the areas of departure of rescue units (taking into account other integral risks of emergencies or dangerous events); placement of rescue units in the united territorial communities (certain settlements). Structural and logical schemes for the implementation of the mathematical model of distribution of rescue units according to the level of integral risk of emergencies (dangerous events) and taking into account its features, namely: existing areas of departure of rescue units and limited resources to create these units. These structural and logical schemes allow for various types of distribution of rescue units, taking into account the integral risk of emergencies or dangerous events). Further research will focus on the implementation of the method for the real region.

References

1. Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction. United Nations, 2016, 41 p. URL: https://undocs.org/en/A/71/644.
2. Implementation guide for local disaster risk reduction and resilience strategies. UNISDR, Geneva, 2018, 94 p. URL: https://www.unisdr.org/files/57399_drrresiliencepublicreview.pdf7.
3. Kravchenko, R.A. Do pytannia pro stratehiiu znyzhennia ryzyku nadzvychainykh sytuatsii pryrodnoho i tekhnohennoho kharakteru v Ukraini. *Investytsii: praktyka ta dosvid*, vol. 3, pp. 104–110 (2019).). *[in Ukraine]*.
4. Kravchenko, R.A. Vykorystannia yevropeiskoho dosvidu orhanizatsiinoho zabezpechennia znyzhennia ryzyku nadzvychainykh sytuatsii pryrodnoho i tekhnohennoho kharakteru. URL: http://www. e-patp. academy. gov. ua/pages/dop/19/files/e7f7dff9-21a4-46f4-b4a1-87e960d9903a. pdf. *[in Ukraine]*.
5. Kravchenko, R.A.. Do vyznachennia orhanizatsiino-pravovykh zasad derzhavnoho upravlinnia znyzhennia ryzyku nadzvychainykh sytuatsii pryrodnoho ta tekhnohennoho kharakteru. *Science Review*, vol. 7 (24), pp. 19-24 (2019). *[in Ukraine]*.
6. Fedorchak, V.V. Naukovi zasady vyznachennia ryzyku vynyknennia NS u systemi derzhavnoho upravlinnia. *Visnyk Natsionalnoho universytetu tsyvilnoho zakhystu Ukrainy. Seriia: Derzhavne upravlinnia*, vol. 2, pp. 379–386 (2017). *[in Ukraine]*.
7. Fedorchak, V.V. Udoskonalennia derzhavnoi polityky u sferi zakhystu naselennia i terytorii vid ryzykiv vynyknennia nadzvychainykh sytuatsii. *Publichne upravlinnia ta mytne administruvannia*, vol. 2 (17), pp. 121–125 (2017). *[in Ukraine]*.
8. Fedorchak, V.V. Intehratsiinyi mekhanizm derzhavnoho upravlinnia ryzykamy vynyknennia nadzvychainykh sytuatsii. *Vcheni zapysky Tavriiskoho natsionalnoho universytetu imeni VI Vernadskoho. Seriia: Derzhavne upravlinnia*, vol. 29 (68), No 1, pp. 167–171 (2018). *[in Ukraine]*.
9. Bielousov, A.V. Naukovi pidkhody do vyznachennia ryzyku nadzvychainykh sytuatsii yak ob'iektu upravlinnia. *Naukovi rozvidky z derzhavnoho ta munitsypalnoho upravlinni*, vol. 1, pp. 224–235 (2015).
10. Kravtsiv, S.Ya., Sobol, O.M., Tiutiunyk, V.V. Otsiniuvannia parametriv vplyvu na intehralnyi pozhezhnyi ryzyk za dopomohoiu faktornoho analizu. *Pozhezhna bezpeka*, vol. 30, pp. 99–104.
11. Kravtsiv, S., Sobol, O., Komyak, V., Danilin, O., Al'boschiy, O. Mathematical Model of Management of the Integral Risk of Emergency Situation on the Example of Fires. *In: International Conference on Information Technology in Disaster Risk Reduction*., vol. 4, pp. 182–195. Springer, Cham (2019, October). URL: https://link.springer.com/chapter/10.1007/978-3-030-48939-7_16.

12. Pro skhvalennya Stratehiyi reformuvannya systemy Derzhavnoyi sluzhby Ukrayiny z nadzvychaynykh sytuatsiy. URL: http://zakon5.rada.gov.ua/laws/show/61-2017-p. *[in Ukraine]*.
13. Pro zatverdzhennya kryteriyi utvorennya derzhavnykh pozhezhno-ryatuval′nykh pidrozdiliv (chastyn) Operatyvno-ryatuval′noyi sluzhby tsyvil′noho zakhystu v administratyvno-terytorial′nykh odynytsyakh ta pereliku sub″yektiv hospodaryuvannya, de utvoryuyut′sya taki pidrozdily (chastyny). http://zakon0.rada.gov.ua/laws/show/874-2013-п. *[in Ukraine]*.
14. DSTU 8767:2018 Pozhezhno-riatuvalni chastyny. Vymohy do dyslokatsii ta raionu vyizdu, komplektuvannia pozhezhnymy avtomobiliamy ta proektuvannia. *[in Ukraine]*.
15. Kravtsiv, S.Ya. Metod minimizatsii intehralnoho pozhezhnoho ryzyku za dopomohoiu optymizatsii pokryttia pozhezhnykh depo. *Naukovi visti KPI*, vol. 30, pp. 30-37. URL: file:///C:/Users/user/Downloads/NVKPI_2018_4_6.pdf. *[in Ukraine]*.

Impact of Industries on the Intensive Application of EU and Republic of Moldova Intellectual Property Rights

Alexandru Gribincea

Academy of Economic Studies of Moldova, Free International University of Moldova, Chisinau, Moldova
agribincea @mail.ru

Abstract. Intellectual Property Rights (IPRs) - the copyrights, patents, trademarks and similar rights on which the share of creative and innovative products and services is based, play a vital role in the development of the economies of developed and developing countries around the world, by stimulating innovation, by offering large and small companies a series of tools to contribute to their success and to the benefit of consumers and society through a continuous flow of innovative and competitive products and services and an extension of the general state of knowledge of the society.

Keywords: Intellectual property, rights, industrial property, patents, designs, application, EU, RM.

1. Introduction

Most of the analyzed information depicts a positive relationship between a company's investment in intellectual property and its business performance and, in particular, its financial performance. Industries whose activity is closely linked to the protection of their intellectual property are viable and developing, and their importance in the European Union economy is growing. In the Republic of Moldova, the enforcement degree of intellectual property rights is not very high, because of the lack of resources, of the society's low awareness level of the intellectual property rights and the existence of conflicts of interest.

The purpose of this research: paper is to investigate whether the companies' accumulation of intellectual property has an impact on their business performance and on the economies of the EU and the Republic of Moldova.

2. Research methods

The methods of systemic, logical and comparative analysis and economic statistics were used to carry out the tasks established in the investigation. Normative and methodological materials from foreign and international regulatory bodies in the field of IPR, as well as other state and international governmental bodies, lie at the methodological basis of the study.

Intellectual property (IP) is now widely recognized as a key factor in modern economies. The emergence of the knowledge economy, globalized markets and the increasing complexity of products and services has contributed to the growing importance of the intellectual property rights. For many companies from advanced economies, their intangible assets are now far more valuable than their physical assets. Empirical evidence suggests that in their attempt to withdraw additional value from their innovations, companies are considering alternative means of IP protection, often complementary. Looking to the future, most policy makers strongly believe that knowledge-intensive industries will drive growth and prosperity not only in Europe but also around the world [1].

As shown in Figure 1, European filings for patents, trademarks, designs and models submitted by European applicants have increased significantly since 2004. Overall, EU trademark and European patent filings have increased over the entire period, witnessing a faster growth for trademarks. The increase in project filings in the EU was comparable to the rate of EU trademarks until 2007, but has been considerably slower since then, and filings stagnated between 2012 and 2015, during the global economic crisis, before taking off again in 2016. Generally, the resilience for the IPR submission is another indicator for the growing importance of IPR.

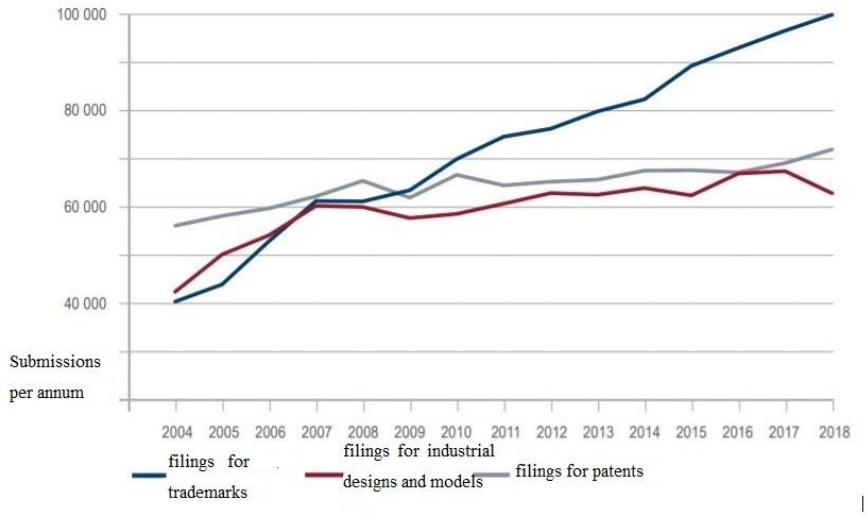


Fig. 1. European filings for patents, trademarks, designs and models submitted by European applicants

Industries whose activity is closely linked to the protection of their intellectual property are viable and developing, while their importance in the European Union economy is growing. This is the overall conclusion that can be drawn from a report presented by the European Union Intellectual Property Office (EUIPO), as well as, the European Patent Office (EPO) published in October 2016. This report, supported by various research and statistical studies, completed in 2011-2013, is a follow-up of a previous report drafted for 2008-2010.

This study meets the need of keeping innovation at the heart of European growth, as specified in the "Europe 2020" growth plan, which should guide the European economy from 2010 to 2020. Indeed, the main goal of any intellectual property is to encourage and protect innovation and creation. Therefore, the study of the industries' importance that fiercely protect their intellectual property, can describe the condition and importance of innovation in the European Union.

All or almost all of the different types of intellectual property rights were considered in this study, also compared with each other, which allows for the estimation of the most active intellectual property sectors. The study mainly considers the gross domestic product (GDP) as an economic indicator, but also studies the employment rate, foreign trade in the European Union, as well as the average wages in different sectors.

Overall, the conclusions of the report are positive: the importance of the studied industries has been constantly growing, accounting for 28% of all jobs in the European Union

and up to 42% of its GDP. This represents 5.7 trillion euros. In addition, these industries make up the majority of the European foreign trade.

However, digging a little deeper and comparing, as far the study allows, the different intellectual property rights, we can observe noticeable differences between these various sectors. For example, the industries in which most trademarks are registered account for 36% of the European GDP, while those that submit most of the patents or generate most of the copyrights account for only 15% and 7%. This difference may be explained by the fact that most industries use trademarks to protect their image, regardless of the targeted economic sector, while patents and copyrights are used only in the sectors where new objects, inventions or works are developed.

The fact that the share of these industries compared to the European GDP is higher than their share in the labor market also means that the employees in these industries generally receive higher wages, the difference in other types of industries is up to 46 %. However, from this perspective, the influence of the different types of intellectual property varies. Indeed, wages are the highest in the copyright and patent industries, where the calculated difference was 64% and 69% respectively, while this difference was 48% for the related industries and trademarks, 31% for geographical indications [2].

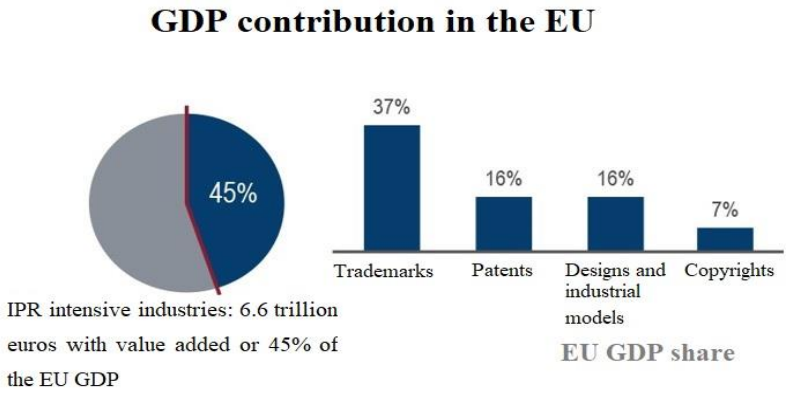


Fig. 2. EU GDP share

Intensive industries in intellectual property rights (IPR) such as patents, trademarks, designs and copyrights generate 45% of the EU GDP per year (€ 6.6 trillion) and account for 63 million jobs. (29% of all jobs in the EU) (Figure 2). These are the conclusions of a joint report currently published by the European Patent Office (EPO) and the European Union Intellectual Property Office (EUIPO), which analyzes the impact of IPR on the European economy between 2017 and 2019.

The EU has had a long tradition for the provision of models, being a world leader. Industries that especially use the design rights have a strong economic impact in the EU. They account for 30.7 million direct jobs and contribute with 16.2% to the total EU GDP. Their exports generated a trade surplus of over 66 billion euros in 2019.

Trademark registrations are often indicators of a company's business success, establishing the degree of brand awareness and emphasizing its distinctive character in the market. Brand industries contribute heavily, by 37%, to the EU GDP and generate 46.7 million jobs. These industries also pay 48% higher wages than those that do not use intellectual property rights [4].

Given the value of intellectual property rights, it is not surprising that criminals try to profit illegally from these assets by producing and distributing counterfeit and pirated products. Indeed, counterfeiting and piracy are not just activities of unfair competitors, but a real criminal scourge. While all types of IPR are infringed, they are often the most innovative and sought-after products and, therefore, the most profitable, being affected, which makes the phenomenon of counterfeiting even more harmful to industries. Having a high return on investment and a relatively lower risk of sanctions compared to other illegal activities, incentives are favorable for counterfeiting and piracy, allowing for the finance of other criminal activities.

Criminal practices have also been rapidly adapting to the technological developments and the international environment. The COVID19 pandemic has been an advantage for counterfeiters, who have quickly exploited citizens' concerns and expanded e-commerce for the sale of counterfeit treatments and counterfeit protective equipment. While e-commerce is facing an unprecedented boom, a third of Europeans have already questioned the product authenticity they have purchased. Almost one in 10 European consumers says they have already been deceived into buying counterfeit products. Thus, consumers are exposed to risks using low-quality counterfeit products that do not meet safety standards and have harmful effects on the environment. Raising the awareness of the IP importance, from an early age and throughout the educational system, can lead to an efficient change in practices.

The magnitude of the threat posed by IPR infringements is of great concern. EUIPO and the OECD jointly estimated counterfeit imports into the EU from the rest of the world at € 119 billion in 2019, or 5.8% of total EU imports. In addition, EUIPO estimated the damage to legitimate industries within 11 sectors in terms of sales losses due to counterfeiting. This deficiency represents over 83 billion euros per year for 2013-2017. In addition, the presence of counterfeits on the market has cost more than 671,000 jobs and deprived the Member States of at least 15 billion EUR mandatory fees per year. As regards the impact of online copyright and related rights infringements, intellectual property infringements must be tackled with determination and cooperation. This fight can only be won if the rightsholders and the law enforcement authorities work in partnership. Rightsholders know their rights, products and services, as well as markets, better than anyone, but the authorities need this information to effectively combat the IPR infringements. The IPEP electronic portal allows rightsholders and public authorities to securely exchange information about the compliance with IPR. Intermediaries, such as e-commerce platforms, payment services or social networks, also have a central role to play in the fight against counterfeiting. It is essential to create the conditions, but also the tools for an enhanced cooperation with them. For this purpose, EUIPO has developed a strategic project in order to improve IPR protection on e-commerce platforms. This is part of a broader concept on the use of the IPEP portal in order to involve all stakeholders in the fight against counterfeiting and piracy.

Likewise, new digital technologies offer promising tools for IPR enforcement. For example, EUIPO will launch a competition to design a blockchain-based authentication platform that would integrate product tracking systems containing risk analysis tools of customs authorities, police, and market surveillance. Developments in artificial intelligence are also being monitored, which would make it possible especially for the development of automated content recognition technologies to detect products or content that would infringe IPR. One project aims namely at listing the existing anti-counterfeiting technologies for guiding companies and, in particular SMEs, on how to use the most appropriate protection tools for their intellectual property rights.

In order to better tackle the threat of intellectual property infringements, the coordination of law enforcement efforts is essential. Close cooperation between EU customs

and police authorities and the exchange of data, taking all the necessary protection measures to safeguard them, are the only ways to avoid the breaches that criminals will surely take advantage of. The recent decision of the EU Council of Ministers to include intellectual property offenses among the priorities of the EU's policy cycle for combating serious international organized crime (known as EMPACT) for 2022-2025, is a crucial step forward. The recognition of the importance of the threat should propel the European and national authorities and facilitate their cooperation [3].

In the Republic of Moldova, the degree of enforcement of intellectual property rights is not very high, because of the lack of resources and the society's low level of awareness of intellectual property rights and the existence of conflicts of interest.

As regards intellectual property rights, most respondents reckoned that they encountered problems with their enforcement, based on the 2019 EESE Survey. 31% of the surveyed stated that intellectual property rights were "defective and not protected by law", while 24.5% considered that they were "well defined, but not well protected by law"; 17% of those interviewed felt that they were "well defined and protected", while 5.5% claimed that these rights "do not exist", and over a fifth (22%) did not want to answer or did not have an answer.

Intellectual property laws cover patents, copyrights and other related rights, the legal protection of industrial designs, trademarks and product designations of origin, the protection of plant varieties and the protection of integrated circuit topographies. The State Agency for Intellectual Property (AGEPI) is a central administrative authority subordinated to the Government, responsible for promoting and carrying out activities in the field of legal protection for intellectual property concerning industrial property rights, copyrights and related rights. The "Intellectual Property Protection Index" (Figure 3) is based on the World Economic Forum (GEF) survey, in which business leaders are asked to express their views on the following: "Measures to protect intellectual property and combat counterfeiting in your country are 1 = inappropriate and not applied, 7 = appropriate and applied". In 2017, compared to all other countries, Moldova did not receive the worst result (3.5 points), as in the same year Albania reached 3.4 points. Also, Moldova registered an improvement, starting with 2014, obtaining 2.8 points. Estonia managed the best performance as regards this, with a value of 5.4 points in 2017, followed by Malta with 5 points. In 2017, Slovenia had a value of 4.7 points, and Romania acquired 4.6 points [5].

Intellectual Property Protection Index Intellectual property protection indices are based on the World Economic Forum survey (WEF) in which business leaders are asked to express their views on the following: Measures to protect intellectual property and combat counterfeiting in your country 1- are inappropriate and not properly applied, 7- appropriate and applied. Source: Global Economic Forum, Global Competitiveness Report.		GC report 2014/15	GC report 2015/16	GC report 2016/17	GC report 2017/18
	Republic of Moldova	2.8	3.1	3.4	3.5
	Estonia	4.9	5.2	5.5	5.4
	Slovenia	4.2	4.5	4.7	4.7
	Albania	2.9	3.2	3.3	3.4
	Romania	3.4	3.9	4.1	4.6
	Malta	4.5	4.6	4.9	5

3. Conclusions

In today's world, the abundant supply of goods and services on the markets has made it very difficult to exist for any business, large or small. In its ongoing effort to stay ahead of competition in this environment, each company strives to create new and improved products (goods and services) that would offer users and customers more value than the products offered by competitors. In order to differentiate their products - a precondition for success in today's markets - companies rely on innovations that reduce production costs and / or improve product quality. In a crowded market, companies must make a continuous effort to communicate the specific value offered by their product through effective marketing that is based on well-thought -out branding strategies.

All businesses, especially the well-established ones, nowadays must rely on the effective use of one or more types of intellectual property (IP) in order to gain and maintain a substantial competitive advantage in the market. Therefore, business leaders and managers need a much better understanding of the IP system tools to protect and exploit the IP assets they own or fancy to use for their business models and competitive strategies in domestic and international markets.

The use of IP is one of the key strategies that businesses can rely on to improve their competitiveness. Innovative undertakings may also rely on delivery time, market speed, contractual arrangements or technical means to prevent copying.

Many companies do not make full use of the IP system, often due to the lack of awareness or understanding, of expertise or concern about costs. The level of understanding how IP is managed and marketed varies among companies, although small and medium-sized enterprises (SMEs) remain namely in this area. It is essential for businesses and their advisers to understand that the IPR legal protection by itself is not enough and a successful IP management strategy must integrate the IP role into the wider business context.

References

1. IPR-intensive industries and economic performance in the European Union. https://euipo.europa.eu/tunnelweb/secure/webdav/guest/document_library/observatory/documents/IPContributionStudy/IPRintensive_industries_and_economicin_EU/WEB_IPR_intensive_Report_2019.pdf.
2. Intellectual Property: Powerhouse for Innovation and Economic Growth. https://iccwbo.org/content/uploads/sites/3/2011/02/Intellectual-Property-Powerhouse-for-Innovation-and-Economic-Growth.pdf
3. An introduction to intellectual property for businesses. https://www.managingip.com/article/b1r8xwrj71rt93/an-introduction-to-intellectual-property-for-businesses
4. Making intellectual property work for https://www.wipo.int/edocs/pubdocs/en/intproperty/956/wipo_pub_956.pdf.
5. Mediu de afaceri favorabil pentru întreprinderi durabile în Republica Moldova https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_ent/documents/publication/wcms_736671.pdf.

Difficulties in Organizing SME Business during Covid-19 Measures and Restrictions in the Republic of Serbia

Ivana Marinovic Matovic [1], Milos Pavlovic [2, 3]

[1] PUC Parking Service, Nis, Serbia
[2] Belgrade Business and Arts Academy of Applied Studies, Belgrade, Serbia
[3] University of Pristina, Faculty of Economics, Kosovska Mitrovica, Serbia
ivana.m.matovic@gmail.com, milospavlovic84@hotmail.com

Abstract. The paper reviews Covid-19 crisis, and the measures implemented at national level to prevent the spread of infection, which affected all business organizations. Small and medium-sized enterprises (SMEs) are particularly vulnerable in crisis situations, as they generally do not have a plan that would led them through the crisis caused by Covid-19 pandemic. SMEs were vulnerable in Covid-19 crisis for several reasons, primarily due to resource constraints and weaker market positioning, compared to large companies. This paper analyzes operational problems in SME business after the pandemic outbreak, and it explores the impact of Covid-19 crisis on certain segments of their business.

Keywords. Small business, SME, Crisis management, Covid-19, Entrepreneurship management.

1. Introduction

The Covid-19 pandemic had a huge impact on macroeconomics and society, and thus represents the most powerful challenge faced by contemporary society. The pandemic began in December 2019 in China and soon spread to other parts of the world. The first case of Covid-19 virus in Serbia was recorded in March 2020, and in February 2022 over 1,750,000 infected and almost 14,000 deaths were registered [1].

The spread of virus and the effect of a pandemic on the economy are very uncertain, making it difficult for policymakers to formulate an appropriate response to the new situation. One of the first studies examining the effects of Covid-19 on macroeconomics simulated the basic and negative scenario on the counties of Asia-Pacific region, North America, Europe and other developed economies [2]. The results of the baseline scenario predicted a sharp decline in global economy in the first half of 2020, followed by a slow recovery. The results of the negative scenario showed a decline in confidence in the measures that countries are taking to reduce the economic effects of Covid-19 virus, and a more pronounced slowdown in growth. Cecchetti and Schoenholtzon [3] and Beck [4] analyzed the Covid-19 impact on financial institutions, primarily banks, and found that these effects primarily depend on fiscal and monetary policy responses and the sensitivity of these institutions to economic shocks. Fernandes [5] argues that the economic effects of Covid-19 are not realistically rated, primarily due to the assumption of similarity with SARS pandemic and the Global Financial Crisis (GFC), since Covid-19 pandemic has both health and economic effects.

In order to limit the virus spread, governments have enforced various measures that include movement restrictions, social distancing, and the like. State measures were rather heterogeneous immediately after the crisis, but differences in variation between countries diminished over time [6]. The Republic of Serbia implemented measures a few days after the

first diseased, introducing a state of emergency and measures that included closing educational institutions and movement restrictions for people over 65. In Covid-19 crisis conditions, institutions have probably the most significant role in preserving the social and economic stability.

2. Challenges of Covid-19 pandemic

In order to restrict the spread of the disease, countries around the world are adopting a number of measures aimed at preventing further infection of the population, including social distancing [7]. The measures also meant companies, schools, public gatherings have been closed; preventive locking measures have been introduced with travel restrictions limited to just necessary [8]. The Covid-19 pandemic has a greatest threat to the health of the oldest citizens [8]. Figure 1 shows the total number of registered Covid-19 infections globally, in the period January 2020 - February 2022.

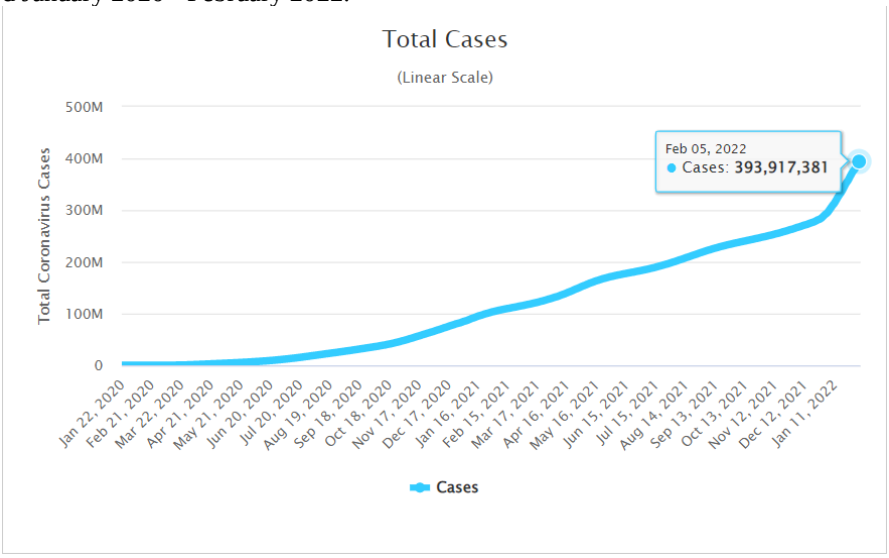


Fig. 1. The total number of registered Covid-19 infections in the world (January 2020 – February 2022) [9]

In response to the appearance of Covid-19 virus, China first applied anti-Covid measures in February 2020. In April 2020, similar was done in many other countries, with some temporal differences and adjustment of measures to the specific conditions of each country [10]. These measures, which are still in force in most countries of the world, include limited movement within countries and across borders, measures of social distance, closing schools, closing parts of companies, restrictions on public gatherings. The measures led to economic inactivity in most countries in an extent not recorded since the GFC, or the Great Depression of 1929-1939 [10]. The timeframe and scope of the applied strategies to combat the Covid-19 pandemic varied significantly from country to country, ranging from efforts to detect and monitor contacts of an infected person, to serious measures of social distance, including national and economic blockades [11]. Figure 2 shows the scope of the nine observed restriction strategies, applied in the world, on a scale from 1 to 100, where 1 represents the absence of restrictions, while the score of 100 is assigned to the strictest applied measures. The nine restrictive measures, which were observed, are: school closure; job loss; disposal of public

gatherings; public transit closure; stay-at-home policy; media campaigns; movement restrictions at local, regional and national level; and international travel-related control measures. Figure 2 shows the scope of preventive measures applied as Covid-19 pandemic control in the world, over a period of one year (March 2020 - March 2021).

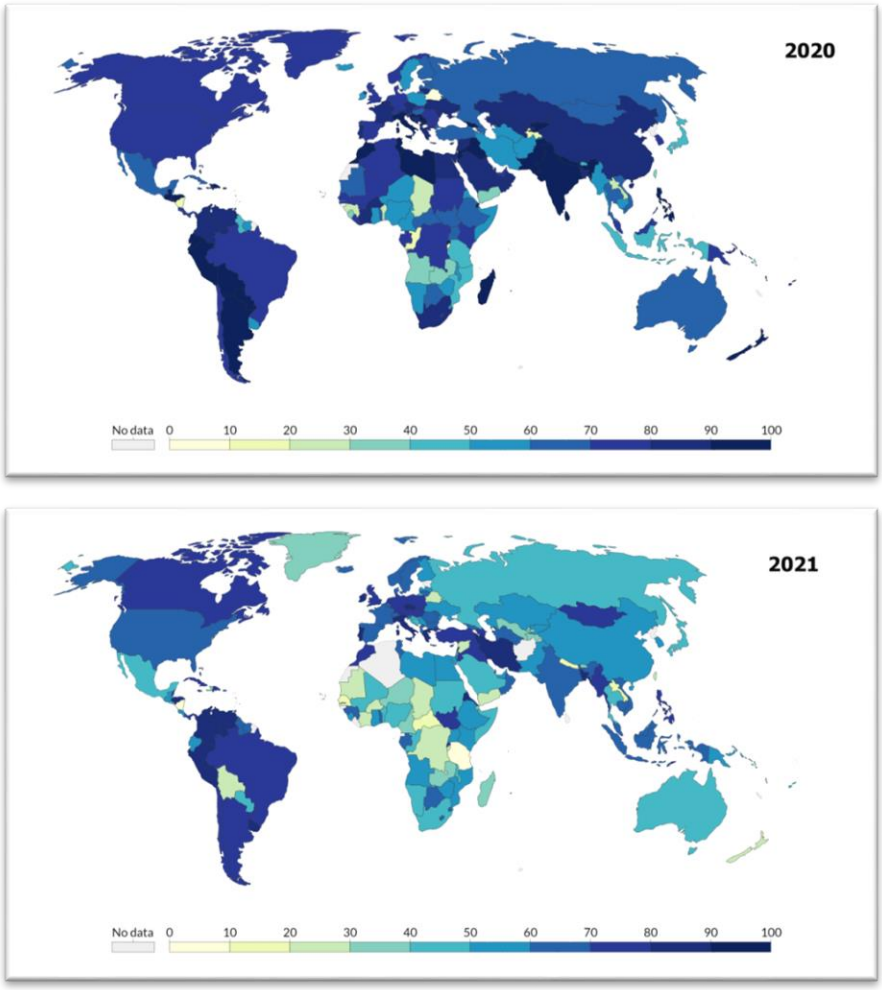


Fig. 1. The scope of anti-Covid measures applied in the world (March 2020 – March 2021) [11]

General measures, related to the suppression of epidemiological emergency, have been imposed with the aim of preventing and protecting from exposure to the virus. Preventive measures included avoiding new infections, or preventing the virus from spreading to locations where it is already present. At the organizational level, the measures included restricting the visitors, limiting access to locations and employees, while isolating those who exhibited symptoms indicating the possibility of infection [12].

Many countries have tried to implement short-term measures, such as increasing external debt and printing money, to alleviate the economic impact of Covid-19. Central banks

have tried to implement financial policies with the goal of cheaper borrowing, in order to compensate for the lack of cash during the pandemic crisis and encourage spending [13]. Measures of social distancing disrupted international supply chains, supply and demand markets. This has led to increased turbulence in the financial market and increased the macroeconomic impact of pandemic. The higher volume of borrowing by business organizations and the population, during the pandemic, caused stronger economic disruptions compared to previously known pandemics [14].

3. Brief overview of anti-Covid measures applied in Serbia

The Covid-19 crisis led to the implementation of a number of monetary measures in the Republic of Serbia. The first set of anti-Covid actions was implemented by the National Bank of Serbia (NBS) in the form of reducing the reference interest rate from 2.25% to 1.75% in March 2020, then reducing its level several times, up to 1% in December 2020 [15]. The NBS also prescribed a three-month moratorium on all bank loans and financial leasing agreements, obligatory for financial institutions, with the consent of the borrower. The moratorium was introduced twice, in April 2020 for a period of 90 days and in August 2020 for a period of 60 days [16].

Fiscal measures included a 10% increase in wages in the health sector, as well as increased budget spending on health. Retired received a one-time financial assistance from the state, while all adult citizens received a financial assistance amounting to 100 EUR [17]. The Serbian government has allocated 200 million EUR to support the development of infrastructure projects, while the European Union has approved a package of financial support of 93 million EUR: 15 million EUR for immediate needs and 78 million EUR for short and medium term needs [18].

The Republic of Serbia is one of the countries of the Western Balkans that has approved state guarantees and direct loans to small and medium enterprises (SMEs) in order to overcome the consequences of the pandemic [19]. Business support have included extending the payment of salary taxes, for three months, for all private-owned companies, with these obligations being repaid in 24 installments during 2021. The state approved wage subsidies, which included the donation of minimum wages for SMEs' employees, in the three-month period, and the contribution of 50% of the net minimum wage for large companies' employees, also for three months. The total economic anti-Covid aid in the Republic of Serbia amounted to 11% of GDP [17].

The Serbian government has implemented a package of fiscal measures to reduce the disruptions that arised from Covid-19 pandemic. The priority was three sets of measures [20]. The first group of measures included prolonging the settlement of tax obligations, and repayment in installments, from the beginning of 2021. The second set referred to direct assistance to SMEs and entrepreneurs, as well as subsidies to large companies to maintain current employment levels. The third set was aimed at maintaining liquidity through programs for granting loans and guarantee schemes to commercial banks [21]. The main motives of the proposed measures were aimed at providing the economic support in the short term, amortizing the decline in macroeconomic indicators, such as unemployment, and boosting demand in the medium term, after lifting the restrictions caused by the pandemic.

At the macroeconomic level, in the hospitality, tourism and passenger transport sectors, SMEs have felt a significant negative impact of the crisis, as a result of limited demand. Given the importance of these sectors for Serbian economy, and the further development in crisis conditions, special measures of state support in the field of liquidity, fiscal and other facilities, have been formed for them. State support included special conditions for these three sectors, additional benefits related to the already active support measures for the economy.

4. Operational problems in SME business in Covid-19 crisis

The survival and sustainable business of SMEs in Covid-19 crisis depends on the duration of the crisis, the applied measures of movement restrictions, business incentive programs, as well as the resources available to the company. The availability of liquid assets is of great importance during the crisis period, especially in a situation of untimely state support [22].

According to the UNCTAD study [23], SMEs have been significantly more affected by declining demand for products or services, lack of necessary raw materials, reduced liquidity and availability of liquid assets, cased by the pandemic crisis. Anti-Covid measures, such as locks and quarantines, have had disastrous effects on the operations of these companies. Although most companies were hit by Covid-19 crisis, the strength of its influence and the possibility to proactively react varied depending on size. The magnitude of the crisis impact on SMEs, as well as the risk of their permanent closure, are illustrated in Figures 3 and 4.

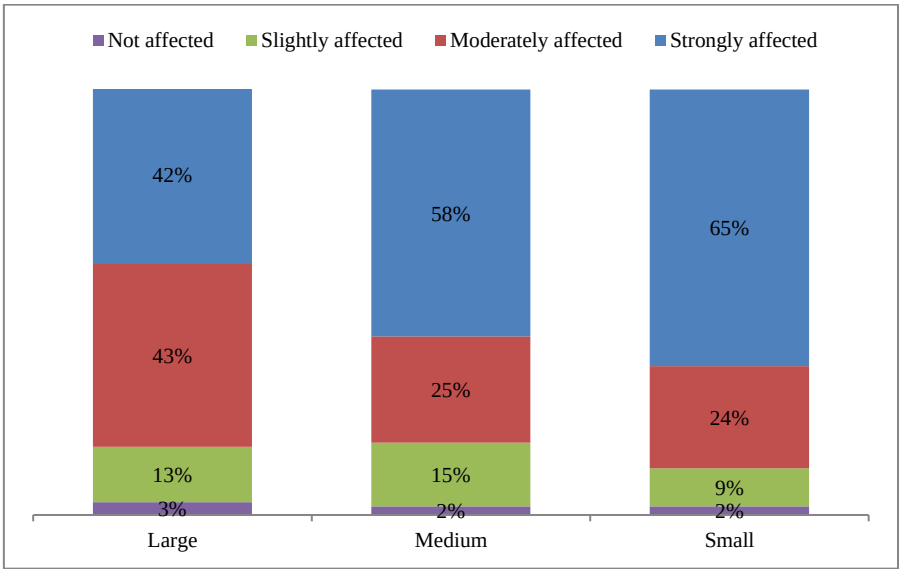


Fig. 3. The strength of Covid-19 effect on SMEs and large companies [24]

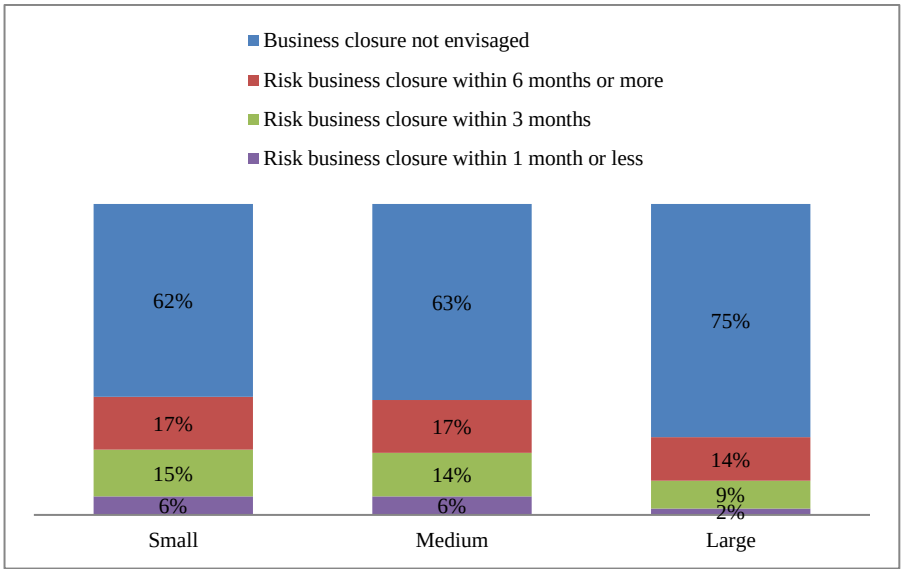


Fig. 4. Risk of permanent closure of SMEs and large companies in pandemic [24]

Levashenko & Koval [25] concluded in their research that risks caused by Covid-19 pandemic were much more pronounced in the SME business, in a comparison with large companies. A survey conducted by the International Trade Center [24] confirmed the strong impact of the crisis on two-thirds of SMEs, compared to 42% of large companies. The impact of Covid-19 crisis on SMEs is particularly strong due to their presence in the most affected sectors [24]. Limited resources and insufficient liquid assets available to SMEs make them particularly vulnerable to crisis conditions, reflecting the risk of their possible permanent closure related to Covid-19 impact. Rathore & Khanna [26] found in their study that up to 70% of small businesses believe that restraint measures have led to the cessation of their business within three months.

Regardless of the threats faced by SMEs, the Covid-19 pandemic has also created new opportunities for profitable business. New needs have emerged in the market for certain products and services, such as protective clothing and other equipment, and other medical devices. Pandemic-accelerated digitization has led to demand for digital products and services, although the digitization process is uneven across national economies. In developed economies, and relatively high-income developing economies, the digitalization process is more intensive, although in these countries there is already a developed digital infrastructure, online trade and distribution network [27].

5. Conclusion

Companies that react quickly and adapt to unpredictable circumstances, are characterized by flexible and direct guidelines in work activities, and properly distributed workforce, which allows them to survive in times of crisis. A complete understanding of the environment changes under the influence of Covid-19 crisis is necessary. Companies should adapt and change according to new circumstances, and in order to reduce the harmful effects.

SME support has been an important part of aid packages in many countries. Some of the measures applied to all business organizations, while others were specifically targeted at SMEs and their employees. The most common measures to support SMEs during the Covid-

19 pandemic were credit financing and employment incentives. In the fiscal field, a number of targeted measures were implemented, in the form of tax breaks, to reduce the liquidity burden on SMEs and help their businesses. Measures to support SME business need to be strengthened in transition countries, especially low-income economies.

Support for SME operations in crisis conditions should include measures to mitigate illiquidity, then measures to support employees to return to work safely, as well as to improve the efficiency of supply chains and regular operations. Profitability and sustainability of business, in addition to the applied measures to support the economy, depends on the specifics of each SME, business model, financial strength, market position and business activities. In addition to relying on economic support measures, it is necessary for SMEs to react in a timely manner and make corrective measures, necessary strategic and operational decisions, in accordance with the conditions of Covid-19 pandemic.

References

1. Institute of Public Health of Serbia. (2022). COVID-19 coronavirus information [Informacije o koronavirusu COVID-19]. https://www.batut.org.rs/index.php?category_id=202
2. OECD. (2020a). Coronavirus: The world economy at risk. https://www.oecd.org/berlin/publikationen/Interim-Economic-Assessment-2-March-2020.pdf
3. Cecchetti, S. G., Schoenholtz, K. L. (2020). Contagion: Bank runs and COVID-19. In: Baldwin, R., Mauro, B. W. (Eds.), Economics in the Time of COVID-19 (pp. 77-80). London: Centre for Economic Policy Research. https://voxeu.org/system/files/epublication/COVID-19.pdf
4. Beck, T. (2020). Finance in the times of coronavirus. In: Baldwin, R., Mauro, B. W. (Eds.), Economics in the Time of COVID-19 (pp. 73-76). London: Centre for Economic Policy Research. https://voxeu.org/system/files/epublication/COVID-19.pdf
5. Fernandes, N. (2020). Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy. IESE Business School Working Paper No. WP-1240-E. https://ssrn.com/abstract=3557504
6. Hale, T., Anania, J., Angrist, N., et al. (2021a). Variation in government responses to COVID-19. BSG Working Paper Series. https://www.bsg.ox.ac.uk/sites/default/files/2021-06/BSG-WP-2020-032-v12_0.pdf.
7. Fong, M. W., Gao, H., Wong, J. Y., Xiao, J., Shiu, E. Y. C., Ryu, S., Cowling, B. J. (2020). Nonpharmaceutical Measures for Pandemic Influenza in Nonhealthcare Settings — Social Distancing Measures. Emerging Infection Diseases, 26(5): 976–984.
8. Brodeur, A., Gray, D., Islam, A., Bhuiyan, S. J. (2020). A Literature Review of the Economics of COVID-19, Discussion paper series, Bonn: IZA – Institute of Labor Economics. https://ftp.iza.org/dp13411.pdf
9. Worldometer. (2022). Coronavirus Worldwide Graphs, https://www.worldometers.info/coronavirus/worldwide-graphs/#total-cases
10. ILO-OECD. (2020). The impact of the COVID-19 pandemic on jobs and incomes in G20 economies, https://www.ilo.org/wcmsp5/groups/public/---dgreports/---cabinet/documents/publication/wcms_756331.pdf
11. Hale, T., Angrist, N., Goldszmidt, R., et al. (2021b). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nature Human Behaviour, 5: 529-538. https://doi.org/10.1038/s41562-021-01079-8
12. Cirrincione, L., Plescia, F., Ledda, C., Rapisarda, V., Martorana, D., Moldovan R. E., Theodoridou, K., Cannizzaro, E. (2020). COVID-19 Pandemic: Prevention and Protection Measures to Be Adopted at the Workplace, Sustainability, 12(9): 3603, 1-18.
13. Adams, J. G., Walls, R. M. (2020). Supporting the health care workforce during the COVID-19 global epidemic, Jama, 323(15): 1439-1440.
14. Boissay, F., Rungcharoenkitkul, P. (2020). Macroeconomic effects of Covid-19: an early review. BIS Bulletins 7, Bank for International Settlements. https://ideas.repec.org/p/bis/bisblt/7.html
15. National Bank of Serbia. (2020a). Reference interest rate reduced to 1% [Референтна каматна стопа смањена на 1%]. https://nbs.rs/sr_RS/scripts/showcontent/index.html?id=16419.

16. National Bank of Serbia. (2020b). The National Bank of Serbia has enabled an additional delay in the repayment of debtors' obligations - a new moratorium [Народна банка Србије омогућила додатни застој у отплати обавеза дужника – нови мораторијум]. https://www.nbs.rs/sr_RS/scripts/showcontent/index.html?id=15628&konverzija=no
17. OECD. (2020b). The COVID-19 Crisis in Serbia. https://www.oecd.org/south-east-europe/COVID-19-Crisis-in-Serbia-archive.pdf
18. European Commission. (2020). Support to the Western Balkans in tackling COVID-19 and the post-pandemic recovery. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0315&from=EN
19. KPMG. (2020). Serbia**:** Government and institution measures in response to COVID-19. https://home.kpmg/xx/en/home/insights/2020/04/serbia-government-and-institution-measures-in-response-to-covid.html
20. Ministry of Economy. (2020). Program of economic measures to reduce the negative effects caused by Covid-19 virus pandemic and support the Serbian economy [Програм економских мера за смањивање негативних ефеката проузрокованих пандемијом вируса Covid-19 и подршку привреди Србије]. https://privreda.gov.rs/program-ekonomskih-mera-za-smanjivanje-negativnih-efekata-prouzrokovanih-pandemijom-virusa-kovid-19-i-podrsku-privredi-srbije/
21. Milanovic, F., Kolar, M. (2020). Measures to combat the economic consequences of the Covid-19 pandemic [Mere suzbijanja ekonomskih posledica pandemije Covid-19]. Belgrade: Center for International Public Policy.
22. Shafi, M., Liu, J., Ren, W. (2020). Impact of COVID-19 pandemic on micro, small, and medium-sized Enterprises operating in Pakistan, Research in Globalization, 2(December), 100018: 1-14.
23. UNCTAD. (2020). Impact of the COVID-19 Pandemic on Trade and Development, New York: United Nations Publications. https://unctad.org/system/files/official-document/osg2020d1_en.pdf
24. International Trade Centre. (2020). Blog: Quantifying the effect of COVID-19 on small business around the world. https://www.intracen.org/covid19/Blog/Quantifying-the-effect-of-COVID-19-on-small-business-around-the-world-the-world/
25. Levashenko, A., Koval, A. (2020). Measures of Financial and Non-Financial Support to Small and Medium-sized Enterprises (SMEs) in the Wake of COVID-19. Monitoring of Russia's Economic Outlook, Trends and Challenges of Socio-economic Development. Moscow: IEP, 9(111): 7-10.
26. Rathore, U., Khanna, S. (2020). From Slowdown to Lockdown: Effects of the COVID-19 Crisis on Small Firms in India. https://ssrn.com/abstract=3615339
27. World Trade Organization. (2020). E-commerce, trade and the COVID-19 pandemic, https://www.wto.org/english/tratop_e/covid19_e/ecommerce_report_e.pdf.

Application of Structural Equation Modeling and Net Promoter Score for Analysis the Effect of the Service Quality and Digital Loyalty Programs on Tourist Satisfaction and Loyalty during the COVID-19 Pandemic

Ellisya Anandya[1], Aisyah Larasati[1], Abdul Muid[1], Nina Aini Mahbubah[2]

[1] Faculty of Engineering , Universitas Negeri Malang, Malang, Indonesia
[2] Universitas Muhammadiyah Gresik, Gresik, Indonesia
ellisyaanandya@gmail, com, aisyah.larasati.ft@um.ac.id, abdul.muid.ft@um.ac.id, n.mahbubah@umg.ac.id

Abstract. This study aims to analyze the effect of service quality and digital loyalty programs on the satisfaction and loyalty of Kusuma Agrowisata tourists and find out the level of loyalty from Kusuma Agrowisata tourists during the Covid-19 pandemic. This study uses the variables of service quality, digital loyalty program, tourist satisfaction, and tourist loyalty. Data is collected using questionary. The data sampling technique is a non-probabilistic sampling technique with a purposive sampling method with 304 respondents from visitors to ripen harvesting fruits, waterpark, and De Tjangkul photo spot. Data are analyzed using Structural Equation Modeling (SEM) and Net Promoter Score (NPS). This study concludes that Digital loyalty program variables have no significant effect on tourist satisfaction. In contrast, service quality variables have a significant effect on tourist satisfaction. (2) Digital loyalty program and service quality variables have a significant effect on tourist loyalty. (3) Tourist satisfaction variable has no significant effect on tourist loyalty. (4) There is no indirect effect of digital loyalty programs and service quality variables on tourist loyalty through tourist satisfaction. (5) Loyalty level of Kusuma Agrowisata tourism visitors is 17% for indicators willing to recommend tourism to others and 20% for an indicator of willingness to visit again.

Keywords. Service Quality, Digital Loyalty Program, Satisfaction, Loyalty, Covid-19

1. Introduction

During the Covid-19, it had a considerable impact on various sectors, one of which was the tourism industry. The tourism industry sector has decreased in purchasing power and people's visiting power because there is a government call for activities at home. One of the tourist attractions in Batu City that experienced a decline in income during the Covid-19 period was Kusuma Agrowisata. During the incident, Kusuma Agrowisata must prioritize health protocols following government recommendations with the principles of CHSE (Cleanliness, Health, Safety, and Environmental Sustainability) so that visitors can still travel. As a company engaged in tourism, Kusuma Agrowisata must pay attention to service quality. Moreover, in order to be able to compete with existing competitors so that visitors still feel comfortable when visiting during a pandemic.

One of the driving factors for the development of the tourism industry is the convenience of tourists. Therefore, good service quality is expected to provide satisfaction and increase visitor loyalty. In addition to service quality factors, some factors can increase visitor loyalty. Namely, provision loyalty programs in the form of rewards for purchasing products

or services [1]. Therefore, further research needs to analyse the effect of service quality. Furthermore, digital loyalty programs on visitor satisfaction and loyalty and how the level of visitor loyalty to Kusuma Agrowisata when the tourism sector situation is during the Covid-19 pandemic.

2. Literature Review

Tourism is a travel activity to enjoy a tourist destination. In addition, tourists are carried out temporarily to get pleasure and satisfaction supported by the best infrastructure facilities and services provided by the surrounding community, local government, and business actors. [2]. There are four essential components of tourism (4A) that need to be developed in developing tourism, namely Attraction, Accessibility, Facilities, and Ancillary service. [2].

After making a tourist visit, tourists will create satisfaction or loyalty to the tourist attraction. Satisfaction, in general, is a feeling condition that a person has after comparing the expected performance results with the perceived results expectations [3]. The previous finding shows that consumers feel pleasure and satisfaction from something that is expected by expectations or exceeding expectations. If consumers feel disappointed or dissatisfied, it indicates a failure to meet expectations. The level of customer satisfaction depends on the feelings received on the company's performance in providing services to consumers. The existence of consumer satisfaction will create consumer loyalty and lead to positive behavior by recommending tourist destinations to others. Several dimensions of tourist satisfaction are product and service features, attributes of service success or failure, and the presence of customer emotions [4].

While loyalty is a customer's trust in a product or service brand that will continue to make purchases in the future [5]. Loyalty is meant in the scope of the tourism industry sector, namely tourists who make repeat visits to a tourist attraction. Indicators that can measure the level of tourist satisfaction are return visits, tourists to recommend tourist attractions to others, and assessments to provide comments to others [6].

Satisfaction and loyalty of tourists will be formed if business actors provide a service of good quality. Quality of service is defined as a subjective assessment of good or bad service by consumers on the provision of services from providers with a benchmark between expectations and visitor expectations. [7]. The quality of services from the tourism sector includes the tourist attractions owned by tourist attractions. The attraction referred to in a tourist attraction is an attraction that attracts attention, facilities, infrastructure, transportation, and hospitality. The following are five dimensions in determining service quality: Tangible, Reliability, Responsiveness, Assurance, and Empathy. [7].

Apart from good service quality, another strategy that tourism owners can apply for customer retention is using a loyalty program [8]. The existence of a loyalty program will provide rewards to consumers for purchases made by consumers. Rewards given by business actors are additional products, free coupons, discounts for unique customers, or cash-back offers. Along with the development of digital technology, where the general public began to switch to the digital version, the loyalty program developed into a digital loyalty program.

The digital loyalty program is a business strategy actor in encouraging digital-based consumer satisfaction positively to create visitor satisfaction and loyalty. The digital loyalty program is a digital strategy for business actors to maintain and increase customer satisfaction and loyalty so that consumers are not attracted and tempted by the offers provided by competitors.

3. Methodology

In this study, using qualitative data were quantified. The data population used is from Kusuma Agrowisata visitors with sample criteria, fruits pick a tour, waterpark, and De Tjangkul visitors who visit at least once in the past year and are more than one or equal to 17 years old.

The data sampling technique used Nonprobability Sampling with purposive sampling. Determination of the number of samples by multiplying the number of indicators used from 5 to 10 [9]. n this study, there were 30 indicators used. The minimum sample used was 150 respondents and the maximum sample used was 300 respondents.

The research variables used were four variables which were divided into three types of variables. 1) The independent latent variables are digital loyalty programs and service quality. 2) The latent dependent variable is visitor loyalty. 3) Intervening variable is visitor satisfaction. The following four variables used are shown in Table 1.

Tabel 1. Variable Indicator

Variable	Sub variable	Indicator
Service Quality [7]	*Tangible*	Public facilities
		Health care facilities
		Employee
		Notice board
	Reliability	Same service
		Accurate operating hours
		Officer professionalism
		Health protocol compliance
	Responsivenes s	Officer alertness
		Officer's response
		Officer readiness
	Assurance	Visitor convenience
		Visitor safety
		Information truth
		Education for visitors
	Empathy	Attention officers
		Understanding visitor wishes
		Officer's concern
		Officer attitude
Digital loyalty program [1]		*Giveaway*
		Special discount
		Free Tickets
Tourist Satisfaction [4]		Service tour guide
		Facilities
		Visitor expectation
		Convenience Tourist attraction
Tourist Loyalty [6]		Visitor comments
		Willing to pay more
		Visitor recommendation
		Return visit

Research hypothesis:

H_1: There is a significant effect of the digital loyalty program on the tourist's satisfaction.

H_2: There is a significant effect of service quality on the tourist satisfaction.

H_3: There is a significant effect of the Digital Loyalty Program on the tourist loyalty.
H_4: There is a significant effect of service quality on the tourist loyalty.
H_5: There is a significant effect of tourist's satisfaction on the tourist loyalty
H_6: There is an indirect effect of the digital loyalty program on tourist loyalty through tourist satisfaction as an intervening variable
H_7: There is an indirect effect of service quality on tourist loyalty through tourist satisfaction as an intervening variable

Data Analysis Stages:

The following are the stages of analysis used as follows;

1. Describe the characteristics of Kusuma Agrowisata visitors using a pie chart.
2. Analyze data with Structural Equation Model (SEM) with several stages as follows;
   a. Doing SEM assumption test
   b. Create a measurement model path diagram and a structural model.
   c. Convert path diagrams into structural and measurement equations.
   d. Selecting the input matrix and estimation technique of the model to be built
   e. Assess the identification of the model being built.
   f. Evaluating the suitability of the model
   g. Modify the model if the model does not fit.
   h. Interpret the SEM model and analyze the value of direct effect, indirect effect, and total effect on the model.
3. Perform NPS Analysis.

4. Result and Discussion

Validity and Reliability Test

From the results of the validity and reliability test on the research instrument, it shows that all of the variable indicators used have a value of $\boldsymbol{r_{count} > r_{table}}$ (0.113) so that it can be said that all indicators used are valid and able to reveal something that will be measured by the questionnaire so that all the questions in the questionnaire will be used for data collection. From the results of the reliability test, Cronbach's alpha > 0.6 so that all statement items were declared reliable and all questions showed the consistency of the observations.

Tourist Characteristics

The following are the results tourists' characteristics of Kusuma Agrowisata based on demographics and tourist behavior which are shown in Tables 2 and 3.

Tabel 2. Demographics Characteristics.

a. Gender

No	Gender	Quantity
1.	Woman	181
2.	Man	123
	Total	**304**

b. Age

No	Age	Quantity
1.	< 20 years	23
2.	21-30 year	227
3.	31-40 year	46
4.	> 41 years	8
	Total	**304**

Tabel 3. Behavior Characteristics

a. Resources

No	Resources	Quantity
1.	Family or siblings	76
2.	Friends	32
3.	Social media (facebook, Instagram)	183
4.	Others (Billboards, Leaflets, etc.)	13
	Total	**304**

b. Frequency of visits

No	Frequency	Quantity
1.	1 time	213
2.	2 time	48
3.	3 time	21
4.	> 3 time	22
	Total	**304**

Based on Table 3(a), the frequency of visitors dominated by a female is 59.54%. Table 3(b) shows the age range of visitors is dominated by ages 21-30 years by 74.67%. Table 4(a) shows that the frequency of visits during the past year was dominated by respondents who visited once by 70.1%. Table 4(b) shows that respondents get information about Kusuma Agrowisata tourism objects from social media by 60.2 %.

Multivariate Normality Test

The assumption test that must be met before performing the analysis with SEM is that it meets multivariate normality. The multivariate normality test can be observed by looking at the value of the critical ratio (c.r) of multivariate kurtosis in the normality assessment of the AMOS. Suppose the critical value is in the range of -2.58 c.r 2.58 with a significance level of 0.01 (1%). In that case, it is concluded that the normal distribution is multivariate [10]. The results of the multivariate normality test in this study obtained that the cr value of all data was 2.137, so that the value was between the value of ± 2.58 so that it could be concluded that the distribution of the observed variable data was customarily distributed multivariate.

Multivariate Outlier Test

Then detect outlier data by observing the Mahalanobis distance (). Mahalanobis Distance is used to show the distance between the observed data and the average value (centroid). It is said to be a multivariate outlier if the value of Mahalanobis d-Squared is greater than the chi-square value at a significance level of p-value < 0.001 [10]. The chi-square value is obtained. The output of the Mahalanobis d-squared shows that all of the observations have the Mahalanobis d-squared below 39.252. It means the observation data used has met the multivariate data outlier test because there is no multivariate outlier data in the data used.

Confirmatory Factor Analysis (CFA).

Confirmatory Factor Analysis is used to present the relationship between indicators and latent variables. CFA is used to confirm the most dominant factors of each indicator variable in compiling latent variables. The following are the results of the CFA analysis presented in Table 4.

Tabel 4. Result Confirmatory Factor Analysis Test

Variable	Indicator	Loading Factor
Digital Loyalty Program	**DLP1**	**0.799**
	DLP2	0.736
	DLP3	0.689
Service Quality	*Empathy*	0.845
	Assurance	**0.846**
	Responsiveness	0.749
	Reliability	0.638
	Tangible	0.697
Tourist satisfaction	**KP1**	**0.807**
	KP2	0.709
	KP3	0.606
	KP4	0.494
Tourist loyalty	LP1	0.735
	LP2	**0.788**
	LP3	0.326
	LP4	0.554

The loading factor results on the digital loyalty program show that the DLP1 has the most significant loading factor value, 0.799. It indicates that the DLP1 has the most significant influence on the Digital Loyalty Program. From the results of the loading factor of the service quality, the Assurance has the most significant value of the other sub-variables, which is 0.846 for service quality. Then, results of the loading factor on tourist satisfaction show that the KP4 has the most significant loading factor value, which is 0.807. The KP4 gives the most significant influence from other indicators on tourist satisfaction. In addition, results of the loading factor on the tourist loyalty show that the LP3 of the tourist loyalty has the most significant loading factor value, which is 0.788. It indicates that the LP3 has the most significant influence on tourist loyalty.

Structural Equation Modeling (SEM) Analysis

Once the measurement model using confirmatory factor analysis (CFA) is calculated, the next step is to analyze the structural model to determine the relationship between latent variables based on the hypothesis used. From the existing hypotheses, it will be developed into a structural model, as shown in Figure 1.

Table 5 shows unfulfilled goodness-of-fit test criteria such as probability (p-value), TLI, and CF1. It is necessary to observe the residual value in the standardized residual covariance column in the AMOS output. Suppose the residual value is obtained outside -2.58 ≤ standardized residual ≤ 2.58, it is necessary to modify the model. The residual values obtained in this study are > 2.58 and < -2.58. It can be interpreted that the model is not acceptable and needs to be modified to the model used.

Next, make improvements to the model to fit and observe the value of modification indices (M.I). The value of the modification indices can be used as a guide to making improvements to the model by looking at the most significant value in the modification indices (M.I). It shows the relationship that needs to be estimated to reduce the chi-square value. Modifying the model can be done by correlating the errors of several indicators or by correlating several indicators with large modification indices values. Modifying the model by estimating the correlation between error terms (measurement errors) does not require theoretical justification, and the value of the resulting error must be greater than or equal to 4.0 [11].

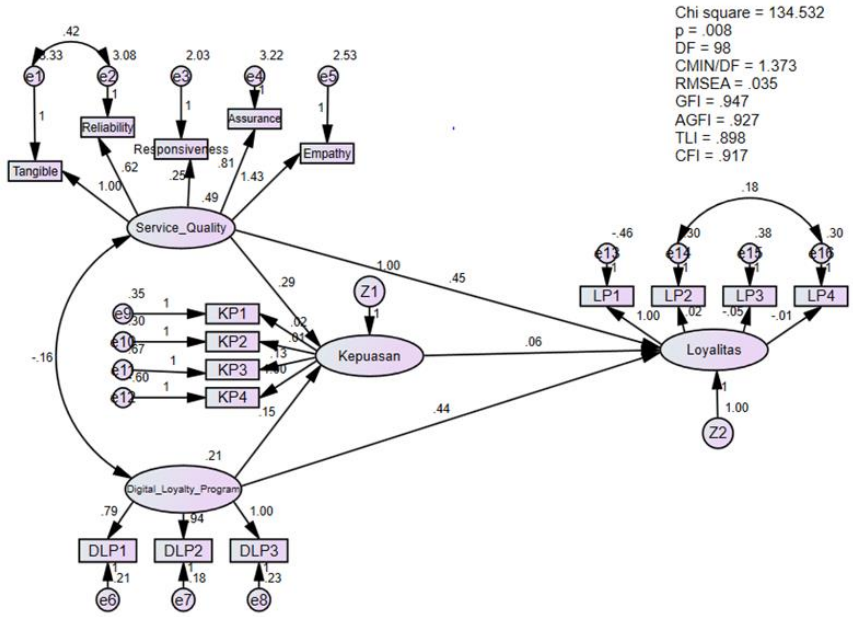


Fig.1. Structural model

Table5. Results structural model goodness of fit test

Goodness of Fit Index	***Cut-Off Value***	**Result**	**Concolusion**
X^2 Chi Square	Expected small (*)	134.532	accepted
Probabilitas	≥ 0,05	0.008	not accepted
CMIN/DF	≤ 2,00	1.373	accepted
RMSEA	≤ 0,08	0.035	accepted
GFI	≥ 0,90	0.947	accepted
AGFI	≥ 0,90	0.927	accepted
TLI	≥ 0,95	0.898	not accepted
CFI	≥ 0,95	0.917	quite acceptable

$*\boldsymbol{X^2}(0.05; 304) = 345.663$

Modifying the model will be carried out in several stages to get the chi-square value to be small and significant. The following summarizes the modification indices values that will be used to modify this model, which can be seen in Table 6.

Table 6. Modification indices forcovariances

			M.I.	Par Change
e4	**<-->**	**e5**	**5.890**	**.449**
e15	<-->	e1	6.245	-.165
e9	**<-->**	**e14**	**6.363**	**.038**
e11	<-->	e2	6.233	.202
e6	<-->	e12	5.456	.042

The steps to be taken in modifying the model are based on the Modification Indices Covariance value as follows:

1. Combining measurement errors from KP1 (e9) and LP2 (e14)
2. Combining measurement errors from Assurance (e4) and Empathy (e5)

The measurement error is not selected because it is illogical to combine. For instance, the suggestion from modification indices to combine the measurement error (e15) from the LP3 Visitor Loyalty variable with the measurement error (e1) from the Tangible service quality variable, which (e15) is an error measurement of the dependent variable while (e1) measurement error of the independent variable.

In the first stage, the measurement error of the indicator is combined with the satisfaction of the service provided. By to the visitor's wishes (KP1), the measurement error of the indicator will not be affected by the increase in the entrance ticket price (LP2). It shows a correlation between measurement errors on KP1 and LP2. It can be interpreted that Kusuma Agrowisata visitors are satisfied with the services provided even though the price of admission tickets increases. Moreover, visitors will not be affected by the increase in admission prices because they are satisfied with the services provided by Kusuma Agrowisata.

In the second stage, the measurement error was combined from the security indicators, cleanliness of facilities, and good tourism education provided by Kusuma Agro wisata employees (Assurance). These indicators were connected with measurement errors from indicators of the attitude of attention and concern of Kusuma Agrowisata employees in providing good service guarantees to tourist Kusuma Agrowisata (Empathy). It shows there is a correlation between measurement error on Assurance and Empathy. It can be interpreted that Kusuma Agrowisata visitors will be guaranteed in terms of security, cleanliness of facilities, and tourism education. It was in line with employees will give an attitude of attention and care to guarantee good service for tourists.

After the model identification process is carried out based on the two stages of combining the measurement error, the model is obtained, as shown in Figure 2.

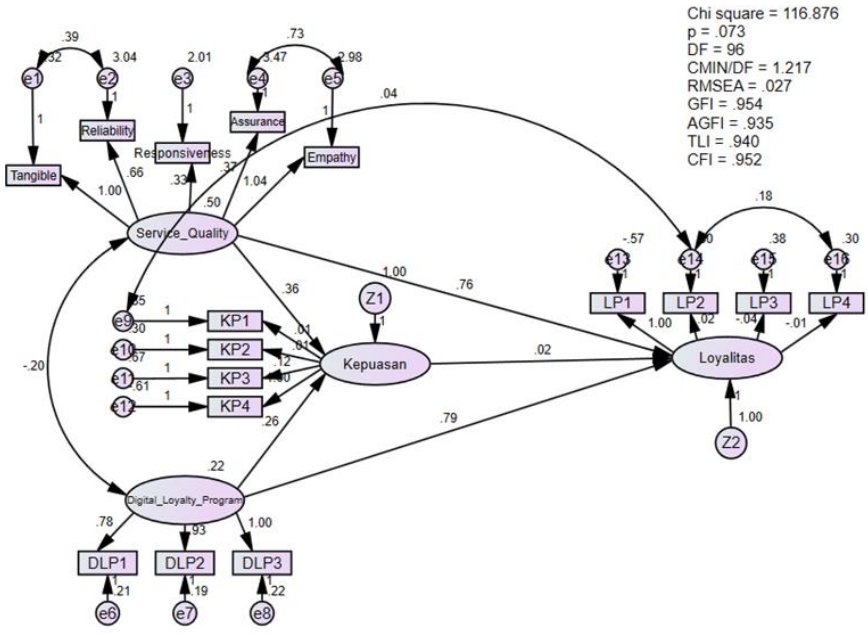


Fig. 2. Model Framework

From the model obtained, then an examination of the goodness test is carried out which is presented in Table 7.

Tabel 7. Results of the goodness of fit structural model

Goodness of Fit Index	***Cut-Off Value***	**Result**	**Concolusion**
X^2 Chi Square	Expected small (*)	116.876	Accepted
Probabilitas	≥ 0,05	0.073	Accepted
CMIN/DF	≤ 2,00	1.217	Accepted
RMSEA	≤ 0,08	0.027	Accepted
GFI	≥ 0,90	0.954	Accepted
AGFI	≥ 0,90	0.935	Accepted
TLI	≥ 0,95	0.940	Quite acceptable
CFI	≥ 0,95	0.952	Acceptable

*$\boldsymbol{X^2}(0.05; 304) = 345.663$

Table 7. shows that the model is good because several criteria for Goodness of Fit have been met. There is no need to modify the model because it shows the model is fit. Thus, this shows that the structural equations produced by the fit model can explain the relationship and influence of each exogenous variable on endogenous variables.

Hypothesis test

Hypothesis testing is used to determine whether the proposed hypothesis is supported by the data used or not. Testing this hypothesis uses the value of Critical Ratio (C.R) on Regression Weights. If | Critical Ratio| which is compared with the value of t_tabel =t⟧ (0.05;304) = 1.968, which is greater, then H_0 is rejected, or the research hypothesis is accepted. The following are the results of the Regression Weights from AMOS presented in Table 8.

Tabel 8. Regression weights

	Estimate	S.E.	C.R.	Decision
KP ← DLP	.097	.141	.685	H_0accepted
KP ← SQ	.285	.128	2.223	H_0rejected
LP ← SQ	.419	.171	2.448	H_0rejected
LP ← DLP	.401	.203	1.976	H_0rejected
LP ← KP	.063	.038	1.652	H_0accepted

Then test the hypothesis of the variables that are influenced by the intervening variable. The following is the critical ratio value on the hypothesis that is influenced by the intervening variable, which is presented in Table 9.

Tabel 9. Critical ratio (c.r)

	Estimate	S.E.	C.R.	Decision
SQ → KP → LP	0.0180	0.0135	1.330	H_0accepted
DLP → KP → LP	0.0061	0.0096	0.635	H_0accepted

In the following, the influence values of each latent variable with a direct effect, indirect effect, and total effect are presented in Table 10.

Table 10. Direct effect, indirect effect, dan total effect

	Direct effect	Indirect Effect	Total Effect
SQ → KP	0.187	0.000	0.187
DLP → KP	0.043	0.000	0.174
SQ → LP	0.269	0.000	0.280
DLP → LP	0.174	0.000	0.177
KP → LP	0.062	0.000	0.062
SQ → KP → LP	0.000	0.012	0.012
DLP → KP →LP	0.000	0.003	0.003

Based on the results of the loading factor obtained in Table 5.4, it will be used to build a combined diagram of the direct effect, indirect effect, and total effect, which will be presented in Figure 3.

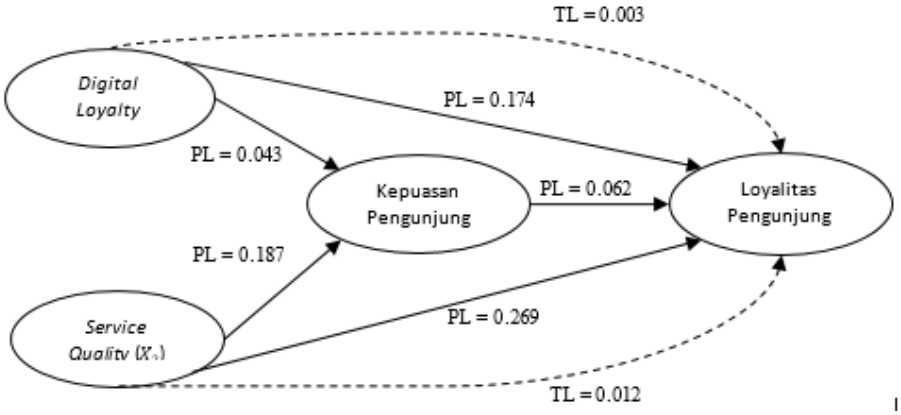


Fig. 3. Direct influence, indirect influence, and total influence

a. **Structural equation I:** Tourist satisfaction = 0.043*Digital Loyalty Program + 0.187* Service Quality + 0.108

b. **Structural equation II:** Tourist loyalty = 0.269*Service Quality + 0.174*Digital Loyalty Program + 0.062*Satisfaction + 0.390

Information:

Error or residual generated from 1 – (Squared Multiple Correlations) obtained from the AMOS.

The effect of digital loyalty program on tourist satisfaction

Hypothesis 1 (KP ß DLP) the critical ratio value obtained is|0.685| < 1.968 accepted. It can be concluded that the digital loyalty program has no significant effect on tourist satisfaction. While the standardized regression coefficient value obtained 0.097, which means that if the digital loyalty program increases by 1 level, the contribution to tourist satisfaction increases by 0.097 times.

It shows that the digital loyalty program presented by Kusuma Agrowisata during the Covid-19 pandemic has not satisfied tourists. Due to the lack of information about the loyalty program conveyed through social media (Instagram) to tourists, there is a need for direct socialization at tourist sites. Besides that, there is no convenience in fulfilling the requirements to take part in the giveaway program or monthly program, so visitors find it difficult to fulfill the requirements set.

The effect of service quality on tourist satisfaction

Hypothesis 2 (KPßSQ) the critical ratio value obtained is|2.223| > 1.968 rejected. The conclusion is that service quality significantly affects tourist satisfaction. At the same time, the standardized regression coefficient value is 0.187, which means that if the service quality increases by 1 level, the contribution to visitor satisfaction increases by 0.187 times.

It shows a relationship between service quality and tourist satisfaction, where the more services provided are good, the more tourists feel satisfied. Kusuma Agrowisata's service quality during the Covid-19 pandemic is good, so tourists feel happy because the expectations and service performance are by the wishes of tourists. Therefore, when the services provided are of higher quality, tourists immediately feel satisfied with what tourists receive to give an attitude with positive comments towards Kusuma Agrowisata.

The effect of digital loyalty program on tourist loyalty

Hypothesis 3 (LPßDLP) the critical ratio value obtained is |1.976| > 1.968 rejected. it can be concluded that digital loyalty program has a significant effect on tourist loyalty. In contrast, the standardized regression coefficient value obtained 0.174, which means that if the digital loyalty program increases by 1 level, it will result in the contribution to tourist loyalty increasing by 0.174 times.

It shows a link between the digital loyalty program and tourist loyalty. Tourists can feel the benefits because of the loyalty program offered. For instance, tourists who have booked hotel tickets through the website will get free tourist entrance tickets.

The effect of service quality on tourist loyalty

Hypothesis 4 (LPßSQ) the critical ratio value obtained is|2.448| > 1.968 rejected. It concluded that service quality significantly affects tourist loyalty, while the standardized regression coefficient value is 0.269. If the service quality increases by 1 level, it will result in the contribution to visitor loyalty increasing by 0.269 times.

Based on the results, the service quality of Kusuma Agrowisata is excellent and successful, so that it can create tourist loyalty. The level of exemplary service performance is evidence in providing services to tourists. Tourists will remember how the attitude of employees in serving tourists makes it possible for visitors to make repeat visits and are willing to recommend tours to others

The effect of tourist satisfaction on tourist loyalty

Hypothesis 5 (loyaltyßsatisfaction) the critical ratio value obtained is |1.652| < 1.968 accepted. It concluded that tourist satisfaction has no significant effect on tourist loyalty. The standardized regression coefficient value obtained is 0.062. If tourist satisfaction increases by 1 level, it will contribute to tourist loyalty by 0.062 times.

In general, tourist satisfaction will affect tourist loyalty, but tourist satisfaction does not affect tourist loyalty in this study. It may reflect tourists who are satisfied but do not want to visit again because some family members or friends do not like the games or activities presented at Kusuma Agrowisata. It could also be that their main goal is not for vacation but to attend a meeting with family or friends.

Effect of digital loyalty program on tourist loyalty through tourist satisfaction

Hypothesis 6 (DLPàKPàLP) the critical ratio value obtained is|0.6354| < 1.968 accepted. It concluded that the digital loyalty program has no indirect effect on tourist loyalty through tourist satisfaction. The standardized regression coefficient value obtained is 0.003. If

the digital loyalty program increases by 1 level, the contribution to tourist loyalty increases 0.003 times.

Based on the results, the digital loyalty program has no significant effect on tourist loyalty through tourist satisfaction. It indicates that tourists have high loyalty to Kusuma Agrowisata. It has proven that tourist satisfaction has failed to mediate the effect of digital loyalty programs on tourist loyalty.

The effect of service quality on tourist loyalty through tourist satisfaction

Hypothesis 7 (*Service quality*àKepuasanàLoyalitas) the critical ratio value obtained is |1.330| < 1.968 accepted. It concluded that there is no indirect effect of service quality on tourist loyalty through tourist satisfaction. The standardized regression coefficient value obtained is 0.012. If service quality increases by 1 level, the contribution to tourist loyalty increases by 0.012 times.

Based on the results obtained that service quality has no significant effect on tourist loyalty through tourist satisfaction, tourists have high loyalty to Kusuma Agrowisata. They have proven that tourist satisfaction has failed to mediate the effect of service quality on tourist loyalty.

Net Promoter Score (NPS) Analysis

The Net Promoter Score is used to answer the fifth research objective, knowing the level of tourist loyalty to Kusuma Agrowisata. The classification results (promoter, Passive, Detractor) from Kusuma Agrowisata tourists are presented in Tables 12 and 13.

Table 12. Indicators recommend tours to others

Type	Num of respondent	Percentage (%)	NPS
Promotor	104	34%	
Passive	149	49%	34% -
Detractor	51	17%	17% = 17%
Total	304	100%	

Table 13. Return visit indicator

Type	Num of respondent	Percentage (%)	NPS
Promotor	100	33%	
Passive	166	55%	33% -
Detractor	38	13%	13% = 20%
Total	304	100%	

The classification results of Promoter, Passive, Detractor of the indicators as factors to recommend tourism to others are as follow. From the results obtained, visitors are dominated by passive type with a value of 49% on the classification results (promoter, Passive, Detractor) of the indicators of the desire to visit again. From the results obtained, most of the visitors are passive type by 55%. From the two indicators, it is dominated by visitors with the passive type. It concludes that visitors have a sense of satisfaction with the services provided by the tourist but may not be willing to recommend tourist objects or return to Kusuma Agrowisata. It also has the potential to move to other tourist attractions that are considered better. However, it is better for business actors always to maintain the quality of service and loyalty programs

provided. It so that visitors with the passive type will decrease and are more likely to become promoter-type visitors.

The NPS value from the two indicators is high when compared to other tourist attractions. For example, the Goa Selomangleng tourism in Kediri has an NPS value of -13% [12]. It can be concluded that, in general, the loyalty level of Kusuma Agrowisata visitors is likely willing to recommend Kusuma Agrowisata to others and willing to visit again. Hence, it is less likely to switch to other tourist attractions.

5. Conclusion

Based on the analysis and discussion in the previous chapter, several conclusions were obtained as follows:

1. The digital loyalty program variable has no significant effect on tourist satisfaction. It can be interpreted that tourist satisfaction has not been reflected by the loyalty program provided. In contrast, the service quality variable has a significant positive effect on tourist satisfaction. It can be interpreted that there is a relationship between service quality and tourist satisfaction which shows tourists have been satisfied with the service quality of Kusuma Agrowisata.

2. The digital loyalty program and service quality variables significantly affect tourist loyalty, so it can be interpreted that the loyalty program provided benefits to Kusuma Agrowisata tourists. The service quality provided by Kusuma Agrowisata has been said to be excellent and successful in creating tourist loyalty.

3. Tourist satisfaction variable has no significant positive effect on tourist loyalty. It is because satisfied visitors may not want to visit again. After all, some people do not like the games or activities presented at Kusuma Agrowisata. The primary purpose of visitors coming is not for holidays but to attend meetings with family or friends.

4. Service quality and digital loyalty programs have no indirect effect on tourist loyalty through tourist satisfaction. Even though Kusuma Agrowisata provides good service quality and rewards in a complete loyalty program, in contrast, tourist satisfaction increases, tourist loyalty does not necessarily increase.

The loyalty level of Kusuma Agrowisata tourists during the Covid-19 pandemic was 17% for the desire to recommend tourism to others and 20% for the indicator of the desire to revisit. It concluded that Kusuma Agrowisata tourists are generally willing to recommend Kusuma Agrowisata to others and are willing to make repeat visits so that Kusuma Agrowisata tourists are less likely to switch to other tourist attractions.

References

1. Anggraeni, D. M. (2019). Pengaruh pelayanan dan promosi terhadap minat kunjung kembali wisatawan dengan kepuasan pengunjung sebagai variabel intervening pada wisata pemancingan dewandaru kecamatan gunung pati Semarang. In Αγαη (Vol. 8, Issue 5).
2. Asmara, Y., & Ratnasari, V. (2016). Analisis Kepuasan dan Loyalitas Pengunjung Terhadap Pelayanan di Kawasan Wisata Goa Selomangleng Kota Kediri dengan Pendekatan Structural Equation Modeling. Sains Dan Seni Its, 5(2), 181–186. https://media.neliti.com/media/publications/132214-ID-analisis-kepuasan-dan-loyalitas-pengunju.pdf
3. Haryono, S., & Wardoyo, P. (2015). Structural Equation. In Dictionary of Statistics & Methodology. https://doi.org/10.4135/9781412983907.n1909
4. Helmi, S. (2020). The effect of customer experience and emotional value on the net promoter score : Middle-class millennial climber category The effect of customer experience and emotional value on

the net promoter score : Middle-class millennial climber category. November. https://doi.org/10.25105/jmpj.v13i2.5763
5. Hermawan, H. (2017). Pengaruh Daya Tarik Wisata, Keselamatan, Dan Sarana Wisata Terhadap Kepuasan Serta Dampaknya Terhadap Loyalitas Wisatawan : Studi Community Based Tourism di Gunung Api Purba Nglanggeran. Wahana Informasi Pariwisata : MEDIA WISATA, 15(1). https://doi.org/10.31219/osf.io/89hqd
6. Paludi, S. (2016). Ananlisis Pengaruh Electronic WOrd Of Mouth Terhadap Citra Destinasi, Kepuasan Wisatawan dan Loyalitas Destinasi"Perkampungan Budaya Betawi (PBB) Setu Babakan Jakarta Selatan.
7. Pratiwi, L. Y. (2020). Analisis Pengaruh Bauran Pemasaran, Loyalty Program, Intentions to Switch Terhadap Kepuasan Pelanggan dan Implikasinya Pada Loyalitas Pelanggan PT. Bank Negara Indonesia 46 Merauke. *Jurnal Penelitian Dan Pendidikan IPS*, *13*(1), 22–31. https://doi.org/10.21067/jppi.v13i1.4752
8. Putra, H. A. (2017). Pengaruh Kualitas Pelayanan terhadap Loyalitas Konsumen Melalui Kepuasan Konsumen sebagai Variabel Intervening. Diponegoro *Journal of Social and Political Science*, 10, 1–8. http://ejournal-s1.undip.ac.id/index.php/
9. Salim, A. K. (2011). Pengaruh Keadilan Prosedural pada Komitmen Afektif Discretionary Service Behavior, dan Kepuasan atas Layanan.
10. Utami, D. M. (2019). Pengaruh Program Loyalitas Terhadap Kepuasan Pelanggan (Studi Empiris Pada Alfamart). 1–25.

The Impact of Gender Inequalities in IT and STEM for the Sustainability of National Social - Economic Security in the Context of the Digital Economy

Diana Ignatiuc

Academy of Economic Studies of Moldova, Chisinau, Moldova
ignatiuc_diana@yahoo.com

Abstract. The aim of this research is to study the impact of the gender gap in the ICT and STEM sectors on the sustainability of socio-economic security in the conditions of digitalization of the economy and its associated transformations.

Keywords. Digital economy, social - economic security, STEM, information and communication technologies, social - economic risks, social - economic threats

1. Introduction

Women represent more than half of the total global population, but they continue to be under-represented in economic activity, which has a significant macroeconomic impact on the sustainable development and socio-economic security of the world's countries. At the same time, the accelerated development of information and communication technologies, which we have witnessed in recent decades, highlights the fact that the digital economy has become a reality.

The digital economy is characterized by the transformation of the ways of creating added values based in particular on the establishment and systematization of digital contents (which have become the object of economic activity), on increasing the intellectualization of algorithms for their automatic processing without human intervention and taking into account the increasing complexity of correlations within the real economy). All this contributes to increasing the productivity of the factors of production and to reducing the transactional expenses within the national economies.

Thus are created the necessary conditions for capitalizing the competitive advantages of national economies as well as the conditions for the accelerated development of enterprises and economic sectors too. But at the same time, the digital transformation of the economy, accompanied by the automation of industries and services, and by the spread of information and communication technologies, is generating new risks and challenges for the sustainability of national social - economic security. The most obvious in this context could be the risk of rising unemployment, inequality and social exclusion.

2. Definition Of Digital Economy

The concept of "digital economy" is widely used in theory and practice, but there is still no unanimously accepted definition of it. It appeared in the last decade of the twentieth century. In 1995 Nicholas Negroponte used a metaphor according to which in his activity humanity goes through the transition from atomic transformation to bit processing, thus noticing the advantages of the new economy. The term "digital economy" is a generic term. It is used to describe the market relations specific to the new economy, which focuses on digital

technologies and reflects the transition from the third industrial revolution to Industry 4.0, which means the replacement of analog electronic and mechanical devices at the end of the twentieth century. According to the World Bank's definition, the digital economy is a system of economic, social and cultural relations based on the use of ICT.

The relevance of researching fundamental and applied issues regarding the digital economy lies both in increasing the scope of communications through social networks and in the accelerated development of digital platforms and tools used in a growing variety of socio-economic activities.

3. Current Digital Economy Characteristics, Indicators And Factors

In the digital economy, the Internet offers major and unlimited opportunities for personal development and value creation.

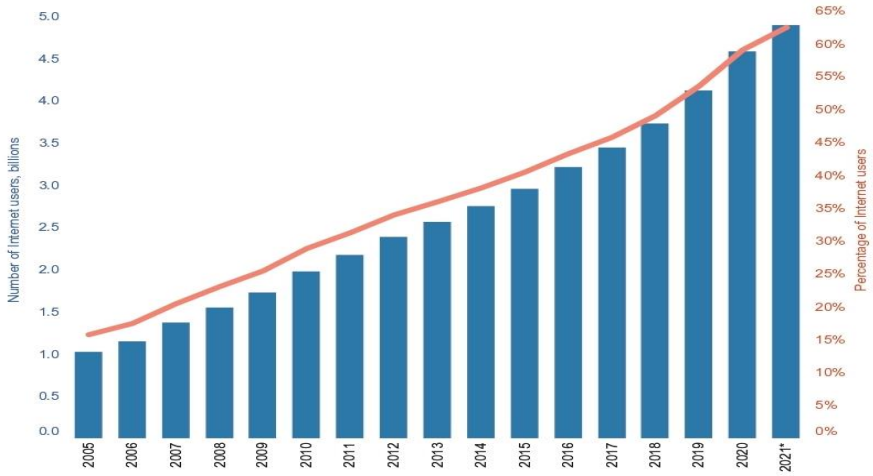
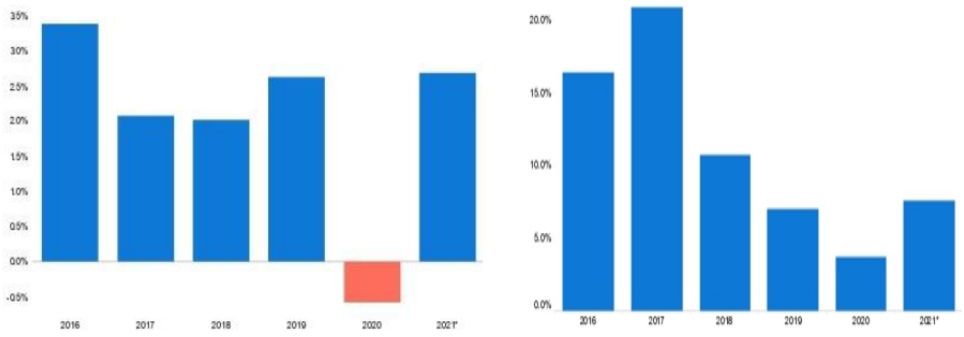


Fig. 1. Dynamics of the number and share of Internet users globally (Source: [1, p.7])

With the COVID-19 pandemic, it became a vital necessity to work, learn, access basic services and communicate. According to data from the International Telecommunication Union (ITU) in 2019, 4.1 billion people (or 54% of the world's population) accessed the Internet. In the last two years, the number of users has increased by 800 million, reaching 4.9 billion people in 2021, or 63% of the population. At the same time, we can conclude that about 2.9 billion people remain offline, 96% of whom live in developing countries. In addition, approximately 390 million people are not even covered by a mobile broadband signal (Figure 1). In 2020, the first year of the pandemic, the number of internet users increased by 10.2%, the largest increase in a decade. This increase is due to developing countries, where internet use has increased by 13.3%. [1]

The digital economy involves the widespread use of mobile telephony. Thus, the using of the fixed telephony is constantly declining in recent years, constituting 11 subscriptions per 100 inhabitants globally, compared to 19 per cent in 2006. At the same time, after a small decrease in 2020, the penetration of mobile cellular subscriptions worldwide it has risen again in 2021, reaching a record 110 subscriptions per 100 inhabitants (Figure 2). [1, pp.12-15]



Global dynamics of the number of mobile subscriptions per 100 inhabitants

Global dynamics of the number of active broadband connections per 100 inhabitants

Fig.2. Dynamics of mobile services and active broadband connections (Source: [1, p.15])

Mobile subscriptions with broadband capacity (3G or better) have followed the same trend, reaching in 2021 83 subscriptions per 100 people (figure 3). [1, pp.13-15]

In most developing countries, mobile broadband (3G or higher) is the main - and often the only - way to connect to the Internet. Ninety-five percent of the world's population now has access to a mobile broadband network. Between 2015 and 2021, 4G network coverage has doubled to 88% globally. [1, p.16]

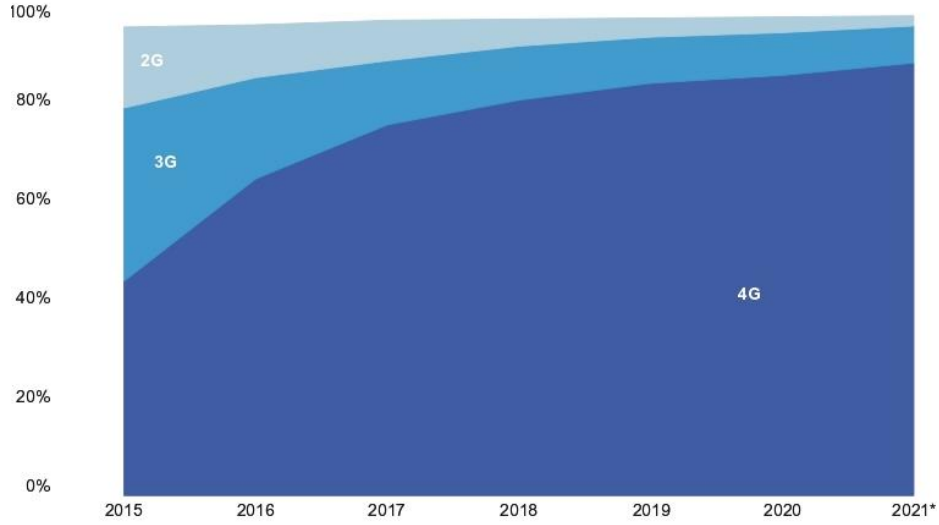


Fig. 3. Share of the population with access to mobile broadband networks in the world (Source: [1, p.15])

The analysis of the share of the population with access to mobile broadband networks by region shows a significant gap between economically developed and underdeveloped countries. According to Figure 4 in the group of developed countries 99% of the population

has access to mobile broadband networks, in developing countries this parameter reaches the value of 85% in 2021. At the same time, we can see that for the least developed countries developed countries (LDCs) and in landlocked developing countries (LLDCs) has not achieved Sustainable Development Goal 9 on - significantly increasing access to information and communication technology and working to provide universal access and at affordable prices in the least developed countries by 2020.

All this data shows the acceleration of digital transformations. The digital economy has become one of the key drivers of economic growth and sustainable development. It contributes to increasing the competitiveness of businesses, sectors and national economies as a whole, providing new opportunities to access markets both nationally and internationally regionally or globally.

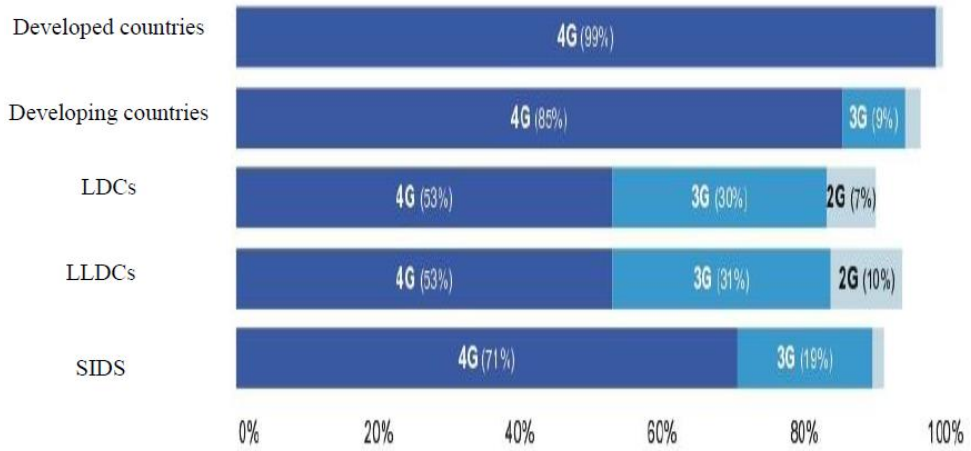


Fig. 4. Share of the population with access to mobile broadband networks in groups of countries (Source: [1, p.15])

The main defining characteristics of the digital economy are:

- the knowledge and people who possess this knowledge are the driving force of sustainable social and economic development (the priority role of the individual in the creation and control of digital content);
- the information is the object of work and key resource of economic activities;
- cooperation between economic agents has them through digital holistic networks as a special type of business networking;
- integrated digital platforms are the foundation of contemporary socio-economic systems and business transformation;
- artificial intelligence - a virtually unlimited source of combinations of means of production and relations of production.

The digital economy involves the creation of new markets, new market laws, new patterns of behavior for both producers and consumers, new types of money, new distribution networks.

The digital economy significantly changes the whole process of socio-economic reproduction, transforming the relationships that develop into the system of production, distribution, exchange and consumption of goods and services. Thus, for example, the implementation of the intelligent management system (EMI) in enterprises aims to reduce the impact of the human factor on the results of economic activity, it involves, first of all, the

transition from resource planning in the ERP (Enterprise Resource Planning) paradigm to management direct business processes by creating a single, multifunctional database and using them in a long-term perspective, implementing the principle of self-organization, automation of economic processes, execution of production and sales operations without the participation and / or involvement of staff [2].

4. The Digital Economy - Social Security - Economic Correlation

In the conditions of the digital economy, the profit of the economic agents is determined by the ability to process efficiently and quickly a large series of data in digital format. Thus, in essence, the digital economy involves the transformation of available information resources and technologies into assets, which can be used to significantly reduce transaction costs and increase the speed and efficiency of economic processes.

In the conditions of the digital economy, *social-economic security represents the totality of current conditions and factors that characterize the efficiency, stability and sustainable development of an organic national economy which is organically integrated in the international economic circuit, and at the same time, provides the necessary conditions for capitalize the available potential for self-development and sustainable progress.* Thus, ensuring the sustainability of national social and economic security requires the consistent implementation of policies and strategies to mitigate global risks and threats that are considered major challenges for the sustainable development of the world's countries. Among the major sources of risks and threats in the context of the digital economy we can highlight the lack or limitation of basic digital skills among citizens. Digital skills are the backbone of the digital economy and society, without which digital technologies cannot be fully benefited. Thus, according to the report on the digitalisation of the economy published annually by the EU Commission for Industry, Research and Energy, although in 2019 before the COVID-19 crisis about 85% of citizens used the internet, only 58% of them owned at least basic digital skills. The pandemic crisis has had a positive impact on the growth of the number of internet users but not on the development of digital skills. [3]

5. The Role Of The Gender Equalities In IT and STEM For The Sustainability Of National Social - Economic Security

These risks and threats for sustainability of national social-economic security are exacerbated by gender disparities, which, despite efforts and policies to combat this phenomenon, continue to persist around the world. It is virtually impossible to capitalize on the available economic potential if all the talents and diversity of a society are not fully utilized. "*Gender equality creates extra jobs and generates increased productivity - a potential that needs to be realized at a time when we are preparing for the transition to a digital economy and facing demographic challenges*". [4, p.2]

In September 2015, 192 of the UN member states adopted the Declaration of the Sustainable Development Summit, committing itself to implementing the 2030 Agenda for Sustainable Development. Thus, by 2030, the world's countries will mobilize their efforts to eradicate all forms of poverty, combat inequality and address climate change issues, ensuring that no one is left behind. Gender inequality remains a key challenge for sustainable development. Gender equality is a fundamental right, a common value promoted by both the European Union and the global international community. This is a necessary condition for ensuring the sustainability of social and economic security.

Women make up just over half of the global population, but they continue to be under-represented in economic activity. In the digital economy, the gender gap has a significant

impact on the sustainability of social and economic security, especially due to the ubiquity of information and communication technologies (ICT) in all areas of activity, which was accentuated during the COVID - 19.

Globally, in 2020, 62% of men are internet users, compared to 57% of women. Gender parity is considered to be achieved when the gender parity rate, defined as the percentage of women - internet users divided by the percentage of men - users, is between 0.98 and 1.02. (figure 5). Thus, the global gender parity rate increased from 0.89 in 2018 to 0.92 in 2020. [1, p.31] Although there is a significant increase in the number of Internet users among women, the risks and threats to basic digital skills and gender stereotypes remain high and have a significant impact on the sustainability of the socio-economic security of the world's countries.

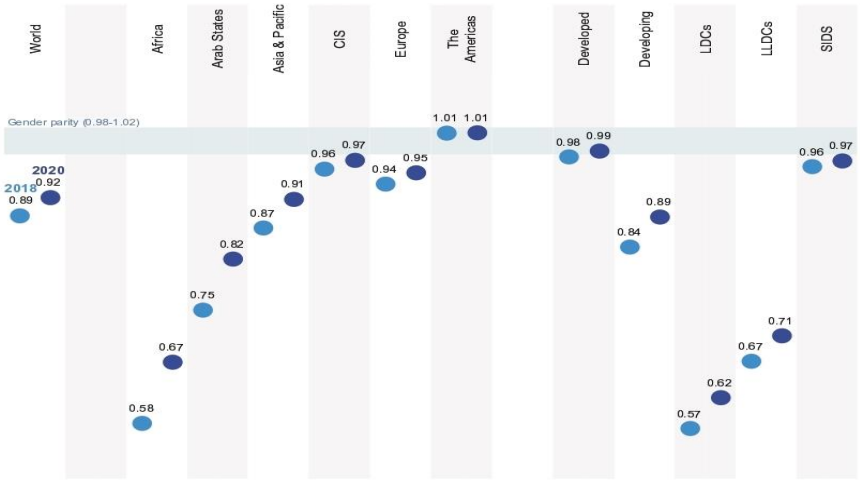


Fig. 5. Gender parity rate of Internet users, 2018, 2020 (Source: [1, p.10])

Therefore, women need to participate in its development as researchers, programmers and users. In this regard, policies and strategies to increase the number of women and girls in priority areas for the digital economy such as ICT and STEM are recommended.

According to recent estimates, more women graduate from university, they remain underrepresented in higher paid professions. There are more women than men in jobs and working in low-paid sectors and in lower positions. Factors that promote this phenomenon include discriminatory social norms and stereotypes about women's and men's skills and the underestimation of women's work.[4, p. 11] In the context of the digitalization of the economy, there is also a rapid transformation of the structure of labor demand, currently 90% of jobs require basic digital skills. At European level, women make up only 17% of those who study and have a career in ICT and only 36% of graduates in STEM. [4, p.12]

6. Gender Equality in the IT and STEM Sectors in the Republic of Moldova

In the Republic of Moldova, according to the analytical study Women and Men in the Information and Communication Technology sector developed within the project "Strengthening the national statistical system" implemented by the National Bureau of Statistics with the support of the United Nations Development Program (UNDP), United Nations Gender Equality and Women's Empowerment (UN Women) and the Government of

Sweden, in high school (grades X-XII), study about 57% girls and 43% boys. However, the analysis of the career guidance options of high school graduates shows that, if 9.3% of boys would opt for a job in the field of IT, and 4.9% - in engineering, then of the girls participating in the study only about 1 % expressed interest in a career in these fields. The share of men graduating from higher education was about 39%, lower than the share of female graduates of 61.2%, but over 1/4 (26.4%) of all men, and only 4.6% of all women, who are studying in higher education, have chosen STEM (science, technology, engineering and mathematics) as their study profile. Women are underrepresented in higher scientific and technical education (STEM specialties) and constitute about 19.5% of the total number of students in this field (only 2 out of 10 students in this field are women).[5]

According to statistical data, in the Republic of Moldova over 20 thousand employees work in the ICT sector, which is over 3% of all employees in the country. Women are much less present than men in the ICT labour market. Thus, only 3 out of 10 jobs in the ICT sector are held by women, about 31%. Women are much less present than men as ICT specialists, accounting for only 19.4% of all specialists in this field. As a priority area in the digital economy, the ICT sector provides decent wages, but with the largest pay gaps between women and men. In the ICT sector as a whole, women earn 33% less than men, and this discrepancy is growing.[5]

Women are also under-represented in the field of ICT entrepreneurship. Out of 100 companies in this field, only 20 are led by women. At the same time, is accentuated the vulnerability for the sustainability of the national social-economic security of the Republic of Moldova generated by the gender discrepancy in areas, considered to be a priority in the context of the digital economy - IT and STEM.

7. Conclusion

From the above we can conclude that, in the context of the digital economy, the fields of ICT and STEM are becoming priority areas for the sustainable development of the countries of the world. At the same time, the international community emphasizes the need for the consistent implementation at national, regional and global level of one of the main principles of democratic states - the principle of equal opportunities in all areas of social and economic activity. Countries will only be able to ensure the sustainability of national socio-economic security if it creates a level playing field for capitalizing on the potential, talents, skills of all members of society and its benefits will be shared equally for both women and men.

References

1. Measuring digital development. Facts and figures 2021. International Telecommunication Union, Geneva, 2021. (https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx)
2. V.V. Ivanov, G.G. Malinetsky. Digital Economy: Myths, Reality, Perspective. Moscow, 2017, 64 p. ISBN 978-5-906906-04-5
3. Digital Economy and Society Index 2020 Thematic chapters. European Commission. (https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1022)
4. Uniune a egalității: Strategia privind egalitatea de gen 2020 RO 2025. Comunicare a Comisiei către Parlamentul European, Consiliu, Comitetul Economic și Social European și Comitetul Regiunilor. (https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A52020DC0152)
5. Nicoară R., Vremiș M. Femeile și Bărbații în sectorul Tehnologiei informației și comunicațiilor. BNS, 2019, 135 p. (https://statistica.gov.md/category.php?l=ro&idc=264&#idc=429&)

National and International Practices of Innovative SME Financing

Daniela Dascaliuc

Academy of Economic Studies of Moldova, Chisinau, Moldova
danieladascaliuc@gmail.com

Abstract. In market economy, the small and medium-sized enterprise (SME) sector is an important segment of the national economy, which contributes significantly to the development of society. Its role is undeniable in stimulating production, growth and diversification, triggering innovation, improving quality and increasing competitiveness, creating the supply and meeting the demand for goods, creating new jobs. The development of small and medium-sized enterprises mainly depends on the company's ability to access and successfully gather its existing financial resources. This option gains even more importance when the lack of funding represents one of the basic problems that entrepreneurs face in trying to organize and expand their business to achieve high performance and to be competitive and innovative.

Keywords. Innovation, financing, small and medium-sized enterprise sector, financing programs, support instruments.

1. Introduction

In the framework of globalization and technologization, the small and medium-sized enterprise (SME) sector constitutes an important segment in the national economy, which contributes significantly to the socio-economic progress of the society. Small and medium-sized enterprises lead to the improvement of the economic situation, the stimulation of the production and quality, register a faster and easier manufacture adjustment to unique and small-lot items, lead to an increase in the new-job creation, trigger the development of technical and scientific progress.

Even though the main path of economic development has been the concentration of production, the versatility of the enterprise, the unification of branches within one unit - small enterprises have not vanished in the mists of time, but on the contrary, by virtue of their unique advantages to adapt to the permanently fluctuating economic conditions and the drastic conditions imposed by the COVID 19 Pandemic in recent years, not only have they survived, but not seldom have been at the forefront of technical, scientific and economic progress.

The organization of economic activities within small businesses is back on track today, taking an important place in national economies, becoming a pillar for social progress and a subject for public policy. Upon these points of view, it can be reckoned that the small and medium-sized enterprise sector is an essential component of the economy and plays a satellite role in the development of large companies. Large companies carry out a complex production process, which includes various industries from different branches, while small companies are narrowly specialized and usually render services to consortia by providing courier services, supplying spare parts and details needed to assemble units and machines, etc.

Materials and methods: The research carried out in this study has been conducted by applying empirical, historical, systemic, analytical, economic and other methods for studying the SME's financing practices.

2. Outcomes and discussions

Recently, there has been an increase in the role of the small and medium-sized enterprise sector in the innovation field. Although, it may seem difficult for small businesses to influence technological development, at present, in partnership with large companies and research centres, they are making remarkable innovations in the fields of information technology, engineering, medicine, pharmaceuticals, environment, etc. However, we must not neglect the state's role in boosting the development of small businesses and their involvement in promoting technical and scientific progress as well as in the economic and social fields.

Currently, in the Republic of Moldova, the state intervenes in the development of the small and medium-sized enterprise sector through the public policies it adopts and promotes, through the facilities it offers to this sector, through bilateral agreements negotiated and concluded with international bodies and through their promotion within various conferences, meetings, exhibitions at local and international level.

Recent studies and the evolution of public policies in the Republic of Moldova have highlighted the need for taking additional measures in encouraging small and medium-sized enterprises to adopt more innovative and greener business practices.

While SMEs generate around 64% of industrial pollution in the European Union, only 16% of them are engaged in environmental actions [4]. Therefore, there is still great potential for the promotion of instruments and good practices to encourage the "greening" and "innovation" of SMEs in the country. Many of these instruments have already been implemented by the Member States of the European Union and can be adapted to the needs of the SMEs in the Republic of Moldova, thus encouraging the companies' and country's economy sustainable development and international economic competitiveness overall.

Statistical data coming from Europe shows that more and more European programs are focused on supporting SMEs that are geared towards innovation in the following areas: sustainable business development; green industry innovation; information and communication technology (ICT), and in Romania, for instance, the companies focused on sustainable development in the marine and maritime sector are also sustainable for financing.

In the conditions of globalization and international competition, the innovation development scenario has no alternative. Encouraging innovation has now become an extremely important requirement for all countries' policies engaged in the building of the knowledge-based society. It has become an important component of modern socio-economic and technological policies, being defined as the state's innovation policy, which shows the state's attitude towards innovation, defines the objectives, directions, activity types of public authorities in the scientific field, technology and the implementation of scientific and technological findings. It represents a system of institutional, normative-legislative, economic-financial and other instruments, meant to boost all the stages of the innovation activity by creating favourable conditions. Their creation, as a rule, begins with the establishment and ratification of the normative-legislative framework adequate to the functioning of an innovative climate favourable for entrepreneurial activity.

The biggest challenge for the EU and the Member States is to take a far more strategic approach to innovation in all its aspects (product, process, organizational and marketing innovation). Specifically, in order to achieve the Innovation Union, it is necessary to improve cooperation between the scientific and business environments, to remove obstacles and barriers from market research findings and to put into practice specific incentives including innovative commercialization.

According to point 2 of the Methodology for financing projects in the research and innovative fields, [2], in the Republic of Moldova, innovation and technology transfer projects

can be financed from the state budget for 2020-2023, in the amount of approximately 8 million lei per year. For 2021, 8 million lei were budgeted in the following sectors:

- Health 1413.3 thousand lei;
- Sustainable agriculture, food security and food safety 1573.3 thousand lei;
- Environment and climate change 1920.0 thousand lei;
- Societal challenges 1253.4 thousand lei;
- Economic competitiveness and innovative technologies 1840.0 thousand lei (Source: ancd.gov.md.)

Concurrently, the information presented by the National Institute of Statistics from Romania, reports that large companies are more innovative (17.9%), compared to medium-sized (11.9%) and (9.3%) small ones [5], a trend also witnessed in the industry and service sectors in the Republic of Moldova.

According to the ranking of the global innovation index in 2020, the Republic of Moldova ranked 59th out of 131 countries and economies, receiving 32.98 points out of 100 possible. For the sub-index innovation inputs, our country ranked 75th, and for the sub-index innovation outputs, it ranked 48th. In terms of infrastructure, we are ranked 88th, human capital and research - 75th place, creative outputs - 51st, market sophistication - 42nd place, business sophistication - 88th place, institutions - 81st place, and for the knowledge and technological results indicator - 51st place. The Republic of Moldova ranked 6th globally among the countries with lower-than-average incomes. It should be noted that Switzerland leads the ranking of the global innovation index, maintaining its position since 2011. It is followed by Sweden, the United States, the United Kingdom and the Netherlands. In 2020, the Republic of Korea appears for the first time in the top 10 most innovative countries. Ukraine is on 45th place, while Romania is on 46th place. The Global Innovation Index is one of the most important reference studies for measuring the country performance in terms of innovation. 131 countries from all around the world have been researched, using approximately 80 indicators, including: human and research capital, infrastructure, loans, investments, interconnections, innovation, outcomes of the creative activity, etc. [6].

In the last 3 years, the state has contributed to the development and implementation of the following programs for the support of small and medium-sized enterprises focused on innovation:

- The Greening Program for Small and Medium-sized Enterprises, approved by the Government Decision No. 592 of 27.11.2019, implemented by the Organization for the Development of the Small and Medium-sized Enterprise Sector, in order to promote, support and develop the capacity of SMEs for the adoption of greening practices in manufacturing processes and service provision. Throughout this program, SMEs can benefit from: training and education in the greening field; financing of greening actions in a small amount of up to 200 000 lei or in a large amount of up to 500 000 lei;
- The backup instrument regarding the digitalization of SMEs, approved by the Order of the Ministry of Economy and Infrastructure No. 100 of 26.05.2020, implemented by ODIMM. The purpose of the instrument is to support the digital transformation and development of SMEs in order to exploit their innovative potential and facilitate their operation in internal and external markets. This tool involves supporting small entrepreneurs through: entrepreneurship training and education; granting financial support by means of a "business voucher" in the amount of up to 20 0000 lei and / or through a grant in the amount of up to 200 000 lei.
- For SMEs that are, planning to export outside the country, to expand on foreign markets and / or to replace product / service imports in the Republic of Moldova, on 01.07.2020, upon Government Decision No.439, the Government of the Republic of Moldova approved the support Program for high growth potential businesses and their internationalization. This

program includes the non-reimbursable financing component by means of a Business Voucher worth up to 50 thousand lei and by means of a grant worth up to 2 million lei. The instrument also provides informational and consultancy support.

All programs and instruments for SMEs, which are non-essentially supported by the state, offer a higher score for the innovative businesses which take part in the competition, but this indicator is not decisive.

In addition to these support instruments, SMEs in the Republic of Moldova can benefit from funding within the applications submitted by the Ministry of Agriculture through the Agency for Payments and Intervention in Agriculture. They can also apply to the bank crediting mechanism, but in order to receive the state support for guaranteeing loans, they can apply to the Guarantee Fund. Domestic SMEs can benefit from financial, technical support and assistance within the Projects financed by the European Union and other foreign donors in the implementation of innovative investment projects, etc.

The institutional framework for supporting small entrepreneurs in the Republic of Moldova is represented by the line ministries; local public administrations; SME support agencies and organizations, such as: the Moldovan Investment Agency, the Organization for Small and Medium-sized Enterprises (ODIMM), the Agency for Payments and Intervention in Agriculture (AIPA), the Chamber of Commerce and Industry of the Republic of Moldova, the State Agency on Intellectual Property (AGPI), Business Incubators; International support organizations; Embassies of European states, etc.

The main role in providing support to small entrepreneurs in the Republic of Moldova belongs to the Organization for the Development of Small and Medium Enterprises (ODIMM), which is currently managing 12 Business Support Programs, through which beneficiaries are offered advice, training, financing and mentoring. (www.odimm.md)

The analysis of the submission documentation for financing projects and of the conditions for contracting state support has resulted in certain **findings**:

a) Over the years, a specific way of applying for financing and financial support for SMEs has emerged, which practically does not differ from one Program / Project to another. The support that SMEs can obtain is in the form of grants, non-reimbursable financing, consultancy and mentoring, and credit guarantee. Small and medium-sized companies can also benefit from such facilities as discounts on the space rental price within business incubators. We emphasize that the non-reimbursable financial support that small and medium-sized companies from the Republic of Moldova can benefit from is very modest compared to the one that SMEs from abroad can obtain. According to the legislation in force (Law No. 139 of 15.06.2012 on state aid), the amount of de Minimis aid granted to an enterprise within 3 consecutive years must not exceed the threshold of 2 million lei. Consequently, the amount of money that can be obtained by an economic entity focused on innovative activities, is often very small to start and develop a business. For example, in the European Union, innovative SMEs can benefit from grants of up to € 2.5 million [7].

b) SMEs that are not eligible for non-reimbursable financial support programs can contract bank loans and benefit from state guarantees under the Guarantee Fund. This mechanism is suitable when no collateral is pledged, but this burdens the activity of economic entities with an additional payment of 0.5% of the guaranteed amount [8]. At the same time, the guarantee procedure does not imply the simplification of the documentation package that must be submitted to the bank for contracting the loan and the total lack of collateral pledged by the loan applicant.

c) The terms and application forms for support programs / projects are similar, varying only depending on the specifics of the implementing organization. In turn, the eligibility criteria vary only depending on the area concerned, and the application forms (business plan)

are similar in content. The business plans / concept note for application to the support projects implemented by ODIMM lack the indicators which reflect the profitability of the entity's activity and the investment recovery period - indicators that are strictly necessary in evaluating a business and creating the right business planning. There is also a lack of information about the risk degree and the causes that lead to its multiplication. The evaluation criteria of the projects submitted for funding are not clearly identified, often the evaluation grid is missing or its content is non-essential.

d) There is no definite program delimitation by eligibility areas, for example, agricultural companies can benefit from support for the procurement of agricultural equipment both on behalf of the projects implemented by AIPA and by ODIMM, the difference lying in the amounts which may be contracted or the purchase period of the investment items.

e) Although all projects and support programs are developed and implemented with the purpose of solving a certain problem, it can still be noticed that there are shortcomings in terms of measuring their impact on the economy. The impact indicators to be achieved are not clearly defined and there is no well-established monitoring mechanism. In turn, entrepreneurs are not always responsible for the actions they take and the goals they set for themselves.

f) The support in the form of training and mentoring often does not meet the entrepreneurs' requirements or their expectations. This is determined by the lack of specialists and experts in the field who could advise and guide entrepreneurs to an efficient business management and the implementation of innovations.

g) Digital transformation requires systemic changes in business processes, business models and economic relationships within and outside the company. The creation of an environment for the SMEs' digital transformation that would operate in the traditional sectors of the economy should provide a range of specialized technological and business consultancy services, which can be offered by centers of excellence, through the collaboration of the public and private sectors in initiatives at national level (e.g., skills development standards and common standards), as well as a comprehensive financial framework to support SMEs in this intricate endeavour.

h) Another apparent problem is that the best human resources are attracted by large companies or lured by the possibilities offered by the economies of other countries, the SMEs being left not with the best workforce, therefore these companies might not manage to provide innovation with their own employees, but they have to resort to the research and innovation outcomes of the academies and universities.

Based on the above findings, the following **conclusions** can be drawn:

- *The international experience in stimulating and promoting the SME sector* shows that in economically developed countries there is now a well-developed, innovative, organically integrated small and medium-enterprise sector into the national economies, with a significant share in GDP, the number of employees in the economy and a range unique products. The establishment of this sector has been possible thanks to the public authorities' awareness of the economic and social role of small and medium-sized enterprises and to the emphasis on promoting mechanisms that would stimulate innovative businesses. The significance of this support has considerably increased along with the deepening of the international financial crisis and the economic crisis triggered by the COVID 19 pandemic.
- *The support for the SME sector in the Republic of Moldova is fundamental* and contributes to the promotion of new innovative initiatives. Innovation promotion and support in our country could lead to the overcome of the economic crisis from the last decade and a faster integration of the SME sector into the business cycle.
- *The lack of an innovation management system* in SMEs in the Republic of Moldova is worsened by the lack of measures for promoting innovation within entities which entail the

loss of benefits. This results in many difficulties at the concept stage and in the process of developing innovations, and these, in turn, lead to the early death of the project or idea.

- The adaptability of the SME sector to the new working conditions is crucial. Consequently, the *awareness of the need to create and implement innovations* ensures their existence in the future.

The conclusions reached include the following **recommendations**:

1) Continuous insurance of state support for the innovative small and medium -sized enterprise sector through non-reimbursable financing, with the possibility to increase the minimis ceiling by at least twice (minimum 4 million lei), fiscal and financial facilities, including the active promotion of innovations nationally and internationally. In this regard, it is recommended to integrate the innovative domestic SMEs into the regional innovation ecosystem by faster joining the European Enterprise Network and receiving support from the European Innovation Council Fund.

2) In order to ensure a higher efficiency for the SME financing programs in the Republic of Moldova, it is recommended to modify the application forms for support programs / projects by the mandatory introduction of the indicators for efficiency, profitability and recoverability of investments supported both by state and donors, investors, as well as by SMEs. Respectively, when developing new financing and support projects for SMEs, the counting of the above-mentioned indicators should be included in the business plan or concept note, and in the evaluation grid, the maximum score should be given to investments in innovation. At the same time, it is recommended to introduce new requirements, thus contributing to the accountability of the applicants, their more detailed argumentation of the support needs for innovations and the deepening of the local entrepreneurs' knowledge and managerial skills. At the same time, the institutions implementing SME support programs are recommended to improve the communication strategy with SMEs and the selection procedures of the final beneficiaries.

3) As a result of the conducted analyses, it has been noticed that the state does not keep an accurate record of the impact of the implemented programs. In this regard, it is recommended to create a single platform for recording all economic agents from the Republic of Moldova, that would offer access to information on the activity field, the entity founders, turnover, investment volume, number of employees, indebtedness level by categories of creditors (banks), the amount of grants received and facilities granted (by categories of financiers), the amount of profits obtained. It is recommended to facilitate secure access to this platform for central and local public authorities, the State Tax Inspectorate (IFS), the Public Service Agency (ASP), banks, organizations and public institutions for financing entrepreneurs. This way, transparency will be ensured in channeling public financial resources and those from foreign donors, thus leading to a reduction in the periods for data collection and analysis of applications. Yet it will be possible to analyze the impact of the resources obtained in the form of support for the beneficiaries' activity in particular and of the economy as a whole.

4) As SMEs are undercapitalized, they cannot do research and innovation, so they have to borrow, which implies extremely much effort, often, without any success, thus we are pointing out that this aspect needs correction. As regards this, it is recommended that Moldovan institutions, including research institutions, should provide information to domestic SMEs and create a favorable framework for them to access budget lines which would target European funding for innovation.

5) Increasing the management capacities of innovative SMEs through various programs and support projects, which would be based on the audit of their activity, elaboration of action plans adapted for these SMEs, including the mapping of the innovation process from idea to

result, as well as leadership, management and cooperation strategies. The action plans provide an in-depth analysis of the innovation management from these companies, investigating possible gaps between innovation capacities and innovation objectives.

6) Integration of SMEs into clusters and value chains. To be able to innovate the SMEs from the Republic of Moldova, clusters should be formed, which would include companies from different fields, not necessarily only from the same activity area, and jointly create innovation and research centers, a structure which already exists in some European states. On the one hand, the SMEs from the Republic of Moldova need more investments in these clusters. On the other hand, SMEs need certain national research and innovation programs. Last but not least, the research carried out by universities and profile institutes must be in agreement with the practical requirement, i.e., to be applicable.

7) The creation of Accelerators in the Republic of Moldova, which will contribute to the support of innovative SMEs, inclusive of consulting and financial support, and to the digital transformation of entities operating in the traditional sectors of the economy. Many small and medium-sized enterprises in our country, which mainly belong to micro and small enterprises, lack consultancy services due to their cost. These companies also need quality consulting, which they often cannot afford, because of its high costs. Entrepreneurs in the Republic of Moldova must be aware of the need for SMEs to allocate a budget for consultancy to manage to take advantage of all the opportunities that arise, therefore it is required to provide support for entrepreneurial consultancy. SMEs are at a disadvantage compared to large companies, at European level as well, having a restricted and limited power in negotiations with these. All these phenomena make SMEs leave innovation in the last place, thinking about the present day, but normally they have to access research, innovate and bring in new technologies to ensure sustainability and longevity as a company.

8) The creation of business incubators within higher education institutions that will represent a symbiosis between the business environment, the academies, the specialized and vocational education institutions focused on innovation and technologizing. It is recommended that universities and research institutes, which have, and have had, sufficiently large funds for research from the state and EU budgets, including SME funding organizations, be focused on partnerships with SMEs and carry out applied research, i.e., to identify what an SME from one field or another needs, so that, based on the identification of financing needs and areas of applicability, to create sustainable partnerships between several players. This would create a value chain and the money used for research, either within the research institutes or universities we already have, should finally be conditioned by a partnership with SMEs and by a research result applied directly to these companies. Solutions can be found and it is certain that, now, if we make a comparison with other European countries, Moldovan SMEs are not generating innovations, because innovation costs. In order to innovate, you must possess human resources who are very good at it, but the best human resources have been attracted by large companies or have been lured by the possibilities offered by the economies of foreign countries, and thus SMEs are not left with the best workforce. Therefore, these companies have not been capable to generate innovations with their own employees, but are forced to resort to the research and innovation outcomes of the academies and universities.

Acknowledgements

This article is part of the project State Program ”Innovative business configuration in the context of regional competition”, no. 20.800020807.42

References

1. Legea Republicii Moldova cu privire la întreprinderile mici și mijlocii nr. 179 din 21.07.2016
2. Hotărârea Guvernului Republicii Moldova Nr.382 din 01.08. 2019 cu privire la aprobarea Metodologiei de finanțare a proiectelor din domeniile cercetării și inovării
3. Dascaliuc D. Managementul politicilor financiare publice de stimulare a businessului mic și mijlociu. Teză de doctor în economie. Chișinău 2009, 173p.
4. Environmental Policy Toolkit for SME Greening in EU Eastern Partnership Countries. OECD Green Growth Studies, 2018. https://read.oecd-ilibrary.org/environment/environmental-policy-toolkit-for-sme-greening-in-eu-eastern-partnership-countries_9789264293199-en#page1.
5. Grapini M. IMM-urile din România nu inovează, pentru că inovația costă. Disponibil: https://www.bursa.ro/maria-grapini-europarlamentar-imm-urile-din-romania-nu-inoveaza-pentru-ca-inovatia-costa-80970534
6. WIPO Prof- un nou instrument lansat de OMPI. In Revista INTELLECTUS nr. 3-4/2020, p. 14. Disponibil: http://www.agepi.md/ro/intellectus/intellectus-3-4-2020
7. Oportunități de finanțare a IMM. https://eic.ec.europa.eu/eic-funding-opportunities/eic-accelerator_en.
8. Programe de suport pentru IMM. https://www.odimm.md.

Challenges of Industry 4.0: Impact for the SME Sector of the Republic of Moldova

Ala Cotelnic, Irina Dorogaia

Academy of Economic Studies of Moldova, Chisinau, Moldova
cotelnic.ala@ase.md, dorogaia.irina.ion@ase.md

Abstract: The modern business environment is characterized by high urgency and uncertainty. The complexity of the operation of enterprises is exacerbated by the situation caused by the Pandemic. In this difficult time, information technology is increasingly penetrating into business and becoming a means and a way of existence. The pervasive impact of information technology, the emergence of artificial intelligence, the impact of Big Data - all of this is fundamentally changing the way businesses operate. Klaus Schwab calls this global process of transition Industrialization 4.0. This process has a serious impact on the SME sector in the Republic of Moldova. The importance of this sector for the economy of Moldova is extremely high, therefore, ensuring the normal functioning of such enterprises in the new conditions is a task of state importance. The impact of transformational changes on the small and medium-sized enterprise sector will be discussed in this article.

Keywords: change, challenges, Industry 4.0, small and medium enterprises, change management

1. Introduction

Constant and dynamic change is an inevitable phenomenon today. In the world of business, processes are beginning to be clearly traced, which until recently were even impossible to imagine. Today it is difficult to imagine our life without the Internet and information technology. These processes affect not only enterprises operating in the field of information technology, but also our whole life, leave an imprint on the way of life of people, relationships in society, and of course, place ever higher demands on the management system and the manager himself.

Science in terms of change also does not stand still. In recent years, many works have appeared that try to foresee the future, outline the trajectory of development, anticipate possible problems and develop mechanisms for their regulation. In this context, the studies of the world-famous organizer of international forums in Davos – Klaus Schwab, who proclaimed the term “Fourth Industrial Revolution” back in 2016, received great publicity. [1, 2]

This concept is increasingly permeating our lives. And if at first the German development plan was associated with the phenomenon of the fourth industrial revolution, today they are talking about a process that affects the whole world. If we use the definition of Thomas Siebel, such processes are called "digital transformation". It is important to note that digital transformation is understood, according to the author, “not a series of changes in the field of information technology and not the transfer of corporate processes, databases and information to a digital platform”, but “the fusion of four technologies: cloud computing, big data, the Internet of things and artificial intelligence”, and “not a single industry remains indifferent [3].

2. Definition of the Fourth Industrial Revolution

In order to delve into the essence of today's changes, Siebel suggests following the previous stages of the processes of transformational change.

So, ***the first wave of transformational processes*** (1980s) begins with the advent, or rather the spread of personal computers, and, as a result, the digitalization of calculations, tables and databases, which significantly increased productivity, which, in turn, led to a wave of economic growth (if in the period from 1989 to 1995 the growth of world GDP was 2.5% per year, then in the period 1995-2003 it increased to 3.5%) [3]

The second wave is associated with the ***spread of the Internet and the optimization of work processes***, which became possible thanks to automated management systems, accounting, and customer relationship management (CRM).

The third wave is ***a combination and influence of the factors of the first two waves***, which begin to radically change the principles of companies' activities. Many companies, realizing the importance and competitive advantage of information technology, have transferred their business to digital platforms. Business transformations, while significantly increasing productivity, did not lead to fundamental changes in work processes.

Today's transformational processes ***(fourth wave)*** are characterized by ***cardinality and irreversibility***. Researchers Eric Brynolfsson and Andrew McAfee call this era the "age of machines", in their opinion, "previously computers always followed well-defined instructions, but now they will be able to learn." [3]

In addition, today's changes are characterized by a high speed, which has not yet been inherent in such processes.

The transformation of society and business, which is developing at lightning speed, is characterized by the following components:

- development of new technologies that are inextricably linked with such concepts as:
  - artificial intelligence,
  - nano- and biotechnologies,
  - blockchain,
  - Internet of things,
  - neurotechnologies,
  - big data,
  - cloud computing,
  - robotization of most business processes;
- change in value orientations;
- reorientation of business models;
- rethinking management models. [1, 2, 3, 4].

It is important to understand that the Fourth Industrial Revolution, or Industry 4.0, affects all systems of society – economic, political and social processes. Also of paramount importance is the awareness of responsibility for the events taking place. Here is the most appropriate quote from Klaus Schwab himself: « The era of new technologies - if its development is guided with sensitivity and responsibility - will be the beginning of a new cultural renaissance that will allow us to feel part of a single whole: a truly global civilization…» [1]. The scientist is aware that the processes associated with the Fourth Industrial Revolution "will change the lives of many people, changing their usual work, living environment, possibly family relationships and even the identity of the person himself", he calls this the term "roboticization of humanity". But, at the same time, these processes “can lead humanity to a new collective and moral state based on a common sense of destiny. This is what we should strive for” [1].

In his research, Schwab asks how to "distribute the benefits of technological breakthroughs equitably, contain their inevitable negative effects, and ensure that new technologies will empower, not limit, the opportunities of all the inhabitants of the Earth"[1].

3. Research problem

Many researchers talk about the significant advantages of such processes in various fields of activity: medicine, namely, early diagnosis, accuracy of surgical interventions, elimination of the "human factor"; in the automotive industry, the introduction of disembodied transport, and, as a result, the reduction of many risks; in the field of production, given the speed and breadth of the process; under the influence of such changes, harmful emissions should also decrease, as well as the consumption of many natural resources. [3]

At the same time, it is impossible not to single out the negative impact of transformational changes. First of all, they will affect the working-age population. The replacement of jobs by machines is obvious, so the working-age population will need to reorient in order to acquire new skills, knowledge and skills. This is often difficult to implement. At the same time, today a large number of studies are reduced to the importance of acquiring not only hard skills (professional information skills), but also soft skills (creativity, the ability to adequately respond to various situations, teamwork).

Already in 2015, the World Economic Forum identified the main problems of the spread of the industrial Internet. The most important of these were incompatibility and lack of standards (worldwide 65%, of which North America 43%, Europe 30%), security problems (64% in total, of which North America 72%, Europe 60%) and uncertainty about return of investments - (total 53%, of which North America 53%, Europe 50%). Next in descending order are issues related to outdated equipment, immaturity of technology, privacy concerns, lack of skilled workers such as data scientists, and social issues. [1]

Considering all of the above, in our opinion, along with global transformations that bring many advantages, there is also a serious problem associated with poor adaptation to the emerging conditions of small and medium-sized enterprises in the Republic of Moldova.

Weak adaptability of small and medium-sized enterprises is associated with the following factors:

- lack of financial support,
- lack of qualified specialists,
- lack of development programs and strategic vision,
- lack of state support programs, etc.

At the same time, we note the significant advantages of such enterprises associated with:

- flexibility,
- adaptability ,
- less "ossified" organizational structures.

It is worth noting that in the Republic of Moldova, the majority of enterprises belong to the SME sector (at the end of 2020 - 98.6% of the total number of enterprises, 60.1% of SMEs from the total number of employees. This is evident from the analysis of the statistical data presented on the official website of the National Bureau of Statistics of the Republic of Moldova (see Fig. 1)

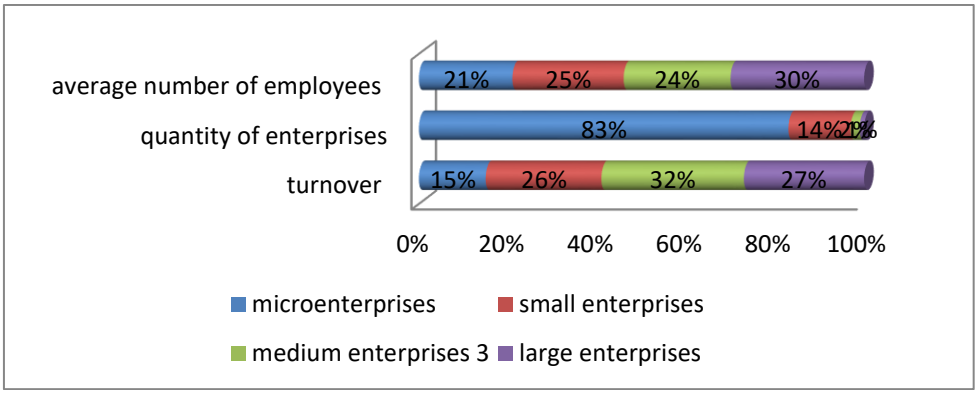


Fig.1. The share of SMEs in the economy of the Republic of Moldova [7]

Note that according to the legislation of the Republic of Moldova, Small and Medium Enterprises (SMEs) are enterprises: micro-enterprises - up to 9 employees inclusive annual turnover up to 9 million lei, or the amount of assets of the enterprise up to 9 million lei; small enterprises : from 10 to 49 employees inclusive, annual turnover up to 25 million lei, or the amount of assets of the enterprise up to 25 million lei; medium enterprises: from 50 to 249 employees inclusive, annual turnover up to 50 million lei, or the amount of assets of the enterprise up to 50 million lei.

Small enterprises are of great importance in the economy of any country, considering the indicators of diagram 1, we can say that for the Republic of Moldova - paramount.

If function successfully, they provide: production of a significant part of products and services; expansion of individual economic opportunities; creation of new jobs; lower needs in terms of institutional formation and market activity; social stability.

Given the dominant importance of this type of enterprise for the economy of the republic, we will analyze the main performance indicators for the period 2014-2020. Basic data are presented in the table.

Table 1. The main indicators of the activity of enterprises by economic activities, 2014-2020 [8]

Indicator	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Number of reporting units	33 027	33 351	33 031	33 718	35 372	36 299	37 228
Average annual number of employees, persons	369 159	368 733	364 721	372 089	387 205	396 332	385 825
Turnover, million lei	235 309,1	252 836,0	262 873,9	290 079,4	310 556,2	344 834	332 607,6
Production value (works and services), million lei	121 373,9	137 423,0	141 492,9	155 769,7	165 333,4	178 600	173 118,3

As can be seen from the table, the number of enterprises increased during the analyzed period, at the same time, the trends in the number of employees were variable - from 2014 to 2016 - had a downward trend, from 2016 to 2019 - increased, and from 2019 to 2020 - decreased. As for the turnover of small and medium-sized enterprises - they increased from 2014 to 2019, in 2020 they decreased, which may be due to the crisis caused by the pandemic. The number of manufactured products and services in monetary terms has the same trend as turnover - from 2014 to 2019 - an increasing trend, from 2019 to 2020 - a decreasing one.

The data presented in the table indicate unstable trends, lack of stable growth and confidence in the future. In such conditions of functioning, it is difficult to make changes, adapt to new trends and develop in accordance with clearly defined plans.

The low innovative activity and the ability to initiate and implement changes are also evidenced by the data of a study conducted by the National Bureau of Statistics in 2021 on the activity of enterprises in the period 2015-2020.

Table 2: Innovative enterprises by economic activities and size classes, 2015-2020 [9]

Period	Total innovative enterprises	Small enterprises	Medium enterprises	Large enterprises
2015-2016	673	418	191	64
2017-2018	605	378	163	64
2019-2020	448	255	134	59

The table shows that the level of innovation activity has dropped sharply over the analyzed period, especially for small businesses - by 49% compared to the level of 2015-2016. The level of innovative activity of medium-sized enterprises decreased by 30% over the analyzed period, the situation in the sector of large enterprises changed least of all - this indicator decreased by 8%. Thus, the problem of uncertainty and the lack of sustainable positive dynamics, most SMEs are characterized by weak innovation activity, which has recently been declining even more. In addition, it is worth noting that the share of organizational innovation is negligible (based on the data of the same study. Data for small enterprises indicate 12% of the total number of innovations in 2015-2016, while in 2019-2020 this figure is 8 % in the total number of innovations. Medium-sized enterprises have similar data: 2015-2016 - 8% of organizational innovations to the total number of innovations, 2017-2018-13%, 2019-2020-7%.There is also a clear downward trend. Thus, we can judge that these enterprises, even being in crisis conditions, are not ready to resort to managerial or organizational innovations.

It is worth noting that the processes of transformation and digitalization, which are an integral part of the Industry 4.0 process, are primarily related to organizational processes, or, in other words, organizational innovations. This once again confirms the problem of the present study.

In addition to these problems, the problem associated with the qualifications of personnel is of great importance. Given the small number of employees in such enterprises, SMEs often lack the right level of specialists, both in the field of management and in the field of specialties related to modern technologies.

Of course, state support is important in such transformations, which consists not only in financing projects, but also in information support.

So, the main problems of SMEs that hinder their development and acceptance of provocations caused by Industrialization 4.0 are related to the lack of:

- sufficient funding,

- qualified labor force,
- appropriate qualifications,
- developed approaches to change,
- readiness to implement organizational innovations,
- necessary support from the state.

Given the crisis, exacerbated by the pandemic and its associated restrictions, the uncertainty and lack of a clear development plan on the part of a large number of SMEs further expose these enterprises to decline and inevitable liquidation. On the other hand, factors that are associated with transformation processes also push these companies to transform. The inevitability of these phenomena is obvious. Thus, SMEs are left to develop a vector for further development in the current extremely difficult conditions.

4. Problem Solving Possibilities

The fourth industrial revolution pushes society to solve the following tasks: to distribute the benefits of technological breakthroughs evenly and fairly, to contain the negative effects that may accompany them, and to ensure that new technologies do not limit the opportunities of the entire population.[1]

The author of the concept, Schwab, identifies four main principles that are important for a new level of leadership. Such a leader must overcome existing and potential problems, anticipate the situation and take into account possible risks. In our opinion, these principles can be the key to building an effective system in the new realities.

Table 3. Using the principles of Industrialization 4.0 for transformational changes in SMEs in the Republic of Moldova (adapted by the authors based on the results of studying the concept of Klaus Schwab and the realities of the activities of SMEs in Moldova)

The name of the principle, according to the concept of Schwab	Essence of principle	The possibility of applying the principle in the context of the transition of SMEs to the transformational changes of Industry 4.0
1. System, not technologies.	In the context of change, it is necessary to focus on a systems approach, which implies the joint interaction of all stakeholders: political systems, investments and joint work of all interested groups. Technologies, on the other hand, create the conditions for building new systems for the implementation of all the tasks set. Applying new technologies to existing systems can exacerbate existing problems	For the SME sector, making changes involves combining the efforts of many enterprises, cooperation with contractors, and even possibly with competitors, interaction with government agencies, non-governmental organizations, universities in order to combine efforts, information, exchange the necessary knowledge and ideas.
2. Expand opportunities, not limit them	Modern technologies are so advanced that some of them can already influence people's behavior in both explicit and implicit ways. It limits our freedom, way of thinking, behavior. The meaning of new technologies	Moldovan SMEs face the problem of poor security and adaptability to new technologies, at the same time, their implementation gives them more opportunities - reduces business processes, reduces errors that often

	should not be limiting, but rather, expanding opportunities. Such systems should provide an opportunity for greater choice, development, freedom, and not vice versa.	occur when using manual work, reduces costs, reduces the time for transferring information. Consequently, SMEs need to consciously approach the model of transformation and organizational change, which will be the way for such companies to survive.
3. **By design, not by default**	In the process of transformation, it is necessary to emphasize the philosophy of the humanistic approach, put the importance of human values at the forefront, develop the principles of justice, equality, humanity and reject those processes that may contradict them. At first glance, it may seem that these processes are too complex, indefinite and self-organized. This mistake cannot be made and a constructive approach to the formation of a future model of existence cannot be applied in time.	For the management of change in SMEs, this principle is of paramount importance. The importance of a humanistic approach and focus on the needs of staff, relationships in the team, professional growth and interaction in a different format, awareness and perception of the need for such changes will be the main factor in the success of the changes. New configurations of the world order, in which the place of SMEs must be determined and fixed is the task of both the leaders of such enterprises themselves and the state system.
4. **Values as a virtue, not a disadvantage**	All technologies carry certain values, they coexist with new ideas and can be named depending on how the implementation of this idea goes. Therefore, values must be traced at all stages of the innovation process. So, tracing all the stages of implementation, it is possible not to miss the initial good idea and not reorient it into a negative return. A lot of research is now being done on the negative impact of progress, progress should be directed in a favorable direction.	When carrying out changes of this magnitude and nature, many are afraid of the negative consequences associated, first of all, with the reduction of jobs, the loss of human identity, and the fatal dependence on technology. In such pessimistic forecasts, SMEs generally risk to cease to exist. Therefore, focusing on existing values, maintaining them and building a model of organizational change in the context of strengthening values is one of the main guidelines for change management.

The above principles indicate the most important guidelines that should form the basis of the upcoming changes. As it becomes clear when diving into the essence of the processes, all systems of functioning will be spent. As the author of the concept says, "they will become part of ourselves."

The good news is that as they spread, technologies become more accessible, that is, not only cheaper, but also understandable. Therefore, in this context, we can talk about the relative "democratic" nature of this process. At the same time, no futurist can accurately determine what awaits business and humanity in general, even in the near future, given the speed and global nature of the process of change.[5,6]

So, the main directions of development of changes in small and medium-sized enterprises will depend on that to what extent the above problems will be solved.

Here are the directions, in our opinion, in which a policy of change for SMEs should be developed:

- governmental support,
- development of a financing policy for such enterprises,

• close cooperation of SMEs among themselves in alliances to achieve a competitive advantage,

• changing the innovation policy of SMEs,

• close cooperation with universities in the field of mutual support (universities: knowledge and qualified personnel, SMEs : the desire to cooperate and employ),

• formation of an understanding among entrepreneurs of the inevitability of transformational transformations and the essence of the nature of changes.

5. Conclusions

The fourth industrial revolution is associated with fundamental changes taking place in the world. All spheres of human existence will be involved in the process of change. Naturally, the business will also be significantly transformed. Already today we feel the power and power of information technology - the use of Smartphone's, the Internet, a computer is an integral part of our lives. Industry 4.0 technologies will make all processes even faster and more radical.

Researchers identify many problems associated with the transformation processes, naming the most significant lack of generally accepted standards, problems associated with information security, lack of necessary qualifications, etc. These problems can be solved by using a systematic and humanistic approach to the ongoing changes.

The object of this study were small and medium-sized enterprises of the Republic of Moldova, which are the core of the Moldovan economy. This sector is developing unstable and in recent years there has been a noticeable decline in all economic indicators. In addition, the innovative activity of these enterprises decreased by almost half over the period from 2015 to 2020. The current situation is aggravated by crisis phenomena, the situation associated with the pandemic and its consequences. Thus, in the current circumstances, the transition to a new mode of activity associated with the provocative factors of the Fourth Industrial Revolution is extremely difficult.

The author of the concept of Industry 4.0, Klaus Schwab, names four principles that should form the basis of change. They are associated with awareness of the introduction of new technologies, observance of the principles and values of society, a constructive approach that considers certainty in the development of scenarios for the use of new technologies and the expansion of opportunities associated with the use of technologies.

Enterprises of the Republic of Moldova need to be guided by these principles, at the same time, taking into account the specifics of SMEs. In addition, it is possible to identify the main areas that need to be taken into account when developing a transformation strategy. This is, first of all, the need for state support and non-governmental organizations in the field of both additional financing of SMEs and information support. Also important is the connection of SMEs with universities for the purpose of mutual support. Universities - in the field of presenting new knowledge, new and promising employees, and enterprises - jobs, information, the desire to cooperate and create new associations. Internal processes at enterprises are also very important - the development of a new concept of change management, which helps SMEs to harmoniously transform, understanding the inevitability of changes and the willingness to introduce innovations. An important role is also played by the possibility of cooperation between SMEs and the same enterprises in order to obtain a synergistic effect from cooperation.

Considering these directions, SMEs of the Republic of Moldova need to develop programs for organizational changes that will be able to survive the crisis phenomena of the provocation of the modern world.

References

1. Schwab, Klaus; Nicholas Davis, Shaping the Fourth Industrial Revolution, 2018, World Economic Forum
2. Schwab, Klaus, The Fourth Industrial Revolution, 2017, World Economic Forum
3. Sibel, Thomas, Digital Transformation: Survive and Thrive in an Era of Mass Extinction, 2019
4. Blommaert, Tjeu, Stephan van Broek, in cooperation with Erik Kolthof, Management in singularity:from linear to exponential management , 2018
5. Harari, Yuval Noah, 21 Lessons for the 21-st century, 2018
6. Mcafee, Andrew Brynjolfsson, Erik, Machine, Platform, Crowd: harnessing our digital future, 2019
7. The activity of small and medium enterprises in the Republic of Moldova, https://statistica.gov.md/newsview.php?l=ro&idc=168&id=6716
8. Main indicators of economic activity, 2014-2020, https://statbank.statistica.md/pxweb/pxweb/ro/40%20Statistica%20economica/40%20Statistica%20economica__24%20ANT__ANT020/ANT020100.px/?rxid=b2ff27d7-0b96-43c9-934b-42e1a2a9a774
9. The activities of innovative enterprises by economic kind of activity and by the size class of the enterprises, 2015-2020 https://statbank.statistica.md/PxWeb/pxweb/ro/40%20Statistica%20economica/40%20Statistica%20economica__24%20ANT__INO010/INO010100.px/table/tableViewLayout1/?rxid=b2ff27d7-0b96-43c9-934b-42e1a2a9a774

Barriers and Solutions to the Integration of Digital Technologies in Schools: A Case Study of School Managers in Moldova

Adrian Andronic

Academy of Economic Studies of Moldova (AESM), Chisinau, Moldova
andronic.adriano@gmail.com

Abstract. This paper presents the results of a study on the adoption and integration of Information and Communication Technology (ICT) in the Moldovan schools from the school managers' perspective. The Internet and digital devices are accessible by most people in Moldova. However, our literature review shows that their adoption and integration in schools is still low. This research adopts the case study method to examine the barriers as well as the solutions from the perspective of the school managers. Twenty school managers from different districts of Moldova participated in this study. A combination of interviews and questionnaires was used to collect data. The results show that the main barriers against the adoption and integration of ICT in the classroom include a lack of adequate and well-trained personnel, poor internet service, as well as high cost of access. The identified solutions include changes to the teaching curriculum to promote and support digital literacy, funding from the Government and external donors, and digital literacy training for the teachers and school managers. These findings provide valuable insights for school managers and policymakers on strategies to improve the adoption and integration of ICT in schools in Moldova.

Keywords. ICT, technology adoption, education, Moldova

1. Introduction

Over the last two decades, education increasingly adopted ICT. Technology can help diversify the students' learning experience in the classroom and help them immerse into knowledge [1-3]. Through technology, students learn in a new way and enhance their knowledge by visualization and deeper understanding [4]. When learning occurs through technology, student engagement improves [5, 6]. Promoting instant access to information through technology enhances students' engagement [7]. The student and the teachers can receive immediate feedback, which can prove useful, especially when the classrooms have a high number of students. The ICT tools for education are being developed at a tremendous speed, however, the adoption, integration, and application of these tools in the educational process in developing countries such as Moldova are still low [8, 9].

2. Technology Adoption in the Classroom

The main users of ICT in schools are the students and teachers. Most of the students embrace new technologies and have a positive attitude towards digital learning. The teachers, however, can be one of the main barriers to ICT adoption in schools [10]. As a result, the teachers were the central theme of the research on technology adoption and integration in the classroom [11-14]. According to Chen [15]"In a classroom, the teacher perceives and defines a teaching situation, makes judgments and decisions and then takes related actions". Therefore,

the teacher plays the most important role in the adoption and application of technology in the classroom.

To fully comprehend technology adoption and integration it is imperative to understand teachers' pedagogical beliefs and attitudes [16-19]. Some studies found that teachers adopt the technology in the teaching process at a slow pace [19-21]. Some of the research was carried out in developed countries, where the rate of digital inclusion and government support are high [22, 23]. Several authors discussed, categorized, and published the barriers to technology adoption and integration [18, 24, 25]. Most of them classify them into two groups [26, 27]. The barriers related to resources and the institutions are defined as "external". The barriers related to teachers and their attitudes are defined as "internal".

According to Ertmer [28], the first-order barriers (extrinsic to the teacher) are the lack of access to technology, insufficient time to plan for integration, lack of training, inadequate technical and administrative support. The second-order barriers (intrinsic to the teacher) are the beliefs about teaching and learning, the beliefs about computers and technology, the beliefs about classroom practices and routines, and the unwillingness to embrace change.

Tsai and Chai [29] proposed the third-order barriers that concern the teacher's ability to create and set learning experiences based on the students' needs and the learning context. They argued that the teachers' access to sufficient resources, facilities is as important as their own beliefs and positive attitude.

3. Motivation and Goal of the Study

In the last decades, there is increasing demand for jobs that are based on digital technologies.

According to the World Banks' Annual Report, there is vital to invest in people so that "everyone can fulfill her or his potential to thrive in the 21st-century economy" [30]. One of the recommendations of the "Connecting to Work" Report is that developing countries need to "bridge education to employment by developing skills for ICT jobs and promoting digital literacy using innovative models" [31]. Half of the world's population still lacks the skills and capabilities to fully, effectively, and equally participate in the global digital world [32, 33].

Achieving several of the United Nations Sustainable Development Goals, namely to "ensure inclusively and quality education for all and promoting lifelong learning" and reduce poverty, achieve gender equality and provide decent work and sustainable economic growth is directly related to closing the digital literacy gap [34, 35].

Introducing technologies in school learning starting with the primary level can help close the digital literacy gap and engage the students in technology throughout their education. This can be done only if the teachers adopt and implement the technologies in the classroom. Unfortunately, this doesn't happen in many developing countries [36-40].

The goal of this study is to examine why the use of technology in schools remains such an issue in the Republic of Moldova. The author employs a user-centered approach to study the barriers and potential solutions to them through the perspective of the school managers, who are also practitioners in the classroom. In addition, this study seeks to understand the environment and the culture in which the school managers are operating.

4. Participants and Research Methods

Data was gathered from twenty public school managers (from different districts) from the Republic of Moldova during September 2021 via phone interviews. Most of the school managers have at least fifteen years of teaching experience.

To explore the barriers with the school managers, the case study approach was chosen.

Case studies are valuable for investigating complex ideas in exploratory research and developing new ones for future study [41]. They are also useful in capturing more contextual information and "lived realities", especially important in studies like education [42].

In this case study the author formulated two research questions:

- "In your opinion, what are the top ten barriers to the successful integration of ICT in the educational process in your school?"
- "Please suggest a solution to the barriers you mentioned in the previous question".

The data from the school managers were examined by thematic analysis. The responses were divided into first and second-order barriers and discussed in the following section.

5. Findings and Discussions

6.1. First-order Barriers

The school managers identified the first-order barriers, divided into three categories. The society-related barriers were unstable government and government policies, unstable educational curriculum, poor funding from the government, lack of materials and relevant gadgets, the high cost of facilities, the lack of good quality internet service, and poverty. The family-related barriers were parental misconceptions, poor educational background of the students, social vices. The school-related barriers were institutional barriers, lack of maintenance, and unconducive environment, inadequate time in the classroom, insecurity.

5.2.1. Society-related Barriers

The instability of the Moldovan government and its policies, the underfunding of education, the lack of minimum conditions like constant electricity and heating in schools, low quality, or insufficient broadband for the Internet access of the students are among the barriers resident in the society. The school managers believe that the government plays the most important role in the integration of ICT in the classroom.

In the Republic of Moldova, in the academic year 2021- 2022, the network of general primary and secondary education comprises 1231 units, consisting of 99 primary schools, 786 gymnasiums, 338 high schools, and 8 schools for children with disabilities in intellectual or physical development. At the beginning of the 2021/22 academic year, 336.7 thousand students were enrolled in general primary and secondary education. In terms of distribution by area of residence, 54.2% of students study in urban areas and 45.8% in rural areas. Of the total units, 98.1% belong to the public sector; 74.3% of the institutions operate in rural areas [43].

Given most of the schools are public, the government's policies and financing shape their educational practices. Although in the past decade various government cabinets have admitted the importance of providing adequate financial and technological support to education, the allocated budget has remained one of the lowest in the world, most of it going for the payment of teacher's salaries [44, 45]. By 2022, the estimated expenditures of the National Public Budget on education will reach 820 million USD [46]. The amount allocated for education in Moldova does not comply with the recommendations of the Organisation for Economic Co-operation and Development [47].

5.2.2. School-related Barriers

Most of the state schools' environment is not conducive for the adoption and application of digital technologies. Creating an enabling environment is crucial. Lim and Khine [48] consider that the classroom environment is both complex and dynamic, the teacher is the key in integrating innovative tools or practices in the process of teaching. For the ICT to

be successfully integrated the schools must be safe, equipped with computers, have consistent electricity and heating, and access to broadband Internet for every student.

At the beginning of the 2021-2022 academic year, out of 1231 primary and secondary general education institutions in the Republic of Moldova, 1214 units carried out the educational process in their premises, and 17 institutions in rented space. Out of a total of institutions, 1018 are in standardized classrooms, and 213 - in refurbished classrooms.

The task of school managers to ensure the necessary conditions is becoming more and more complicated because the schools depend financially on each enrolled student, but their number has been decreasing for several years. The principal of the high school in the village of Chircăiești, Căușeni, says that in such conditions, the financing per student is no longer a fair solution and consequently many of the school's needs are still postponed: "The class is ten students, or the class is 35 students, it must be heated anyway, it must be lit anyway. It is not so fair that according to the number of students the funding is allocated. Yes, we try, based on the financial means, we make optimizations, we closed one floor at a time, we closed some classes. You can't do lessons with illustrative material if the wind blows and, in 2018, we put an end to all this. "

In most of the educational institutions, which have their classrooms in old buildings, the maintenance and repair expenses absorb a good part of the budget. School managers say they would prefer to equip the classroom with teaching materials, renew the book fund, and activities that would involve and interest students, but for now these needs are gradually being met at a low level.

General and secondary education institutions have 37.0 thousand computers used for educational purposes (or about 23% more than in the 2020-2021 academic year), of which 18.1 thousand (48.9%) are connected in the common school network, and 31.4 thousand (84.9%) have an Internet connection. At the same time, 20.1 thousand computers are used by students, 15.9 thousand by teachers and a thousand computers are used in libraries. Compared to the 2020-2021 academic year, both the number of computers used by students (by 17.5%) and by the teaching staff (by 31.4%) increased [43].

According to a study on the economic and social impact of the COVID-19 pandemic, developed by UNDP Moldova, about 150 thousand students did not have access to education during the restrictions due to lack of necessary equipment, knowledge, or Internet connection of students or teachers [49].

High-speed internet access in study classes, as well as the integration of digital technologies in schools in the Republic of Moldova, are essential for the continuity and efficiency of the current study process, but also the long-term development of competent professionals. According to the data presented in the Integrated Management System in Education, 16,575 students and 212 teachers stated that they do not have access to the Internet [50].

One of the school managers stated: "If we go online, I tell you that I don't know what it will be. Why give the student a laptop if he doesn't have internet? Simple. What will he/she do with it? Play games? Some of the pupils are from the neighboring village, which is over the hill in the center, in the cradle. Not even a cell phone picks up a signal there. Please tell me, what should be done with these laptops? ..... Crack the nuts?" [51].

5.2.3. Family-related Barriers

The digital divide can also be caused by the presence or the lack of computers at home [52, 53]. In the Republic of Moldova, the insufficient number of computers in schools is like the situation of the families, especially in the rural areas. The main reason is the lack of financial means to purchase them.

Most of the school managers consider that having a laptop at home, in addition to using one at school would increase the adoption of the ICT in education. However, many families believe the time spent in front of the computers at school is sufficient for their children, and the sole responsibility of education resides on the teacher's shoulders.

Some parents believe that owning a laptop at home does not help their children, because they might get distracted by computer games or even get involved in cyber-crimes. Many believe a smartphone is enough for their child.

Also, there are a lot of families where parents work abroad, and the children are left with the grandparents. Unfortunately, among the elder generation, digital literacy is low, and they do not see the future in digital technologies. Many families involve children in agricultural work and household duties from an early age, limiting their access to digital technologies and knowledge in general.

6.2. Second-order Barriers

The school managers identified five second-order barriers divided into two categories. In the first category "Beliefs about computers and technology", were the attitudes and beliefs of teachers and other relevant stakeholders, the misinformation, and poor orientation about computers. In the second category, "Competence and digital literacy" were included the failure to apply/use digital tools properly, the lack of well-trained personnel, the laziness of teachers.

5.2.1. Beliefs About Computers and Technology

The adoption and integration of technology in the classroom by the teachers are directly impacted by the information they possess regarding technologies and their pedagogical beliefs [54-56]. The main determinant of the teachers' readiness to adopt and integrate ICT in the educational process is their personal beliefs about the usefulness of digital technologies.

The availability of computers, support from the school management, and government policies regarding the digitalization of education are not important unless the teachers are willing to adopt the ICT in the classroom.

Some of the teachers think the traditional methods are the best and the computers are just a waste of time, others are even afraid to break them [11, 57, 58]. The belief that the use of the computer is not for any "serious" teaching process and the fear of being held responsible if students damage the computers during their class, make the teachers avoid using them at all [59, 60].

5.2.2. Competence and Digital Literacy

In the academic year 2021-2022, teachers and managers in primary and secondary education were 26.5 thousand people (or 1.5% less than the previous year of study), including 3.4 thousand people – staff management (12.8%) and 23.1 thousand people – teaching staff (87.2%).

The distribution of teachers and management by age groups reveals that an important share is held by those aged 50-59 (26.8%), being practically at the level of the 2020-2021 academic year, while the share of teachers and driving young people under the age of 30 decreased by 0.5%, and those aged 60 and over increased by 0.6%.

Female teachers have a share of 86.8%, female management staff being 92.5%. More than half of the teachers and management have seniority in pedagogical work of 20 years and over (56.9%).

Previous research has shown that in developing countries, especially in the countryside the teachers lack the necessary skills and competencies to integrate the ICT in their teaching process [61-63]. Some researchers mentioned that this could be a major barrier in the digitalization of education [61, 64, 65].

In this study, the school managers pointed out the lack of training and expertise for using the new technologies in the teaching process. They added that most of the older teachers also lack confidence in using computers.

Digital education is a compulsory subject for primary school students in all educational institutions in the country. The initiative, launched in 2018, aimed to teach children about computers, technology, and robots, to cultivate their passion for STEM education, and to encourage them to choose one of the professions of the future. The program was launched following the signing of a Memorandum of Understanding on the development of digital education in general education between the Ministry of Education, Culture and Research, the Ministry of Economy and Infrastructure, the National Association of ICT Companies, and the Tekwill project [66].

Between July 26 and August 20, 2021, approximately 1,200 teachers from all over Moldova attended the "Digital Education" training program so that in the fall, they could develop students' computer skills, technologies, and robotics and develop a passion for information technology [66].

In Moldova, the Ministry of Education together with the colleges of education and local administration are yet to integrate the ICT in their teacher training programs.

6. Solutions Proposed by School Managers

After identifying the barriers, the school managers proposed solutions to encourage the integration of ICT in the classroom educational environment. These solutions involve three types of stakeholders: government, teachers, and school management.

6.1. Government

Solid infrastructural requirements drive the adoption and integration of ICT in the classroom [65]. The school managers identified essentials like electricity, internet, computers, and comfortable spaces. The government must allocate enough funds to make sure all the schools have the proper infrastructure for ICT use in the classroom.

The Ministry of Education and Research must develop policies that highlight the importance of enhancing teaching with technologies and stipulate practical ways to motivate the teachers to adopt the ICT in the classroom. In addition, the curriculum must be updated and extended to make sure the ICT is a part of the teaching process as well as the school management.

Apart from the standard courses offered by the Institute of Continuous Education (Methodology of preparing the modern multimedia lesson, computer-assisted assessment, conducting school experiments in digital laboratories, the technology of using interactive whiteboards), the Ministry of Education and Research must commission additional courses to encourage all the teachers to attend specialized workshops and seminars to make them feel more comfortable using ICT and developing their digital teaching content for using in the classroom. Given the consequences of the ongoing pandemic, the Ministry of Education and Research must strengthen teachers' skills as they move from traditional teaching to "reverse classroom", which involves online teaching and learning.

6.2. School Management

The school management must use the support from the government and provide a motivating environment for the teachers to adopt the ICT in the classroom. The way this adoption takes place is as equally important as the funding from the government.

The change that comes from introducing the ICT in the classrooms must be managed by the school principals. According to Ritchie and Rodriguez, the school management is

responsible for "providing and selling the vision to the community, obtaining resources such as time, personnel, knowledge, materials, and facilities, and providing encouragement and recognition for teachers successfully making the transition" [67]. The school management must create "the role of a technology leader – someone who is comfortable with ongoing change and can keep themselves abreast with products, procedures, policies related to digital technologies while supporting and empowering teachers to adopt technology" [67].

Finally, the teachers recommended that other teachers who have practical knowledge of digital tools should be encouraged to train their colleagues who do not have these skills. This community of practice may help reluctant and slow adopters share their concerns, increase their confidence and efficacy, and eventually make them willing to integrate digital technologies in their teaching practice. The school management has the responsibility for using the support of the government in providing an enabling environment for teachers to adopt and integrate digital technology in the classroom. Although funding is vital to technology adoption and integration, the manner and process in which this adoption and integration take place are equally as important. School management is responsible for driving and managing the change that comes with introducing technology-enhanced learning. Ritchie and Rodriguez [46] highlight various methods through which the school management can support the process of technology adoption. This includes "providing and selling the vision to the community, obtaining resources such as time, personnel, knowledge, materials, and facilities, and providing encouragement and recognition for teachers successfully making the transition" [46]. The school management should also create "the role of a technology leader – someone who is comfortable with ongoing change and who can keep themselves abreast with products, procedures and policies relating to digital technologies while supporting and empowering teachers to adopt technology" [46].

The school managers also recommend that all the teachers who have enough knowledge and practice in using ICT in the classroom should be motivated to train their colleagues, organize workshops improve other teachers' digital literacy and skills. Learning from colleagues could help them feel less reluctant to try new ways of teaching the subjects and increase their confidence in front of the students. The younger teachers should support the older ones in making the first steps into using ICT in the classroom.

6.3. Teachers

The teacher is the central actor in the teaching process, therefore is up to them to maximize the support from the government and the school management in successfully integrating the ICT in the classroom. The lack of skills of the teachers should also be addressed by committing to continuous skills development that would make them comfortable using the new technologies and saving their time in the process of student assessment and result reporting. First, they must understand the value the technology brings to the teaching process and the learning opportunities it gives to the students to be prepared for the jobs of the future. Second, the teachers must understand that to stay competitive and grow professionally and personally they must embrace the new educational technologies.

7. Conclusion

This paper examines the main barriers and solutions to the integration of ICT in the classroom from the perspective of school managers. The society-related barriers were unstable government and government policies, unstable educational curriculum, poor funding from the government, lack of materials and relevant gadgets, the high cost of facilities, the lack of good quality internet service, and poverty. The family-related barriers were parental misconceptions, poor educational background of the students, social vices. The school-related

barriers were institutional barriers, lack of maintenance, and unconducive environment, inadequate time in the classroom, insecurity. The intrinsic barriers related to beliefs about computers and technology and competence and digital literacy. The proposed solutions are aimed at the government, school managers, and teachers. This study is focused on the Republic of Moldova and is based on the feedback from 20 school managers. Further research can approach quantitatively all the school managers in primary and secondary education as well as vocational training institutions. The insights drawn from this paper can be compared to other developing countries and the results can be used to improve the level of ICT adoption in schools.

References

1. Baş, G. and M. Baştuğ, Teaching-learning conceptions, teaching motivation, and perceptions towards ICT: A research in Turkish public high schools. Education and information technologies, 2021. 26(2): p. 1607-1625.
2. Terzieva, V., et al. Implementation of ICT-based teaching in Bulgarian schools. in Proceedings of the 6th International Conference on Education and New Learning Technologies. 2014.
3. McRae, H., T. McKenzie, and R. Toia, ICT Tools Supporting Teacher Practice in Language Acquisition. Ubiquitous Learning: An International Journal, 2010. 2(3).
4. Hernández-Lara, A.B., A. Perera-Lluna, and E. Serradell-López, Game learning analytics of instant messaging and online discussion forums in higher education. Education+ Training, 2021.
5. Morais, P., M.J. Ferreira, and B. Veloso, Improving student engagement with Project-Based Learning: A case study in Software Engineering. IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 2021. 16(1): p. 21-28.
6. Ahshan, R., A framework of implementing strategies for active student engagement in remote/online teaching and learning during the COVID-19 pandemic. Education Sciences, 2021. 11(9): p. 483.
7. Cheng, P.-H., et al., 5E mobile inquiry learning approach for enhancing learning motivation and scientific inquiry ability of university students. IEEE Trans. on Education, 2015. 59(2): p. 147-153.
8. Billon, M., J. Crespo, and F. Lera-Lopez, Do educational inequalities affect Internet use? An analysis for developed and developing countries. Telematics and Informatics, 2021. 58: p. 101521.
9. Kaushik, M.K. and D. Agrawal, Influence of technology readiness in the adoption of e-learning. International Journal of Educational Management, 2021.
10. Nair, I. and V.M. Das, Using Technology Acceptance Model to assess teachers' attitude towards use of technology as teaching tool: A SEM Approach. International Journal of Computer Applications, 2012. 42(2): p. 1-6.
11. Brzycki, D. and K. Dudt, Overcoming barriers to technology use in teacher preparation programs. Journal of Technology and teacher education, 2005. 13(4): p. 619-641.
12. Dougherty, K., Understanding factors that influence college faculty in deciding to adopt digital technologies in their practice. 2015: University of Ontario Institute of Technology (Canada).
13. Pareja Roblin, N., et al., Practical considerations informing teachers' technology integration decisions: the case of tablet PCs. Technology, Pedagogy and Education, 2018. 27(2): p. 165-181.
14. Scherer, R., et al., The importance of attitudes toward technology for pre-service teachers' technological, pedagogical, and content knowledge: Comparing structural equation modeling approaches. Computers in Human Behavior, 2018. 80: p. 67-80.
15. Chen, C.-H., Why do teachers not practice what they believe regarding technology integration? The journal of educational research, 2008. 102(1): p. 65-75.
16. Sang, G., et al., Student teachers' thinking processes and ICT integration: Predictors of prospective teaching behaviors with educational technology. Computers & Education, 2010. 54(1): p. 103-112.
17. Liu, S.-H., Factors related to pedagogical beliefs of teachers and technology integration. Computers & Education, 2011. 56(4): p. 1012-1022.
18. Burke, P.F., et al., Exploring teacher pedagogy, stages of concern and accessibility as determinants of technology adoption. Technology, Pedagogy and Education, 2018. 27(2): p. 149-163.
19. Bano, M., et al., Mobile learning for science and mathematics school education: A systematic review of empirical evidence. Computers & Education, 2018. 121: p. 30-58.

20. Hu, X. and N. Yelland, An investigation of preservice early childhood teachers' adoption of ICT in a teaching practicum context in Hong Kong. Journal of early childhood teacher education, 2017. 38(3): p. 259-274.
21. Lawrence, J.E. and U.A. Tar, Factors that influence teachers' adoption and integration of ICT in teaching/learning process. Educational Media International, 2018. 55(1): p. 79-105.
22. Tan, C., Globalisation, the Singapore state and educational reforms: Towards performativity. Education, Knowledge & Economy, 2008. 2(2): p. 111-120.
23. Aldunate, R. and M. Nussbaum, Teacher adoption of technology. Computers in Human Behavior, 2013. 29(3): p. 519-524.
24. Porter, W.W., et al., A qualitative analysis of institutional drivers and barriers to blended learning adoption in higher education. The internet and Higher education, 2016. 28: p. 17-27.
25. Reid, P., Categories for barriers to adoption of instructional technologies. Education and Information Technologies, 2014. 19(2): p. 383-407.
26. Bingimlas, K.A., Barriers to the successful integration of ICT in teaching and learning environments: A review of the literature. Eurasia J. of Math., science and technol. educ., 2009. 5(3): p. 235-245.
27. Makki, T.W., et al., When first-order barriers are high: A comparison of second-and third-order barriers to classroom computing integration. Computers & Education, 2018. 120: p. 90-97.
28. Ertmer, P.A., Addressing first-and second-order barriers to change: Strategies for technology integration. Educational technology research and development, 1999. 47(4): p. 47-61.
29. Tsai, C.-C. and C.S. Chai, The" third"-order barrier for technology-integration instruction: Implications for teacher education. Australasian Journal of Educational Technology, 2012. 28(6).
30. Bank, W., Annual Report 2017 End Extreme Poverty Boost Shared Prosperity. 2017, World Bank Washington DC.
31. Raja, S., et al., Connecting to work: How information and communication technologies could help expand employment opportunities. 2013.
32. Varis, T., European and global approaches to digital literacy. Nordic Journal of digital literacy, 2008. 3(1): p. 53-60.
33. Knobel, M. and C. Lankshear, Digital literacy and digital literacies: Policy, pedagogy and research considerations for education. Nordic Journal of digital literacy, 2006. 1(1): p. 12-24.
34. Westerman, S. and W. Barry, Mind the gap: staff empowerment through digital literacy. 2009, The Higher Education Academy.
35. Nations, U., United Nations Sustainable Development Goals: Education. New York, NY, 2019.
36. Richardson, J.W., Technology adoption in Cambodia: Measuring factors impacting adoption rates. Journal of International Development, 2011. 23(5): p. 697-710.
37. Pischetola, M., Exploring the relationship between in-service teachers' beliefs and technology adoption in Brazilian primary schools. International Journal of Technology and Design Education, 2020: p. 1-24.
38. Kamau, L.M., Technology adoption in secondary mathematics teaching in Kenya: An explanatory mixed methods study. 2014, Syracuse University.
39. Lopez-Perez, V.A., P.E. Ramirez-Correa, and E.E. Grandon, Innovativeness and factors that affect the information technology adoption in the classroom by primary teachers in Chile. Informatics in Education, 2019. 18(1): p. 165-181.
40. Garcia, M.B., E-learning technology adoption in the Philippines: an investigation of factors affecting filipino college students' acceptance of learning management systems. The International Journal of E-Learning and Educational Technologies in the Digital Media, 2017. 3(3): p. 118-130.
41. Amerson, R., Making a case for the case study method. 2011, SLACK Incorporated Thorofare, NJ. p. 427-428.
42. Hodkinson, P. and H. Hodkinson. The strengths and limitations of case study research. in learning and skills development agency conference at Cambridge. 2001. University of Leeds Leeds, UK.
43. National Bureau of Statistics of the Republic of Moldova, Activity of the primary and general secondary education institutions, beginning of the academic year 2021/22 National Bureau of Statistics of the Republic of Moldova, Editor. 2021, National Bureau of Statistics of the Republic of Moldova,: Chisinau, Republic of Moldova.
44. Ministry of Education and Culture. It is time for everybody to learn. To learn from the lesson of life. 2020 Dec. 20, 2021]; Available from: https://mec.gov.md/sites/default/files/sarov-interviu.pdf.

45. Ursu, V. Maia Sandu: We Intend to Make Consolidated Investments in the School. 2012 December 20, 2021]; Available from: https://moldova.europalibera.org/a/24734334.html.
46. Parliament of the Republic of Moldova, State Budget Law for 2022, Parliament of the Republic of Moldova, Editor. 2021, Official Gazette of the Republic of Moldova: Chisinau, Republic of Moldova.
47. Organisation for Economic Co-operation and Development, Finance & Funding. 2022, Organisation for Economic Co-operation and Development: Paris, France.
48. Lim, C.P. and M. Khine, Managing teachers' barriers to ICT integration in Singapore schools. Journal of technology and Teacher Education, 2006. 14(1): p. 97-125.
49. Botezatu, S. Digital Transformation of Moldova: there is no way back. 2021 December 20, 2021]; Available from: https://www.md.undp.org/content/moldova/en/home/blog/2021/transformarea-digital-a-moldovei--nu-mai-e-cale-de-intoarcere.html.
50. Tekwill, High Speed Internet and Digital Technologies for Schools In Moldova, Tekwill, Editor. 2021, Tekwill,: Chisinau.
51. Șevciuc, L., Digitally, we're doing badly, in Cu sens. 2021, Cu sens: Chisinau.
52. Peña-López, I., From laptops to competences: Bridging the digital divide in education. RUSC. Universities and Knowledge Society Journal, 2010. 7(1): p. 21-32.
53. Cruz-Jesus, F., et al., The education-related digital divide: An analysis for the EU-28. Computers in Human Behavior, 2016. 56: p. 72-82.
54. Leem, J. and E. Sung, Teachers' beliefs and technology acceptance concerning smart mobile devices for SMART education in South Korea. British J. of Educational Technol., 2019. 50(2): p. 601-613.
55. Baturay, M.H., Ş. Gökçearslan, and F. Ke, The relationship among pre-service teachers' computer competence, attitude towards computer-assisted education, and intention of technology acceptance. International Journal of Technology Enhanced Learning, 2017. 9(1): p. 1-13.
56. Koh, J.H.L., et al., Technological pedagogical content knowledge (TPACK) and design thinking: A framework to support ICT lesson design for 21st century learning. The Asia-Pacific Education Researcher, 2015. 24(3): p. 535-543.
57. Opeyemi, D.-A., et al. Barriers and identified solutions to the integration of digital technologies in the classroom: A case study of teachers in Nigeria. in 2019 IEEE Global Engineering Education Conference (EDUCON). 2019. IEEE.
58. Teo, T., C.B. Lee, C.S. Chai, Understanding pre-service teachers' computer attitudes: applying and extending the technology acceptance model. J. of comp. assisted learning, 2008. 24(2): p. 128-143.
59. McNierney, D.J., Case study: One teacher's odyssey through resistance and fear. TechTrends, 2004. 48(5): p. 64-69.
60. González, A., M.P. Ramírez, and V. Viadel, ICT learning by older adults and their attitudes toward computer use. Current gerontology and geriatrics research, 2015. 2015.
61. Ghavifekr, S., et al., Teaching and Learning with ICT Tools: Issues and Challenges from Teachers' Perceptions. Malaysian Online Journal of Educational Technology, 2016. 4(2): p. 38-57.
62. Albugami, S. and V. Ahmed, Success factors for ICT implementation in Saudi secondary schools: From the perspective of ICT directors, head teachers, teachers and students. International Journal of education and development using ICT, 2015. 11(1).
63. Adebayo, O. and M. Fagbohun, An assessment of computer and ICT skills among secondary school teachers in Ota Ogun State. Library Philosophy and Practice, 2013.
64. Raman, K. and H. Yamat, Barriers teachers face in integrating ICT during English lessons: A case study. Malaysian Online journal of educational technology, 2014. 2(3): p. 11-19.
65. Williams, D., et al., Teachers and ICT: Current use and future needs. British journal of educational technology, 2000. 31(4): p. 307-320.
66. Tekwill. 1,200 teachers in the country start training to teach the professions of the future to primary school students! 2021 November 8, 2021]; Available from: https://tekwill.online/ro/tifs-content/1200-de-profesori-din-tara-incep-instruirea-pentru-preda-profesiile-viitorului.
67. Ritchie, D. and S. Rodriguez, School administrators and educational technologies: narrowing the divide. Journal of Information Technology for Teacher Education, 1996. 5(1-2): p. 107-114.

Assessing Nurses' Emotional Exhaustion from a Regional Teaching Hospital in Taiwan by a Longitudinal Study

Yu-Ning Chao[1], Chih-Hsuan Huang[2], Yii-Ching Lee[3], Hsin-Hung Wu[1]

[1] Depart. of Business Administration, National Changhua University of Education, Taiwan
[2] School of Business Administration, Hubei University of Economics, Hubei, China
[3] Department of Health Business Administration, Hung Kuang University, Taichung, Taiwan
a83227@gmail.com, tititacer@hbue.edu.cn, yiiching.lee@gmail.com, hhwu@cc.ncue.edu.tw

Abstract. This study evaluates nurses' emotional exhaustion from a regional teaching hospital in Taiwan through the internal datasets from 2014 to 2020 based on the Chinese version of the safety attitudes questionnaire. The results show that nurses have statistically different perceptions in emotional exhaustion from 2014 to 2020. Four items have been improved with less fatigue from nurses' viewpoints. In contrast, five items have been deteriorating to result in nurses' higher burnout. Therefore, hospital management needs to place these five items in a high priority to reduce nurses' emotional exhaustion in order to enhance patient safety culture as well as patient healthcare quality.

Keywords. Chinese version of the safety attitudes questionnaire, emotional exhaustion, burnout, longitudinal study.

1. Introduction

The Chinese version of the safety attitudes questionnaire (SAQ) based on the six dimensions with 30 questions from the short form of the SAQ developed by Sexton et al. [1] has been the instrument to assess the patient safety culture from medical staff's viewpoints for healthcare organizations in Taiwan since 2009 [2,3]. The first version of the Chinese version of the SAQ developed by the Joint Commission of Taiwan used the forward and backward translation to check the quality of the translation and the pilot-testing and discussion by an expert panel for intelligibility and applicability of the items from the SAQ developed by Sexton et al. [4]. These six dimensions are teamwork climate, safety climate, job satisfaction, stress recognition, perceptions of management, and working conditions [5]. Later, the second version of the Chinese version of the SAQ incorporated three additional dimensions from the Agency for Healthcare Research and Quality including hospital management support for patient safety, teamwork across hospital units, and hospital handoffs and transitions [6]. Therefore, the instrument became nine dimensions and 41 questions [3].

In 2014, the third and current version of the Chinese version of the SAQ was developed by removing three dimensions in the second version and further added two dimensions, namely emotional exhaustion and work-life balance, to become eight dimensions and 46 questions [3,7]. Emotional exhaustion is to assess the degree of fatigue toward the work from staffs' perceptions, whereas work-life balance is to evaluate the work-life balance for each staff each week in terms of frequency [8,9]. Emotional exhaustion borrowed from one of three dimensions of Maslach burnout inventory-human services survey evaluates the degree of fatigue toward the work from medical staff's perceptions [10]. Work-life balance adopted from the work-life climate developed by Sexton et al. [11] assesses the balance between individual's

work and life in terms of frequency per week by understanding medical staff's degree of burnout [12].

Chi et al. [13], Huang et al. [14], and Huang et al. [15] pointed out that physicians and nurses are the core medical staff who directly contact patients very often and their attitudes toward patient safety determine the quality of care. Moreover, nurses are the biggest workforce in each healthcare organizations such that assessing their attitudes toward patient safety is more essential because their attitudes are a result of all other contributory features of the working environment [8]. Furthermore, nurses are the frontline workers to face problems, and their perceptions in patient safety might indicate whether the safety culture and quality of healthcare are positive [7].

Lee et al. [16] conducted a case study to analyze emotional exhaustion in a medical center from all of medical staff's viewpoints. Shanafelt et al. [17] stated out burnout affects workers in all fields including people interactions, and burnout impairs both personal and social functioning and reduces the work quality and health both physically and psychologically [18]. Moreover, physicians' burnout increases medical errors and decreases patient care [19,20]. Furthermore, nurses' burnout increases infection rates and decreases patient healthcare quality [21]. However, nurses' perceptions in emotional exhaustion might be more important than the other medical staff in healthcare organizations. There is a need to concentrate on nurses' perceptions. Besides, Lee et al. [16] used a cross-sectional study to assess how medical staff perceive in 2017. In contrast, using a longitudinal study to evaluate emotional exhaustion from nurses' viewpoints enables hospital management to trace the performance and trends on a timely basis. Therefore, this study uses nurses' perceptions in emotional exhaustion from a regional teaching hospital in Taiwan to figure out their degree of fatigue from a longitudinal viewpoint.

2. Emotional Exhaustion

Maslach burnout inventory-human services survey is a valid and reliable instrument to measure burnout in individuals who work with people, i.e., human services and medical professionals [10]. The instrument has three dimensions including emotional exhaustion, depersonalization, and personal accomplishment [18]. In 2014, the Chinese version of the SAQ borrowed emotional exhaustion to assess medical staff's burnout with nine questions as shown in Table 1. Each medical staff in healthcare organizations is asked to answer each question by a Likert's scale with five levels from strongly agree to strongly disagree with the numerical value of 5 to 1, respectively [14]. Each question uses negative wordings. Thus, the smaller value indicates lower burnout for individuals.

Table 1. Nine questions in emotional exhaustion

Dimension	Question
Emotional exhaustion	1. I feel like I'm at the end of my rope. 2. I feel burned out from my work. 3. I feel frustrated by my job. 4. I feel I'm working too hard on my job. 5. I feel emotionally drained from my work. 6. I feel used up at the end of the workday. 7. I feel fatigued when I get up in the morning and have to face another day on the job. 8. Working with people all day is really a strain for me. 9. Working with people directly puts too much stress on me.

3. Research Method

The annual internal data gathered from a regional teaching hospital in Taiwan from 2014 to 2020 from nurses' viewpoints are used. The nurses' demographic information regarding gender, age, supervisor/manager, respondents reporting events in the past 12 months, job status, experience in organization, experience in position, education, and direct patient contact is summarized in Table 2 (from 2014 to 2017) and Table 3 (from 2018 to 2020). Female nurses are the majority with their ages in 21-30 and 31-40 years old and most of them are not in charge of supervisors/managers. More than 90% of nurses report zero or 1-5 events in the past 12 months. In addition, more than 85% of nurses are full time employees in this regional teaching hospital. On the other hand, nurses' experience in this organization are distributed from 1 to 2 years to 11 to 20 years, and their experience in position has a similar pattern, i.e., from 1 to 2 years to 11 to 20 years. More than 90% of nurses have college/university degrees. Finally, nearly 90% of nurses are very often to contact patients directly.

Table 2. Nurses' demographic variables from 2014 to 2017 in a regional teaching hospital

Demographic Variable	2014	2015	2016	2017
	Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)
Gender				
Male	12 (3.1)	19 (5.5)	14 (3.5)	25 (5.3)
Female	369 (96.9)	328 (94.5)	384 (96.5)	451 (94.7)
Age				
Less than 20	8 (2.1)	8 (2.3)	11 (2.8)	7 (1.5)
21-30	160 (42.0)	146 (42.1)	174 (43.7)	220 (46.2)
31-40	149 (39.1)	130 (37.5)	137 (34.4)	145 (30.5)
41-50	55 (14.4)	54 (15.6)	64 (16.1)	90 (18.9)
51-60	9 (2.4)	9 (2.6)	12 (3.0)	14 (2.9)
61 and above	0 (0)	0 (0)	0 (0)	0 (0)
Supervisor/Manager				
Yes	31 (8.1)	28 (8.1)	40 (10.1)	38 (8.0)
No	350 (91.9)	319 (91.9)	358 (89.9)	438 (92.0)
Respondents reporting events in the past 12 months	198 (52.0)	217 (62.5)	234 (58.8)	292 (61.3)
No	169 (44.4)	114 (32.9)	155 (38.9)	174 (36.6)
1-5	12 (3.1)	14 (4.0)	7 (1.8)	4 (0.8)
6-10	0 (0)	2 (0.6)	2 (0.5)	6 (1.3)
11-15	2 (0.5)	0 (0)	0 (0)	0 (0)
More than16				
Job Status				
Full Time	334 (87.7)	313 (90.2)	358 (89.9)	433 (91.0)
Contract	20 (5.2)	11 (3.2)	14 (3.5)	14 (2.9)
Part Time	7 (1.8)	6 (1.7)	5 (1.3)	6 (1.3)
Agency	20 (5.2)	17 (4.9)	21 (5.3)	23 (4.8)
Experience in Organization	42 (11.0)	58 (16.7)	51 (12.8)	49 (10.3)
Less than 6 months	15 (3.9)	10 (2.9)	39 (9.8)	16 (3.4)

6 to 11 months	65 (17.1)	53 (15.3)	71 (17.8)	128 (26.9)
1 to 2 years	62 (16.3)	48 (13.8)	48 (12.1)	67 (14.1)
3 to 4 years	90 (23.6)	79 (22.8)	86 (21.6)	95 (20.0)
5 to 10 years	96 (25.2)	90 (25.9)	90 (22.6)	104 (21.8)
11 to 20 years	11 (2.9)	9 (2.6)	13 (3.3)	17 (3.6)
21 years or more				
Experience in Position				
Less than 6 months	56 (14.7)	65 (18.7)	62 (15.6)	51 (10.7)
6 to 11 months	16 (4.2)	11 (3.2)	40 (10.1)	24 (5.0)
1 to 2 years	69 (18.1)	59 (17.0)	77 (19.3)	135 (28.4)
3 to 4 years	66 (17.3)	57 (16.4)	57 (14.3)	77 (16.2)
5 to 10 years	100 (26.2)	93 (26.8)	94 (23.6)	104 (21.8)
11 to 20 years	71 (18.6)	59 (17.0)	63 (15.8)	80 (16.8)
21 years or more	3 (0.8)	3 (0.9)	5 (1.3)	5 (1.1)
Education				
Junior High School and below	0 (0)	0 (0)	0(0)	0 (0)
Senior High School	4 (1.0)	0 (0)	5 (1.3)	4 (0.8)
College/University	363 (95.3)	336 (96.8)	377 (94.7)	456 (95.8)
Graduate School and above	14 (3.7)	11 (3.2)	16 (4.0)	16 (3.4)
Direct Patient Contact				
No	7 (1.8)	13 (3.7)	16 (4.0)	11 (2.3)
Rare	25 (6.6)	20 (5.8)	27 (6.8)	30 (6.3)
Very Often	349 (91.6)	314 (90.5)	355 (89.2)	435 (91.4)

Table 3. Nurses' demographic variables from 2018 to 2020 in a regional teaching hospital

Demographic Variable	2018	2019	2020
	Frequency (%)	Frequency (%)	Frequency (%)
Gender			
Male	24 (5.3)	30 (6.3)	34 (7.8)
Female	426 (94.7)	449 (93.7)	404 (92.2)
Age			
Less than 20	3 (0.7)	8 (1.7)	0 (0)
21-30	215 (47.8)	225 (47.0)	189 (43.2)
31-40	127 (28.2) 86	136 (28.4)	120 (27.4)
41-50	(19.1)	84 (17.5)	104 (23.7)
51-60	19 (4.2)	24 (5.0)	25 (5.7)
61 and above	0 (0)	2 (0.4)	0 (0)
Supervisor/Manager			
Yes	45 (10.0)	40 (8.4)	47 (10.7)
No	405 (90.0)	439 (91.6)	391 (89.3)
Respondents reporting events in the past 12 months	260 (57.8)	286 (59.7)	264 (60.3)
No	175 (38.9)	176 (36.7)	160 (36.5)
1-5	13 (2.9)	13 (2.7)	9 (2.1)

6-10	0 (0)	2 (0.4)	2 (0.5)
11-15	2 (0.4)	2 (0.4)	3 (0.7)
More than16			
Job Status			
Full Time	407 (90.4)	447 (93.3)	420 (95.9)
Contract	18 (4.0)	13 (2.7)	11 (2.5)
Part Time	6 (1.3)	3 (0.6)	0 (0)
Agency	19 (4.2)	16 (3.3)	7 (1.6)
Experience in Organization	40 (8.9)	52 (10.9)	17 (3.9)
Less than 6 months	21 (4.7)	35 (7.3)	3 (0.7)
6 to 11 months	101 (22.4)	85 (17.7)	79 (18.0)
1 to 2 years	67 (14.9)	78 (16.3)	82 (18.7)
3 to 4 years	96 (21.3)	100 (20.9)	121 (27.6)
5 to 10 years	102 (22.7)	93 (19.4)	87 (19.9)
11 to 20 years	23 (5.1)	36 (7.5)	49 (11.2)
21 years or more			
Experience in Position			
Less than 6 months	49 (10.9)	61 (12.7)	21 (4.8)
6 to 11 months	27 (6.0)	39 (8.1)	9 (2.1)
1 to 2 years	108 (24.0)	98 (20.5)	94 (21.5)
3 to 4 years	70 (15.6)	78 (16.3)	88 (20.1)
5 to 10 years	104 (23.1)	107 (22.3)	117 (26.7)
11 to 20 years	79 (17.6)	75 (15.7)	80 (18.3)
21 years or more	13 (2.9)	21 (4.4)	29 (6.6)
Education			
Junior High School	0 (0)	2 (0.4)	0 (0)
and below	4 (0.9)	7 (1.5)	3 (0.7)
Senior High School	433 (96.2)	460 (96.0)	420 (95.9)
College/University	13 (2.9)	10 (2.1)	15 (3.4)
Graduate School and above			
Direct Patient Contact			
No	8 (1.8)	8 (1.7)	12 (2.7)
Rare	29 (6.4)	39 (8.1)	30 (6.8)
Very Often	413 (91.8)	432 (90.2)	396 (90.4)

The mean values in terms of nurses' perceptions in each of nine questions of emotional exhaustion from 2014 to 2020 are provided in Table 4 and Figure 1. The smaller values indicate nurses have less burnout. From the trends of the descriptive statistics, nurses feel less burnout in Questions 1, 3, 8, and 9. In contrast, nurses feel more fatigue in Questions 2, 4, 5, 6, and 7. From a hospital managerial viewpoint, it is of interest to observe if nurses feel less or more burnout in emotional exhaustion steadily by a longitudinal study statistically. Therefore, analysis of variance is employed with $\alpha = 0.05$ to test if nurses have different perceptions in emotional exhaustion in terms of nine questions from 2014 to 2020.

Table 4. Mean values of nurses' perceptions in emotional exhaustion from 2014 to 2020

Question	2014	2015	2016	2017	2018	2019	2020
1. I feel like I'm at the end of my rope.	3.08	3.11	2.82	2.88	2.79	2.82	2.72
2. I feel burned out from my work.	2.76	2.71	3.19	3.20	3.20	3.21	3.08
3. I feel frustrated by my job.	3.04	3.07	2.85	2.81	2.82	2.81	2.71
4. I feel I'm working too hard on my job.	2.64	2.71	3.29	3.29	3.30	3.27	3.29
5. I feel emotionally drained from my work.	2.76	2.76	3.16	3.29	3.22	3.20	3.01
6. I feel used up at the end of the workday.	2.45	2.46	3.37	3.54	3.43	3.46	3.28
7. I feel fatigued when I get up in the morning and have to face another day on the job.	2.79	2.85	3.04	3.13	3.14	3.14	2.98
8. Working with people all day is really a strain for me.	3.30	3.37	2.61	2.67	2.66	2.66	2.57
9. Working with people directly puts too much stress on me.	3.39	3.44	2.48	2.54	2.51	2.54	2.46

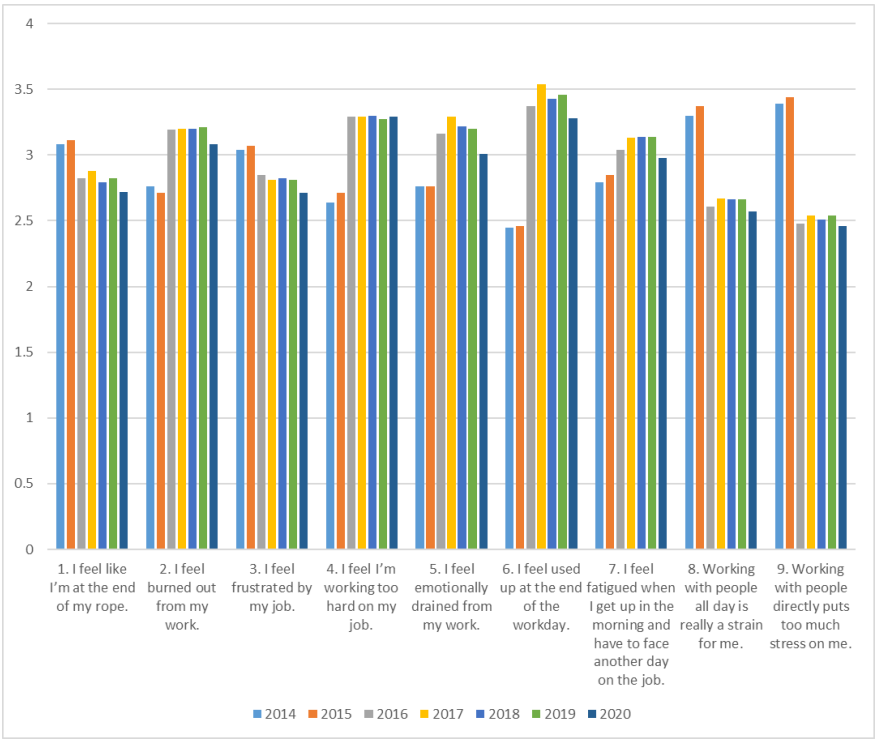


Fig. 1. A histogram of nine questions in emotional exhaustion from 2014 to 2020

4. Results

The ANOVA results depicted in Table 4 show that nurses from 2014 to 2020 have statistically different perceptions in emotional exhaustion in terms of nine questions. Some questions show similar patterns in post hoc analyses. For instance, Questions 1, 3, 8, and 9 indicate that nurses feel more fatigue in 2014 and 2015 but feel less burnout later. In fact, it is a good sign for nurses that they don't feel like they are at the end of their ropes; they do not frustrate their jobs; they don't feel working with people all day is really a strain for them; and they don't feel working with people directly puts too much stress on them. However, some questions have opposite directions. For instance, Questions 2, 4, 5, 6, and 7 imply that nurses are more fatigue in recent years when the comparative basis is set in 2014 and 2015. That is, nurses feel more burned out from their work; they feel they are working too hard on their jobs; they feel emotionally drained from their work; they feel used up at the end of the workday; and they feel fatigued when they get up in the morning and have to face another day on their jobs. Therefore, hospital management can place Questions 2, 4, 5, 6, and 7 in a higher priority to reduce nurses' burnout.

Table 5. ANOVA results of nine questions in emotional exhaustion from 2014 to 2020

Question	F	Sig.	Post hoc analysis
1. I feel like I'm at the end of my rope.	9.556	< .001	2014 > 2016; 2014 > 2017; 2014 > 2018; 2014 > 2019; 2014 > 2020; 2015 > 2016; 2015 > 2017; 2015 > 2018; 2015 > 2019; 2015 > 2020
2. I feel burned out from my work.	19.263	< .001	2016 > 2014; 2016 > 2015; 2017 > 2014; 2017 > 2015; 2018 > 2014; 2018 > 2015; 2019 > 2014; 2019 > 2015; 2020 > 2014; 2020 > 2015
3. I feel frustrated by my job.	8.072	< .001	2014 > 2017; 2014 > 2018; 2014 > 2019; 2014 > 2020; 2015 > 2016; 2015 > 2017; 2015 > 2018; 2015 > 2019; 2015 > 2020
4. I feel I'm working too hard on my job.	48.848	< .001	2016 > 2014; 2016 > 2015; 2017 > 2014; 2017 > 2015; 2018 > 2014; 2018 > 2015; 2019 > 2014; 2019 > 2015; 2020 > 2014; 2020 > 2015
5. I feel emotionally drained from my work.	20.774	< .001	2016 > 2014; 2016 > 2015; 2017 > 2014; 2017 > 2015; 2017 > 2020; 2018 > 2014; 2018 > 2015; 2018 > 2020; 2019 > 2014; 2019 > 2015; 2020 > 2014; 2020 > 2015
6. I feel used up at the end of the workday.	90.754	< .001	2016 > 2014; 2016 > 2015; 2017 > 2014; 2017 > 2015; 2017 > 2020; 2018 > 2014; 2018 > 2015; 2019 > 2014; 2019 > 2015; 2020 > 2014; 2020 > 2015
7. I feel fatigued when I get up in the morning and have to face another day on the job.	8.43	< .001	2016 > 2014; 2017 > 2014; 2017 > 2015; 2018 > 2014; 2018 > 2015; 2019 > 2014; 2019 > 2015

8. Working with people all day is really a strain for me.	47.020	< .001	2014 > 2016; 2014 > 2017; 2014 > 2018; 2014 > 2019; 2014 > 2020; 2015 > 2016; 2015 > 2017; 2015 > 2018; 2015 > 2019; 2015 > 2020
9. Working with people directly puts too much stress on me.	84.303	< .001	2014 > 2016; 2014 > 2017; 2014 > 2018; 2014 > 2019; 2014 > 2020; 2015 > 2016; 2015 > 2017; 2015 > 2018; 2015 > 2019; 2015 > 2020

5. Conclusions

Studies such as Shanafelt et al. [17], Shanafelt et al. [19], and West et al. [20] showed that physicians' burnout can increase negative impacts including medical errors and the decrease of patient care. Cimiotti et al. [21] further stated that nurses' burnout results in a higher infection rate and a poor patient care quality. As a result, assessing nurses' emotional exhaustion is essential in healthcare organizations. This study uses a longitudinal approach (from 2014 to 2020) to analyze nurses' perceptions in emotional exhaustion based on the Chinese version of the SAQ from a regional teaching hospital in Taiwan. The results show that four of nine questions in emotional exhaustion have been improved statistically based on post hoc analysis. However, five of nine questions in emotional exhaustion have been deteriorating, including 2. I feel burned out from my work, 4. I feel I'm working too hard on my job, 5. I feel emotionally drained from my work, 6. I feel used up at the end of the workday, and 7. I feel fatigued when I get up in the morning and have to face another day on the job. Therefore, hospital management needs to focus on these five items to reduce nurses' emotional exhaustion in this regional teaching hospital.

References

1. J. B. Sexton, R. L. Helmreich, T. B. Neilands, K. Rowan, K. Vella, J. Boyden, P. R. Roberts, and E. J. Thomas, "The safety attitudes questionnaire: Psychometric properties, benchmarking data, and emerging research," BMC Health Services Research, vol. 6, pp. 44, 2006.
2. C. Y. Chi, C. H. Huang, Y. C. Lee, C. F. Wu, and H. H. Wu, "Understanding critical demographic variables affecting patient safety culture from nurses' viewpoints," The TQM Journal, vol. 32, no. 3, pp. 429-440, 2020.
3. Y. F. Li, Y. C. Lee, J. I Shieh, C. H. Huang, K. K. Huang, C. H. Yu, and H. H. Wu, "Studying patient safety from viewpoints of physicians and nurses: A case of a district hospital in central Taiwan in 2017," Journal of Quality, vol. 28, no. 5, pp. 322-348, 2021.
4. C. Y. Chi, C. H. Huang, Y. C. Lee, and H. H. Wu, "Critical demographic variables on affecting patient safety culture from medical staffs' viewpoints," Engineering Letters, vol. 27, no. 2, pp. 328-335, 2019.
5. J. I Shieh, C. H. Huang, Y. C. Lee, and H. H. Wu, "Using structural analysis to construct causal relationships of the patient safety culture," Engineering Letters, vol. 27, no. 3, pp. 482-489, 2019.
6. Y. C. Lee, J. I Shieh, C. H. Huang, C. Y. Wang, and H. H. Wu, "Analyzing patient safety culture from viewpoints of physicians and nurses - A case of a regional teaching hospital in Taiwan," Journal for Healthcare Quality, vol. 39, no. 5, pp. 294-306, 2017.
7. Y. C. Lee, C. H. Huang, C. F. Wu, H. W. Hsueh, and H. H. Wu, "A longitudinal study of identifying critical variables influencing patient safety culture from nurses' viewpoints in Taiwan," Journal of Testing and Evaluation, vol. 47, no. 5, pp. 3387-3398, 2019.
8. C. Y. Cheng, Y. C. Lee, C. H. Huang, and H. H. Wu, "Assessing nurses' overall satisfaction of patient safety culture from a regional teaching hospital in Taiwan," International Journal of Industrial and Systems Engineering, vol. 36, no. 4, pp. 537-548, 2020.

9. C. Y. Cheng, Y. C. Lee, C. H. Huang, and H. H. Wu, "Assessing the medical staff's overall satisfaction of patient safety culture from a regional teaching hospital in Taiwan," Journal of Testing and Evaluation, vol. 48, no. 6, pp. 4153-4162, 2020.
10. C. H. Huang, H. H. Wu, Y. C. Lee, I. Van Nieuwenhuysea, M. C. Lin, and C. F. Wu, "Patient safety in work environments: Perceptions of pediatric healthcare providers in Taiwan," Journal of Pediatric Nursing, vol. 53, pp. 6-13, 2020.
11. J. B. Sexton, S. P. Schwartz, W. A. Chadwick, K. J. Rehder, J. Bae, J. Bokovoy, K. Doram, W. Sotile, K. C. Adair, and J. Profit, "The associations between work-life balance behaviours, teamwork climate and safety climate: Cross-sectional survey introducing the work-life climate scale, psychometric properties, benchmarking data and future directions," BMJ Quality & Safety, vol. 26, no. 8, pp. 632-640, 2017.
12. M. H. Lee, Y. C. Lee, C. H. Huang, C. F. Wu, H. W. Hsueh, and H. H. Wu, "A longitudinal study of nurses' work-life balance: A case of a regional teaching hospital in Taiwan," Applied Research in Quality of Life, https://doi.org/10.1007/s11482-020-09881-8, 2020.
13. C. Y. Chi, H. H. Wu, C. H. Huang, and Y. C. Lee, "Using linear regression to identify critical demographic variables affecting patient safety culture from viewpoints of physicians and nurses," Hospital Practices and Research, vol. 2, no. 2, pp. 47-53, 2017.
14. C. H. Huang, H. H. Wu, and Y. C. Lee, "The perceptions of patient safety culture: A difference between physicians and nurses in Taiwan," Applied Nursing Research, vol. 40, pp. 39-44, 2018.
15. C. H. Huang, H. H. Wu, C. Y. Chou, H. Dai, and Y. C. Lee, "What we should know about patient safety culture: The perceptions of physicians and nurses in a case hospital," Iranian Journal of Public Health, vol. 47, no. 6, pp. 852-860, 2018
16. Y. C. Lee, C. H. Huang, D. H. Rahman, C. H. Yu, and H. H. Wu, "A case study of analyzing emotional exhaustion from a medical center in Taiwan," Jurnal Kajian Bimbingan dan Konseling, vol. 5, no. 1, pp. 1-8, 2020.
17. T. D. Shanafelt, L. N. Dyrbye, and C.P. West, "Addressing physician burnout: The way forward," JAMA, vol. 317, no. 9, pp. 901-902, 2017.
18. C. Maslach, M. P. Leiter, and G. Fink, G. Stress: Concepts, Cognition, Emotion, and Behavior. Handbook of Stress Series, 1, 351-357, 2016.
19. T. D. Shanafelt, C. M. Balch, G. Bechamps, T. Russell, L. Dyrbye, D. Satele, P. Collicott, P. J. Novotny, J. Sloan, and J. Freischlag, "Burnout and medical errors among American surgeons," Annals of Surgery, vol. 251, no. 6, pp. 995-1000, 2010.
20. C. P. West, L. N. Dyrbye, P. J. Erwin, and T. D. Shanafelt, "Interventions to prevent and reduce physician burnout: A systematic review and meta-analysis," The Lancet, vol. 388, no. 10057, pp. 2272-2281, 2016.
21. J. P. Cimiotti, L. H. Aiken, D. M. Sloane, and E. S. Wu, "Nurse staffing, burnout, and health care–associated infection," American Journal of Infection Control, vol. 40, no. 6, pp. 486-490, 2012.

Objective Difficulties in Evaluating "Covid" Statistics

Olga V. Pugacheva

Francisk Skorina Gomel State University, Gomel, Belarus
OPugacheva@gsu.by

Abstract. The article discusses important features of the statistics of the distribution of Covid-19 around the world. The reasoning is based on the well-known data of operational information on the countries, as well as the comments of experts. Particular attention is paid to countries with a large economy. This is the name of the country, for each of which GDP was more than 1% of world GDP for at least one year from 1980 to 2019. The World Health Organization is the only organization that could establish common formats and data requirements for all states with different health system organizations. The article shows how great the difference in the formation of data on a pandemic in different countries is. This makes it difficult to compare situations using standard statistical processing.

Keywords. Demographic statistics, operational information, tests redundancy, excess mortality rate.

1. Introduction

The widely available and fairly detailed daily statistics on the spread of the COVID-19 pandemic across the countries of the world in 2020-21 had a great impact on human behavior. In Russia, daily statistics came from the Information Center for Monitoring the Situation with Coronavirus (ICC), created in March 2020 to collect and analyze information on the development of the situation with the spread of COVID-19 infection. Such information has enabled health systems in all countries to promote the necessary culture of behavior in the context of the global epidemic. With the rapid growth of those infected with a new and practically unexplored virus, doctors and all medical personnel made tremendous efforts to provide possible assistance to those who fell ill. In almost all countries, at the initial stage, there was a need for additional equipment and medicines.

Note that about a year before the pandemic was announced, the Global Health Security Index for 195 countries (GHSI) was published This Index was created by a solid group of international experts [1]. The Index is the result of a joint project by three organizations: the Johns Hopkins Health Security Center, the Nuclear Threat Reduction Initiative (NTI) and the Economist Intelligence Unit (EIU). The general conclusion of the invited panel of experts is: "The indicators show serious deficiencies in the ability of countries to prevent, detect and respond to disease outbreaks." And, unfortunately, this general conclusion was fully confirmed by the development of the situation with the COVID-19 pandemic.

A detailed analysis of the achievements and mistakes in the struggle for the health of citizens in extreme conditions is the field of activity of experimental scientists and medical practitioners. They gained invaluable experience, which will undoubtedly form the basis for subsequent practical recommendations. In this work only issues related to the statistical description of such processes are considered. There are several such questions.

2. Registration of Infected with COVID-19

The daily registration of the infected, the recovered and the dead, as well as the collection of this information from many medical units and the further publication of the resulting data on the Internet - all this turned out to be quite a difficult task. Technical information problems were largely resolved within a few months. But still, the harsh impossibility of correcting mistakes made in the past when collecting data from various sources sometimes led to the fact that negative values appeared in the reports, while by definition, these values could only be positive or equal to zero. We are talking about the daily number of people infected, recovered and died.

Recall that modern international statistical standards used during a pandemic include daily information on five indicators for countries (cities, regions). One of them - Inf (Infected with coronavirus) - characterizes the total number of people infected with coronavirus on a given day on a cumulative basis. And the other value - $dInf$ - characterizes the number of confirmed cases of infection on the given day, respectively. The relationship between the values of the indicated quantities for two consecutive dates $t-1$ and t is determined by the ratio.

$$dInf(t) = Inf(t) - Inf(t-1) .$$

Three other indicators determine the further detailing of the infected into three groups: patients with Ac (Active cases), deceased D (Deaths) and recovered R (Recovered). Of course with this method of keeping statistics, the D value does not mean the number of deaths due to the coronavirus. Obviously, on every fixed day, the identity

$$Inf(t) = Ac(t) + D(t) + R(t)$$

is realized.

Thus, in order to get into the published statistics, it is necessary to obtain the infected status. However, on the recommendation of WHO, a person receives such a status if his situation refers to the so-called confirmed case, when "there is laboratory confirmation of the presence of COVID-19 infection, regardless of the clinical manifestations and symptoms of the disease." And this means that if a person for some reason did not pass the test, or if he received a false-negative test result, and, unfortunately, died, then he does not fall into the global statistics on coronavirus.

Today, polymerase chain reaction (PCR) tests are used to determine whether a person is sick with COVID-19. The widespread use of such tests in the world allowed a group of scientists from the Johns Hopkins School (USA) to evaluate how the accuracy of the tests changes depending on the number of days since infection with COVID-19. The researchers analyzed seven articles with clear data on the time interval between infection, symptom onset, and testing. After excluding unreliable data, the study included 1.3 thousand test samples of patients who were sick with COVID-19. The results are as follows [2]. In the first three days, almost all tests in SARS-Cov-2 carriers were negative. On the fourth day, a third of those infected received a positive test result. On the fifth day, the first symptoms of COVID-19 usually appear - and here almost two-thirds of the tests have shown the truth. After another three days, only a fifth of the results were false, and then the probability of the test's truthfulness began to fall. Three weeks after the onset of infection, two-thirds of the tests were negative.

However, even earlier, on April 8, 2020, the Moscow Health Department warned the capital's medical institutions about a large number of false-negative PCR test results and

recommended sending patients with pneumonia for computed tomography if complications were suspected, calling the combination of CT and the clinical picture of the disease the only reliable analysis. Deputy Director of the Research Institute of Pulmology of the Russian Academy of Sciences, Doctor of Medical Sciences SN Avdeev referred to the results of a study involving 1.4 thousand patients. 300 of them had a negative laboratory test for SARS-CoV-2, but the results of computed tomography of the lungs showed changes characteristic of the new coronavirus infection [3].

In a later review dated 03.10.2020 [4], the director of the Institute of Medical Parasitology, Tropical and Vector-borne Diseases of the Sechenov University, Alexander Lukashev, explained that the technical sensitivity of the PCR test is usually higher than 98%. But its practical sensitivity in patients with COVID-19 can be 60 and 70 percent. This is possible both for physiological reasons - there is simply no virus in the oropharynx, and for technical reasons: this smear itself is not so easy to take. This requires a certain preparedness of both the doctor and the patient. And these problems are the same for the tests of any manufacturer around the world. "The pathogen usually first multiplies in the oropharynx and then descends into the lungs, affecting them. Therefore, at a later date, it is more difficult to detect it in the oropharynx, where the smear is taken from."

Table 1 contains data for 26 countries [5]: 24 countries belong to countries with a large economy [6], and two countries (Belarus and Israel), as representatives of small countries with approximately the same population. Countries are sorted in descending order by the fourth column - the amount of test redundancy as of 20.05.2021. Here, the redundancy of tests is understood as the value of the ratio of the total number of tests to the total number of registered infected. The last column shows data on vaccination with at least one dose of vaccine as of 25.06.21

The second and third columns of data are provided to illustrate the economic and human potential of the countries in question. The second column shows the share of the country's GDP in total global GDP for 2019, according to the International Monetary Fund. This data is referred to as wGDP2019 in the Table 1.

The number of tests among the countries under consideration varies greatly. Of course, the recommendations on the need for laboratory tests to declare the number of infected in general are being followed: as of 20.05.2021, the redundancy of tests lies in the range of 2.9 (Mexico) - 1759.8 (China). The spread in the number of tests per million population is also large: from 5.26×10^4 (Mexico) to 2.52×10^6 (Great Britain).

Table 1. Country data on the number of tests performed, the number of infected and the percentage of the population vaccinated

Country	wGDP 2019, %	Population, million people.	Tests redundancy (Tests/Inf)	Number of tests per million people	Number of infected per million people	% of the population vaccinated on 25.06.21
China	16,29	1404,44	1759,79	$1,11*10^5$	64,7	43,21
Australia	1,62	25,97	589,95	$6,85*10^5$	1154,9	23,79
Korea	1,90	52,08	70,44	$1,84*10^5$	2575,2	29,82
Saudi Arabia	0,87	34,54	41,80	$5,16*10^5$	12668,5	
United Kingdom	3,24	67,25	38,69	$2,52*10^6$	66208,6	64,93

Russia	1,85	146,75	27,06	$9,20*10^5$	33900,6	14,82
Canada	1,99	37,82	25,14	$8,85*10^5$	35494,1	67,52
Japan	5,93	125,75	19,27	$1,05*10^5$	5552,7	20,21
Israel	0,44	8,71	17,30	$1,57*10^6$	88296,8	63,98
Belarus	0,07	9,51	16,13	$6,49*10^5$	43805,5	7,40
Germany	4,54	82,96	15,99	$6,91*10^5$	43788,0	52,95
Italy	2,32	60,72	15,27	$1,05*10^6$	68811,9	54,77
USA	24,46	331,81	13,87	$1,41*10^6$	101881,7	54,04
France	3,16	65,32	13,86	$1,25*10^6$	90590,9	50,47
Spain	1,64	46,86	13,47	$1,05*10^6$	77500,2	52,03
Belgium	0,61	11,53	12,92	$1,15*10^6$	90070,6	59,58
India	3,41	1369,56	12,39	$2,30*10^5$	19000,6	18,86
Switzerland	0,81	8,69	10,82	$8,48*10^5$	78958,8	48,53
Turkey	0,81	84,04	10,02	$6,05*10^5$	61404,4	38,22
Indonesia	1,26	269,86	9,00	$5,70*10^4$	6517,8	9,52
Sweden	0,63	10,42	8,99	$9,34*10^5$	101264,2	44,63
Netherlands	1,05	17,29	8,37	$7,88*10^5$	93435,5	54,39
Iran	0,56	84,15	6,61	$2,17*10^5$	33329,0	4,60
Argentina	0,55	45,55	3,73	$2,77*10^5$	74888,3	34,55
Brazil	2,25	211,21	2,98	$2,20*10^5$	74879,0	32,08
Mexico	1,42	127,09	2,88	$5,26*10^4$	18786,0	23,15

Source: number of tests for coronavirus by country: URL: https://covid-stat.com/ru/kolichestvo-testov-koronavirus/; number of infected per 1 million people: URL: https://coronavirus-monitor.info/country/russia/; number of vaccinated from coronavirus in the world, there is no data on vaccination in Saudi Arabia: URL: https://gogov.ru/covid-v-stats/world.

The first three leaders in terms of test redundancy (China, Australia, the Republic of Korea) are also leaders in the lowest number of infected per 1 million population. Japan demonstrates moderate redundancy, however, in terms of the number of infected, it is closest to the aforementioned leaders. For the rest of the countries, there is no clear dependence of the number of infected people on the value of test redundancy. Although statistically, the correlation between the redundancy of tests and the number of infected is nonzero at the 5% significance level and is -0.39.

Moreover, it is easy to see that the difference between countries in the data on the number of infected per 1 million of the population is large even with close values of test redundancy. These facts suggest that a clear effect on the number of tests is achieved with redundancies greater than 70. Perhaps the cause and effect can be reversed: namely, having achieved a small amount of infection by strict local quarantine measures, countries can afford to conduct a large number of tests.

The level of vaccination in countries is not statistically related to the number of tests, but the correlation with the number of infected is 0.62 and is determined not only by the availability of vaccine in the required amount, but also by the attitude of the population towards vaccination and the persistence of the country's leadership. There is another difficulty in detecting the real number of infected. Already in April 2020, an article appeared by a large group of Chinese scientists [7], who reported that the highest concentration of viral particles was observed in patients in whom symptoms were just beginning to appear and the peak of infection occurred at this time or even earlier. About 44% of patients were infected from people who have not yet shown any symptoms of the disease. Moreover, most often at this time they were among relatives or outside the home.

To find out the proportion of asymptomatic transmission of the virus, scientists do not randomly test any potentially infected people in the population, but analyze the chains of transmission statistically. The researchers analyzed the course of the disease in 94 patients with confirmed coronavirus infection (none of them were critically ill). Throat swabs were taken from them within 32 days of the onset of symptoms. It turned out that the peak in the concentration of the virus in the mucous membranes begins approximately at the time of the onset of the first symptoms of the disease, and then decreases and practically disappears by about the 21st day. This was not influenced by gender, age, or the severity of the person's illness.

The constructed mathematical model showed that people begin to be infectious, on average, 2-3 days before the onset of symptoms, and the peak of the ability to infect occurs in the period from 14 to 16 hours from the moment of the first manifestation of symptoms. Further, within a week, the ability to infect the virus decreases until it disappears completely.

These data cannot be considered completely complete. For example, the moment of onset of primary symptoms was noted retrospectively: the patients themselves spoke about this after they were diagnosed. In addition, patients received treatment, which could affect the concentration of the virus in the mucous membranes, and, accordingly, the data on the rate of its decrease could also be underestimated relative to the drug-free variant of the development of events.

So, we can fix the following factors that make the daily published data about the number of infected, objectively not reliable enough:

- difficulties in organizing daily collecting data from numerous medical organizations;
- rather vague criteria for including citizens who have applied for help in the category of infected with the COVID-19 virus, taking into account the possibility of identifying the virus by tests;
- the existence of a fairly large number of asymptomatic patients who do not even go to doctors.

If the first factor presents mainly difficulties for building a mathematically adequate model according to these time series and ultimately is integrally corrected by correcting fictitious negative values4, then two other factors can lead to a significant decrease in the number of deaths from among those infected with the COVID-19 virus. In addition, these factors make it difficult to estimate the number of people who are already immune and do not

yet need vaccination. Examples include the opinions of people who are largely responsible for information about the pandemic in their countries.

The first example. US chief infectious disease specialist Anthony Fauci believes that the number of deaths from coronavirus in the country is underestimated. According to Johns Hopkins University, more than 581 thousand people died from COVID-19 in the United States (as of 23.04.2021), the country ranks first in the number of cases (over 32.7 million people). However, according to a study by the University of Washington, the number of deaths from the virus in the United States may exceed 900 thousand people (which is 54% more than official current statistics).

The second example. The head of Rospotrebnadzor Anna Popova on September 28, 2020 on the air of the Komsomolskaya Pravda radio said that 24-25% of Muscovites have antibodies to coronavirus [8]. However, in Moscow, according to official data, on this day, 287 993 infected with the population of Moscow about 12.647 million. Consequently, the officially registered infected were about 2.3% of the population.

In many countries, there have been many such assessments from officials. But it is difficult to find out the true picture of the number of people infected with the COVID-19 virus, as the only official source of this information is the data provided by each country. And this data is practically unrelated to other indicators not included in the operational summary. The situation with mortality data is somewhat different.

3. Mortality Statistics

At the initial stage of the pandemic, the lack of approved protocols for the treatment of a new disease required constant experimentation from doctors and the involvement of various specialists for consultation. The rapid spread of the pandemic to large segments of the population required the mobilization of the entire health system. This led, in many cases, to postponement of routine care for other chronic diseases, which also led to additional deaths that could have been avoided in the absence of a pandemic. Back in April 2020, Academician V.M. Polterovich [9]:

“Most recent data on mortality in small Italian towns show that it has increased enormously, and this increase is several times higher than the death toll from coronavirus according to official figures. The Economist magazine reports a similar situation in a French city and in three Spanish ones. This can be partially explained by the underestimation of deaths from the coronavirus, but there are other reasons as well. The measures taken against the pandemic cause panic, deprive people of the possibility of regular treatment, habitual eating, walking, playing sports, communicating with friends and family. Alcohol consumption rose sharply in Russia, especially in the capital, in March: constrained by the need to drive, people are resorting to standard means of suppressing fear and melancholy. And there is a series of stresses ahead as a result of large-scale unemployment and a decrease in income. By reducing deaths from coronavirus, we unwittingly contribute to the death of citizens from other causes. The longer the "quarantine" regime lasts, the greater losses (including in human lives) we will suffer from the upcoming economic recession. For this reason, political, unconfirmed declarations demanding the Russian government to immediately introduce a nationwide quarantine and immediately spend all the reserves on helping citizens do not make me sympathetic. The harm from them - fear mongering - is obvious, but the benefits are questionable."

On the Internet, in some countries, you can find a lot of criticism about the published figures on the number of deaths among those infected. At the same time, as a rule, the authors had in mind precisely the fact that the deceased, who had symptoms of COVID-19, did not always fall into the "like" statistics, the openness and publicity of which had no previous

analogue. But these new qualities of the "like" statistics made it possible to notice that the number of deaths in these statistics does not always correspond to the statistics of total mortality, which is also available in many (but, unfortunately, not in all) countries.

In many countries, statistics on total mortality are in a weekly or monthly format and are published with some delay. The level of detail of these data varies. Eurostat has adopted a weekly format disaggregated by sex and age. At the same time, official national mortality statistics are provided weekly from 29 European countries or subnational regions within the framework of the joint EuroMOMO (European mortality monitoring) network. This network is supported by the European Center for Disease Prevention and Control (ECDC) and the World Health Organization (WHO) and is hosted by the Statens Serum Institut, Denmark.

Some countries publish periodically a multi-year compendium of all fatalities by week, month, by area of registration, sex and age. In particular, such statistics for Germany are available on the Internet [10].

Regular information on total mortality from African countries, as well as from India, Indonesia, China, Saudi Arabia, Turkey, is generally unavailable. In Russia, Rosstat publishes monthly statistics showing reported deaths by cause of death, but with a delay of about 1.5 months. Statistics are presented by regions and three cities (Moscow, St. Petersburg, Sevastopol), but there is no breakdown by sex and age. For 2020, information on causes of death not clearly related to the pandemic is not provided, but there is monthly information on deaths associated with COVID-19. Information is presented for the country as a whole, for regions and three cities [11].

This statistic includes cases from four causes of death:

1) COVID-19, the virus has been identified;

2) COVID-19 is possible, but the virus has not been identified;

3) COVID-19 is not the main cause of death, but had a significant impact on the development of fatal complications of the disease;

4) COVID-19 is not the main cause of death, and did not have a significant impact on the development of fatal complications of the disease.

A detailed analysis of the distribution of the increase in mortality by regions of Russia for the first year of the pandemic was carried out in the work of E.M. Andreev [12].

In the previous section, it was shown that the data on those infected with the COVID-19 virus is difficult to verify. But it turned out that the data on the number of deaths from COVID-19 is not a very adequate indicator, because in different countries the deaths from COVID, COVID and presumably from COVID are taken into account in different ways, even the WHO recommendations on this about somewhat contradictory. In the article by I. Danilova back in the spring of 2020, it was shown that the published data on the number the number of infected and deaths may not be comparable between countries, as countries use different criteria for both testing for the virus and determining deaths from COVID-19 [13].

The first statements appeared that the mind the indicator in this case can be a great rank of excess mortality. And now many see her as the only one adequate overall metric. Sometimes it is even called the "gold standard". Various options are offered for finding this indicator. The simplest way to calculate excess mortality is the difference between do monthly data on total mortality of the studied year and the corresponding data of the previous year. We denote such a model calculating excess mortality as M1. Consider the situation in Russia for 2015–2021, using data from Rosstat [11].

In Figure 1 shows graphs of total mortality in Russia for 2018–2021.

From the data in Figure 1 it follows that the total mortality in Russia from April 2020 to March 2021 significantly exceeds the total mortality for previous 2015–2019, and total mortality for 2015–2019 differs little over the years, but has a general downward trend. Let's

calculate excess mortality in 2020 in the simplest models as the difference between the values of the total mortality 2020 and 2019 (model M1) and compare it with daily published data on the coronavirus of the Information Center for Monitoring the Situation with Coronavirus (ICC) for 2020 and detailed data from Rosstat, indicating the impact of the COVID-19 virus on some of the total reported deaths.

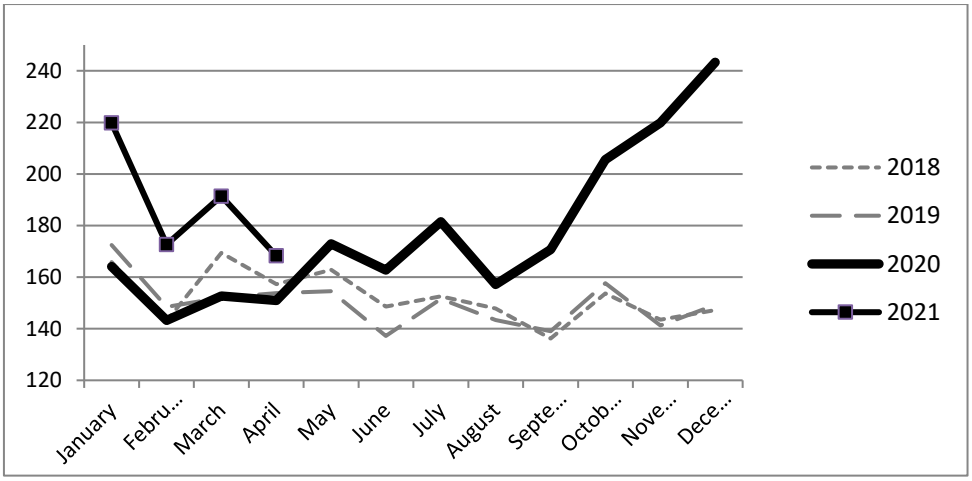


Fig.1. The total number of deaths in Russia (thousand people) by months for 4 years

Without going into details of the data by month, we will focus only on the resulting data for 2020, given in the line ≪Total for 2020 ≫ of Table 2. Excess mortality for the year according to the M1 model is 323.8 thousand people. At the same time, according to the operational statistics of the ICC (the last column of Table 2), only 57 thousand are shown. And this naturally comparison gives rise to many unpleasant questions. Of course, the excess mortality data includes all sorts of causes of death, but Rosstat reports on the number of deaths related to COVID-19 (columns 1-4 of table 2 in the line "Total for 2020") somewhat clarified the situation.

It became clear that of the 323.8 thousand people included in the excess mortality, only 162.4 thousand can be directly related to COVID-19, as the sum of columns 1 + 2 + 3 + 4.

Table 2. Excess mortality, mortality with the presence of covid according to Rosstat data and operational statistics of the ICC for 2020 and three months of 2021 (thousand people)

Month	Excess mortality	Cause of death				Sum 1	Sum 2	Operational statistics
2020	M1	1	2	3	4	1+2+3+4	1+3+4	Internet
January	-8,4	0	0	0	0	0	0	0
February	-5,2	0	0	0	0	0	0	0
March	0,7	0	0	0	0	0	0	0,017
April	-2,9	1,35	0,398	0,435	0,642	2,825	2,427	1,056
May	18,3	5,926	1,677	1,609	3,457	12,669	10,992	3,62

June	25,5	5,825	1,492	1,484	3,534	12,335	10,843	4,627
July	29,9	5,063	1,021	1,237	3,05	10,371	9,35	4,643
August	13,8	3,436	0,582	1,184	2,471	7,673	7,091	3,213
September	31,7	4,579	0,859	1,428	3,313	10,179	9,32	3,546
October	47,8	13,077	2,026	1,794	7,436	24,333	22,307	7,268
November	78,5	21,262	3,845	2,288	10,214	37,609	33,764	11,905
December	94,1	25,98	5,57	2,065	10,82	44,435	38,865	17,124
Total for 2020	323,8	86,498	17,47	13,52	44,937	162,429	144,96	57,019
2021								
January	47,319	22,747	4,708	2,337	8,086	37,878	33,17	16,163
February	24,181	14,791	2,329	1,956	5,708	24,784	22,455	12,94
March	39,4	15,003	2,454	1,401	4,857	23,715	21,261	12,728
Total for 04.20 - 03.21	448,3	139,04	26,961	19,22	63,588	248,806	221,85	98,85

Legend in Table 2: M1-model of excess mortality as the difference between the data for 2020 (2021) and 2019 from the data in Figure 1; column 1 - COVID-19, virus identified; column 2 - possibly COVID-19, but no virus identified; column 3 - COVID-19 is not the main cause of death, but had a significant impact on the development of fatal complications of the disease; column 4 - COVID-19 is not the main cause of death, and did not have a significant impact on the development of fatal complications of the disease; column Sum1 - the sum of columns 1-4; column Sum2 - the sum of columns 1,3,4; the Internet column is from daily ICC data. All data are in thousand people.

This represents approximately 50% of excess mortality. However, in the deaths listed in column 2, COVID-19 was not identified and therefore, formally, only the sum of the columns (1 + 3 + 4) should have been included in the operational statistics, and this is only 144 thousand. But only 57 thousand are reflected in the operational information.

Two conclusions follow from these comparisons. First, the ICC data were 2.5 times underestimated. And secondly, the aforementioned prediction of Academician V.M. Polterovich about the possible impact of restrictive measures, expressed by him back in April 2020, unfortunately, was realized: about half of the excess mortality was not directly associated with the virus. And this discrepancy began in June 2020.

If we correlate the numbers M1 and ICC in the line <<Total for 2020≫ of Table 2, it turns out that the excess mortality in Russia as a result of the pandemic for 2020 is 5.7 times exceeds the ICC data. It is this fact and determined the title of D. Kobak's article “Excess mortality shows the true losses of Russia from Covid-19”, published in the journal “Royal Statistical Society” [14]. The last row of Table 2 characterizes the situation with data on deaths for the year from April 2020 to March 2021. During that time excess mortality according to the M1 model is 448 thousand (rounded off). To cases associated with COVID-19, column data can be referred to 1 + 2 + 3 + 4, which is 248 thousand (55% of the excess mortality). However, only 221 thousand of them had the virus confirmed. But again, according to data on deaths from ICC, only 99 thousand cases were confirmed, that is, 2.24 times less. Thus, we

can state that the ratio of excess mortality to the operational data of the ICC began to change rapidly after January 2021 and for the annual period from April 2020 to March 2021 was already only 4.5 instead of former 5.7.

Is this situation common to many countries? No, it is not, as the example shows Germany. Calculations show [5] that excess mortality in Germany for 2020 is 37% higher than the mortality data from the operational reports, but for the annual interval (April.2020–March 2021) it is only 82% of the operational data. In other words, German statistics are likely to not only cover all COVID-19 cases, but also indicate an increase in deaths from non-COVID-19 causes during 2020, and vice versa, a decrease in deaths from other causes in during the first three months 2021

There are different methods for calculating the value of excess mortality. The general calculation scheme is based on comparing the mortality rate of the studied year with the predictive mortality model - the so-called baseline. The latter is based on the analysis of mortality for several previous years. Model M1, which was used here for Russia and Germany, assumed that the predictive model of mortality for 2020 is mortality in 2019.

With information on total mortality in Russia over the previous five years (see Figure 1), it is possible to build a baseline for 2020 with fixed effects model that used constructing linear regression for panel data [5, 15, 16]. Let's call it model M2. The advantage of the M2 model over the model M1 is the use of statistical information over several years, which allows you to establish a confidence interval for the values of predicted mortality.

A more sophisticated method for calculating excess mortality in 2020 is based on an analysis of mortality over several previous years with the availability of weekly statistics and a smooth baseline curve for 2020. This method was used by the Institute for Health Metrics and Evaluation based at Washington State University in Seattle [17]. Obviously, the obtained values of excess mortality will depend on the applied calculation method, but most often the results of different methods differ within the calculated error.

4. Conclusion

Comparing data on today's pandemic in different countries is significantly different from comparing data from various laboratory experiments on which all knowledge in the scientific and technical field of knowledge is based. And this difference is due to the absence of uniform standards for recording data in different countries. And as a consequence of this, great difficulties arise in the analysis of these data by conventional statistical methods, which have proven themselves in the natural - scientific field. Therefore, the use of such methods, with such a wide variety of methods for generating data on a pandemic, can lead to doubtful conclusions.

References

1. https://www.ghsindex.org/about/.
2. Kucirka L.M. et al. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure [Preprint] 2020.
3. https://medvestnik.ru/content/news/Chastota-lojnotricatelnyh-rezultatov-testov-na-SARS-GoV-2-20-30.html.
4. https://iz.ru/1080158/olga-kolentcova/odnim-mazkom-test-na-covid-19-pravdivee-na-vosmoi-den-posle-zarazheniia.
5. Chetverikov V.M., Pugacheva O.V., Vorontsova T.D. Challenges to Generating Reliable COVID Statistics:Domestic and International Experience. *Voprosy Statistiki*. 2021;28(4): 45-66.
6. Chetverikov V.M. Unique Features and Intensity of COVID-19 Spread in Large Economies. *Voprosy Statistiki*. 2020;27(6):86–104.

7. He X**.** et al. Temporal Dynamics in Viral Shedding and Transmissibility of COVID-19. *Nature Medicine*. 2020;26: 672–675.
8.: https://www.rbc.ru/rbcfreenews/5f71f0b49a794764a1996fca.
9. Polterovich V.M. What will the Russian economy have after returning to normal life. https://www.kommersant.ru/nauka/131713.
10. Sonderauswertung Sterbefalle 2016 bis 2021 (Stand: 31.5.2021). URL:https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Sterbefaelle-Lebenserwartung/Tabellen/sonderauswertung-sterbefaelle-pdf.pdf?__blob=publicationFile.
11. Natural movement of the population in the context of the constituent entities of the Russian Federation L: https://rosstat.gov.ru/storage/mediabank/D4uPozCj/edn05-2021.htm.
12. Andreev E.M. The first year of the covid-19 pandemic in Russia (April 2020 – March 2021). URL: https://demogr.hse.ru/seminar.
13. Danilova I**.** Morbidity and Mortality from COVID-19. The Problem of Data Comparability. *Demographic Review*. 2020; 7(1):6–26. (In Russ.)
14. Kobak D**.** Excess Mortality Reveals Covid's True Toll in Russia. Significance 2021; 18(1):16-19. Available from: https://doi.org/10.1111/1740-9713.01486.
15. Balash V.A., Balash O.S. *Models of Linear Regression for Panel Data. Textbook.* Moscow: MESI; 2002. (In Russ.)
16. Ratnikova T.A., Furmanov K.K. *Panel Data Analysis and Data on the Duration of States. Textbook.* Moscow: HSE Publishing House; 2014. (In Russ.)
17. Karlinsky A., Kobak D. The World Mortality Dataset: Tracking Excess Mortality Across Countries During the COVID-19 Pandemic. medRxiv [Preprint]. Available from: https://doi.org/10.1101/2021.01.27.21250604.

State Cybersecurity Governing Principles and Structures

Valeriu Cernei

Academy of Economic Studies of Moldova (AESM), Chisinau, Moldova
Valeriu.Cernei@bsd.md

Abstract. In the new reality, the Governments face significant difficulties in managing cyber risks. They have to focus on both, external relations and geopolitical aspects as well as on the internal aspects of cyber security. The risk management methodologies have to be reviewed and rebuilt in order to align with the post pandemic reality: a reality where the attack vector is switching from Critical Infrastructure to end-user. There is a need to re-consider information and cyber security and make sure that every citizen is protected against cyber-attacks as it states the Constitution. By protecting citizen, the Government will protect, indirectly, its State Critical Infrastructure and other assets. To ensure a secure environment states have to implement efficient cybersecurity risk governing bodies based on relevant principles. Every state is to select the best governing approaches and principles that apply at a moment in time and that are in line with Cyber Security Strategy. By selecting them, states have to take into consideration a number of characteristics like country culture, dimensions, regional position and neighbors, the level of public services digitalization, and so on. The paper reviews approaches and trend from different countries in terms of cybersecurity governance and concludes on the approaches that may be implemented by Republic of Moldova in order to ensure the implementation of a cyber security strategy, based on strong and continuous cybersecurity risk activities.

Keywords. Risk, Cybersecurity, Model, Management, Government

1. Introduction

For the R. of Moldova and for the entire world, the ICT development imposed a new way to operate for both, businesses and governments. Specifically, the countries shift their processes from manual processing to automated one. Industries including banking, communication, manufacturing, utilities are implementing complex information systems to keep the pace with the new reality. The world became a very big network of smaller networks, which are, in their majority, interconnected and inter-dependent.

The Covid-19 pandemic had a speed-up effect on the digitalization of processes. The government and businesses started to work remotely, request significantly more online services, implement new and modern technologies and information systems.

The national economy is becoming more and more dependent on ICT. This is a necessity. The new government of Moldova, for the first time in the history, appointed a minister, responsible for the government services digitalization. It is a clear strategic initiative and shall be accepted by every citizen as the only way to move forward in successful economic development of the country.

By the other side, by implementing new and modern technologies, every organization has to understand and manage related risks. Cyberattacks take place at all levels of competence: citizen, state institutions, geopolitical. To ensure the cyber security of a country it is necessary to ensure the security of the person and business, of state enterprises as well as to have very good and efficient mechanisms to protect the country at a macro level. [2].

Experts state that there is an activation of services related to the dark digital economy, offered with the help and through the cyberspace. The criminal world is organized in such a way that even in times of crisis, it tries to obtain maximum profit. The impact of the COVID-19 crisis on the digital shadow economy continues to evolve. Alternative platforms are being created, such as social networks, instant messaging and secure communications applications, which are increasingly being used to facilitate the spread of illegal goods, including drugs, child pornography and even weapons. In response to these trends, governments have, in various ways, mobilized significant national and international resources to create a broad cybersecurity framework [1]. It should be noted that governments, especially those representing digital forces, often prefer to organize their activities in isolation, as part of the activities of intelligence services, thus limiting the possibilities for cooperation.

The Republic of Moldova does not represent and is not considered a digital force from geopolitical point of view. At the same time, the Republic of Moldova must ensure a very flexible strategic approach. Cooperation with international organizations such as ENISA, the European Competence Center on Cybersecurity (created in Bucharest), Interpol, Europol will provide a significant advantage in Moldova's regional positioning in cybersecurity. That cooperation impose reviewing and renewing current programs and strategies and align them with current requirements and realities. The importance is also given by the results of the report issued by the World Economic Forum, which, along with pandemic and environmental risks, appreciates at a high level the risk of failure of cybersecurity and digital inequality - 39% and 38% of respondents, who consider that these risks will be realized in the next 2 years.

In the context of the COVID-19 pandemic situation, the ordinary Moldavian citizens remained and still remain vulnerable to diverse cyber risks and attackers, who, taking advantage of the created situation, directed and intensified their efforts towards the ordinary internet users. Although no statistics have been identified in the Republic of Moldova on the number of reported attacks, the FBI Report (Internet Crime Report 2020) concludes that in 2020 the number of attacks made through the Internet (phishing, scamming, wishing, etc.) and directed at the citizen has doubled (241,342) compared to 2019 (114,702) in the United States. According to information provided by the cyber security service of "Sberbank", the largest Russian bank, the number of fraudulent e-mails increased by 30%, with one-fifth of e-mails containing information about Coronavirus.

Human resources play a central role within the process of cyber and information risk management at every level. Each individual, on a daily basis, manages various risks: when passing by the street, when going up / down stairs, when driving the car, etc. Regardless of the activity, we are exposed to risk and we must be fast enough to manage it properly.

Cyber and information security risk management is at the core of any cyber security activity. Moreover, the risk management process is aimed at focusing efforts on risks with a high level of probability and / or impact. Risk management involves strengthening state assets, the most important and critical, but also ensuring a risk level which is in line with the accepted risk appetite. A strong risk management process involves the use of a variety of specific information, obtained from different sources. The risk-oriented approach to security can facilitate and optimize the effort to ensure cyber security.

These days, when the level of dependence on IT and information systems is increasing, organizations tend to implement complex and modern solutions. At the same time, most organizations don’t pay sufficient attention to a specific group – human resources risks. The risk management models usually are being focused on technical measures and less on organizational aspects.

The current paper has an objective to study the mechanisms that states may apply and implement in order to facilitate and ensure the protection of its assets.

2. Core principles in building a cybersecurity risks management model

The approach to managing Cybersecurity risks is different from country to country. Strategies and policies with respect to the cybersecurity risks are differently structured, and this is normal. In managing risks, there is a need to keep a balance between the potential impact and costs required to minimize risks. Also, we are willing to know what the impact of our decisions will be and whether we are elaborating the correct decisions.

When defining the cybersecurity risk management principles we have to consider the strategies for governing risks. There is a number of known risk treatment actions, which are being applied in managing risks. Each of these are usually being selected depending on the risks probability and impact. They can be grouped into 2 categories: risk prevention/recovery (control, terminate, transfer, take more, tolerate, contingency), risk-info (check compliance, tell someone, communicate, commission research).

The risk prevention/recovery strategy involves activities meant to minimize identified risks to the acceptable level. This is done by implementing security measures and controls, implementing recovery mechanisms and procedures, passing the responsibility to a third party, etc.

The risk-info strategy deals with continuous monitoring, communication and research to increase knowledge, awareness, confidence and trustworthy through clarification of facts and details, involvement of third parties that may be/had been affected, etc.

It is to mention that in the world of cyber security, the second strategy (risk-inform) is rarely applied. This is because often the risk realization affects sensitive data and information.

Based on the above, a government has to build its Risk Governance model on a set of principles, which will assure that the model is functional and efficient in time. Some of basic principles (not limited) are:

- Responsibility – responsibilities for managing cybersecurity risks will be assigned to the right bodies and control mechanisms will be implemented for monitoring
- Openness – the government will be open about the risks to the citizen and its structures.
- Transparency – the government will be transparent about its understanding of the nature of risks to the public as well as the actions it takes to manage them.
- Involvement – the government will involve the parties in the risks mitigation decision process
- Centralization – the government will make sure that all relevant risks will be registered in a database (e.g. National Risk Register or Database), analyzed and reported properly.
- Consistence - implemented measures will be consistent when dealing with risks affecting public resources
- Informed decisions – decisions will be elaborated by analyzing all relevant information and evidence

As the entire Risk Management domain, we operate with subjectivity. It is not excluded that other researchers and experts could propose other principles and / or exclude some specified in the list above. The objective of the current paper is to highlight some important, as we consider, principles, that the governments have to adopt. Moreover, our objective is to land the idea that a consistent and efficient cybersecurity risk management process is to be built on a set of principleS.

3. Cybersecurity risk management governance body and responsibilities

Governments needs to be able to handle risk at three levels: strategic, tactical, and at the operational.

At the strategic level, decisions are related to formulation of strategic objectives, the resource allocation decisions that are meant to implement the strategic objectives, and defining policy options in response to changing circumstances.

The tactical level will make sure that detailed policies are being developed for the implementation and the delivery of public services. Decisions are made on managing risks related to procurement or acquisition, funding, organization, establishing projects, service quality and business continuity.

At the operational level, the government will manage technical related risks, managing resources, schedules, providers, partners and infrastructure.

In order to ensure a comprehensive cybersecurity risk management process, the government has to assume the following groups of responsibilities:

- Identification and inventory of information and systems of state importance
- Coordinating the effort to secure data and systems and protection against cyber risks at national level
- Coordination of risk management activities both of state institutions and in relation to international partners (security awareness)
- Ensuring the secure operational activity of information systems and critical state infrastructure segments
- Coordinating crisis management and resilience activities

From our point of view, it is imperial that government follows the above specified principles and establishes relevant bodies to make the process efficient. There is a need that the right organizational structure is created and implemented.

Clear and detailed roles have to be appointed for each of the abovementioned levels: strategic, tactical and operational.

One cannot identify all the risks at a time. In some cases, the cybersecurity risks can be identified only by analyzing them in complex. Accordingly, registering all the identified risks in a unified risks database would help identifying other, interconnected cybersecurity risks.

4. Governing cybersecurity risks elements

If we are together nothing is impossible. If we are divided all will fail –

Winston Churchill.

States might have difficulty contributing to cybersecurity policy if they cannot secure their own information assets. The information-security community long ago identified the best practices that can meaningfully reduce risk to the confidentiality, integrity, and availability of state-owned and -controlled data and related systems. Modern software and hardware offerings reduce the burden of integrating those best practices with IT management and adapting them to new threats. As a result, the core challenge for state cybersecurity professionals today is not technical; the cutting edge of cybersecurity is governance.

Defining and establishing of a governing risk management mechanism is key to ensure an environment that follows the same principle, the same objectives and acts as a semi-closed system. We specify that the system is semi-closed because the state risk-management system

cannot be closed – there are many risk strategies that require sharing of information. At the same time, the system is to remain with limited access, as the information processed within is, most of the time, critical and sensitive.

We further provide foundational elements that may form the Cyber-security Risk management system at the government level.

Stakeholders (Public/Private sector, civil society, academia, other states/foreign actors/international bodies/partners). Stakeholders play a central role in ensuring an adequate cyber security level. They are those who are the ultimate responsible for risk analysis and mitigation.

Threat actors (insiders, criminals, hacktivists, script kiddies, state proxies, terrorists, etc.). It is to mention that each threat actor has a strict relation with a state (external and/or internal).

Attack technics and vulnerabilities. We define two main attack technics categories: technical and social. Threat actors will always use a mix of the two, which offer a greater attack area.

Impact – Stakeholders should know that the impact of a risk is characterized by the quality and magnitude, which could be minimal or devastating. Also, the impact may have diverse forms, specifically physical, social, economic, psychological, as well as reputational.

5. Challenges of managing national cybersecurity risks

The digitalized world brings growing complexity of managing the cybersecurity area, specifically national cybersecurity. It is seen that the **skill gap** represents the key challenge. When we say the skill gap, it is meant skills at every levels: from basic (user related knowledge and skills) up to professional level (experts in the cybersecurity area). Some states have already established government programs in order to prepare experts and increase the awareness of its citizen. This are core actions which would close the specified gap during the years that follow.

The **Recourses** challenge is characterized by the fact that they are exposed to damages. The constraint here is to define the level of resources allocation in order to allocate sufficient resources for ensuring a balanced approach in threating cybersecurity risks. Not every actor is able to decide on the above.

The growing number of assets to be protected characterizes the increased **attack surface** challenge. The cyber threat landscape is an immense challenge for every country. Our society, economy, and critical infrastructures and industrial networks (e.g. SCADA) have become largely dependent on computer networks and information technology solutions. The exponential growth of the Internet interconnections has led to a significant growth of attack surface.

Threat actors are more sophisticated and represents a challenge for Governments. To deal with them there is a need to implement straight and functional security measures, which usually affects citizen's daily routine. The globalized nature of the Internet allows these threat actors to be physically located anywhere in the world and still affect the security of a country.

Another challenge is the **resistance to change**. Security measures are to be implemented in a way that does not significantly influence the citizen's life. Complex security awareness campaign will help Governments dealing with this challenge.

6. Conclusions

Security is a very complex domain. Governments and businesses, in the current interconnected and interdependent environment, cannot handle cybersecurity alone.

When developing national cybersecurity strategies they will take into consideration all

the component environments: internal, regional and global. The matter becomes of a national importance and it should be, among others, a responsibility of national intelligence services. Governments will have to share intelligence with its partners, they will have to align cyber education with current and future needs, organize specific bodies and ensure incident-response capabilities.

In other words, the cyber security is a domain, which becomes, and will remain for many years further, central and important area of the digitization agenda. That is why building an efficient and sustainable governance represents the key success factor.

References

1. Michael Garcia, David Forscey, Timothy Blute, Beyond the Network: A Holistic Perspective on State Cybersecurity Governance, 2017
2. Ryan Leirvik, Understand, Manage, and Measure Cyber Risk, 2022
3. Nazli Choucri, Stuart Madnick, Priscilla Koepke, Institutions for Cyber Security: International Responses and Data Sharing Initiatives, 2016
4. Malcom Harkins, Managing Risk and Information Security. Protect to Enable, 2013
5. Mal'ko, Aleksandr V.; Soldatkina, Oksana L, Informacionno-pravovaya politika v sovremennom obshchestve sravnitelnyj analiz (Information Legal Policy in the Modern Society: Comparative Analysis) // Comparative Politics Russia, 2019, No. 1, pp. 42-58.
6. Erdal Ozkaya, Cybersecurity Leadership Demystified, 2021
7. World Economic Forum, The Global Risks Report 2021,
8. Gabriele Jacobs, Ilona Suojanen, Kate E. Horton, Petra Saskia Bayerl, International Security Management. New Solutions to Complexity, 2021
9. John Maddison CMO & EVP Products, 2022 IT Trends and Cybersecurity Challenges, 2021.
10. Nazli Choucri, Stuart Madnick, Priscilla Koepke , Institutions for Cyber Security: International Responses and Data Sharing Initiatives, 2016
11. International Telecommunication Union, Digital Skills Insights, 2021

The Role and Place of Time Management in the Mechanism of Economic Security Management of the Enterprise

Nataliia Zachosova, Dmytro Kutsenko

Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine
natazachosova@gmail.com, lawagens@gmail.com

Abstract. The paper reveals the possibilities and feasibility of using techniques and instruments of time management in the process of managing the economic security of modern enterprises. The place of time management in the mechanism of economic security management is determined. The advantages of efficient allocation and use of time resources to counteract the risks of economic activity and increase the level of economic security of the business structure are established. The expediency of using certain indicators of time distribution efficiency in the process of complex assessment of the state of economic security of industrial enterprises is substantiated. The issues of corporate time management in the process of economic security management and in the process of functioning of the business entity as a whole are considered. The peculiarities of time management of employees of enterprises in the conditions of remote work during a pandemic and lockdown are outlined and the consequences of its inefficient organization for the efficiency of enterprises are determined.

Keywords. Time management, economic security management, risks, personnel, management mechanism, industrial enterprise.

1. Introduction

Time management is a set of approaches, methods and tools, the use of which is aimed at the optimal distribution of time of a person or team to ensure the achievement of individual or corporate goals. Time has long been recognized by experts as one of the important corporate resources of enterprises at the level of labor, capital, land, finance. Therefore, the management of this resource deserves close attention in the scientific, methodological and applied areas. At the same time, time management techniques and technologies for its effective distribution are used mainly by individuals seeking to improve their own results in professional and/or personal life. At the enterprise level, time management processes are not studied, and in scientific circles time management is sometimes not considered a science. However, studies of numerous information resources and data sources show that time management has its own concepts, methodology, numerous tools for those who seek to use their time resources to maximum advantage to meet their own needs and achieve goals. Therefore, the scientific basis for the study of time management in general is formed, however, is not formalized as specifically and rigidly as the terminological and methodological apparatus of other sciences. We believe that time management in the enterprise should be given the same important role as financial management, risk management, personnel management, innovation and investment management, and, accordingly, as economic security management, although this management does not consider the management of many enterprises mandatory. Management of economic security is often reduced by the management of companies to control the levels of income and expenditure, superficial monitoring of risks, monitoring the achievement of regulatory values of economic indicators. At the same time, it is a much broader in nature and content

management area, the rational integration of which into the corporate structure of enterprise management can help align the strategic goals of owners and top management of the entity with the risks and threats associated with their achievement; to balance the costs of achieving the planned indicators with the losses that may arise when choosing too risky ways to do so; to optimize the use of enterprise resources in accordance with its plans and needs of financial and economic activities and development in the future; to form reserves and protective mechanisms in case of intensification of existing internal and external threats and risks and emergence of their new forms and species over time.

During the global pandemic, and especially during lockdowns, there are changes in the conditions of enterprises operation. Organizing activities remotely required managers to take new approaches to work with staff, as well as redistributing resources to ensure business continuity. The issue of economic security management has gained new relevance given the emergence of new challenges and risks that have not been known to business so far, and the readiness to counter them at both micro and macro levels has been zero. To avoid a recurrence of this situation, it is necessary to introduce permanent mechanisms of economic security management, which would be characterized by a comprehensive structure, take into account the current realities of enterprise organization, as well as ensure adequate protection of economic resources. tactical goals of the business structure.

The purpose of this study is to establish the place, role and importance of effective time management in the mechanism of economic security management of business structures.

2. Vectors of Research of Problematic Aspects of Economic Security Management in the Scientific and Methodological Plane

Various aspects of economic security, including from the standpoint of management and building a mechanism for its management to effectively achieve their strategic goals in terms of influencing their activities numerous risks and threats, external and internal, are studied in detail by both Ukrainian and foreign scientists. The authors of this study have also previously made attempts to address at the scientific and methodological level of such problematic issues as the search of innovative approaches in the estimatology of financial institutions economic security, possibilities of its use in management and regulatory activity within the means of provision of the state financial security [1], research and methodological framework for managing the economic security of financial intermediaries in Ukraine [2], establishing a list of human risks of project management involving remote teams in the context of the enterprise economic and financial security [3]. Also a popular mainstream in security science today is the project approach to the formation of economic security system and its management, one of the supporters of which is Zanora V.O. [4]. Important for the effective management of economic security is the personnel component, which is directly a tool of such management, and therefore in the context of this study it is worth mentioning the publications of Kovalenko A.: «Personnel security as an element of human resources (personnel) policy in the economic and financial security system of the business entity» [5] and «Determinants of personnel policy in the process of management of financial and economic security of business entities» [6]. As for the time limits of security-oriented management, scientists and practitioners distinguish between tactical (operational) management of economic or financial-economic security and strategic management [7, 8], while the second of these time horizons is given much less attention by the scientific community. This does not correspond to global practices, according to which strategic management is an important aspect of enterprises development and their self-determination in the market and successful competition. Among

foreign publications in the context of this study we consider Dominitz J., Manski C.F. «Perceptions of economic insecurity: evidence from the survey of economic expectations» [9] and Scheve K., Slaughter M. «Economic insecurity and the globalization of production» [10] works worthy of attention.

However, the issues of time management in the context of the organization and implementation of effective management of economic security of both industrial enterprises and other economic entities have not yet been reflected in the modern economic literature. Given this, it was decided to conduct this study.

3. The Place of Time Management in the Mechanism of Economic Security Management of Enterprises

Determining the possibilities of integrating time management into the structure of economic security management mechanism, it should be noted that not every company uses or plans to implement such a mechanism in practice. However, those businesses whose owners and managers are serious about ensuring their own economic security develop well-detailed and structured mechanisms for security-oriented management. In them, the integration of time management mechanisms is possible in several segments.

Thus, from the functional point of view, the mechanism of economic security management should cover all components of a complex or fragmentary system of economic security, force them to interact effectively and not contradict each other. Personnel or intellectual-personnel component is one of the priorities for modern systems of economic security, and it is for its state that time management is extremely important.

The integration of time management techniques to manage the personnel security of the enterprise will allow:

- increase the speed of tasks fulfilment by employees of the enterprise, including those related to economic security (calculations of economic and financial indicators, risk identification, budget allocation for further activities and security measures, response to the impact of the threat, the manifestation of risk, challenges of the external environment, compensation for the consequences of such influence, the formation of a set of management decisions and actions to return to the normal functional state of the entity). Indeed, most time management tools are designed to minimize the time a person needs to perform certain tasks. At the enterprise level, we can talk about increasing the level of personal efficiency of an individual employee, and group work – the implementation of collective tasks by a group of employees, department, shop, other structural unit;
- to achieve the correct goal setting and setting priorities in the implementation of tasks. One of the vectors and an important component of time management is the development of the ability to analyze one's own goals and identify those that should be a priority and deserve to spend most of the effort and resources; their achievement should take place as soon as possible, as they are the basis for further steps towards the realization of strategic personal or professional goals. In addition, in the context of economic security management, the definition of target priorities may indicate that for an individual employee's own goals are more important than the company's goals, which can pose a conflict of interest, which can result in staff fraud and inefficient use of working time. Using a time management tool such as the Eisenhower matrix (represents four squares with common sides – for this purpose the sheet of paper is divided by two lines at an angle of 90 degrees; each square is filled with a list of cases – the first or square A – urgent and important cases, the second or square B – important but non-urgent cases, the third or square B – unimportant but urgent cases and the fourth or square D – unimportant and urgent cases), allows you to track those actions that the employee might not

perform, or perform last. The important thing is objectivity, because most time management tools should be used directly by a person who has decided to manage their time professionally, and it may have priority in some tasks, while in the workplace priority should be given to other tasks. We emphasize once again that it is important to separate the private from the professional, otherwise the picture will be subjective, and attempts to use time management techniques – a waste of time. Therefore, at the corporate level, the use of time management tools should be analyzed by the company's management. In this way, it will be possible to make adjustments to the employee's work schedule, define tasks for the first half of the day, week, etc., as well as see which tasks are non-urgent and optional, delegate them or replace them with more urgent cases for the company;

- effectively plan daily activities – Franklin's pyramid is known to time management practices (it is a pyramid where the main values of life are located at the lower level – for use from the position of an employee can be replaced by the main professional values, career ambitions, etc.; levels are global goals (personal or professional), the next level – the master plan to achieve the goal, then – long-term goals (for a year, for three years, for five years, etc.; then a short-term plan – for a month, a year – this segment of the pyramid can be divided into several more levels depending on the time intervals chosen for planning, and at the very top of the pyramid is a plan for the day – to be done within one working day) is the tool that would be appropriate in context. With its help, it would be easy for management to reconcile the vital strategic goals of the enterprise with daily, short-term, annual action plans to achieve a proper state of economic security. The structure of the pyramid will allow you to clearly track at which of its levels there is a contradiction of tasks; which plans need to be implemented quickly, and which can be postponed for a certain period without the threat of loss of financial and economic stability, reputation, market position, customers and partners;

- to form reserves of free and "conditionally free" time – getting rid of irrelevant, optional and urgent tasks will free up a certain amount of employee time. This time can be spent on rest – time management experts have proven the need to alternate work and rest, and this rest should be real, not be the time when the employee is trying to complete work on tasks or solve some personal problems. Only a person who really has a good rest is able to work effectively. Unfortunately, neither the capitalist model inherent in the modern economy, which dictates profit at any cost, and therefore allows the maximum load on human resources to achieve high financial performance, nor the planned economy, which encouraged the overfulfillment of plans on the basis of enthusiasm at the cost of processing and violation of all possible labor standards, did not form a tradition of decent and sufficient rest of employees and respect for employers to this their right. Only now, modern companies, such as Google, begin to introduce practices and leisure technologies of their employees between periods of their active work, creating a comfortable environment for breaks between tasks. Therefore, if the employee has free time more quickly and efficiently due to well-established time management tactics, the direct management should encourage such an approach to work. It is also possible to talk about "conditionally free" time, when the employee after optimizing the workflow sees opportunities to perform additional tasks during the free time or at its discretion transfers the task from one square of the Eisenhower matrix to another to fill the gap. This free time can be spent on rest, so we suggest calling it "conditionally free";

- detect, reduce and get rid of chronophages – an analysis of the time employees use to perform their duties can help determine how much time and when is being misused. Chronophages in time management are called "absorbers" of time, i.e. things on which time is spent uselessly to achieve the employee's personal or professional goals. The most common chronophages in the workplace are smoking breaks, surfing social networks, correspondence in various applications, watching videos, channels, chatting, personal conversations on the

phone without urgency, and so on. It is believed that such forms of pastime can have a positive effect on the employee – he is distracted from routine affairs, is socially active and can even establish business contacts and develop a network of acquaintances, including professional ones. However, in most cases, such activities are a waste of time for which the employer pays the employee wages, i.e. there are overspending, irrational spending, losses for enterprise. The presence of chronophages can be traced by recording the activity of the employee, monitoring his activities, however, for a normal working atmosphere, it is important for the manager not to become a supervisor, whose main function is to force employees to work without any distraction. Time management techniques are aimed at making the employee realize how big his time loss from chronophages can be, and that as a result of giving up the habit of wasting time, you can quickly perform work tasks, receive bonuses and other forms of material and moral encouragement.

Thus, the place of time management in the mechanism of economic security management is clearly traced in the subsystem of personnel security. However, it is possible to notice importance of rational distribution of time and for system of information and analytical safety. Information, like no other resource of the enterprise, has the ability to become obsolete and lose its relevance, and hence the value for making financial, investment, economic and other decisions. Therefore, in this subsystem, time management will: gradation and build a rating of the importance of tasks to search, verify, process and use information, optimize the time to search for the necessary data in closed or open sources, build a time frame of analytical work with information, structure goals information, to form time reserves for the use of information in a timely manner in order to bypass competitors and gain market advantage.

4. The Role of Time Management in Managing the Economic Security of Enterprises

The main positive effect we expect from the integration of time management technologies into the mechanism of economic security management of a modern enterprise is to increase its efficiency by reducing the time required to neutralize risks and threats or prevent them by making timely management decisions. At the same time, one can expect additional benefits from the integration of time management practices into the work of business structures:

- increasing the level of staff motivation;
- saving corporate resources;
- establishment of communication channels;
- clarification of strategic guidelines and development of rational, possible for implementation in modern economic conditions, strategies;
- optimization of economic processes;
- providing the ability to work remotely (maintaining the ability to continuity of financial and economic processes, even during the spread of the global pandemic and during lockdowns).

To assess the effect of the introduction of time management in the economic security management mechanism, it is necessary to introduce time saving indicators as sub-indices to assess the level of economic security that has been achieved as a result of security-oriented management. We can suggest indicators such as:

- speed of reaction to the threat (minutes, hours, days, weeks, months);
- speed of cessation of the negative impact of risk, threat, danger (minutes, hours, days, weeks, months);

- speed (duration) of recovery of safety after violation by the influence of a negative factor (minutes, hours, days, weeks, months, years);

- speed (duration) of compensation for the effects of a negative factor on the resources of the enterprise and its economic security system (minutes, hours, days, weeks, months, years);

- savings (financial, economic) from timely safety-oriented measures (savings from the fact that the negative factor did not affect or minimally affected a particular corporate resource of the enterprise).

This list is not exhaustive and may be extended. Note that the proposed indicators are quite specific, do not have established formulas for their calculation, as well as normative or oriented values in order to evaluate the result obtained after the calculation at least on a scale of "normal"–"not normal". Given this, only an expert method can be recommended to evaluate the proposed indicators. It is advisable to involve internal experts in the analysis, as the information needed to determine the values of the proposed indicators is quite specific and can only be accessed by insiders.

If the company has its own method of assessing the level of economic security, the proposed sub-indices should only be added to the list of indicators that are traditionally calculated by analysts and determine the weights in the final indicator (if the method provides such technology). In the absence of a comprehensive assessment, you can apply the approach to assessment, which offers in her monograph Yu. Kasperovich to assess the level of fiscal security of the state [11, c.85-92].

It is based on an integrated assessment of fiscal security in five stages. The first is the formation of a set (list) of indicators. In our case, at this stage, experts should systematize the existing indicators of financial and economic condition of the enterprise in order to create a list of such indicators that can characterize the state of a component of economic security, and supplement them with indicators that will diagnose additional signs of economic security in one or another of its functional plane. It is at this stage that it is necessary to plan the introduction of time management indicators in the list of indicators for assessing the level of economic security. In the second stage, the characteristic values of the indicators are determined, again with the involvement of experts or analysts of the enterprise (who, without a doubt, are also among the experts in this matter). At this stage, information resources are also identified that will be used to obtain information on the final values of individual indicators or to establish the data and information needed to calculate them. At the third stage, the rationing of indicator values is envisaged, which will allow in the future to easily interpret, understand and explain the results obtained in the evaluation process. At the fourth stage, the weights of groups and indicators are determined by an expert method. The choice of a specific expert method should be left to the discretion of a particular enterprise and the creators of the methodology for assessing the level of its economic security. In the fifth stage, it is proposed to calculate the integrated security index by multiplying the normalized values of the indicator by its specific weight [11, p.87-89]. Each specialist approaches the issue of determining weights differently. We are impressed by the method of Fishburn weights, which weights are determined using the opinions of experts who, according to their experience and knowledge, give priority to a particular indicator in the final indicator, and the sum of all weights is 1. According to another popular methodology for assessing economic security indicators in the integrated indicator are assigned the same weight values (thus recognizing their equal importance in assessing the level of economic security). The final choice should be left to the management of enterprises and developers of methods for assessing the level of their economic security, however, we think a more rational approach to the variation of weight values of indicators depending on their significance for the state of economic security.

The role of time management in the mechanism of managing the economic security of the enterprise lies in accelerating the operation of all its systems and subsystems, ensuring the appropriate speed of reactions and management decisions to take security-oriented measures to protect corporate resources, in rapid but prudent risk management and the use of tools to prevent, minimize, eliminate, compensate for the effects of negative external and internal factors on the level of economic security of the enterprise and its overall financial and economic condition.

5. Conclusion

The study of the role and place of time management in the mechanism of economic security management of the enterprise allowed to draw the following important conclusions.

First, time management as one of the areas of enterprise resource management – time resource – is not given due attention in the management process of modern enterprises. Corporate time management has not found its application in Ukrainian enterprises; there is no reliable evidence that its tools are widely used in foreign companies, except in the practice of large corporations, which have strong financial and human resources. Today, businesses are facing new threats and risks, and traditional tools and traditional approaches to economic security management are showing less and less efficiency, and at the same time, more and more attention is being paid to the resources of time in competition and the digital economy. With this in mind, we propose to integrate techniques of individual and corporate time management into the mechanism of economic security management of enterprises, institutions and organizations.

Secondly, the place of time management in the mechanism of economic security management is best and most rationally allocated in the processes of personnel or intellectual and personnel security management. It is during personnel management that it is possible to use most time management tools, even those that are designed for individual use and the use of which for work needs can lead to a loss of objectivity in the interpretation of the results of such use. However, the resources of the personnel security subsystem allow the use of various forms of employee motivation to effectively manage their own time in order to ensure a balance between personal and professional interests of staff during their stay at work or remotely from home.

Third, the role of individual and corporate time management in managing economic security increases during the pandemic and lockdown period. The remote work of the staff requires a greater level of organization, discipline, and, consequently, attention to the distribution and use of their own time in order to perform all professional tasks without unnecessary reminders, control and monitoring. Delays in the implementation of professional functions due to the inability to manage their time without being at work and control by management can intensify the risks of suspension of financial and economic processes and threaten the company with economic losses and loss of business reputation.

Prospects for further research are to develop a concept for the functioning of the mechanism of economic security management system with integrated time management tools and to assess the effectiveness of such a mechanism in comparison with its traditional counterparts.

References

1. Zachosova N. Innovative approach in the estimatology of financial institutions economic security: possibilities of use in management and regulatory activity within the means of provision of the state financial security. Baltic Journal of Economic Studies, 2019, 5/2, 45-56.

2. Zachosova N., Babina N., Zanora V. Research and methodological framework for managing the economic security of financial intermediaries in Ukraine. Banks and Bank Systems, 2018, 13/4, 119-130.
3. Zachosova N., Zyvko Z., Zanora V. Human risks of project management involving remote teams in the context of the enterprise economic and financial security. Financial and credit activity: problems of theory and practice, 2020, 4 (35), 514-521.
4. Zanora V.O., Silchenko B.I.: Economic security system management on the basis of the project approach. Economic Herald of Zaporizhzhia State Engineering Academy, 2017, 5(11), 130-133.
5. Kovalenko A. Personnel security as an element of human resources (personnel) policy in the economic and financial security system of the business entity. European journal of economics and management, 2020, 6, 1, 100-106.
6. Kovalenko A. Determinants of personnel policy in the process of management of financial and economic security of business entities. Bulletin of the Cherkasy University. Economic Sciences, 2019, 3, 70-77.
7. Koval O.V. Strategic management of economic security system of business entities: theoretical aspects. Bulletin of the Cherkasy National University. Economic Sciences, 2020, 1, 40-47.
8. Koval O.V., Shevchenko A.R. Strategic risk management in the system of business entities financial and economic security. Evropský časopis ekonomiky a managementu, 2020, 6, 118-123.
9. Dominitz J., Manski C.F. Perceptions of Economic Insecurity: Evidence From the Survey of Economic Expectations. Public Opinion Quarterly, 1997, 2(61), 261-287.
10. Scheve K., Slaughter M. Economic insecurity and the globalization of production. American Journal of Political Science, 2004, 4(48), 662-674.
11. Kasperovych YU.V. State fiscal security in a hybrid war: analytical report. Kyiv, Phoenix, 2019, 128.

Digital Components of Developing a Modern Business Model in an Innovative Economy

Tetiana Bochulia

State Biotechnological University, Kharkiv, Ukraine
t.bochulia@gmail.com

Abstract. The evolution of approaches to the development of a business development strategy has been considered, as a result of which the methodology of constructing business models has been changed from the perspective of the strategist and innovator, who understand the need for transformations and changes. The concept of digital business transformation is substantiated and the matrix of digital transformations of business model is proposed to form an individual roadmap for digital change. A framework for developing a digital business transformation model is proposed that based on the basics, points of impact and actions required to innovate an enterprise through digitalization tools.

Keywords. Digital innovations, business model, concept, development

1. Introduction

The trend of digitalization has become an objective and sought-after means of moving to a new level of business, increasing its adaptability to new rules and market requirements with an understanding of key customer values and necessary competencies of activity. Today, are working on various progressive areas and tools of business digitalization, in particular, using innovative digital transformation results for their implementation in business strategy and processes. The new business formula focuses entrepreneurial structures on changes in management priorities with the emphasis on digital technologies, which allow achieving efficiency in business decisions and sustainability of competitive positions.

One of the differences in the digital economy is the change in the position of business leaders, who today is determined not by the history of past success and achievements, but by the ability to change and transform, the willingness to move to the next level, understanding the need for training and adaptation to new market demands and conditions. Choosing a change strategy is the choice of a strategy that works and guarantees the result for the formation of new business competencies.

Today, the re-understanding of the traditional concept of business model formation with the substantiation of digital organization of business structures as an innovative tool for acquiring key competencies of strategic development is of particular importance.

2. Digital Context of Business Development

Modern economy is the result of comprehensive transformations of economic, information, communication processes with a change of the paradigm of strategic management and the concept of business development. Today, business is the opportunity for change for the economy as a whole, and involves the involvement of technological, behavioral and intellectual innovations that collectively form the basis for quality change. The transition to the digital economy is a logical consequence of the development of a society in which digital

innovation and behavioral technologies are trending. A hallmark of modern business is its initiative in change and its willingness to transform in new directions. First of all, such changes and transformations are possible after developing a detailed development plan based on business strategies and new type of information.

The digital stage of economic and social development, characterized by the intensity of the introduction of digital technologies and the digital reorientation of business, puts forward new requirements to the approaches used in the formation of strategies and models of development. Strategy is an expression of the company's aspirations and the goals for which it exists; a statement by the company about what it intends to become and what it wants to achieve in the future; a perfect statement of the reasons why the company is in business [1].

First of all, this is due to the fact that today the external environment is changing much faster and deeper than before, and the degree of uncertainty of the influence of various factors on the future is increasing. Approaches to the development of innovative strategies have evolved under the influence of new conditions of world development, the internal source of which is a person with his psycho physiological characteristics and ability to transform thinking (Figure 1).

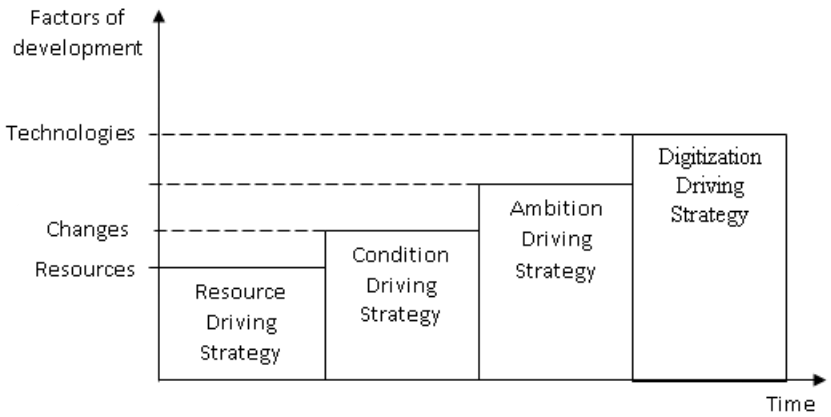


Fig. 1. Evolution of approaches to business strategy development [2]

Lower level of modern development takes strategy that based on available resources, that is, an option in which the strategic thinking of a management system is limited to assets that are available for use. Such an approach could be considered effective at the beginning of the last century, when land, capital and assets were the priority factor in production. Owning a larger volume of assets guaranteed the enterprise a competitive edge, and the strategy was to save capital and provide increase of asset. In today's economic conditions, such an approach is not only ineffective, but also dangerous for business. The asset framework is a narrower approach to development that is unable to meet the current requirements of compliance with competitive dynamics in the external economic environment.

The next step towards the transition to innovation management is the approach to developing a strategy that based on the conditions of activity. In contrast to the previous approach, it is anticipated that achievement of competitive stability is ensured not through the ownership of assets but by the speed of adapting the enterprise to new environments with the ability to develop assets and create new products and services. Such approach allows timely response to past and current changes in the external economic environment, taking into account

the internal conditions inherent in business. Accordingly, the planning process involves the collection of Big Data on the state of the market and the trends in its change, as well as on the activities of the enterprise at the current time and in the recent past. However, one should not forget that the changes taking place rapidly on the market make the future increasingly uncertain and less predictable.

The traditional process of strategy development is, first of all, analytical activity, the result of which is the choice of one of the most acceptable scenarios for the development of the enterprise, which will then be transformed into a business plan and specific planning indicators. The strategy should be how to achieve more by being different from other companies [3]. At the same time, the risks inherent in such a strategy are not at the development stage, but in the process of its implementation, and the main obstacle to its successful implementation is that before the management system, there are two complex tasks, which have to be solved simultaneously, – risk management in conditions of uncertainty and support of strategic decisions by personnel of the company.

The following approach (Ambition Driving Strategy) does not reject the analytical methods of the traditional approach, based on the conditions of development. At the level of strategy development based on ambition, analytical methods and procedures are integrated with the processes and tools that provide creativity, dynamism and interest from the staff of the enterprise. Distinctive features of this innovative approach are: the ambition of goals; behavioral management context; continuity of change management. Enterprises that are focused on ambitious development strategies determine the competence that is a complex of developed skills, knowledge, technologies, patterns, ownership, which guarantees the company the achievement of a stable competitive position in comparison with other economic entities. A combination of core competencies provides the uniqueness (individuality) of business models and draft decisions, repeat (simulate) which competitors cannot, which is provided not only by a set of technological solutions, but by a number of knowledge, ideas, professional competences, skills, information.

It is the low dynamic ability, the inability to quickly adapt to changes in the market and manage knowledge is one of the most significant reasons for the weak competitive position of enterprises. A company that strives for leadership should provide a high innovation potential in developing a strategy that responds appropriately to the development of new products and services, human resources management, customer relationship management policy, implementation of innovation in the management model.

Today, the most effective approach is to develop a strategy based on digital technologies, since exactly technologies are promising competence of an enterprise that is oriented towards innovative development. Enterprises have discovered that new relationships are not limited to their employees and consumers – the real opportunity was to become part of many new global networks that unite companies, people and things in various industries around the world. This situation and the extent of its impact on the transformation of modern business opens up a new era of economic and social relations – «The Economy of All of Us», when enterprises become part of a single «digital fabric» that combines all aspects of business [4].

Intelligent equipment becomes a kind of bridge connecting the digital enterprise and the physical world. Today, within the framework of the concept of Industrial Internet of Things, many companies implement «smart» sensors in a variety of digital devices and bind them together.

After that, customers can get more than just new products and services, namely the benefits and the results they want. In this sense, the market is fundamentally changing under the influence of the development of new type companies with an innovative strategy, which

involves offering the customer not a products or services, and the result of using a combination of digital technologies of the enterprise with intelligent end devices that used by the client. This is evolution not only in the strategy of development, but in the economy – the economy of the final result.

Thus, the next level of production efficiency will be the result of latest discoveries in the field of intelligent software. This can contribute to a new stage of evolution in all areas of business structures. In the short run, cognitive technologies contribute to evolutionary change, and in the long run, their impact is likely to be revolutionary [5].

Digitization Driving Strategy is a serious step towards radical changes. Such strategy cannot be defined by a modernization tool that integrated with general business strategy. Successful digital strategy sets the right direction and allows company executives to implement digitization programs, evaluate their progress and, if necessary, redirect efforts [6]. Digital changes are global transformations that are based on a digital basis, with the incorporation of their basic provisions into the development strategy, business model, operational activities and corporate culture of the enterprise. The digital platform connects the ecosystem and accumulates data from many sources [7].

The construction of the Digitization Driving Strategy is based on traditional factors of competitive growth, with the strengthening of key competencies that are updated in accordance with the requirements of the digital sphere (Figure 2).

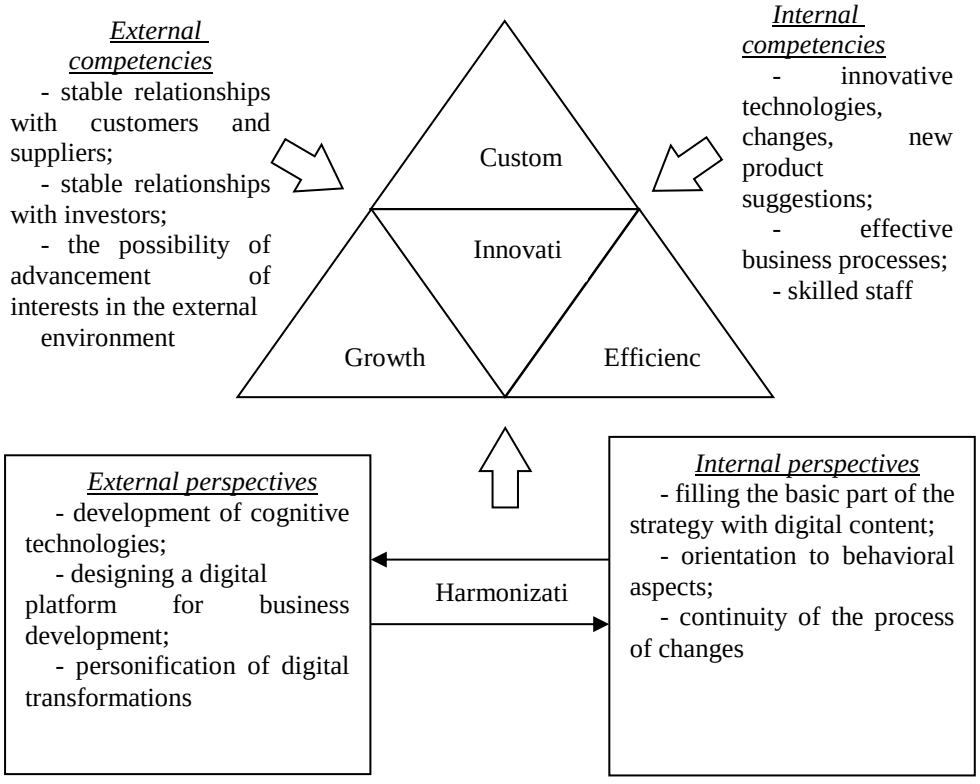


Fig. 2. Totality of factors, competencies and perspectives in shaping the model of digital transformation

The central factor in digital transformation is innovation, which today is an obvious feature of a digital-oriented business, and therefore increases the amount of information processing by expanding the networks of interconnections and promoting new products through the digital platform.

Digital innovations are not just about the introduction of digital technology, it is a complex process of change and adaptation to new economic development conditions and trends.

Today business is based not on corrected management methods but on rethinking the business concept and changing the business model with new approaches to business competencies, information management and the digital development culture. Today, business is capable of development provided it is transitioned to digital ways of organizing activities and engaging with counterparties and stakeholders.

3. Concept of Digital Business Transformation

Digital reality focuses business on rethinking of the mission and strategy of operations with transformations and changes in organization of activities, primarily in the business model, which is the basis for the formation and development of key competencies, business culture, activity management and organization of information and analytical decision provision.

Leading experts note the need not only to adjust and develop the business model of the company to suit the variability of the external economic environment, but a radical change, that is, a rethinking of the way, means and processes of doing business [8]. Certainly, such cardinal actions are associated with the risks and challenges of the planning, development and implementation stages, but decisive action and a willingness to change will succeed in sustainable business development and gain sustainable competitive positions.

At the decision-making stage, business leaders should determine which type of digital transformation most closely matches the chosen development concept and strategy:

1) Type «New features» – a radical change in the value proposition for the market in order to strengthen own competitive position. This proposal is based on the focus on digital capabilities and the development of a new business model. This type of business transformation involves abandoning the traditional concept of activity and spreading drastic changes in structure, organization of activity and management. This type of digital transformation has both high levels of risk and success.
2) Type «Adaptation benefits» – transformation of the business model based on the principle of adaptability, ie integration of existing key competences with digital capabilities. This is not a rejection of the chosen concept of activity, but its adaptation to market realities and new user requests. It is a combination of the best business opportunities with the benefits of a new economy that focusing on new trends.
3) Type «Advanced competencies» – having advanced competencies allows an enterprise to be a leader in a particular industry and strengthen its position. The refocusing on digital tactics allows businesses to embrace new industries and thus to reach a new level of development. Digital changes are needed for such companies to expand previously developed core competencies. This is not an adaptation policy, not a rejection of the traditional model, but its modernization in line with the requirements of digital reality.

Most organizations do not create clearly focused strategies – instead, they make long lists of desired outcomes, completely ignoring the need to clearly coordinate forces and focus on core resources [9].

A timely developed business model makes for a simplified approach that allows you to get an overall strategic picture of the business, minimizing the discussion of specific tactical steps [10].

Businesses that have chosen digital transformation as a strategy need to remember that digitization does not only involve technological change – it is also analytical, cloud, social, information, and managerial transformations that are capable of delivering large-scale change, as they are implemented simultaneously and qualitatively change core competencies. Digital technologies should be seen as a means of achieving the goal of development and gaining competitive advantage based on business differentiation and the launch of a new promising offering, made possible by the digital reorientation of activities.

Digital transformation of business involves the formation of a new concept, which is a matrix of changes with the identification of key objects of transformation:

1) Integrated communication – using Big Data system to create an attractive client proposal; establishment of effective channels of communication with consumers; develop an expanded proposal based on an evaluation of customer requests; using of analytics applications for a better understanding of customer requests; integrated circuit of value added.
2) The gradual transformation – willingness to partially lose competitive positions in the initial phase of digital modernization of business; willingness to master of new markets; willingness to change the core competencies of an enterprise in line with the realities of digital transformation; gradation of innovations in strategy, business processes, information management, business activity.
3) Digital driver – organizing the digital environment as a new type of system of company functioning; organization of digital culture of activities with transformation of information management and the decision-making process; using the digital experience as a new tactical line to gain core competencies; using opportunities analytical applications for data formation about customers and anticipating their requests and preferences; integration of delivery model with a proposal model.
4) Innovation offers – offering new services and products as a result of digital transformation of business; development of a model of adaptability to the realities of the external ecosystem of business; continuity of innovation in the supply of goods, products and services.

When deciding on digital business transformation, the company should determine which direction of the transformation concept corresponds to the business model and what strategy it chooses for further development.

4. A Model of Digital Business Transformation

Business dynamics are becoming more integrated with effective management and communication, support for partnerships and the overall business network with a focus on digital development opportunities and entry into a single digital space. With the transition to the digital concept of activity, the company ceases to be exclusively the supplier of goods, products, services, and acquires a new status – an element of the overall value chain, ie an effective model of overall economic development with opportunities to effectively influence the market and partnership between business leaders.

There is no universal model of digital transformation, but there are basic provisions that allow extending the methodical techniques of organizing activities (Table 1).

The choice of digital transformation of the business model is a stage of transition of the company to a new level of development, which is a new formula of successful leadership in the modern market.

Table 1. Basic provisions of development of model of digital business transformation

Basic provision	Object of the impact	Action
Digital capabilities	Capabilities and risks	Evaluation of company activity concerning capability to expand core competences through digitalization and be prepared for the risks of potential losses as implementation of changes take place
Digital priorities	Directions and tools of digitalization	Choosing a digital transformation scenario with an appropriate set of principles, methods and procedures that best suits the previous stage of evaluation of company activity
Digital competencies	The propensity to digitalization	To evaluate the potential of developing digital competitive advantage among market participants with the expansion of innovative competencies related to digital transformation of activity
Digital changes	Willingness to changes	Transformation of all aspects of business without priorities and constraints. Formation of digital culture, using digital resources to create value and strengthening competitive positions, implementing digital business philosophy as a sustainable development strategy
Digital experience	Digital skills	Assessment of own skills and competitor skills with identifying opportunities to leverage the best experiences of market leaders, which demonstrate the successful outcome of digital business transformation. Adapting the competencies of leaders to own experience
Digital staff	Staff skills and competencies	Investing in staff training and digital competency development in accordance with the chosen business transformation strategy. Creating a digital partnership environment not only within the company, but also establishing partnerships with the external business ecosystem

Forming a digital transformation model is based on the concept that assumes availability of partnerships, thus, creating a new ecosystem of entrepreneurship, and this is what gives the best result. Digital changes make it possible to shape behavioural models and modify processes in an environment where interaction between market leaders is gaining new value. Therefore, enterprises should choose the strategy of development of digital initiatives as the most promising for the modern stage of economy and business relations.

5. Conclusion

Thus, the choice of the type of digital transformation involves the further development of a strategy for innovative development. This is justification of possibility of forming a map of decisions on digital change and a basic template for real evaluation of the enterprise activity, which decided on the need for digital transformation.

Own view on the concept of digital business transformation is presented in the research, which identifies four basic provisions that allows the company at its discretion to choose the original direction of change according to the opportunities and limitations of the enterprise. The proposed options for digital transformation can be used to develop an individual roadmap, in which will take into account the current state of activity of the enterprise and its expectations in accordance with the chosen strategy, which is subject to development according to the needs of both the internal and external environment. Such options allow to make adjustments to the digital transformation of business in a timely manner and to modernize the enterprise development model to meet the objective needs of changes.

Basis for developing model of digital business transformation are offered, which describes the objects of influence and actions that required achieving the goal of sustainable development. Digital transformation is based on the basic methodological provisions, which necessarily include business processes, information management, personnel policy, business activity and the formation of core competencies of activity. It is on these aspects that emphasis is placed because their digital transformation enables the company to move to the level of digital business culture. If downplay or ignore the importance of one of the aforementioned basic provisions, it will only be a partial digital transformation that will not achieve the goal of development and transition to a new level of leadership. Therefore, it is important to be guided by the basics described that enabling the enterprise to design measures and identify tools for digital transformation without asymmetry and unnecessary risks. Enterprises at the current stage of development have all the potential to drive digital innovation, so the success of business modernization depends on the willingness to changes.

References

1. Tracy B., TurboStrategy: 21 Powerful Ways to Transform Your Business and Boost Your Profits Quickly, AMACOM; Special edition, 2003.
2. Bochulia T., Innovative changes in development of enterprise information system: digital transformations, behavioral technologies and new decisions, Consilium Sp. z o.o., 2019.
3. Johnson M. W., C. M. Christensen, H. Kagermann, HBR's 10 Must Reads on Business Model Innovation, Harvard Business Review Press, 2019.
4. Daugherty P., P. Banerjee, M. J. Biltz, Digital Business Era: Stretch Your Boundaries (https://www.accenture.com/t20151117T010853__w__/nl-en/_acnmedia/Accenture/Conversion-Assets/Microsites/Documents11/Accenture-Technology-Vision-2015.pdf)
5. Davenport Thomas H., The AI Advantage: How to Put the Artificial Intelligence Revolution to Work (Management on the Cutting Edge), The MIT Press; 1st edition, 2018.
6. How to Go Digital: Practical Wisdom to Help Drive Your Organization's Digital Transformation (The Digital Future of Management), The MIT Press; Illustrated edition, 2018.
7. Charan R. Rethinking Competitive Advantage: New Rules for the Digital Age, Currency, 2021.
8. Linz C., G. Müller-Stewens, A. Zimmermann, Radical Business Model Transformation: Gaining the Competitive Edge in a Disruptive World 1st Edition, Kogan Page; 1 edition, 2017.
9. Edge D., Good Strategy Bad Strategy: The Difference and Why It Matters, Currency; 1 edition, 2011.
10. Osterwalder A., Y. Pigneur, Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, John Wiley and Sons; 1 edition, 2010.

The Role of Information Technology in Auditing Financial Statements in Modern Conditions

Liudmila Lapiţkaia

Academy of Economic Studies of Moldova - University, Chisinau, Moldova
liudmila@ase.md

Abstract. The priority of the strategic direction of any country and economy is the development of information technologies, especially digital ones, which ensure rapid economic growth and an increase in the country's rating in the global information space. In this regard, methods for auditing financial statements are improving in the digital economy: from traditional methods to the use of artificial intelligence. Thus, in the context of the fourth industrial revolution, a wide range of digitalization opportunities (cloud technologies, artificial intelligence, etc.) contributes to improving the quality of audit of financial statements, reducing the labor intensity of the processes of collecting and analyzing information. At the same time, from the point of view of scientific research, it is important to understand the nature of the influence of information technologies on the audit methodology, taking into account the fact that these technologies, including digital ones, are developing very quickly, offering new opportunities for modernizing the organizational and technological support for auditing financial statements.

Key words: Cognitive auditing, information technologies, combined task

1. Introduction

The digitalization of the economy presents tremendous opportunities for the effective development of its various industries, allowing it to provide a wide range of information services, including services for the audit of financial statements.

Developing information technologies in the field of artificial intelligence, robotization, big data processing, cloud technologies can find an applied aspect in the digitalization of audit activities, which will help to improve the quality of auditing financial statements. Accordingly, we can talk about the modification of audit methods from traditional methods to the use of artificial intelligence.

The audit methodology is an integral part of the intellectual capital of the audit firm, just as the methods used in the audit are the auditors' know-how. The auditor's opinion regarding the reliability of financial statements is influenced by the analysis of evidence that can be obtained using traditional methods and using information technology. It should be noted that changes in the volume of economic, production information, management systems, including information and investment flows, lead to the need to modify the financial statements of the enterprise, to more complete disclosure of market valuations, contingent assets and liabilities, etc.

Thus, the auditor needs to check and evaluate the data and information presented not only in the financial statements, but also in managerial accounting and in a number of internal documents of the audited enterprise. Thus, in conditions when large volumes of data, information must be verified in the shortest possible time, there is an urgent need for the use of information technologies, especially digital ones.

2. Description of the Studies

One of the options for using digital information technology is the use of artificial intelligence. Below are data on revenue received from the use of artificial intelligence, as well as forecast data reflected in the article of a research expert for the global software industry and emerging tech trend Shanhong Liu.

Artificial intelligence software market revenue worldwide 2018-2025

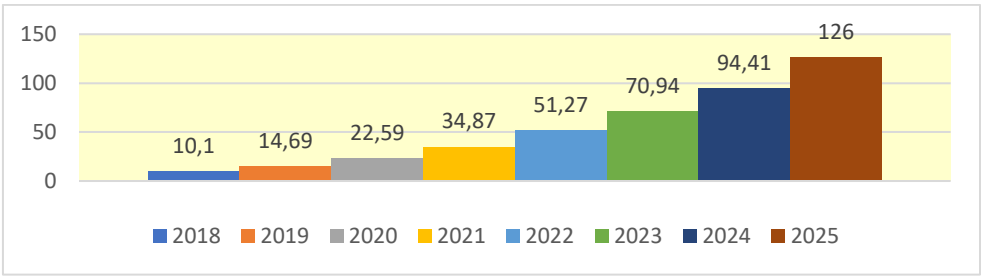


Fig. 1. Revenues from the artificial intelligence (AI) software market worldwide from 2018 to 2025(in billion U.S. dollars)

Source: data presented at site https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/ [1]

When conducting an audit of financial statements, auditor's judgment plays an important role. Most audit procedures are unstructured, containing many alternative options and solutions, which requires the use of professional judgments. Thus, it is important for the auditor to develop: rationality of thinking, algorithms for developing opinions and ways of coordinating activities.

In her article, the researcher V. Yakimova "Opportunities and prospects for the use of digital technologies in auditing" examines the types of tasks facing the auditor during the audit of financial statements, classifying them into:

- ✓ Standard (formalized),
- ✓ Slightly structured (contoured),
- ✓ Unformalized [2,3].

It should be noted that researcher V. Yakimova, after analyzing the research materials of Abdolmohammadi, S.E. Bonner, N. Pennington, made the above classification. S.E. Bonner and N. Pennington identified 28 tasks in the audit and described cognitive processes in relation to solving each of them, and also determined the appropriate knowledge base [3]. For example, to understand the business of the audited organization, the auditor must have information about the state of the economic situation in the country where the audited organization operates, about external and internal factors that affect the business, and to perform analytical procedures and identify potential problems, information about the frequency of errors in financial statements is required.

Having analyzed this classification, from the point of view of the use of information technology in auditing, the author suggests classifying tasks as follows:

1) Tasks (standard), which, taking into account the use of digital information technologies, are easily formalized,

2) Tasks (combined), for the solution of which it is necessary to use an audit judgment based on the data presented taking into account the use of digital information technologies,

3) Tasks (special) that are difficult or almost impossible to solve with the help of information technology, in this case it is necessary to apply an audit opinion.

When performing standard tasks, the auditor uses computerized audit programs (Audit XP, IT Audit, Seal, etc.), which can work effectively, in this case information technologies are effective:

- ✓ at the stage of planning the audit: when determining the level of materiality, determining the audit risk, carrying out analytical procedures,
- ✓ at the stage of conducting an audit: when determining the sample, applying analytical procedures, etc.

In the case of a combined task, it is necessary to use the capabilities of information technology, primarily digital, as well as audit judgment.

At the same time, it should be noted that when performing special tasks in the audit, the development of information technology has reached such levels that it suggests making decisions based on generating possible options, choosing and proposing the best option to the auditor and analyzing the consequences of the proposed solution. For the rapid search and processing of information in a space that is poorly structured, algorithms have been developed that are a kind of cognitive technologies and simulate the process of natural selection based on tools for assessing and mutating possible solutions.

It can be stated that there are certain algorithms for modeling the auditor's decision-making about the facts of fraud. The principle of this technology is to find the best possible solution. At the present stage of the development of cognitive technologies, the most promising are information technologies with the use of neural networks, the difference between which from expert systems and genetic algorithms is the ability to self-organization and self-learning. Informational neural systems imitate the work of the human brain and consist of interconnected blocks that process information by responding to the arrival of external data with the transfer of information between blocks.

It should also be noted that when conducting an audit using artificial intelligence, it becomes possible to recognize documents and images, work with information and intelligent analysis of big data, while using information technologies of a hierarchical neural network to analyze data and extract complex and abstract functions underlying the raw data. In their article "Decision-making capabilities of a hybrid system applied to the auditor's going concern assessment. Intelligent Systems in Accounting", authors Mary Jane Lenard, Pervaiz Alam, David Eugene Booth proposed applied genetic programming using fuzzy set theory, expert assessments and statistical methods for making predictive models and assessing business continuity [4], there are a number of developments, in particular hybrid systems integrating neural network and elements of expert systems, for risk assessment.

When conducting an audit of financial statements, the auditor has to work with a large amount of data that needs to be classified, structured, checked and evaluated. In this case, the use of big data analysis technologies (Big data) plays a great role for the auditor. To extract data from various sources representing audit evidence, processing technologies are used, which allows analyzing test documents and web pages, understanding the context and content, including emotional coloring.

R. Fay and E. Negangard describe the application of big data analysis for risk research and fraud detection. The use of recognition methods obtained from various media (video, audio) allows to identify an object, organize and classify events into a logical class, create a single information base for further comparison with typical situations in similar companies.

A promising area of research in the field of the introduction of artificial intelligence during the audit is the consideration of the creation of robotic systems and intelligent assistants that combine a set of technologies to solve various tasks in the audit.

Thus, in modern conditions, we can talk about intellectual audit, in her article "Opportunities and prospects for the use of digital technologies in auditing" V. Yakimova defines: that "intelligent audit is an audit, the organization and methodology of which is based on the use of AI technology to carry out cognitive processes and decision-making at all stages, from planning to forming an opinion in the auditor's report. Artificial intelligence can generate hypotheses based on judgments, collect data from disparate, including unstructured sources of information, analyze large amounts of information and visualize the results of the analysis, as well as evaluate the collected evidence. " [2].

In this regard, such a definition as the Cognitive ecosystem of audit activity has appeared in the world research literature in the field of audit. As noted in a 2016 published study by KPMG, Harnessing the power of cognitive technology to transform the audit: *The audit cognitive ecosystem is a set of robotic process automation (RPA) technologies, data and analytics, cognitive technologies and add-ons. - predictive analytics.* [5]

In her article "Opportunities and prospects for the use of digital technologies in auditing" V. Yakimova defines that the cognitive ecosystem of audit activity consists of three subsystems: an intellectual subsystem (a basic complex of software and hardware technologies that does not take into account the applied nature), a subsystem of cognitive technologies (for example, expert systems and genetic algorithms) and knowledge management subsystems (aimed at generating knowledge about the subject area). [2] Analyzing the above definition of the cognitive ecosystem of audit activity, the author proposes to include one more hybrid subsystem at the decision-making level, which is a combination of artificial intelligence technologies and auditor decisions.

When conducting an audit of financial statements, the auditor conducts a sample check of the general population, however, the use of digital information technologies will allow the auditor to conduct a continuous sample, thereby minimizing the risk of non-detection.

Big data mining and predictive analytics are essential technologies for identifying risk areas and assessing the level of risk as part of a financial statement audit.

An important point when using artificial intelligence in auditing is the need to apply it in cases where there is a multipolarity of choice, uncertainty, ambiguity of data or information.

The task of the auditor in an information technology environment is to assess the adequacy and materiality of the information selected and to analyze the identified misstatements. Artificial intelligence, as it obtains evidence in real time, can evaluate and analyze information, automatically select or suggest alternative procedures, adjust the audit plan as needed, and respond in a timely manner to increased levels of audit risk.

In the context of the use of artificial intelligence in the audit of financial statements, there is an urgent need to amend the regulations and legislation in this part, especially in the International Standards on Auditing, for example: by the types of audit procedures that can be applied by the robotic auditor and the separation of responsibilities during conducting an audit.

3. Conclusion

The research has shown the possibilities and prospects of using digital information technologies as an independent field of scientific knowledge. Thus, the main conclusions based on the materials of this article are as follows:

1) The use of information technologies in auditing has a certain history, currently the use of computerized audit programs Audit XP, Audit IT, etc. is widespread. At the same time, in modern conditions, prerequisites have been formed for the use of digital information technologies when conducting an audit of financial statements,

2) The use of digital information technologies (robots, artificial intelligence, etc.), modification of the theory and methodology of auditing, leads to the formation of interdisciplinary areas of scientific knowledge - cognitive auditing,

3) the use of digital technologies, and in terms of a continuous sample of the general population, can improve the quality of auditors' services, increase the degree of confidence in the audit report, which in turn will improve the auditor's business reputation. Artificial intelligence may be more objective than the auditor's judgments,

4) it is necessary to modify the regulatory framework in the field of auditing, namely the International Standards of Auditing. In this case, there are two options: either to introduce paragraphs on the use of information technology in each of the relevant standards, or to develop a new international standard on auditing regarding the use of information technology, including digital (artificial intelligence, robotics, etc.)

References

1. Shanhong L., Artificial intelligence software market revenue worldwide 2018-2025. https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/
2. Yakimova V. "Opportunities and prospects for the use of digital technologies in auditing." Bulletin of St. Petersburg State University. Economy. 2020, vol. 36, issue. 2B. https://dspace.spbu.ru/bitstream/11701/19112/1/287-318.pdf
3. S. E. Bonner, N.Pennington. Cognitive processes and knowledge as determinants of auditor expertise. Journal of Accounting Literature Vol.10.1991 p.1-50. https://www.marshall.usc.edu/sites/default/files/sbonner/intellcont/Bonner%20and%20Pennington%201991-1.pdf
4. M. J. Lenard, P. Alam, D. E. Booth, Decision-making capabilities of a hybrid system applied to the auditor's goingconcern assessment. Intelligent Systems in Accounting https://www.researchgate.net/publication/312984304_Decision-making_capabilities_of_a_hybrid_system_applied_to_the_auditor's_goingconcern_assessment_Intelligent_Systems_in_Accounting
5. KPMG, Harnessing the power of cognitive technology to transform the audit, 2016 https://assets.kpmg/content/dam/kpmg/us/pdf/2017/02/harnessing-the-power-of-cognitive-technology-to-transform-the-audit.pdf

Technology for Organizing the Security of Information and Communication and Information Systems

Veselina Aleksandrova[1], Milena Ivanova [2], Violeta Vasileva[3]

[1] G. S. Rakovski National Defense College, Sofia, Bulgaria
[2] Council of Ministers of the Republic of Bulgaria, Sofia, Bulgaria
[3] Artshare, Estarreja, Portugal
alexandv@yahoo.com, milena.ivanova80@gmail.com, violetta.ziv@gmail.com

Abstract. The protection of communication and information systems is considered as a process that applies technology to organize the security of information and communication and information systems (CIS). This process is implemented in a system for the protection of CIS. For this purpose, models of the CIS protection system are proposed, considering the methods and means of information and CIS security, as tools of processing and transferring the information.

Keywords. Information Security, Communication and information systems, Technology, Models.

1. Introduction

In general, the model for security of information and communication and information systems (CIS) includes elements such as: Objects of threat; Threats; Sources of threats; Objectives pursued with the threats; Sources of information; Ways of illegally acquiring classified (confidential) information; Methods (methods) for information protection; Means or tools of protecting the information.

Activities (measures) to protect information in computer networks are implemented through security procedures. Security procedures are a detailed description of the order and responsibilities for the implementation of the activities in the application of the established security measures.

The set of procedures for ensuring safety is in accordance with the sections of ISO 17799 that will be considered [1].

The purpose of this report is to review the methods and means of information and CIS security

2. Security Procedures in Accordance with the Sections of the ISO 17799 Standard

The ISO 17799 standard is organized into ten main sections (security procedures), each covering different security topics or areas:

Continuous functional planning

The objectives of this section are to counter major system crashes that would result in the disruption of activities and the occurrence of critical processes.

Access control system

The objectives of this section are:

• to control access to information;

• to prevent unauthorized access to information resources;
• to ensure the protection of network services;
• to prevent unauthorized access to a given computer;
• to disclose the unauthorized actions;

System development and maintenance

The objectives of this section are:
• provision of built-in protection in operating systems;
• to prevent loss, alteration or misuse of user data in application systems;
• to preserve the confidentiality, reliability and integrity of the information;
• to ensure control and security of information technology projects;
• to maintain the security of application system software and data.

Physical and environmental protection

The purposes of this section are to prevent unauthorized access, damage and interference with functionality and information, to prevent the loss, damage or risk of interference and interruption of functions.

Compliance

The objectives of this section are:
• to avoid violating the law, regulatory or contractual obligations and security requirements;
• ensure compliance of systems with security and protection policy and standards;
• to maximize efficiency and minimize conflict with the systemic audit process.

Protection of personnel

The objectives of this section are to reduce the risk of human error, theft, fraud or misuse of the equipment [2]; the users of the information are aware of the dangers for information protection and are equipped to maintain the departmental security policy in the process of their work [4]; to minimize damage from safety-related incidents and malfunctions and to educate people from such incidents.

Protection of the Organization

The objectives of this section are:
• to manage information security and protection within the Organization;
• to maintain the security of the flow of information to which third parties have access;
• to maintain the security of the information when the responsibility for the generation of the information is assigned to an external Organization[3].

Management of computers and operations

The objectives of this section are:
• to ensure the correct and secure use of information resources;
• to minimize the risk of system crashes;
• to protect the integrity of the software and information;
• to maintain the integrity and usability of information and communication;
• to ensure protection of information in the networks and protection of the supporting infrastructure;
• to prevent damage to valuables and disruption of activities;
• to prevent the loss, alteration or misuse of information exchanged between the Organizations.

Asset classification and control

• The objectives of this section are to appropriately maintain assets and to ensure that information assets have an appropriate level of protection.

Security and Protection Policy

The objectives of this section are to provide management guidance and support regarding information security and protection.

Methods and means of information security and CIS are considered from an organizational perspective.

In organizational terms, the purpose of the protection of CIS and circulating information is to build a system for their guaranteed functioning, in the centers of storage and processing of information, communication networks and data transmission networks, as well as their monitoring and control.

The protection of CIS covers the protection of both the communication and information subsystems. The requirements that this protection must meet from an operational point of view are: national identity; guaranteed durability; adaptability to changing conditions, access control and mobility. These requirements are achieved through the complex use of methods, means or tools, bodies (organizational structures and personnel) and events.

3. Methods and Means of Protection of Communication and Information Systems.

As a result of the accumulated experience, it is possible to make the following most general classification of methods and means of protection.

CIS protection methods may include:

- limiting and managing access to the information and elements of CIS;
- coding of the information;
- regulation of access to information and elements of CIS.

Some of the methods and means of protection are considered separately.

The protection methods are designed to ensure the protection of the information and elements of the CIS in order to prevent access to them.

The first method of protecting CIS is to restrict and manage access. The essence of the method consists in the physical obstacle to the information and elements of CIS located in separate areas / organizational structures /, premises and others. Access management is related to regulating the use of all CIS resources / technical, programmatic, data, documents, bodies /. For the use of part of the resources, a list is made for the persons who have been granted the right to use, by identifying the users and checking the credentials (authorization). Of particular value here may be the construction of a module for checking credentials, logging and alarming when attempts are made to access unauthorized information.

This method is implemented through the organization, individualization and construction of a system for limiting access to CIS resources. It is a subsystem of the CIS protection system.

The second method of protecting CIS is encoding the information. It is carried out by cryptographic hiding / way of writing with conditional characters - encoding / of the information. The substitution methods, the permutation method, the analytical transformation method, and the gamming method are used. The methods of detecting and protecting against viruses and hackers are of particular value in recent times.

These methods can also be singled out as a separate system of protection, within the CIS, and a separate subsystem for cryptographic protection and protection from computer viruses and hackers is formed.

The third method of CIS protection is regulation of access to CIS information and elements and user control.

The method is related to strict regulation of the hierarchy and structural construction of

the data in the created databases (DB) and forcing users to comply with the rules for the processing and use of the protected information, for which they bear material, administrative or other type of responsibility.

These methods may be integrated within the proposed access control subsystem.

The indicated methods of protection are realized by the use of various means of protection. The means of protection can be: technical; programmatic and organizational-administrative activities.

Of the software tools, the most widespread are cryptographic tools, the system of passwords and keys / for protecting files, directories, technical media, databases, system resources /, as well as a firewall. In addition, secure operating systems, secure network software, and a network file system with data labeling and access control are used. This achieves:

- identification of terminals and users;
- registration and control of the operation of technical means, tasks and others;
- program protection, backup and recovery;
- destruction of residual information;
- antivirus prevention and protection.

Technical means of protection include:

- systems for the protection of territories and premises in which specific information is processed and a separate node is located, as an element of CIS;
- means of protection, implementing access control to objects containing CIS resources;
- means of physical protection;
- combined systems.

Systems for the protection of territories and premises include a variety of systems, devices and means of protection.

More typical of them are sensors and converters, laser and optical systems, television systems, cable systems, electromagnetic radiation protection systems and others.

Security measures that control access to CIS resources and identify users include:

- identification cards;
- voice, image, fingerprint recognition systems;
- handwriting and smart card recognition systems;
- hand geometry recognition systems, electronic keys and others.

Means of physical protection are the various mechanical, electromechanical or electronic - mechanical devices and facilities, specially designed to create physical obstacles such as: window and door bars, electronic locks and means of recognition, alarm and security systems, fire alarm systems and various means of Protection.

The combined systems combine the advantages of the mentioned and other technical systems and means of protection.

The specified software and technical means can be distinguished as a system that is a subsystem of the CIS protection system.

4. A Model of the System for Protection of Communication and Information Systems

The CIS protection system model includes two separate CIS protection models developed in organizational and functional aspects.

Organizational-administrative activities for the protection of CIS cover all aspects of protection in the process of development and operation of CIS. This includes the following activities:

• Determination of the strategy for construction and management of the protection depending on:

- the structure, type and degree of secrecy of the classified information / top secret, secret, confidential, for official use /;

- the likely channels for information leakage / risk assessment / and the possibility of providing the planned means of protection;

- the necessary degree of security of the CIS and the economic expediency.

• Implementation of technology for data processing, archiving and recovery;

• Normatively ensuring compliance with the technology of electronic processing and technological prescriptions, work rules, instructions and others, by regulating the powers and obligations of users;

• Organization of control for compliance with protection mechanisms;

• Periodic change of keys, passwords and data;

• Normative-legal activities for the protection of CIS.

The specified protection activities make it possible to build a system of administrative-legal and organizational activities functioning as a subsystem within the CIS protection system.

Subsystems that are distinguished are:

• Administrative-legal and organizational activities for the protection of CIS;

• Program-technical and technical subsystem for protection of CIS;

• Limiting and regulating access to information and CIS;

• Crypto protection and protection from computer viruses and hackers.

These subsystems are components of the information protection system and CIS, in an organizational plan. From a functional point of view, these subsystems are realized within the framework of the following three, relatively independent aspects of protection:

• Protection of information from unauthorized access;

• Protection from information broadcasting;

• Protection of information circulating in CIS.

This is because in modern conditions not only information is protected. Computational and communication resources, designated as means of communication and information support, are also subject to protection, along with the organizations and structures that build, maintain, plan and manage CIS as a whole.

Each of the proposed aspects is sufficiently complex and at the same time important. Here, an approach is taken where CIS protection is taken as a process to which a given system corresponds. The above security aspects are considered in interrelationships and dependencies, as a system of processes. As a result of ongoing synergistic actions related to the various aspects, a process of "protection of CIS and the information circulating in it" is identified. This process is implemented only in cases where the processes related to the above mentioned aspects are carried out at the same time. Therefore, the process "protection of CIS and the information circulating in it" / for short - protection of CIS /, is provided or realized by the processes of protection from unauthorized access, protection from information broadcasting, protection of information circulating in CIS, as well as by the process of protection of CIS and the information from physical destruction.

Using the scientific approach for the correspondence between processes and systems, it follows that the process-adequate "system for the protection of CIS and the information circulating in it" is collectively formed by components such as: protection from unauthorized access, protection from information broadcast, protection of information circulating in CIS, as well as from the component protecting CIS and information from physical destruction.

The protection of CIS covers the different levels of the organizational structure, national, regional, etc. This suggests that the CIS protection system has a hierarchical structure, and in terms of structure and composition, it stands out as a complex system. Components of the CIS protection system are:

- Information protection subsystem from unauthorized access;
- Subsystem for protection of information from broadcasting;
- Subsystem for protection of circulating information;
- Subsystem for protection of CIS from physical destruction;
- CIS protection system at a lower organizational level.

Lower rank CIS protection systems define a hierarchical structure of the proposed CIS protection system in the organization. This suggests that the proposed CIS protection system is a complex system with horizontal components and vertical components.

So far, the structure and composition of two separate models for a CIS protection system have been proposed, considered in organizational and functional terms. However, in order for the subsystems of these models to function as a single mechanism in the interest of CIS protection, the two models must be integrated.

Each of the thus proposed components for information protection and CIS in the organizational aspect is implemented through the components in the functional aspect. For the implementation of the information protection system defined in this way and CIS, relevant bodies or organizational structures are created and function. Applying certain administrative-legal and legal means, these bodies control the construction and functioning of the unified mechanism for the protection of information and CIS.

Based on what has been said, a detailed plan for the activities of the security authorities for information protection and CIS can be developed. The plan may include four sections. These are the subsystems in an organizational plan, and the content of each of the sections includes the protection aspects of CIS and circulating information, separated as subsystems in a functional plan. Part of these sections related to the achievement of information security and CIS in administrative management structures represent a separate subject of research.

5. Conclusion

In conclusion, it can be summarized that CIS protection is a process that is implemented, through technology, and realized in a CIS protection system. The CIS protection system, as a complex system, consists of components that are its horizontal subsystems. These subsystems, organizationally, technically and functionally, are provided and consist of means and organizational structures. In addition, it contains in its composition, as components, the systems for the protection of CIS from a lower level of management. The CIS protection system is a complex system, dynamically functioning, with a hierarchical structure. It is a material environment for the CIS protection process and can be planned, deployed, operated and restored. For this purpose, a specialized working body is created, such as the Commission for the Protection of CIS.

References

1. ISO/IEC Standard 17799 - Information technology -- Security techniques -- Code of practice for information security management (http://www.iso.org/iso/en/CatalogueDetailPage).
2. Vasileva, V. (2022). Application of a Human-Centric Approach in Security by Design for IoT Architecture Development. In: Gelenbe, E., Jankovic, M., Kehagias, D., Marton, A., Vilmos, A. (eds) Security in Computer and Information Sciences. EuroCybersec 2021. Communications in Computer and Information Science, vol 1596. Springer, Cham. https://doi.org/10.1007/978-3-031-09357-9_2,

Published30 June 2022, Print ISBN978-3-031-09356-2, Online ISBN978-3-031-09357-9John Elkington, Cannibals with Forks: The Triple Bottom Line of 21st Century Business, 1998.

3. Vasileva V. (2019). Information Sharing Processes in Public Sector Cybersecurity Structures, Proceedings of the 9th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2019), volume: 100, Publishing Complex – UNWE, Sofia, Bulgaria, p. 136-147, ISSN 2367-7635 (PRINT), ISSN 2367-7643 (ONLINE), Issued for Publication: August 24TH, 2020, URL: http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2019.pdf Leading Corporate Sustainability Issues in the 2012 Proxy Season: Is Your Board Prepared? Ernst & Young, 2012, (http://www.ey.com/Publication /vwLUAssets/2012_proxy_ season/ $FILE/ 2012_proxy_season.pdf)
4. Vasileva V., J. Hlebarov (2020). Methodology for Improving Cyber Awareness Competencies of Employees in Public Administration, Proceedings of the 10th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2020), volume: 100, Publishing Complex – UNWE, Sofia, Bulgaria, p. 278-285, ISSN 2367-7635 (PRINT), ISSN 2367-7643 (ONLINE), Issued for Publication: August 21ST, 2021, URL:http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2020.pdf

Increasing the Power of Mobile Computers for Specialized Models

Georgi Pavlov[1], Vania Ivanova[2]

[1] University of National and World Economy, Sofia, Bulgaria
[2] Bulgarian Defence Institute, Sofia, Bulgaria
gpavlov@unwe.bg, v.ivanova@di.mod.bg

Abstract. The authors discuss the construction of a mobile device with high computing power, suitable for security and defense applications as well as for complex economic models related to COVID 19 measures. The methodology of the computational setting is presented and the pseudocode of a universal software solution suitable for operation of a wide range of hardware devices of mobile type is given. The results of experimental work on several types of mobile devices with different computing capabilities are presented. The author's criterion for selecting a suitable microcomputer device and as the most suitable for a mobile platform for intensive calculations for the specific numerical example is indicated.

Keywords. Computer calculations, ordinary differential equations, embedded mobile system

1. Introduction

Common tasks in practice are related to intensive mathematical calculations. The question is trivial when it comes to office work or laboratory conditions. In field conditions it is difficult to apply computer equipment of general type, because there are typical difficulties such as lack of unlimited power supply, the need for reduced weight and size of the computing device and last but not least the cost of the device because it will work in atypical conditions. high probability of damage or destruction.

A classic computational problem found in security and defense activities is the modeling of the motion of a body thrown at a certain angle to the horizon in order to obtain a certain effect at the end point of the trajectory. In the recent past, the initial conditions for starting the movement were determined exclusively by applying a reference to printed on paper. In our time, in order to increase the accuracy of the hit and not least, reduce the preparation time, various mobile computer systems are used.

In practice, two approaches are known to solve this problem. One is the use of an existing and commercially available general purpose mobile computer device and its equipment with a specially designed software product. The other, non-trivial approach applied by leading manufacturers in the military field is the development of specialized computing modules, the design of which generally remains a company and trade secret.

In this paper, the authors consider possible solutions for organizing a highly efficient mobile computing module and present experimental data obtained during a specific computational experiment.

2. Statement of the computational problem

The model of motion of a body thrown at a certain speed and at a certain angle to the horizon, with certain characteristics such as mass and drag coefficient, is presented as a system

of differential equations. Detailed descriptions of the applied mathematical model can be found in the sources [3, 6].

Figure 1 in the form of a diagram presents the physical and mathematical dependencies that are followed in the construction of the motion model.

The system of differential equations examines the change in the current values of the following quantities:

X (t) - current horizontal coordinate of the movement;

Z (t) - current vertical coordinate of the movement.

The individual symbols in the figure have the following physical entities:

V (t) - change in speed;

λ (t) - change of the angle in the momentary direction of movement relative to the horizon;

$\beta = c_i S \rho_{n0} / 2m$,

where: ci - coefficient of body shape;

$S = \pi d^2/4$ [m²] - characteristic body area;

$\rho_{n0} = 1{,}225$ [kg/m³] - standard air density at sea level;

m [kg] - body weight.

Cx (M) - law of resistance;

H (z) - normalized normal air density at altitude z [m] above sea level;

λ (t-1), V (t-1) - change of the angle in the momentary (t-1) direction of movement relative to the horizon, change of the speed in the moment (t-1);

g - ground acceleration;

M = V/a(z) - Max number;

a(z) - speed of sound at altitude z[m] above sea level;

ρn(z) - density of air at altitude z[m] above sea level.

Dependencies, designated as lookup tables, are characteristics of the traffic environment, set in tabular form.

3. Program implementation

We use a method for solving ordinary 4th order differential equations, known as the classical Runge-Kuta method [1, 5].

The task is formulated as follows:

$$\dot{y} = f(t,y), \quad (1)$$

$$y(t_0) = y_0 \quad (2)$$

In dependence (1) **y** is denoted by a scalar or vector variable, which is a function of the time denoted by **t.** With **y**. the rate of its change is indicated, also a function of the current value of the variable **y** and the time **t**.

The dependence (2) determines the initial conditions. At the beginning the time is e $\mathbf{t_0}$, and the initial value of the studied variable y is denoted by $\mathbf{y_0}$..

The 4th-order Runge-Kuta method for this problem is described by the following dependences:

$$y_{n+1} = y_n + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad (3)$$

$$t_{n+1} = t_n + h, \quad (4)$$

:

c y_{n+1} the approximation of $y(t_{n+1})$, and

$k_1 = hf(t_n, y_n)$, (5)

$k_2 = hf(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1)$, (6)

$k_3 = hf(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2)$, (7)

$k_4 = hf(t_n + h, y_n + k_3)$. (8)

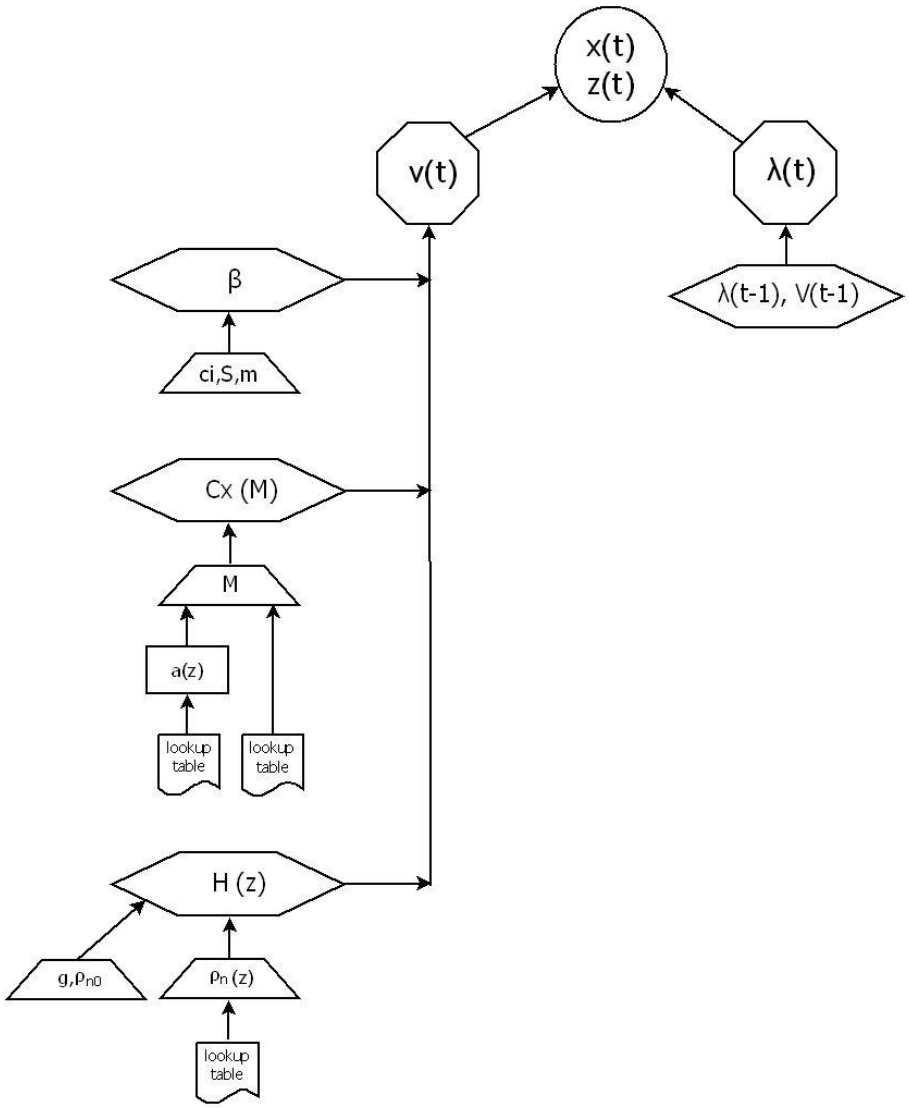


Fig. 1. Computational physical-mathematical model of motion

The new value
$\mathbf{y}_{n+1}$ is calculated using the current value

$\mathbf{y}_n$ plus the value of the four increments of certain weights, where each increment is the product of the magnitude of the interval **h** and the estimated slope given by the function **f** on the right-hand side of the differential equation

$\mathbf{k}_1$ is an increase based on the slope from the beginning of the interval using y_n, Euler's method;

$\mathbf{k}_2$ is an increase based on the slope of an intermediate point of the interval using $y_n+½k_1$;

$\mathbf{k}_3$ is again an increase based on the pressure from the intermediate point, but using $y_n+½k_2$;

$\mathbf{k}_4$ is an increment based on the slope at the end of the interval using $y_n+ k_3$.

Averaging the four increments gives more weight to the increments at the intermediate points. In the case under consideration for the study of the motion of a body thrown at an angle to the horizon (Fig. 1) the functional change of the required variable is described by representation as a system of equations as follows:

$$\dot{y} \equiv \begin{cases} \dot{x} = f(v,\lambda) \\ \dot{z} = f(v,\lambda) \\ \dot{v} = f(v^2,z,\lambda) \\ \dot{\lambda} = f(v,\lambda) \end{cases} \quad (9)$$

The dependencies (1… 9) are applied to create a universal programming code in the C ++ programming language. The algorithm of the software implementation is shown below.

ALGORITHM 1: Ordinary Differential System Equation Solver

Input: x_0, z_0, v_0, λ_0 , t_0 ← beginning conditions
z_k, t_k ← ending conditions
m, d, c ← moving object physical characteristics
Output: x_t, z_t, v_t, λ_t , t ← endpoint results
Constants: g, ρ, lookup_tables(air_dest_table,
air_vel_table,
air_drag_table) ← environment characteristics
h ← solver increment
Variables: Y_temp[4], Y_current[4],
K1[4], K2[4], K3[4], K4[4]
Subroutine: ODE(Y_tmp, K)

```
z = Y_temp[1]
v = Y_temp[2]
λ = Y_temp[3]
K[0] = v*cos(λ)
K[1] = v*sin(λ)
K[2] = -β(m, d, c, v^2, z, lookup_tables) – g*sin(λ)
K[3] = -q* cos(λ)/v
```

Endsub
While($z_t > z_k$) || ($t < t_k$)

```
ODE(Y_current, K1)
Y_temp = Y_current + K1*h*0.5
ODE(Y_temp, K2)
Y_temp = Y_current + K2*h*0.5
ODE(Y_temp, K3)
```

Y_temp = Y_current + K3*h
ODE(Y_temp, K4)
Y_current = Y_current+h/6(K1+2*K2 + 2*K3 + K4)
t = t+h
x_t = Y_current[0]
z_t = Y_current[1]
v_t = Y_current[2]
λ_t = Y_current[3]
Endwhile
Print(x_t, z_t, v_t, λ_t, t)

4. Results of the conducted experimental work

The computational task was performed on several mobile platforms. All devices used in the experiment are commercially available. Exclusively open-source software tools and methods were used to conduct the experiment.

In order to perform an analysis of a certain practical value, in comparative Table. 1 are indicated weight, price and speed indicators of the studied mobile platforms.

Table 1.

CPU	Weight [kg]	Price [USD]	Boottime [s]	Calctime [s]	Alltime[s]
Notebook Intel Core I-7 6500U @2500MHz	1.650	850	31.29	0.04	31.33
Single Board Computer Jetson Nano, ARM Cortex A57 @1430 MHz	0.120	180	10.15	0.15	10.3
System on Chip Raspberry Pi Zero W, ARM1176@1000MHz	0.021	50	46.45	0.68	47.13
Development Board ESP32, Tensilica Xtensa LX6 @240 MHz	0.007	20	~0.00	12.95	12.95
Development Board STM32F103, ARM Cortex M3@72 Mhz	0.006	17	~0.00	61.45	61.45

The comparative data were obtained from freely available sources, or with measurements performed by the authors of the material. The price values are averaged using data from the site amazon.com and are by the fall of 2021. The data on the weight of the studied mobile platforms were obtained by direct measurement of the available copies of the equipment. Minimum fasteners and attachments are also included in the weight of the samples.

The tested devices are of general purpose or potential application for security and defense purposes, as can be seen from known publications [4, 2].

The time-relevant characteristics of the experiment were obtained in two ways. The time to start the device from off mode is determined for Notebook, SingleBoardComputer and SystemonChip devices using the built-in Linux operating system monitoring tool called systemd-analyze. The time interval for performing the computational actions for solving the ballistic task, under the same initial and final conditions for all devices is determined using a system function in the C ++ programming language.

If the authors adopt a criterion for the price of the mobile device below 100 USD and the time for performing the calculations below 1 s, a device of practical interest is the so-called Raspberry PI Zero W with a current price of about 50 USD and a calculation time of 0.68 s. As a total start-up and calculation time, the specified device is located in the middle segment of the time interval of the fastest and slowest device. The disadvantage of this as an object of interest is the high start-up time from the off state. This start-up time of 46.45 s is the highest compared to other tested devices.

Future research on the topic can be focused on reducing the total time for activation and calculation. The calculation time has the possibility to be reduced by applying actions to optimize the calculation procedures, in the currently tested software solution no such measures have been applied. The boot time of the device is the subject of study of the processes at boot the operating system. Here, high hopes for optimization can be sought in the exclusion of unnecessary procedures for the specific task, and even the transition to a simplified version of the operating system without a human-machine user interface.

5. Conclusions

In the article, the authors discuss selected solutions for building a mobile device with high computing power, suitable for applications in the interests of security and defense. This type of processing would contribute to speeding up the work with both COVID 19 data [7] and airport security data processing [8]. A classical ballistic problem is adopted as a basic computational example. The methodology of the computational setting is presented and the pseudocode of a universal software solution suitable for operation of a wide range of hardware devices of mobile type is given.

Temporal numerical data on the performance of the individual tested mobile devices are presented. In terms of practical applicability in field conditions, the comparative table gives the same price and weight characteristics. The author's criterion for selecting a suitable microcomputer device is indicated and Raspberry PI Zero W is indicated as the most suitable for a mobile platform for intensive calculations for the specific numerical example.

An analysis of the strengths and weaknesses of this possible solution has been performed. The following research steps are indicated in order to improve the time characteristics of the selected mobile platform for intensive mathematical calculations.

Acknowledgment

This publication was made under the project" Modeling, analysis and optimization of socio-economic measures to reduce the negative consequences of the pandemic COVID-19 ", contract № КП-06-ДК2/2 of 30.03.2021, funded by the Fund "Scientific research" at the Ministry of Education and Science. The content of the publication is the sole responsibility of the authors and should not be construed as an expression of the opinion of the Research Fund.

References

1. Butcher, J. (May 1963), "Coefficients for the study of Runge-Kutta integration processes", Journal of the Australian Mathematical Society, 3 (2): 185–201, doi:10.1017/S1446788700027932;
2. Mahlyanov D. and N. Stoianov, "Security of Internet of Things – IoT," In: Proceedings of the Eighth International Scientific Conference 'HEMUS-2016', pp. III-217 - III-225, ISSN 1312-2916 (in Bulgarian).
3. Kolev A., Determining the probable zone of damage when activating an improvised explosive device in urban conditions, Information bulletin Counteraction of IED, issue 2, year MMXVV, Military Geographical Center Troyan, pp. 59 - 70, ISSN 2367-8844;
4. Kolev A., and P. Nikolova. "Instrumental Equipment for Cyberattack Prevention."Information & Security: An International Journal 47, no. 3 (2020): 285-299. https://doi.org/10.11610/isij.4720;
5. Delin T., C. Zheng (2012), "On A General Formula of Fourth Order Runge-Kutta Method", Journal of Mathematical Science & Mathematics Education, 7 (2): 1–10;
6. Afanasyev A., E. Mayshev, K. Anisimov, Elaboration of mathematical model of flight trajectory of material point in atmosphere, Vibroengineering Procedia, September 2018, volume 19, p. 246-252, DOI https://doi.org/10.21595/vp.2018.20121.
7. Пенчев Г., К. Пудин„Изследване на икономическите и пространствени ефекти на пандемията COVID-19: данни и модели" (Izsledvane na ikonomicheskite o prostranstveni efekti na pandemiata COVID 19: danni i modeli), сб. доклади от Втора международна научна конференция „Иновации и конкурентоспособност" Светът през призмата на пандемията COVID-19(sb. dokladi ot Vtora mezhdunarodna nauchna konferentsiya „Inovathcii i konkurentosposobnost" Svetŭt prez prizmata na pandemiyata COVID-19), 2021
8. Gechkova T. and T. Kaleeva., "The McKinsey 7S Model in the Airport System Protection", Knowledge International Journal Vol. 42, pp. 843-848, 2020

Applicability of ASP.Net Frameworks for Developing Web-based Business Information Systems

Monika Tsaneva

University of National and World Economy, Sofia, Bulgaria
mtzaneva@unwe.bg

Abstract. This paper defines the most significant functional requirements to the frameworks and tools for development of business software main components. Modern .NET platforms that can be integrated into a single standalone software project are discussed and an analysis of their major pros and cons is made. On this basis, a set of requirements imposing the selection of the most suitable ASP.NET framework for the implementation of the various types of components of a typical web-based business application is proposed.

Keywords. Component-based framework, ASP.NET, development of business information systems, software project management, integration of business applications.

1. Introduction

.NET is one of the most popular open-source development platforms for developing a huge variety of business software solutions, including web, desktop, native cloud and mobile applications, APIs, services and microservices etc. [13]. Such a popularity of .NET is a fact due to the combination of powerful functionality and usability from developers' point of view, nice user experience from consumers point of view and predictability and manageability from managers point of view. .NET, and especially its' extension, ASP.NET, provides many development frameworks which are basically equally beneficial for end users, sponsors, and managers, but the choice of the most suitable one has significant influence on the time, efforts, knowledge, and experience that software engineers must invest during development and implementation process. Applicability of different ASP.Net frameworks that can be integrated into a single standalone software project is the main topic discussed from business software developers' point of view.

2. Functional requirements to the frameworks and tools for development of business software

Each business information system has three types of functions - organization and storage of data, business logic and dialogue with the end user. Established good practices in software design imply the implementation of the principle of "Separation of concerns" [15], according to which different in type and purpose functions - data management, business logic and visualization should not be mixed. Thus, source code must be structured as a set of interoperable software components organized into tiers (see 0). Each separate software component encapsulates only one of the above types of functions and the minimum required by them to solve a specific atomic problem. Components from each software tier can interact, including communicate with other components from the same or an adjacent software tier. The idea is to automate business functions through the orchestration of autonomous reusable

software components, thus ensuring maximum flexibility of the software product and maximum speed and accuracy of the development process. Separation of concerns reduces the complexity of the application by making it easier to program, debug and deploy each component individually, instead of dealing with functionality that is scattered across many components, making them interdependent. Practice shows that the user interface often changes regardless of business logic, but if their program code is common, then when changing the presentation, it is easy to introduce an error in the business logic and vice versa.

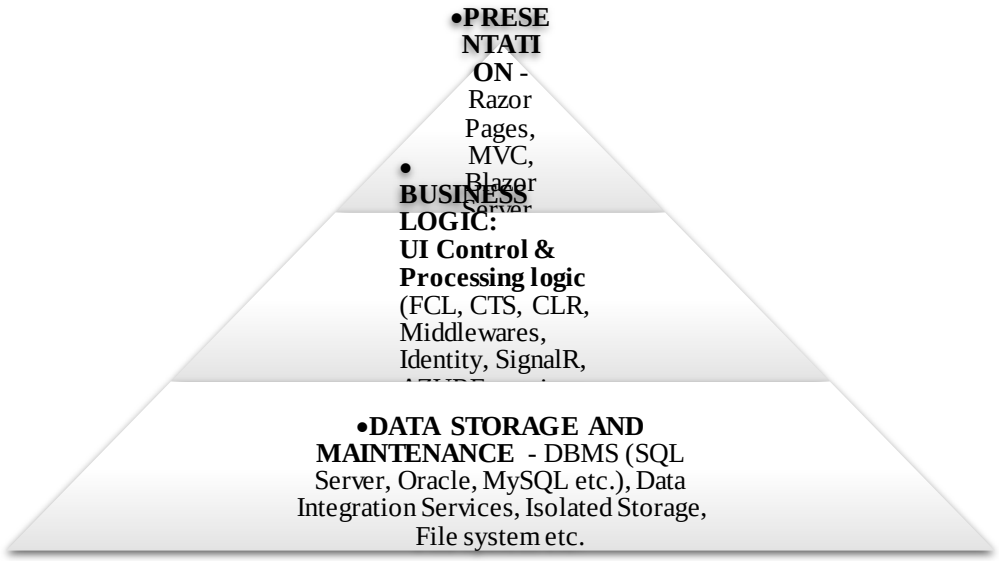


Fig. 1. Multitier ASP.NET Business Software Architecture

The basic requirements for development frameworks largely stem from the opportunities they provide to facilitate and even enforce compliance with this principle. Virtually all development frameworks that have been used in recent years, even those used for database programming, are object-oriented, but modern requirements for the quality of business software, the speed of its development and change suggest much more:

- It is natural to object-oriented program code to provide reusability. In modern conditions, however, the limits of usability are vastly expanded and go beyond corporate systems, allowing interoperability by integrating several internal business applications and platforms [16], and information systems of business partners. Therefore, development and runtime environments should facilitate the integration of software modules ensuring uniform and secure communication [3] and by automating the process of exporting functionality in the form of different types (web, cloud, micro, etc.) of services [4].
- A significant part of the business applications' functionality is standard. For example: the need to organize an access control system [5] based on roles, claims or a combination of both, the need to inform users about the results of real-time processing, compliance with European rules and regulation for personal data processing [17] etc. This requires the development environment to provide developers with a diverse and functionally complete set of ready-made software components for the implementation of such type of standard functionality or at least to ensure a high degree of automation

of the process of creating standard software modules through ready-to-use project templates and built-in or custom libraries. Specific software development [12] can also significantly benefit from the ease of custom libraries integration through dependency injection.

- The process of digital transformation is global [1] and already covers virtually all activities of an extremely wide range of users possessing a variety of smart devices [7]. Development environments should allow for such an organization of program code, in which the same business logic is exposed by different types of presentation (desktop, web, mobile) or non-graphically as a service.
- The organization of data processed by business information systems is becoming more diverse given the emergence of big data [2] with their dramatically increased processing requirements, as well as expanding the traditional functionality of relational databases with capabilities for integration with big data or BI platforms [11]. This process should not require continuous redesign of existing business software, but instead the development framework should include a metadata-based [8] "isolation layer" for data retrieval and propagating of changes back to the data source, which as much as possible hides the specifics of the organization and the storage of data from the components implementing business logic. In addition, data binding support is a must in order to facilitate easy and reliable processing of data entries.
- The creation of business software is often carried out by a large team, including specialists with various types and degrees of qualification. Staff turnover is regular, so easy team members replacement is an essential prerequisite for project success. In these conditions, it is important that the development environment on the one hand imposes discipline and compliance with good practices and standards in programming, and on the other hand allows a combination of different programming paradigms in one project, depending on the type of software component, qualification and experience of the developer

The main requirements for a modern development and runtime framework for the implementation of business software can be summarized as follows (0):

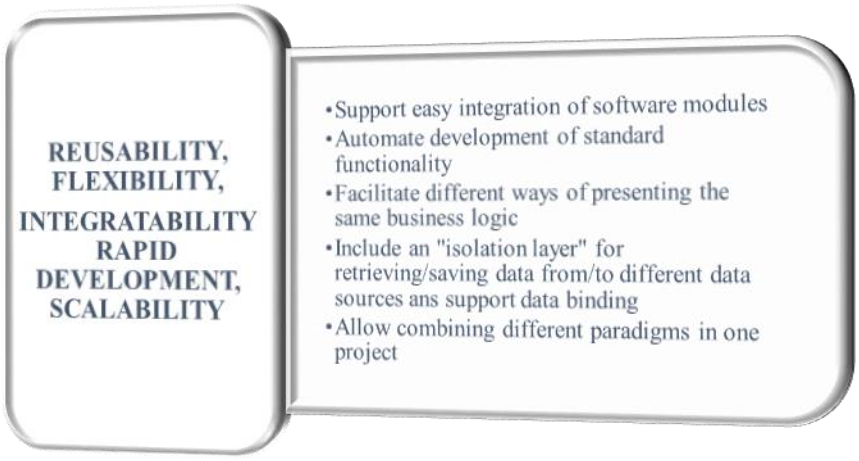


Fig. 2. Functional requirements to the frameworks and tools for development and execution of business software

3. ASP.NET frameworks and tools for developing basic components of web-based business software

One of the most popular frameworks, that meets all the functional requirements for building business software listed above is .NET - an open-source development environment comprising of programming languages and libraries for building different types of software including Web apps, web APIs, and microservices, cloud-based serverless functions and cloud native applications [6], mobile apps [9], desktop solutions, Internet of Things (IoT), machine learning, console applications, Windows services and, games. The base platform provides software components that can be used in practically all types of applications. ASP.NET is an additional framework that extends .Net, and is a platform-independent, open-source framework for creating modern, software systems including cloud and web applications and services, Internet of Things (IoT) applications, and backends for mobile apps using development tools on Windows, macOS, and Linux and deploy them both on-premises or to the cloud and run them on .NET CLR [13]. It adds to the .NET platform basic framework for processing web requests using modern powerful programming languages, general use Web Model Libraries (MVC), syntax (Razor) for making web page templates and dynamic web pages, complete authentication system (database models, libraries, and pages with templates for user entry), including multifactor authentication and integration with most popular external identity providers.

ASP.NET supports MPA (traditional or multi-page web applications) and SPA (one-page applications) and Web APIs [10].

- Multi-page web applications include little client-side functionality and perform post back (web query, resulting in a whole page render and load in the end user's browser) to the web server for all the data manipulations, business logic and user interface control. Multi-page applications can be developed using classic MVC pattern or an alternative programming model named Razor Pages. In both cases some web page operations can be enhanced with AJAX functionality implementing some client-side logic (e.g., validation) and partial updates, but the overall architecture of the application is based on MVC views and URL endpoints. A third option for MPA development is Web Forms, but due to poor performance and some conceptual design issues it became deprecated and will not be discussed in this paper.
- Single-page (SPA) applications use web APIs for data manipulations and can provide better user experience. They occasionally include dynamically rendered downloads on the server page. Usually, SPAs are designed as static HTML files that load some JavaScript libraries needed for application start and execution. Blazor is the most modern and powerful platform for developing interactive SPA web applications with rich client-side functionality for a wide range of browsers (including mobile), which allows sharing application logic written on .NET by server and client using C # instead of JavaScript:
- Web APIs is a platform for easy creation of HTTP services that will be used by various clients, incl. browsers and mobile devices. The ASP.NET Web API is used to develop RESTful applications for .NET [14]

It is typical for business software that it contains various modules and is used by users with various information problems. From this point of view, one of the most significant advantages of ASP.NET is the ability to combine all the software paradigms supported by the framework in the same project and build them using the same set of tools and core platform libraries, most important of which is Entity Framework – the preferable way to organize application's data access layer and make it accessible in a uniform manner as a service.

Development in MVC compared to Razor Pages requires more effort due to the need to comply with many conventions related to dynamic routes, the correct naming of multiple files, etc. The project organization in Razor Pages is simpler as each page comprises of a file with Razor View (same as MVC) and only one other file with the whole program code (page model) instead of a bunch of controllers and models dispersed in different folders. With Razor Pages responsibilities are clearly defined, as each page is self-sufficient with its own look and program code which simplifies code understanding, testing and documentation. Eventually, both these frameworks require the same basic set of knowledge from software engineers and the choice between them is based more on personal experience and preferences than on major principal differences so they can be easily combined in a single solution built by different team members.

Blazor differs from other platforms with the ability to control both the client and server part of the application with C # code without the need for JavaScript as a second technology. Instead of sharing the source code between the two parts of the application - server and client, in Blazor one and the same model is defined once and reused in both parts of the application. The application takes full advantage of the server, including the use of the .NET Core API. Blazor offers two development and hosting models [13]: Blazor server and Blazor Web Assembly.

With Blazor server An ASP.NET Core server is needed to control the application because the program code, including the application component code, is not sent to clients and the application takes full advantage of the server, including the use of the .NET Core API (e.g., EF if data access is required) and all.NET tools (e.g., debuggers), work as expected. Blazor Server applications work with browsers that do not support WebAssembly and on devices with limited resources, but there is a higher latency, multi-user application scaling uses server resources to handle multiple client connections and client status and every interaction with the user involves network traffic [10].

The Blazor WebAssembly hosting model when an ASP.NET Core web server is not required to host the application, offers several advantages: Once the application is downloaded it becomes independent from the .NET server, so it remains functional even when the client has no Internet connectivity, the work is unloaded from the server to the client, client resources are fully utilized. The Blazor WebAssembly hosting model has some drawbacks like the limitation of the application to the capabilities of the browser, requirements of appropriate client hardware and software, downloaded pages are larger and applications take longer to load [14].

4. Framework selection criteria for developing typical business application components

As mentioned above, in the most common scenario business software comprises various modules that target different users performing different kinds of activities therefore have different information problems (see 0). On the other hand, software is designed and programmed by developers possessing different knowledge, qualification level and experience. Thus, many web applications need to include a combination of traditional web application behavior (usually for content, business logic, and data processing), SPA (for interactivity) and web APIs for integration and interoperability with other business solutions.

All platforms utilize the benefits of ASP.NET, but each involves a different style of programming and requires different knowledge and experience. The choice of platform basically depends on the type and requirements of software module being developed, the style and qualification of the developers.

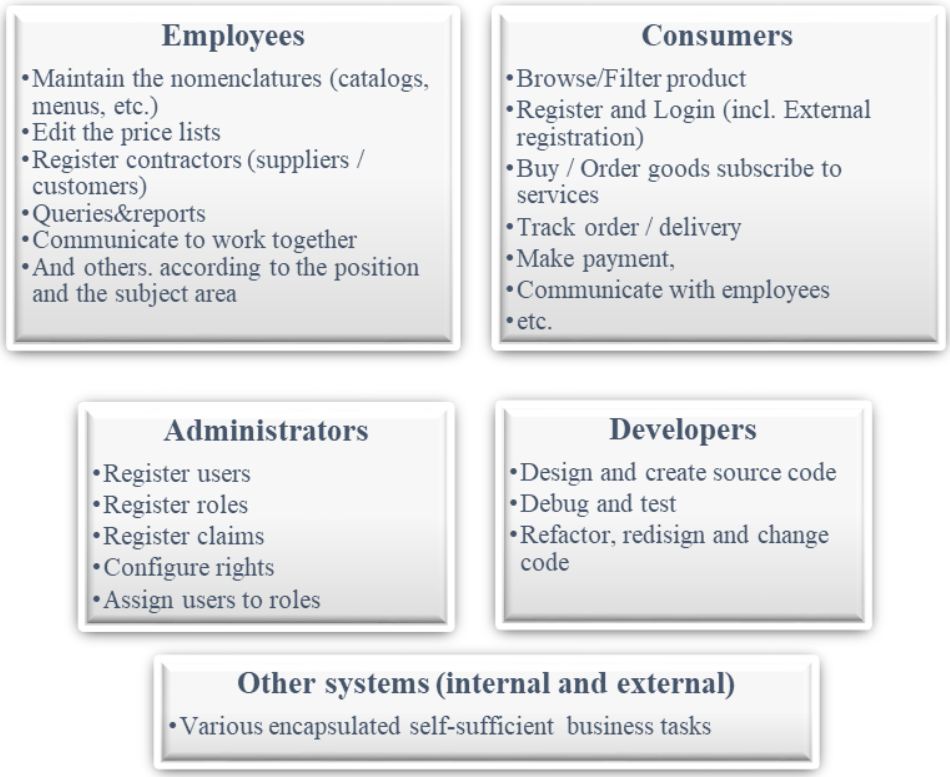


Fig. 3. Business software stakeholders and their typical activities

Based on the principle of "Separation of care" established according to world standards and good practices, the business logic of the application should be encapsulated as much as possible in a separate layer of separate classes (or microservices), accessible by user interface management components, or be exported via web APIs. In this manner, for example, all components for data retrieval, manipulation, and recording (data access layer components), the standard logic for user access organization and control, communication hubs for real-time processing, etc. should be implemented.

The first choice related to the system's presentation tier (see 0) to be made, is about the type of web application - SPA or MPA.

Basically, the modules of the system intended for use by end users, external to the company, should be designed and built with the idea to be more interactive and more attractive. Their typical functions are related to frequent search, filtering and sorting of information, visualization and resizing of graphic images, work faster if implemented with more client-side then server-side logic. Thus, SPA platforms are more suitable in this case.

On the contrary, the modules designed for use by company employees include more business logic, larger and more specific data structures and operations, more data retrieval, processing, and propagation of changes back to the data source. In this case, the emphasis is on the simplicity and productivity required in the daily operational activities and the user interface of the system must closely follow the way of performing business functions according to the approved procedures in the company. These features make server-side implementation and in particular MPA the more suitable platform.

The case of administrative modules is specific, as their functionality on the one hand is associated with a lot of data manipulations, which is suitable for server-side processing, and on the other hand is very typical, which suggests using a development environment offering many standard user interface components (like pages for log in, log out, change password, register, 2FA, edit user profile etc.) and related generators

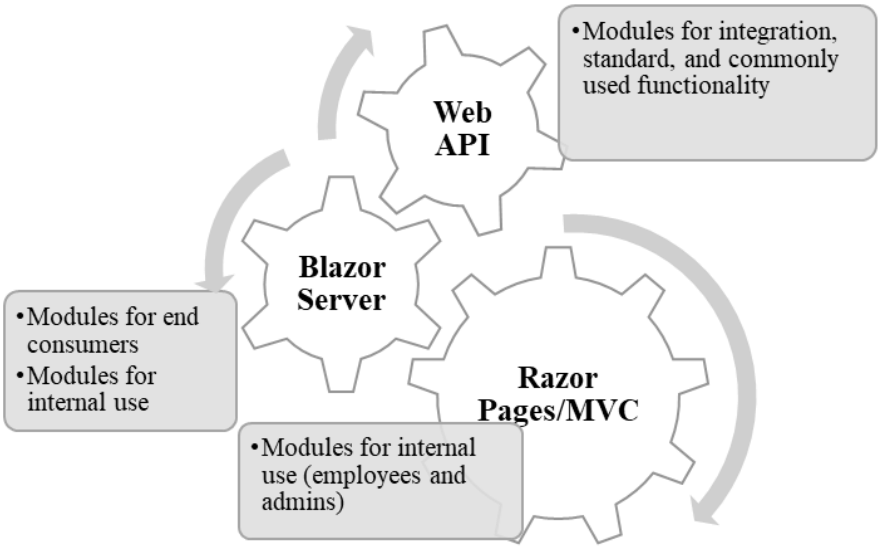


Fig. 4. Selection of ASP.NET development and implementation framework for developing business software components

The requirements of business software, which have the major influence on the choice of ASP.NET development and implementation framework, can be summarized as follows:

- Data base connectivity and data operations – server-side processing is a must which makes Blazor WebAssembly unapplicable. All the other frameworks are equally applicable because in a well-structured application all the data base logic is encapsulated in a separate data access tier accessible via API calls to the ASP.NET Core server
- Real time communication and page content updates – It is recommended to use the Azure SignalR service for Blazor Server applications. The service allows scaling of the Blazor Server application for a large number of simultaneous SignalR connections and is much more natively embedded in Blazor compared to other (MPA) frameworks. SignalR hubs are supported in both MVC and Razor Pages, but their utilization is done manually and requires a large amount of JavaScript code implemented on the client.
- Rich end user experience – undoubtedly from this point of view Blazor is the most powerful framework
- The type and computing power of the client workstations - the components running server-side practically do not impose requirements for computing power of the client machines, therefore they are preferable in conditions where the number and types of workstations for access to the application are not known
- Need to use a software component offline - this is not typical for business software modules, but, if necessary, the only alternative is Blazor WebAssembly

- Rapid development – Razor pages and Blazor provide the richest set of predefined and ready to use software components. Razor pages is slightly ahead, especially with those based on standard data binding models (e.g., Identity).

- Developers' knowledge of Java script and derivate libraries – Blazor applications do not involve any Java script at all, and Razor Pages requires just a general idea of jQuery in order to build standard client-side logic like validation

Based on these considerations, MPA frameworks like pure MVC and Razor Pages look more suitable for developing software modules that target company employees and system administrators, and SPA frameworks like Blazor web assembly and Blazor server facilitate building rich UI modules designed for end consumers (see 0).

From developers' point of view, MVC and Razor pages frameworks ensure equally full control over HTML and CSS for page content rendering, but letter platform provides higher degree of automation based on its' richer set of page scaffolding templates and data binding via page model. Blazor is the choice of developers who are not comfortable with java script and libraries built on top of it like jQuery, TypeScript, Angular etc.

5. Conclusion

The review clearly shows that each .NET framework has its pros and cons and none of them can be pointed out as the panacea or best solution of all the problems, but from developers' point of view, Blazor server and Razor Pages look like the most perspective to use. The approach proposed is to combine different ASP.NET frameworks for building the business application presentation tier in an altogether software solution, making their selection based on functional requirements and considerations that were discussed above and minding both end users' and developers' concerns.

References

1. Boyanov L., The Digital World - The Change, The global digital transformation - enriching or impoverishing humanity, ISBN 978-619-239-637-4, Avangard Prima Publ., Sofia 2021, 188 p.
2. Delchev D., V Lazarova, Big Data Analysis Architecture, Economic Alternatives, 2021, 2, pp. 315-328 (https://www.unwe.bg/doi/eajournal/2021.2/EA.2021.2.09.pdf)
3. Kouzmanov, S., Analysis and Classification of Business Software Systems Integration Approaches, Economic Alternatives, 2015, 2, (https://www.unwe.bg/eajournal/en/journalissues/article/7770)
4. Kouzmanov S., M. Tsaneva, Impact Of Business Problem Characteristics On The Architecture And Specification Of Integration Framework, 3-rd International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE-2013), 2014. (http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2013.pdf).,
5. Lazarova, V., Managing User Access to Cloud Services by Company Administrators. 2016 TEM Journal, 5(3), p.289-293
6. Lazarova, V., Principles of Transactional Approach in the Classical Web-based Systems and the Cloud Computing Systems - Comparative Analysis, Economic Alternatives, 1/2013, (https://www.unwe.bg/uploads/Alternatives/Article9_01.2013.pdf)
7. Marzovanova M., Building Multi-Touch User Interface, 4TH International Conference on Application of Information and Communication Technology and Statistics in Economy And Education (ICAICTSEE-2014), 2014, (http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2014.pdf).
8. Mihova V., Murdjeva A. Metadata for generating a specific data warehouse. International Conference on Application of Information and Communication Technology ans Statistics in Economy and Education (ICAICTSEE-2012), Sofia, Bulgaria, 2012. (http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2012.pdf)

9. Mihova, V., Stefanov, G., Marzovanova, M., Cognos Mobile -dashboards design and implementation technology, International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE-2016), 2016. (http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2016.pdf).
10. Pickett W, Choose an ASP.NET Core web UI, 2021, https://docs.microsoft.com/en-us/aspnet/core/tutorials/choose-web-ui?view=aspnetcore-6.0
11. Radoev M., A Comparison Between Characteristics of NoSQL Databases and Traditional Databases, Computer Science and Information Technology, Vol. 5(5), 2017, San Jose, CA, USA, (http://www.hrpub.org/download/20171130/CSIT1-13510351.pdf).
12. Radoev M., Specific Requirements for the Accounting Software in Budget Organizations, Ikonomiceski i Sotsialni Alternativi, University of National and World Economy, Sofia, Bulgaria, 2016, issue 3, pages 93-98.
13. Richard, Announcing .NET 6 — The Fastest .NET Yet, 2021, https://devblogs.microsoft.com/dotnet/announcing-net-6/
14. Roth D, R. Anderson, and Sh. Luttin, Introduction to ASP.NET Core,2021, https://docs.microsoft.com/en-us/aspnet/core/introduction-to-aspnet-core?view=aspnetcore-6.0
15. Smith S. Et al , Architect Modern Web Applications with ASP.NET Core and Azure, 2021, https://docs.microsoft.com/en-us/dotnet/architecture/modern-web-apps-azure/
16. Tsaneva, M., A practical approach for integrating heterogeneous systems. Business Management 2/2019 (http://bm.uni-svishtov.bg/title.asp?title=1383).
17. Tsaneva, M., Implementation of GDPR Requirements in the Information Systems of Consumer Financing Companies, Serdica Journal of Computing, Volume 13, Number 3-4, 2019, pp 155-170 (http://serdica-comp.math.bas.bg/index.php/serdicajcomputing/article/view/346/315)

Development of Cybersecurity Policies in Higher Education Institutions

Rosen Kirilov

University of National and World Economy, Sofia, Bulgaria
rkirilov@unwe.bg

Abstract. The paper addresses cybersecurity issues in higher education. The global Covid-19 pandemic has significantly increased the use of online platforms in the educational process. In addition, the requirements for these platforms in terms of user data have increased significantly. The analyzes are the basis for building comprehensive cybersecurity policies. Finally, conclusions and recommendations are outlined.

Keywords. Information technologies, cybersecurity, education.

1. Theoretical background and regulations

The last few years have seen a significant increase in the use of Internet resources by citizens and businesses. Several new applications, online learning platforms, remote work systems, new social networks and more have appeared.

All this has forced public organizations to increase the quantity and quality of digital services provided [1]. These services should be available through the organizations' website. Presence on the Internet, as well as maintaining an up-to-date and highly detailed website, is an important part of the activities of any organization [2]. The expansion of digital services requires the creation of new concepts regarding data security.

Creating comprehensive cybersecurity policies is now part of the proper functioning of any public institution. The normative regulation of these processes is largely determined by the Cyber Security Act and the Ordinance on the minimum requirements for network and information security [3]. These documents set out network and information security measures, such as:

- Organizational;
- Technological;
- Technical [3].

The application of concrete measures depends on the specifics of the given organization. For these reasons, finding a balance between them ensures the creation of an effective cybersecurity system.

Another essential element in the creation of the system is the availability of adequate data from measurements and observations of traffic. This is a rather complex process and requires monitoring the operation of all systems for months. Both external traffic outside the organization and the behavior of the internal network are observed. Surveillance as a tool of analysis is very important for the creation of security policy and the construction of an adequate protection system.

2. Creating a cybersecurity policy

The development of cybersecurity policy should be based on current regulatory requirements. According to them, such a document is mandatory for every organization. The review of its topicality is performed once a year.

In the present study it is proposed to develop a cybersecurity policy for the field of higher education to include (Fig. 1):

- Classification of information;
- Classification of user types;
- Information assets management;
- Security of newly created applications;
- Risk management.

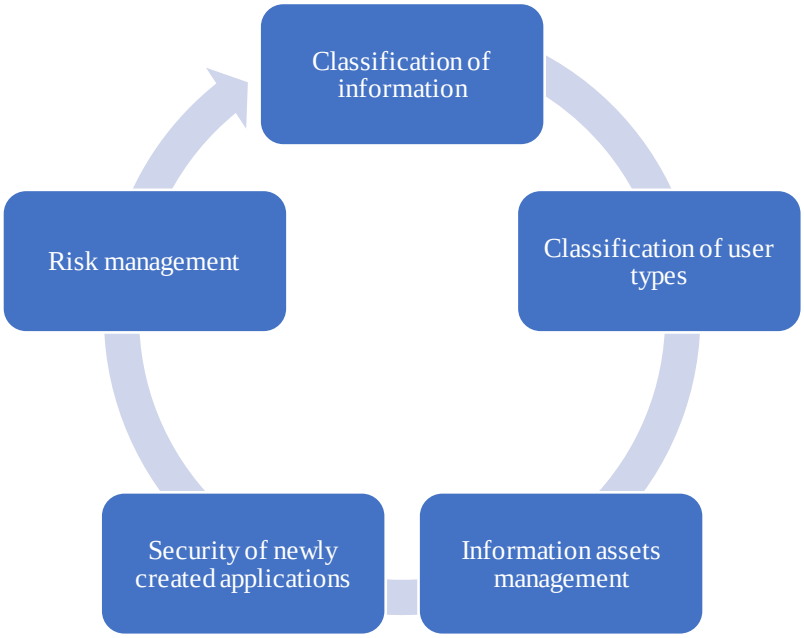


Fig. 1. Stages in developing a cybersecurity policy

The starting point for creating a cybersecurity policy is to find a match between the types of information in the organization and the types of users. Building such a matrix is key to the success of cybersecurity policy.

Another very important thing is to find a match between the users of the organization and their access to certain information resources. This compliance should always be kept up to date and allow easy identification of network access types.

Finally, the key role in building and implementing a successful policy is the role of human capital in the organization. Adopting and adhering to the measures that result from cybersecurity policy is key to its success. The success of the policy developed depends to a large extent on the way in which the new commitments and obligations are presented.

3. Implementing a cybersecurity policy

The implementation of cybersecurity policy is related to the introduction of measures to monitor the network of higher education. The implementation of such software allows centralized monitoring (Fig. 2).

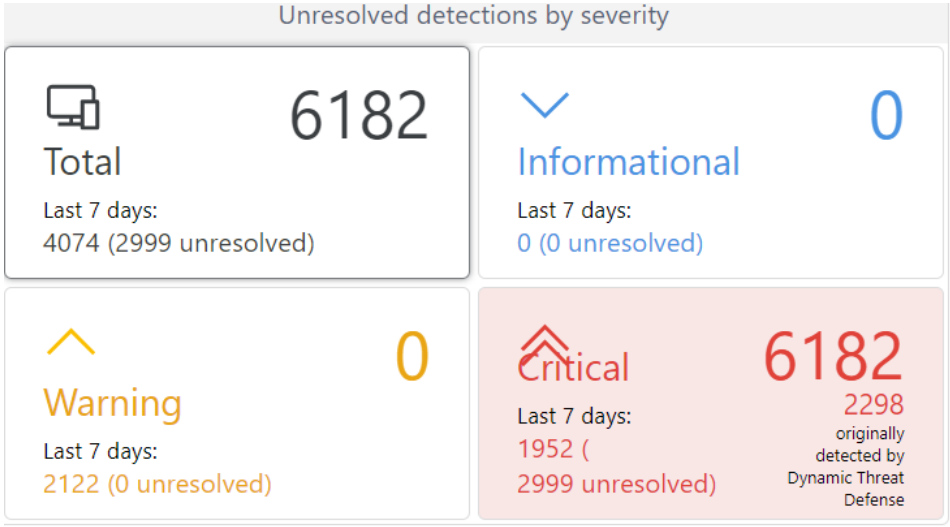


Fig. 2. Centralized monitoring

As a result of the measures taken, the dynamics of the load on the assets and the problems with them can be monitored (Fig. 3).

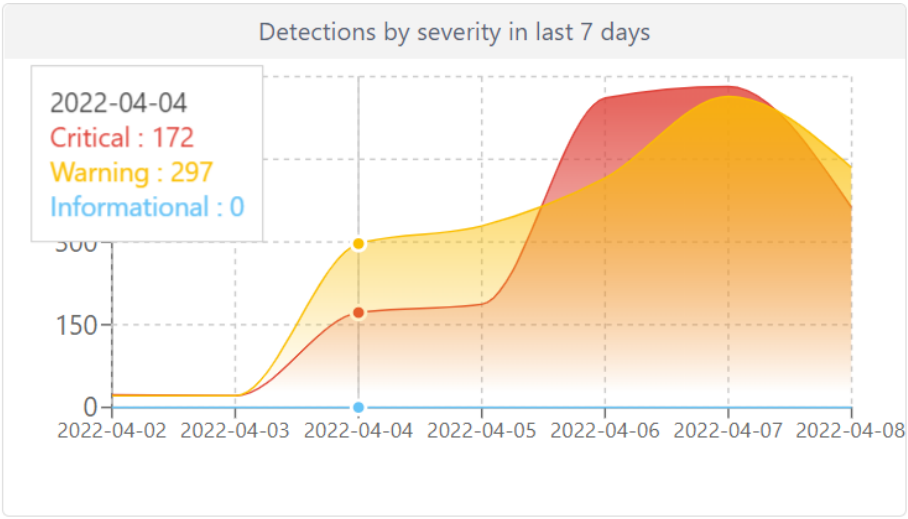


Fig. 3. Identification of problems

Based on such observations, it is good to install specialized software for:

- Reporting the status of workstations and servers.
- Possibility to impose different security policies.
- Malware outbreak notification.
- Creating various dynamic groups to perform security tasks.
- Role-based access of administrators.
- Detailed reports on the status of workstations and servers, etc.

Following the introduction of a cybersecurity policy, a stage of analysis of the achieved results follows. The main directions for future improvement should be:

- Placing all public applications under protection and monitoring.
- Introduction of an IT asset inventory system.
- Introduction of a ticket system for traceability, management and accountability of actions performed by employees.
- Gradual removal of obsolete and unsupported applications and platforms by manufacturers.
- Rewriting with the latest technologies of software applications in accordance with security requirements.
- Additional network segmentation to achieve better access management.

4. Conclusions

The development and implementation of cybersecurity policy in public organizations can allow the identification of:

- Coin miner, "digging" for cryptocurrency.
- Attacking public services (websites) of the organization.
- Use of weakly protected protocols and applications.
- Publicly available services with weak security mechanisms, such as passwords, administrative panels, etc.
- Many seriously vulnerable and obsolete IT assets.
- Alarms for potentially unwanted or dangerous applications installed on the organization's workstations.

Acknowledgement

This work was supported by the UNWE Research Programme (Research Grant No. NID NI-7/2020).

References

1. Kirilova, K., Status of digital administrative services in the Republic of Bulgaria, SHS Web of Conferences, Vol. 120, 2021, ISSN 2261-2424.
2. Milev, P., Opportunities for implementing Internet monitoring in public organizations, Economic and Social Alternatives, Issue 2, 2018, ISSN 1314-6556.
3. Ordinance on the minimum requirements for network and information security, 2022, 30.04.2022, https://www.mtc.government.bg/sites/default/files/nar_minimalnite_iziskvaniq_mrejova_info_sigurnost-072019.pdf

Approach for Analysis and Comparison of Search Query Results in Web Publications

Plamen Milev

University of National and World Economy, Sofia, Bulgaria
pmilev@unwe.bg

Abstract. The paper addresses issues related to analysis and comparison of data in different forms. The main concept of the research is focused on the possibilities for analysis and comparison of search queries in web publications in the form of different charts. The specifics of a traditional approach for analysis and comparison of the popularity of search queries are presented. A modern approach for analysis and comparison of search query results in web publications is proposed. Promising directions for future research are outlined.

Keywords. Data analysis, Internet, search query, web publication, data science.

1. Introduction

Presenting the results of data analysis in comparable forms is not an easy task. Typically, different analytical tools rely on their own approaches for visualization of the results of analytical data processing. However, there are some imposed standards, due to which most users are accustomed to seeing certain graphs in the study area. At the same time, organizations engaged in analytical data processing are discovering new ways to visualize the results of data analysis, which in many cases can be defined as innovative. These new ways are also a result of the development of information technology, which provides specialists with new opportunities to present data. Some authors also emphasize the fact that delivering data is not enough, even if it is delivered in modern, highly graphical ways [6]. Despite modern visualization tools, a key factor in data presentation is the focus on data. The ability to compare data from different sources or for different entities is also essential in the context of today's digital economy. Some authors define data analysis as one of the areas with the most significant impact of digitalization [1]. In the public sector, research is constantly increasing in the field of presenting trends in the development of various services, which need appropriate means for graphical comparison of data, which may relate to different regions or countries [4]. The education sector is no exception, where the introduction of modern information technologies creates preconditions for improving the management of business processes [3].

The purpose of this research is to review the traditional approaches for data analysis and comparison in order to propose an approach for analysis and comparison of search query results in web publications. The most significant consideration is that this approach should be built on the basis of the specifics of the data in web publications.

2. Traditional approaches for analysis and comparison of the popularity of search queries

There are classic examples of approaches for analysis and comparison of data against relevant criteria. A well-established approach in the subject area is that of Google Trends [2].

Their tool analyzes the popularity of search queries in different regions. Charts in different forms are used to show the results of the analysis, which compare the volume of data over time, as well as the regional distribution of results. For the purposes of the present study, the search terms presented in Table 1 are used.

Table 1. Search terms, used in the research

1	Apple
2	Facebook
3	Google
4	Microsoft

The chosen terms point to some of the largest companies in the IT industry [7]. Comparing their popularity involves the use of different graphics. For the purposes of this study, we will use the following graphical distributions of data:

- Line chart for time distribution of results.
- Bar chart for summary presentation of average or total results.
- Map chart for distribution of results by regions.

The study period covers data for the whole year of 2021.

Fig. 1 presents a traditional line chart for analysis and comparison of the popularity of search queries.

Fig. 1. Traditional line chart for analysis and comparison of the popularity of search queries

In this graph, as in all other graphs in the present study, the results for the individual search terms are displayed in different colors that are randomly selected. The blue color is for Apple, the red color is for Facebook, the yellow color is for Google, and the one for Microsoft is green.

The presented data show a relative consistency in the results for the studied period. The most searched term of all time for the period is Facebook, followed by Google. The terms Apple and Microsoft share and exchange third and fourth place.

Fig. 2 presents a traditional bar chart for analysis and comparison of the popularity of search queries.

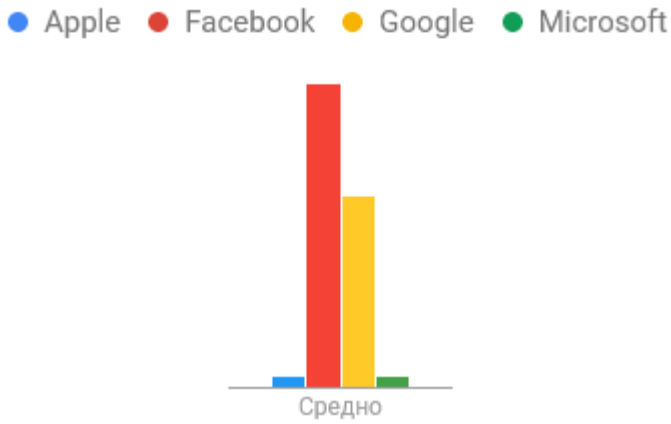


Fig. 2. Traditional bar chart for analysis and comparison of the popularity of search queries

The presented data show that on average for the period the most common term in searches is Facebook, followed by Google. The advantage of Facebook and Google is significant compared to the third place with equal indicators for the period Microsoft and Apple.

Fig. 3 presents a traditional map chart for analysis and comparison of the popularity of search queries.

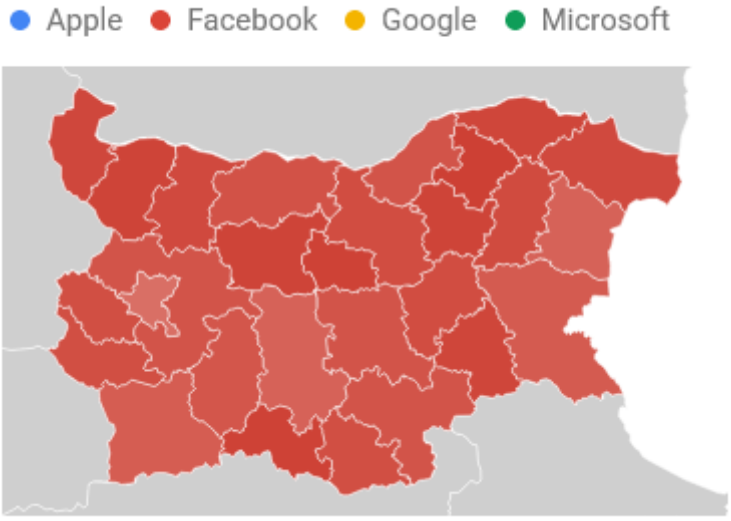


Fig. 3. Traditional map chart for analysis and comparison of the popularity of search queries

The presented data show the absolute regional dominance of Facebook for the study period. This term is dominant in all regions in Bulgaria. None of the other three terms participates in search queries more than the term Facebook in any region.

After researching the possibilities of Google Trends and conducting research on selected terms for analysis and comparison of the popularity of search queries, we should consider the possibilities for analysis and comparison of search queries in web publications.

3. Analysis and comparison of search queries in web publications

To provide data for the study of the possibilities for analysis and comparison of search queries in web publications, we will use a platform for media monitoring and analysis, which provides opportunities for online tracking of publications and frequency on certain topics and keywords in the Bulgarian Internet space [5]. The profile specialization of the platform includes web publications from news websites, news portals, regional online media, online information agencies, institutional websites, websites of political parties, blogs, etc. For the purposes of this paper, we will repeat the study with the selected terms (words), but this time the focus will be on the possibilities for analysis and comparison of search queries in web publications. The approach proposed includes similar graphical data distributions, but with a new focus. The study period is the same so that there can be some comparability of the results.

Fig. 4 presents an approach for analysis and comparison of search queries in web publications in the form of line chart.

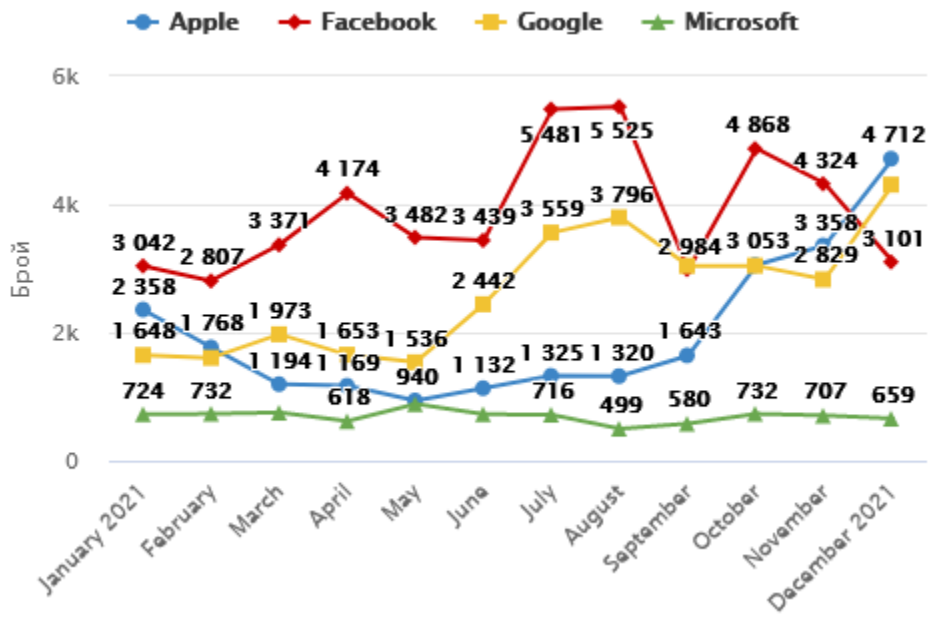


Fig. 4. Approach for analysis and comparison of search queries in web publications in the form of line chart

The data presented show that Facebook has been the dominant word for almost the entire study period. At the same time, there are shifts in the positions of the searched words. The least mentioned word in web publications is Microsoft. There are some similarities with the traditional line chart for analysis and comparison of the popularity of search queries.

Fig. 5 presents an approach for analysis and comparison of search queries in web publications in the form of bar chart.

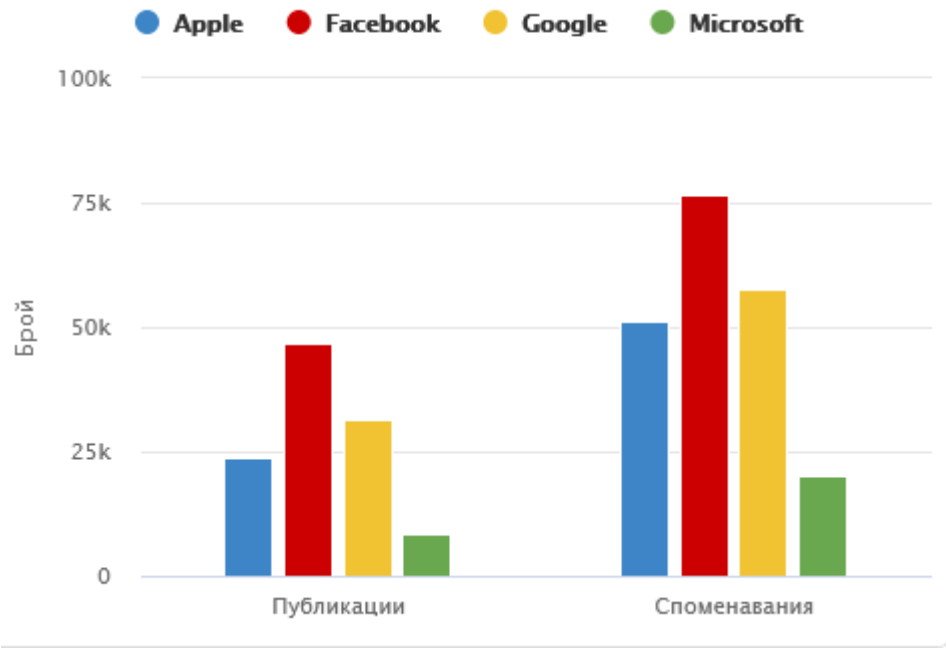


Fig. 5. Approach for analysis and comparison of search queries in web publications in the form of bar chart

The graphic consists of two parts. The left part (publications) shows a comparison of the four words according to the number of web publications in which they are present. The right part (occurrences) shows a comparison of the number of occurrences of the four words in the respective web publications. Usually both parts are interesting in the context of the opportunities for analysis they provide. For some types of analysis, the number of web publications is important, no matter how many times the words occur. For other types of analysis, the number of occurrences of the respective words is important. Usually, most web publications suggest more occurrences of the words, and vice versa – the presence of more occurrences of a word implies more web publications in which it occurs. Very often the two graphs represent the same ranking of the searched words, but with a difference in the specific numbers, which is logically in favor of the occurrences.

The presented data show that in total for the period the most common word in web publications is Facebook, followed by Google. The advantage of Facebook and Google is not as significant as it was in the traditional bar chart for analysis and comparison of the popularity of search queries. In third place is Apple, and in fourth place is Microsoft. The ranking is the same both in the context of the number of web publications and in the context of the number of occurrences of the search terms in the respective web publications.

Fig. 6 presents an approach for analysis and comparison of search queries in web publications in the form of map chart.

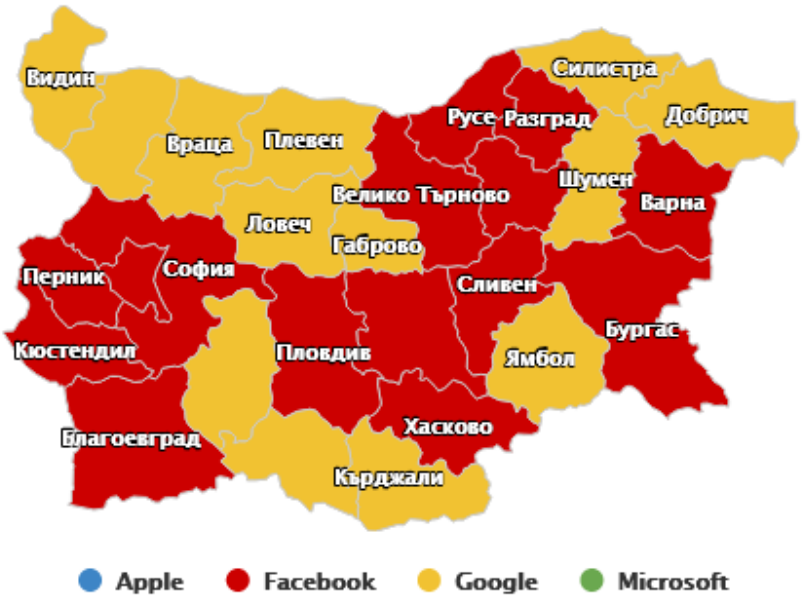


Fig. 6. Approach for analysis and comparison of search queries in web publications in the form of map chart

The presented data show a serious regional dominance of Facebook and Google over Apple and Microsoft in the surveyed web publications. Facebook is the most common result in web publications in the four largest regions in Bulgaria – Sofia, Plovdiv, Varna, and Burgas. It can be said that these results are somewhat expected given the data from the traditional map chart for analysis and comparison of the popularity of search queries, which showed that Facebook is the most searched term compared to the other three.

4. Conclusion

Based on the presented study the following conclusions can be outlined:

- The possibilities for analysis and comparison of data in different forms is crucial for the organizations concerned.
- Both traditional analysis and comparison of the popularity of search queries and analysis and comparison of search queries in web publications are of interest for research.
- More and more organizations are interested in data from web publications and in this sense approaches for analysis and comparison of this type of data is an actual research topic.

For these reasons, the presented approach is a promising area for future research.

References

1. Belev, I., Comparing Business Process Management and Case Management, Proceedings of the 10th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2020), UNWE – Sofia, 2021, pp. 451-457, ISSN 2367-7635.

2. Google Trends, 2021, 01.02.2022, https://trends.google.com
3. Kirilov, R., Current Trends in the Career Development of Students, Proceedings of the 9th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2019), UNWE – Sofia, 2020, pp. 219-222, ISSN 2367-7635.
4. Kirilova, K., Trends in the Development of E-services in Bulgaria, Proceedings of the 10th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2020), UNWE – Sofia, 2021, pp. 441-445, ISSN 2367-7635.
5. MediaBoard, 2021, 01.02.2022, https://www.mediaboard.bg
6. Mihova, V., Stefanov, G., & Marzovanova, M., Cognos Mobile – Dashboards Design and Implementation Technology, Proceedings of the 6th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2016), UNWE – Sofia, 2019, pp. 514-519, ISSN 2367-7635.
7. Statista, 2021, 01.02.2022, https://www.statista.com/topics/4213/google-apple-facebook-amazon-and-microsoft-gafam/

Business Processes in Accounting Using Big Data

Vanya Lazarova

University of National and World Economy, Sofia, Bulgaria
vlazarova@unwe.bg

Abstract. Big data in accounting is mostly related to the processing of complex data coming from multiple sources such as various databases, websites, payment systems and more. In accounting, it is important to link and compare data coming from different sources. The problem is that the huge amount of the big data is unstructured and semi-structured. Primary documents are used in accounting and existing information must be extracted from them, whether it comes from an unstructured source, such as an invoice received by e-mail or a card transaction notice. The business processes in accounting most closely related to big data are the processes of recording business operations related to customers and suppliers. This report will be considered on description of a business process using big data - settlements with clients. This business process was chosen because in accounting, it most clearly represents the relationship with big data. The accountants who prepare the entries in the current account for the clients verify the data from the primary documents and the bank data, ensuring the accuracy of the entries. An important part of the work of these accountants is also the analysis of receivables from customers. By analyzing and interpreting financial information, accountants can provide control and solutions to uncollected receivables.

Keywords. Accounting, Big Data, business processes with Big Data, settlement with clients, Big Data in Settlement

1. Big data in settlement with suppliers and customers

Big data in accounting is mostly related to the processing of complex data coming from multiple sources such as various databases, websites, payment systems and more. In accounting, it is important to link and compare data coming from different sources. The problem is that the huge amount of the big data is unstructured and semi-structured. Primary documents are used in accounting and existing information must be extracted from them, whether it comes from an unstructured source, such as an invoice received by e-mail or a card transaction notice.

Settlement with suppliers and customers is one of the business processes in accounting that uses big data. Settlement relations have two sides - settlements with suppliers and with customers. In Bulgaria there is a recommended chart of accounts, which has separate accounts for the two types.

The individual recommended chart of accounts for reporting settlement supply relationships provides for accounts from group 40 - Suppliers and related accounts. Depending on the suppliers and the conditions of delivery to this group the respective synthetic and analytical accounts are kept. We need to clarify that in the case of budget organizations there are specific requirements in the chart of accounts [2], but nevertheless, the business process using big data remains the same.

Business transactions resulting from relationships with suppliers of the enterprise usually occupy a large relative share in the total volume of settlement operations. They depend on the activity that the company develops.

The relationship with the supplier and the company is proved by a simplified invoice or by a tax invoice.

Settlements between the supplier and the company usually occur when there are differences between the time of the transaction or the events giving rise to the estimated relationship and the payment on them.

The relations of the organizations with clients from the country and abroad are reported through accounts from group 41- Clients. In this group are provided accounts, both for reporting receivables from customers and for reporting the liabilities of organizations on advances received from customers. The first group of accounts are active and are debited when receivables from customers arise and are credited when they are settled. Accounts for reporting received advances from customers are passive and are credited upon receipt of advances from customers and are debited when providing tangible assets.

Business transactions that result from customer relationships can be varied and numerous and can become big data.

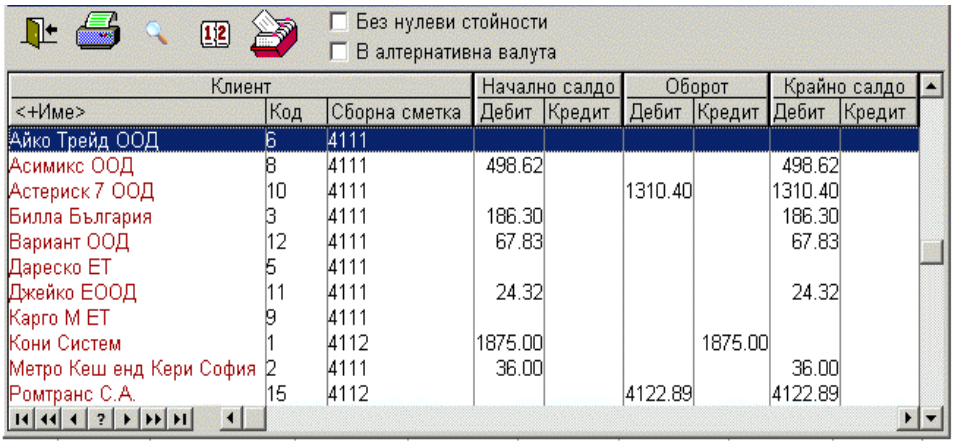
2. Description of a business process using big data - settlements with clients

The business process "Settlements with Clients" will be considered. The accounting relationship of the organization with customers expresses the relationship of sales of products, goods and services, as well as liabilities received from the organization in connection with such sales by third parties. The amount of receivables is formed from the sale value of the production, goods or services, including all costs invoiced to third parties, incurred by the executor of the transaction, i.e. the performed loading and unloading, transport and other expenses at the expense of the client. The accounts from group 41-Clients and related accounts are intended for reporting the settlements with clients. When reporting the settlements with the client, a number of accounting documents are used, such as tax invoice and simplified invoice, tax-credit notice, receipt, order, customs declaration, specification, order for shipment of goods, etc. At the end of the reporting period, the accounts of group 41 may be closed without a balance (if all receivables are repaid) or remain with a debit balance, which shows the amount of receivables not received from customers.

This business process was chosen because in accounting, it most clearly represents the relationship with big data. The accountants who prepare the entries in the current account for the clients verify the data from the primary documents and the bank data, ensuring the accuracy of the entries. An important part of the work of these accountants is also the analysis of receivables from customers. By analyzing and interpreting financial information, accountants can provide control and solutions to uncollected receivables (Fig. 1).

In accordance with the primary documents, the accountants make records, and the accounting should be able to be done automatically, using big data processing technologies.

As an example of big data processing can be the data that enter the accounting from the various transaction systems - online payments with bank cards, cash payments upon receipt of goods, etc. These are mostly payments with credit and debit cards with large frequency. For example, payments in an e-shop, payments on betting sites, online payment of fees, subscriptions, etc. If the company has a large online store, a betting site, a site for buying movies, music, etc., hundreds of thousands of transactions can be generated. day to be processed, as the data from the transaction systems enter the accounting, as well as the data from the various bank documents and the data from the clients' orders. The following figure can be used to track the movement of transactions (Fig. 2).



☐ Без нулеви стойности
☐ В алтернативна валута

Клиент			Начално салдо		Оборот		Крайно салдо	
<+Име>	Код	Сборна сметка	Дебит	Кредит	Дебит	Кредит	Дебит	Кредит
Айко Трейд ООД	6	4111						
Асимикс ООД	8	4111	498.62				498.62	
Астериск 7 ООД	10	4111			1310.40		1310.40	
Билла България	3	4111	186.30				186.30	
Вариант ООД	12	4111	67.83				67.83	
Дареско ЕТ	5	4111						
Джейко ЕООД	11	4111	24.32				24.32	
Карго М ЕТ	9	4111						
Кони Систем	1	4112	1875.00			1875.00		
Метро Кеш енд Кери София	2	4111	36.00				36.00	
Ромтранс С.А.	15	4112			4122.89		4122.89	

Fig. 1. Turnover statement for settlement with clients

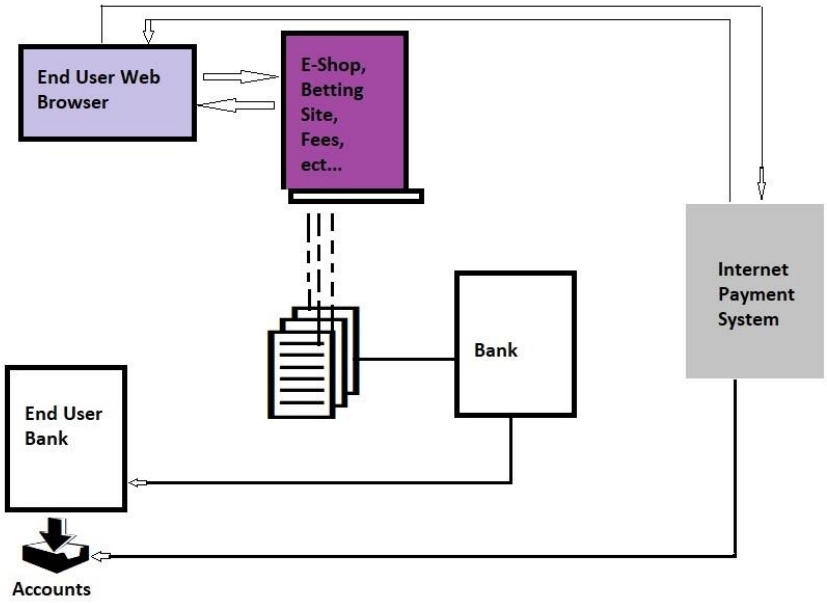


Fig. 2. Participants in online payments in settlement with clients

In this operation the End User Bank contains the current account of the end user (buyer). The bank issues cards and is a guarantor for the fulfillment of the buyer's financial obligations. The payment system also has its own bank, as well as the company, and the movement of transactions between them is carried out via the Internet. In just one day, many such transactions can occur. Recently, with the increase in the use of mobile phones and various technologies [3] related to them, transactions have increased enormously. The ability of the end user to analyze and make decisions in real time, increased online banking. This has led to an even greater amount of data to be accounted for.

In order to record these business operations, each of them must be done individually by the accountants, and a lot of work is required. Linking end user payments to their accounts and record operations in accounting for them quickly and accurately is the cornerstone of the settlements process.

Traditionally, this process consumes a lot of time and effort of accountants, and the return does not match the effort. Here come to the aid of big data processing tools - robotic accounting procedures.

3. Robotic procedures for reporting settlements with customers

Many robotic procedures have already been developed around the world to handle this kind of large number of transactions. Example - Blackline [5] – it's robotic procedures can automatically post over 7 million subscriber transactions with 99.9% accuracy and is fully automated. Robotic procedures are already loading data into specialized robotic software, such as banking transactions, credit card payments and general ledger and journal data. The system automatically checks transactions and compares them with their log entries.

UiPath [6] and Blue Prism [7] are also other examples of Robotic Process Automation (RPA) procedures designed to help accountants outside Bulgaria. The power of these systems is in the work with e-mail. Upon receipt of an e-mail invoice, using RPA they can retrieve the data from invoice and enter same data into an accounting program. Robotic procedures ensure accuracy and security, especially for vulnerable data and financial services. There is a lot of data in accounting that needs to be protected because data has to be processed, but it is not owned by the company [8].

In Bulgaria, various accounting products have the beginnings of such procedures, but the process must be fully automated to be fully helpful to accountants.

One of the companies that develops software for automating business processes is Controlisy [9]. The software is aimed at processing customer invoices. Through OCR and other data retrieval technologies, CONTROLISY accounting automatically retrieves number, date, amount, currency, exchange rate, VAT rate, counterparty and automatically creates an accounting account. The extracted information is very easily imported into accounting software.

Almost all large companies, such as Oracle [10] and Microsoft [11, 12], which develop complete ERP systems, have accounting modules for processing business processes for big data. However, these modules must comply with the specific ways of processing accounting data in the country.

4. Conclusion

By applying robotic procedures to automate business processes in accounting using big data, company makes the routine accounting process faster, and accurate. Company relieves accountants of the stress that accompanied these processes of finding records and payments to different subscribers or retrieve data from emails. It also helps free up the time for an accounting team to focus on activities that bring more added value, such as transaction verification and data analysis. There is still time and space for accountants to master new data analysis technologies [1,4] in order to transform their main function from business registrars to analysts and active participants in business processes.

There is another part of the business process settlement with suppliers and customers that needs special attention and time. In the accounting process, accountants often come across problematic documents that cannot be accounted for due to lack of clarity, lack of supporting documents or other reasons. When an accountant works with many companies, a number of

accounting documents accumulate that need to be clarified before they can be accounted for. The problem is usually clarified by phone or email and makes it difficult to keep track of all open cases. As these activities require a lot of time, freeing accountants from routine work is a means of resolving these problematic documents.

The role of accountants as participants in management, and not just as event registrars, is growing.

References

1. Ivanov, M. Modern methods for intelligent data analysis. https://core.ac.uk/download/pdf/78377667.pdf
2. Radoev, Mitko. Specific Requirements for the Accounting Software in Budget Organizations. Ikonomiceski i Sotsialni Alternativi, University of National and World Economy, Sofia, Bulgaria, 2016, issue 3, pages 93-98.
3. Mihova, Veska; Stefanov, Geno; Marzovanova Maria. Cognos Mobile - Dashboards Design and Implementation Technology. Proceedings of International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE - 2016). Sofia, 2016.
4. Stefanov, Geno. Analysis of Cloud based ETL in the Era of IoT and Big Data. Proceedings of International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2019); Sofia, 2019.
5. Blackline. (2022,01,30). Retrieved from https://www.blackline.com/accounts-receivable-automation/
6. UiPath. (2022,01,30). Retrieved from https://www.uipath.com/
7. Blue Prism. (2022,01,30). Retrieved from https://www.blueprism.com/
8. Tsaneva, Monika. Implementation of GDPR Requirements in the Information Systems of Consumer Financing Companies. Serdica Journal of Computing, issue 3-4, 2019. Institute of Mathematics and Informatics, Sofia.
9. Controlisy (2022,01,30). Retrieved from https://accounting.controlisy.com/
10. Oracle / NetSuite. (2022,01,30). Retrieved from https://www.netsuite.com/portal/products/erp/financial-management/finance-accounting.shtml
11. Microsoft / Dynamics 365 Business Central (2022,01,30). Retrieved from https://dynamics.microsoft.com/en-us/business-central/accountants/
12. Team Vision/ Microsoft Dynamics 365 Business Central. (2022,01,30). Retrieved from https://team-vision.bg/resheniq/microsoft-dynamics-365-business-central/

Manageable Mature Maturity Model (M-MMM) for Evaluation of Cybersecurity

Nedko Tagarev

University of National and World Economy, Sofia, Bulgaria
ntagarev@unwe.bg

Abstract. The problem addressed in this article is maximizing the effect of the cybersecurity management in organizations. The object of the article is the models for Mature Maturity Model (MMM). The methodology is based on implementation and analyses in the Mature Maturity Model (MMM) that is used in project management. According to this, we can look at cybersecurity as a project with its economic and technical aspects(soft and hard). In this article will be found previous applications and theoretical principles of MMM. Proposed Manageable Mature Maturity Model (M-MMM) is based on data, evaluation and quantity metrics. The M-MMM can provide method for distribution/allocation of the resources for cybersecurity with measurable results. In addition, we will try to provide some experiments on the application of the methodology and without it. These examples are based on problems that we can met in management of cybersecurity.

Keywords. Mature Maturity Model, Organisation culture, Cybersecurity, Evaluation

1. Introduction

In cyberspace, systems and processes should be manageable, whether simple or complex. Cyber as analog for cybernetic, which comes from the Greek word kubernētēs (κυβερνᾶν), "steersman", from kubernan "to steer"[1]. Problems arise in the presence of other non-technological factors. Regardless of how much technology and other resources, including finances, we "input" into cybersecurity systems, we still fall victim to attacks and incidents. ***We can solve these problems by maturing the management system.*** In other words, *when the problem is systematic, we can solve it by changing the system.*

Thesis: purely technological factors are easily detectable and easily solvable if is given the availability needed of resources.

A maturity model is a set of characteristics, attributes, indicators, or patterns that represent capability and progression in a particular discipline. The model provides a benchmark against which an organization can evaluate the current level of capability of its practices, processes, and methods.

The problem with the introduction of the Maturity Assessment of the Cybersecurity System is relevant due to the introduction of national standards for the United States(CMMC)[2] and in the next decade and in the EU. "These requirements were implemented to address a myriad of unintended consequences to domestic and international partners. The Defense Department's CMMC standard is no different and will likely disrupt its massive public procurement market by disqualifying small, medium and even large businesses in addition to international contractors who are unfamiliar with this standard."[3]

Another *major problem is that we rely on technology instead of our own knowledge* when making cybersecurity decisions. This is generally due to uncertainty, poor preparation,

ignorance of the terms, object and subject of cybersecurity. These shortcomings can only be compensated by quality training.

The main ***problem in the cybersecurity*** and education in cybersecurity is the *obsessive concentration on attack and attack mechanisms*. From a tactical point of view, this is justified "know yourself and know the enemy"[4], but this leads to creating the defence on the capabilities of the employees or auditors based on the knowledge of how to attack the system. We apply the "hard" measures/defence mechanisms on attack as a *reactive*(situational) *management*. This is very illogical, because if somebody attack our organisation we cannot counterattack because it is illegal. In education in cybersecurity, we study how to attack not to defend systems. Most of the people know how to attack or penetrate the system, but do not think how to protect it. In practice, we cannot find very often cybersecurity (defence) as a proactive process. It is not build in terms of information assets, their benefits and goals for the organization.

The cybersecurity have to come from defence to attack and not form attack to defence. This article address this issues, as it put the defence activities in front of the attack. "However, the detailed cybersecurity activities are usually listed in standards, guidelines, and practices.

They have the following nature:

- Focus on information security technical and functional controls (customizable).
- Applicable for developing checklists and conducting compliance/audit assessments."[5]

This basic problem leads to several tasks that require a solution:

- Choice between reactive not proactive approach / Cybersecurity concepts;
- Choice between stakeholders interest vs assets management;
- Choice of methodology of evaluation;
- Management of human behaviour;
- Choice between cybersecurity management concepts and models.

"Reactive cybersecurity methods are aimed at preventing known malware from entering your network. However, if a virus does get through, reactive methods defend against attacks that have already happened to keep it from corrupting your business databases."[6] Proactive cybersecurity management is about building the *cybersecurity system(CSS)* with digital infrastructure and training staff. This is called by design approach. In most cases, it is built without paying attention to cybersecurity. I.e. Cybersecurity is not a design element of either hardware or software components. Unfortunately, the only results that can measure potential investment in security are those of the reactive approach. I.e. we must have the presence of an incident to quantify the losses we have suffered. Based on incidents we can make case studies on MMM with CSS.

One of the existing fallacies in the theory is that penetration testing is a proactive action. In fact, it takes place after the digital infrastructure is built, so it is reactive. Any action taken as a result of identifying a vulnerability that can be exploited by a threat is a reactive action. An example of this is communication protocols that are not built with security in mind. Accordingly, to protect them we use protection protocols, which is a reaction to the corresponding pre-existing vulnerability. To be a little more proactive, we can consider cybersecurity as a project with the corresponding programs, goals. That's when we don't build it by design or accept that security is a "never-ending process" that requires continuous monitoring and improvement.

On the other side are the interests of the stakeholders and the assets of the organization. Stakeholders have an interest in asset value increasing. In this way, the value of their capital also increases. Assets are any single item that can be financially quantified and capitalized at

a certain point in time. Asset value has a direct relationship with the attack potential for benefits(APB) to the system. The higher the value of these assets is, the greater is the attacker's motivation to carry out an attack. An example of this is the types of attacks that can develop as incidents. They depend on the resources and knowledge of the attacking agent. Phishing attacks are the most common, but specific spear and whale phishing are applied to objects with higher asset values. SQLi or XXS are applied to objects with higher APB. They are rarer and not widespread, because the knowledge of the attacker is required to be above average and, accordingly, the attacked object must have a greater potential to implement an attack.

In terms of methodology, security investments are difficult to determine how they affect an organization's assets. For this reason, a methodology must be chosen that clearly represents the effectiveness and efficiency of security investments. Such an approach is proposed in this publication by M-MMM.

Last but not least, perhaps the most important factor is the human factor. It has two aspects that can lead to the seizure of the organization's assets. In the first place is the culture of behavior. This culture is managed through so-called personal security. In second place is the level of training and availability of staff. Regardless of whether we have built the most technologically advanced CSS, it cannot operate effectively if we do not have enough motivated and trained personnel. When we talk about human factors there are four main stakeholder groups:

- Investors
- Employees
- Clients
- Suppliers

Accordingly, each of these stakeholder groups relates to the core assets of the enterprise that may be affected in a cyber incident:

- Finances
- Reputation
- Legality

In its essence, cybersecurity is a reactive process of exploiting vulnerabilities by threats. A large part of the concepts of cybersecurity - their essence and their aspects - are built on this principle. Table 1 shows an example structure of some of the concepts.

Table 1. Cybersecurity Concepts

Concept	**Essence**	**Aspects**	**Example/Larger implementation of**
Threat concept	Technical process related to software and hardware;	Responding to vulnerabilities that can be exploited by threats; Vulnerability detection; Creation of a defensive mechanism against discovered vulnerabilities.	Creating "patches".
Risk evaluation concept	A Quantification-Based	Setting the context; Risk analysis;	Allocation of resources according

	Cybersecurity Resource Allocation Management Process.	Risk Assessment; Risk management;	to the environmental assessment and the principle of "acceptable risk";
Antivirus concept	Proactive system protection of hardware or software technical solutions.	Building databases from existing malware; Building additional security mechanisms for network and Internet security.	Antivirus software; Firewalls; IDS/IPS; SOC; SIEM;
Zero trust concept (in fact, there is no unified theory of this concept)	A system based on system segregation and enhanced access control.	Assessing the importance of assets; Defining trust zones; Establishing the rights of information asset owners;	VPN; Enhanced Cryptography; Security zones; PLC/PLD; Short work sessions;

One of the things that can make some asset owners nervous is the concept of cybersecurity resilience. If one reads more seriously in the literature on the matter a large part of the documents and standards talk about resilience and inevitable breach in the cybersecurity systems. But nowhere in these publications is there any mention of a methodology for evaluating or quantifying the economic effect of these breakthroughs. Anticipating such an inevitable event should not reassure asset owners because the financial, legal and reputational damages could run into the hundreds of millions or even billions. We can define *the concept of sustainability as a cult with followers with a vague idea of what they are aiming for*.

2. Background and review of the literature

The background of this publication is related to the previous work of the author in connection with the development of his dissertation work, work on a scientific project “A modern security assessment toolkit”(NID NI 1-9/2013)(ver. 2 of proposed M-MMM), where elements of the present publication can be seen[7], as well as the implementation of this evaluation methodology in the discipline "Information technologies" since 2011(ver. 1 of proposed M-MMM). Also, elements of the medel are published in the monograph “Modern approaches in cybersecurity management”(ver. 3 of proposed M-MMM).[8]

For the purpose of this publication, more than 300 related topics have been reviewed. They mainly focus on the review of other similar publications, standards and models. In essence, the main difference is in the number and authors of the cited sources.

In the following lines we have taken out some essential and original elements from some of the reviewed publications. “It was observed that the cybersecurity capability maturity models have similar elements because they use processes and levels of maturity, they also manage the risk, although at different levels of depth. Finally, it has been observed that each model due to its particularity has different fields of application.”[9] In most of the cases we can find a complication and variation of the elements so we can create and produce a new model or a new standard. *Complicating the system in most cases leads to increased costs, which may deny the organization's management to invest in cybersecurity*.

"Information Technology (IT) and Information Security (InfoSec) audits that were efficient in the past, are trying to converge into cybersecurity audits to address cyber threats, cyber risks and cyberattacks that evolve in an aggressive cyber landscape. However, the increase in number and complexity of cyberattacks and the convoluted cyberthreat landscape is challenging the running cybersecurity audit models and putting in evidence the critical need for a new cybersecurity audit model."[10]

"Since there may be more threats than solutions, it is important to be able to prioritize efforts. For this, the mitigation efforts will have to be based on what is important according to the value system and worldview of the entities – individual, community, organisation and nation. Thus it is imperative to be able to measure the cybersecurity status in terms of the priority of values defined by sustainable high-level objectives of the concerned entities. Resources necessary to mitigate risks as result of threats exploiting vulnerabilities are limited. Priority of values allows each entity to optimize the allocation of its mitigating resources based on agreed value system and their worldview – based on what is important or critical to sustain their way of life."[11]

"The Cybersecurity Capability Maturity Model C2M2 focuses on the implementation and management of cybersecurity practices associated with information technology IT, operations technology OT, and information assets and the environments in which they operate. The model can be used to strengthen organizations cybersecurity capabilities, enable organizations to effectively and consistently evaluate and benchmark cybersecurity capabilities, share knowledge, best practices, and relevant references across organizations as a means to improve cybersecurity capabilities, and enable organizations to prioritize actions and investments to improve cybersecurity capabilities."[12] The main idea for implementation of such a standard is *to provide some level of quality control* on the digital assets of the organization. In most of the cases regardless of the product, such control does not exist. In most of the cases, the reasons are lack of skills, resources or goals of the companies for fast profit. Addressing cybersecurity vulnerabilities is usually left in the background. "Additionally, as technology evolves, it is important that agencies manage information systems in a way that addresses and mitigates security and privacy risks associated with new information technologies and new information processing capabilities."[13]

The CMMC is a framework designed to provide the Department of Defence in USA(DoD) with verification that Defense Innovation Board (DIB) members can adequately protect Federal Contract Information (FCI) and Controlled Unclassified Information (CUI) flowing through the supply chain from customer to prime contractors to sub-contractors. It builds upon existing regulations, other models' best practices, and combines multiple existing cybersecurity standards both from within the U.S. government and internationally.[14]

Maturity Modelling is a method for representing domain specific knowledge in a structured way in order to provide organizations with an evolutionary process for assessment and improvement. This is in full force for Small and Medium Enterprises (SMEs) where "SMEs mostly have limited capabilities regarding cybersecurity practices. Given the increasing cybersecurity risks and the large impact that the risks may bring to the SMEs, assessing and improving the cybersecurity capabilities is crucial for SMEs for sustainability"[15].

In rare cases, the companies and organizations have the time to implement the Return on Security Investment (ROSI) methodology[16]. "Applied to security, a Return On Security Investment (ROSI) calculation can provide quantitative answers to essential financial questions:

- Is an organization paying too much for its security?
- What financial impact on productivity could have lack of security?

- When is the security investment enough?
- Is this security pr oduct/organisation beneficial?"[17]

"NIST CSF is a toolkit of 98 "pointers" to guidance provided by five standards or collections of best practices. The pointers, called "subcategories," are classified into five functions: Identify, Protect, Detect, Respond, and Recover. The five functions are further broken down by tasks in cybersecurity risk management, which are referred to as categories. The table of functions, categories, subcategories, and informative references is called the Framework Core."[18]

"Frameworks such as the cybersecurity capability maturity model (C2M2) and the NIST Cybersecurity Framework (CSF) are often used by the critical infrastructure owners and operators to determine the cybersecurity maturity of their facility. Although these frameworks are exceptional at performing qualitative cybersecurity analysis and identifying vulnerabilities, they do not provide a means to perform prioritized mitigation of those vulnerabilities in order to achieve a desired cybersecurity maturity."[19]

Other types of models are based on "questionnaire model for assessing and improving cybersecurity capabilities based on industry standards"[20]. Such a models require a lot of testing and require the full implementation of the cycle of change management for the object of critical infrastructure[21].

"The B-C2M2 is based on the C2M2. B-C2M2 uses the same security domains and practices for evaluating the maturity of cybersecurity programs for building control systems but the questions are altered to fit the needs of the building control system community. Further, the B-C2M2 is offered in a "Lite" version that shortens the assessment process and requires fewer questions. A major element of shortening of the assessment process was achieved by cascading the questions. If a "not implemented" answer was received for some key questions, quite a few questions could be skipped, because those questions would also have to produce a "not implemented" response."[22]

"Cloud Security Alliance (CSA) developed security guidance for critical areas of focus in cloud computing including various versions… The latest version focused on meeting the demand of security changes. It also introduced better standards for organizations to manage cybersecurity for cloud by implementing security domains. The guidance can be applied to cloud service model (IPSaaS) and four deployment models (Public, Private, Community, and Hybrid Cloud) with derivative variations that address specific requirements. The guidance included thirteen (13) different domains, which are divided into two general categories: governance and operations. The governance domains focus on broad and strategic issues as well as policies within a cloud computing environment, while the operations domains focus on more tactical security concerns and implementation within the cloud architecture."[23]

The Information Security Focus Area Maturity (ISFAM) model is a focus areaoriented maturity model, originally proposed as a method for incremental progression. It consists of a fixed number of maturity levels, each process identified by a focus area/domain, is assigned its own number of progressively more mature capabilities. The model is able to determine the current information security maturity level. ISFAM model has 12 maturity levels and 13 focus areas. In these focus areas, 64 capabilities are assigned at the various maturity levels. The assessment of the maturity level is executed through a survey or a directed interview with an expert.[24]

"PNNL has developed a Secure Design and Development Cybersecurity Capability Maturity Model (SD2-C2M2). The SD2-C2M2 is an integrated platform that can correlate management priorities with the cybersecurity maturity, vulnerabilities, and related best practices. SD2-C2M2 allows hardware and software designers to assess the maturity level of their design and development processes, allows management to

determine desired maturity levels in seven domains, and allows developers to monitor process maturity improvements against management goals. The tool can be used by commercial developers as well as internal development organizations." [25] "The Secure Design and Development Cybersecurity Capability Maturity Model (SD2-C2M2) provides a browser-based tool that allows hardware and software developers to assess the maturity level of their design and development processes, allows management to determine desired maturity levels in seven domains, and allows developers to monitor process maturity improvements against management goals. The tool can be used by commercial developers as well as internal development organizations."[26]

Electrical Subsector Cybersecurity Capability Maturity Model (ES-C2M2) ESC2M2 defines ten domains of Cybersecurity performance: Risk, Asset, Access, Threat, Situation, Sharing, Response, Dependencies, Workforce, and Cyber. Each domain in ESC2M2 encompasses several objectives. Each objective, in turn, consists of a set of Cybersecurity practices. ES-C2M2 is reasonably uncomplicated, an organization can classify the practices vital for each objective in the related ES-C2M2 domains to progress towards the needed maturity levels. ESC2M2 confirm Nothing Exists, Basic, Progressed, Advanced, Risk Management, Governance, Access control and Incidence Management.[27]

Systems Security Engineering Capability Maturity Model (SSE-CMM) was design with six maturity levels, namely ; not Perform, Performed Informally, Planned and Tracked, Well Defined, Quantitatively Controlled, and Continuously Improving.[28]

Community Cybersecurity Maturity Model (CCSMM): The CCSMM is design to address the requirements of U.S communities to develop a practicable and sustainable plan for Cybersecurity. The model defines five maturity levels; Initial, Established, Self-assessed, Integrated, and Vanguard.[29].

3. Mature Maturity Model and Digital Infrastructure Methodology

First of all, it should be noted that at each stage of cybersecurity, we use cyclical processes. I.e. building the cybersecurity system is a one-time and permanent action. I.e. once built, it needs constant analysis and maintenance.

"Decisions and activities within an IT project are likely to have a lasting impact on cybersecurity. Procurement and supply chain management are one example… Quality management is a second example. Defects in design, deployment, or provisioning during the IT project could be exploitable until detected and corrected – potentially throughout the active lifecycle of the IT system. The security of the project office and the project infrastructure is also of lasting impact. The tools and processes used for project work, document management, and communication within the project team are all components of information security and integrity."[30]

Interesting case is Hyper-convergence infrastructure(HCI). HCI is a type of infrastructure approach that is largely software-defined with tightly-integrated compute, storage, networking, and virtualization resources running on commodity hardware. This stands in contrast to a traditional converged infrastructure, where each of these resources is typically handled by a discrete hardware component that serves a singular purpose.[14]

HCI enables a shift in management paradigm from a hardware approach to an applicationfocused one, with centralized management, policies, and mobility conducted at the virtual machine level. HCI's flexible infrastructure makes it easier to launch new cloud services…HCI enables operations to scale up quickly to meet new demands without disrupting existing applications. Simply put, HCI focuses on the user and helps cut down deployment of new services to the user to a matter of minutes.[14]

The concept of capability maturity models has been extended to the domain of cybersecurity and can be applied to the protection of critical infrastructure. In lieu of simple checklists, managers now have well-defined criteria against which to measure the maturity of their preparedness against cyber-threats[31], with models shifting from early examples such as the International Organization for Standardization's Systems Security Engineering Capability Maturity Model (SSE-CMM), Citigroup's Information Security Evaluation Model (CITI-ISEM) and Computer Emergency Response Team / CSO Online at Carnegie Mellon University (CERT/CSO) around the turn of the century to modern initiatives such as the current International Organization for Standardization (ISO/IEC) standards, the National Institute of Standards and Technology (NIST) Cybersecurity framework, the U.S. Department of Energy's Cybersecurity Capability Maturity Model (C2M2), and the U.S. Department of Homeland Security's NICE-CMM released in 2014. These modern cybersecurity capability maturity models provide the stages for an evolutionary path to developing policies and processes for the security and reporting of cybersecurity readiness of critical infrastructure.[32]

We can acknowledge these main project management maturity models:

1). PjM3(Portfolio, Programme and Project Management Maturity Model)[33] as part of P3M3 and PRINCE 2 (Projects IN Controlled Environments, version 2);

2). OPM3, the Organizational Project Management Maturity Model, developed and maintained by the Project Management Institute. This model differed from the staged model in that this revised model recognizes that an organization can possess differing levels of maturity across the various knowledge areas. [34]

3). ISM3, O- ISM3 model. This model allows dividing the organizations into five levels of security maturity. The classification level depends on the type of controls and the level of automation supported in the organization.[35]" First step specifies a set of processes such as the document management, system, and business audit, design, and evolution. Second step addresses the strategic management. It consists of a group of processes such as stakeholder's report, coordination of information security and physical security, strategic vision, resources allocated for information security. Third check targets the tactical management. It consists of 11 processes such as background checks, service level management, insurance management, security personnel selection and training, and security awareness. Finally, the fourth check states the operational management processes. It consists of 25 processes such as inventory management, environment patching, security measures, change control, access control, information quality, compliance probing, forensics, etc."[36]

When building a cybersecurity system(CSS) with an existing information infrastructure, in order to have a positive result(some level of resilience), we must follow the following steps:

1) Recognition of information assets;*(management)*
2) Classification of information assets; *(management)*
3) Building trust zones; *(management)*
4) Assessment of the maturity of the cybersecurity system; *(management)*
5) Construction of protection mechanisms according to the principles of one of the models or concepts.*(technical)*

We can build this concept according to the PDCA cycle model (Fig.1).

"Consequently, cybersecurity decision makers have significant problems in identifying the appropriate set of cybersecurity models that is relevant to them. Furthermore, the lack of a

uniform classification of cybersecurity models aggravates the communication regarding cybersecurity issues."[37] most models are technically oriented and do not take into account the problems of management and business goals of the organization. *We base technical aspect on the core of the model and many speculations and legends connected to the cybersecurity and information technologies mythos.* Such myth is that implementation of standards, security mechanisms and concepts will provide some level of cybersecurity resilience in organization. "Determining cybersecurity resilience maturity level requires measurement of events, and/or measurement of levels of occurrence of variable factors and effects. We determined that utilising relevant components from the five frameworks cited above and discussed in more detail below (hereafter referred to as the "combined core") would provide an appropriate foundation for a CRMM framework, noting that any identified gaps could be filled progressively as the framework is implemented and tested in actual organizations."[38]

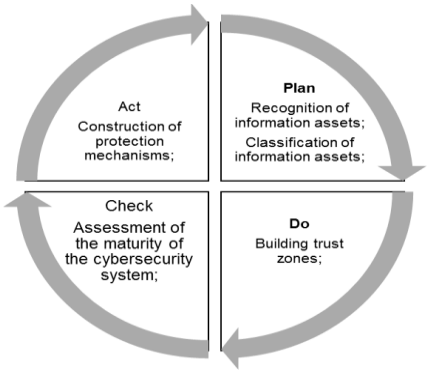


Figure 1. Building CSS in PDCA cycle

On the other hand, we base the construction of the cyber security system on the tasks of protecting the information security (IS) values and the change of the organization's assets against the change of the information security assets (Fig.2).

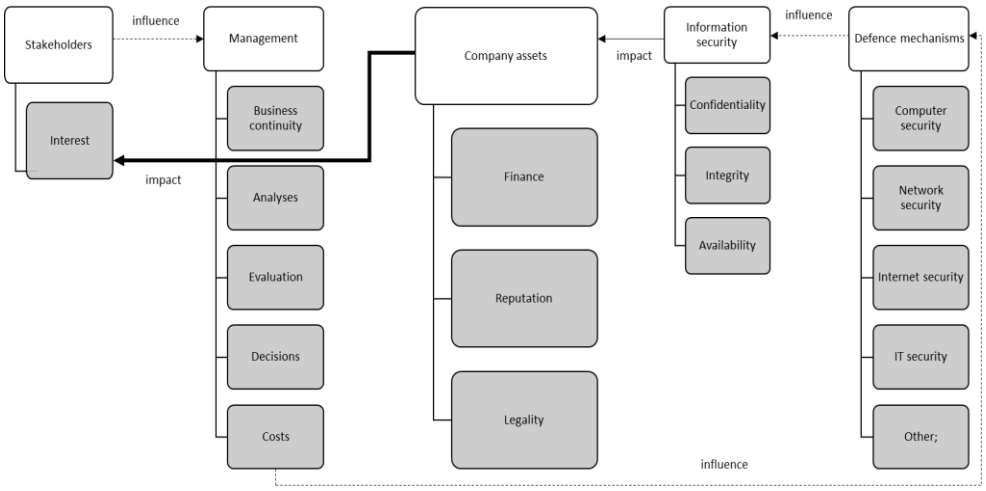


Figure 2. Information security effect on Organization assets

4. Alternative Mnageable Model

We base the core of the model on Cobit 4.1 standard for maturity evaluation of maturity of digital infrastructure[39] and Cobit 5 for the levels of maturity[40].

To achieve the goal of maturity assessment we have to go through five stages[7]:

4.1. Defining the tasks

A **goal** is a **desired state** for which we do not have the resources to achieve. **Tasks** are the **vehicle for acquiring** these resources. To achieve the goal, we have to achieve some tasks to gather the minimum resources needed for achieving the goal. Tasks and objectives are cyclical due to the volatility of the security environment. On the other hand, there is the physical and moral obsolescence of technologies, which necessitates the replacement and change of information infrastructure and assets.

To define the goals and objectives of cybersecurity management we need to define the relationships between(Fig. 3):

- the risks to the information system(we must have accurate threat measurements for each information asset.);
- the complexity of the information system/ domains, infrastructure, etc.;
- investment efficiency in cybersecurity.

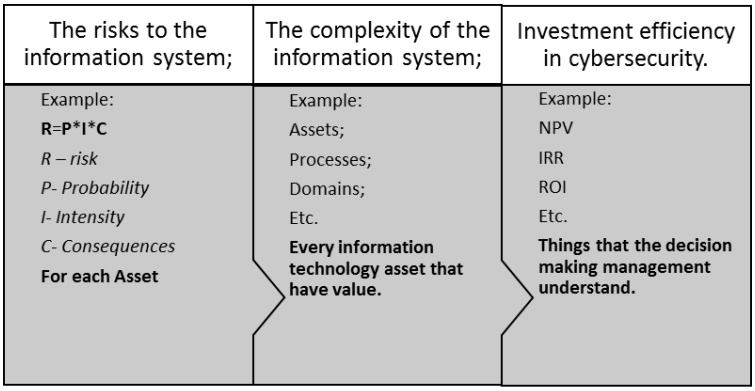


Figure 3. Relations between risk evaluations, complexity of the system, efficiency of investments.

The main tasks of this undertaking are:

- Determination of the maturity of the management of cybersecurity system (Fig. 4);
- Tasks and fulfillment of the set goals;
- Determination of input and output resources;
- Guidelines for managing processes to achieve the set goals.

4.2. Determination of the maturity of the management of cybersecurity system

The maturity levels we take from the maturity assessment system in Cobit 4.1[39] and we present them in Table 2. The purpose of responsible management is to transfer the maturity level of management to a higher level.

Table 2. Maturity Levels [8]

Level-Matu-rity	Maturity of Management	Management of the processes in cybersecurity	Notes
0	Does not exist	*There are no management processes.*	In most of the cases of incidents, there are no list of assets and assets owners.
1	Initial	*Management is situational and largely disorganized.*	When an incident occurs, it is reacted or not reacted to. There is no pre-planning or allocation of resources.
2	Repeatable and intuitive	*Management processes follow an established pattern.*	There is some management experience. Largely based on the culture of the organization. In many cases, this is experience from another line of business and does not take into account the specifics of cybersecurity.
3	A system with defined processes	*Management processes are documented and communicated.*	Cybersecurity management is based on a certain standard. This is usually an information, computer or network security standard. Control is carried out at every stage of the processes.
4	A system that can be managed and evaluated	*Processes are observable and can be measured. Qualitative and quantitative criteria can be applied.*	Cybersecurity management has all Input and Output information. Management that makes decisions can rationally allocate resources.
5	Optimized system	*Good practices are followed. Good practices are automated as a management process.*	A large part of the processes are pre-programmed and planned

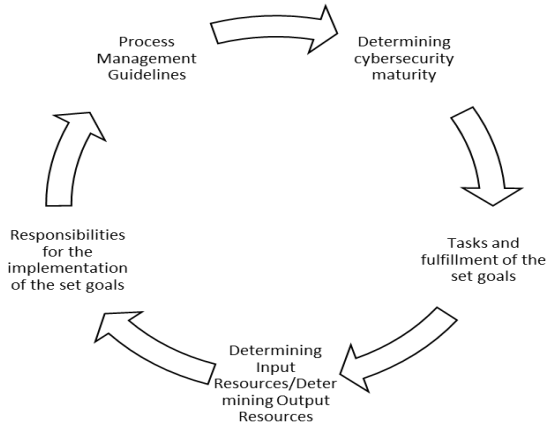


Figure 4. Cycle of goals

To determine the maturity of a cybersecurity management system, we need to choose an appropriate and comprehensive metrics that does not conflict with our goals and objectives. For example if we do not have industrial production, we can not put industrial management metrics. The most important thing is that it can be quantified, so that it can be compared, through mathematical abstraction.

On the basis of the maturity model, we can determine the following indicators determining the maturity level of the system[8]:

- Degree of awareness and communication;
- Availability of policies, plans and procedures(project management);
- Tools and automation (meaning software and hardware);
- Skills and experience(cyber workforce);
- Responsibility and accountability(owners of the information assets);
- Clearly set objectives and tools for measuring indicators(*metrics for all of the previous*).

If we stick to the requirements of standards and good management practices, we are required to create a baseline profile and documented requirements for determining system maturity. The following documents would help us get enough comprehensive information to be able to make an objective assessment:Набор от изисквания и точки за „развитие"(преминаване) за всяко ниво на зрялост;

- "A scale in which differences can be easily measured;
- A scale in which the comparison is pre-programmed;
- Basis for establishing the current and future state;
- Systematic analysis of gaps, establishing what is needed to move to a selected level;
- Overall assessment of cybersecurity management."[8]

4.3. Determination of the tasksfor the plans and processes necessary to achieve the set goals

First of all, the processes related to the **technological infrastructure** should be defined. Regarding it, there are three main processes:

- Acquisition,
- Implementation,
- Maintenance.

Accordingly, two plans are prepared for them:

- Acquisition,
- Maintenance.

GOAL OF PROCESSES AND PLANS: TO ENSURE CONTINUITY OF BUSINESS PROCESSES! Regardless of when we first realized we needed cybersecurity, the goal of security is to protect and enhance an organization's assets. Accordingly, ensuring the continuity of business processes. Thus, cybersecurity is present in each of the three core processes, as well as in both plans.

In terms of cybersecurity, we measure the following indicators - effectiveness, efficiency and reliability, and in terms of information security, we monitor confidentiality, integrity and availability of information. Regarding the process of organization of the research, it goes through the stages described in Fig. 5:

Cybersecurity monitoring and evaluation requires control and maintenance of the technological infrastructure (software and hardware) that meets the requirements for a complete and continuous business process. It is about maintaining an appropriate business platform that meets information systems standards and can be supported and protected by a cybersecurity system.

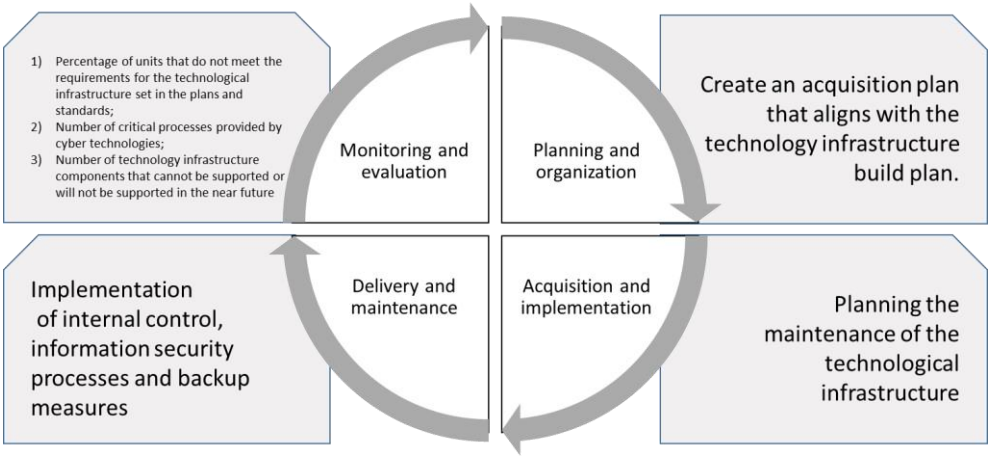


Figure 5. Process of organization

In accordance with the process, cybersecurity management should develop four plans for acquisition and maintenance, for information protection and change management of digital assets (Fig. 6).

The technology infrastructure acquisition plan

- is concerned with the acquisition, implementation and maintenance of technology infrastructure. It is in accordance with the functional and technological requirements established by the organization for working software applications and hardware. They meet the technology goals of the organization. For this reason, they can vary, with some organizations having processes that don't in others. Here, the assessment of the complexity of the digital infrastructure made in the previous stages plays a major role in the effectiveness of cybersecurity.

System maintenance plan

- is about a strategy and plan to ensure that changes are controlled in accordance with the rest of the organization's procedures. Continuous monitoring of new risks, improvement of existing systems, system weaknesses, security threats and system upgrades should be carried out continuously. This plans include the responsible persons.

A plan for the protection and availability of information resources

- which implements control, protection and archiving in the processes of implementation, integration and maintenance of systems in order to preserve the values - integrity and accessibility of information. The responsibility and the owners of the assets in the maintenance of sensitive technological infrastructure should be clearly defined and announced already at the integration of such type of components. Their use should be systematically monitored and evaluated. This plan also regulates the application of access control. This plan includes the allocation of information assets.

Managing change plan for digital infrastructure.

- creating a systems testing environment where compatibility, weaknesses, effectiveness and efficiency of the system are checked. Testing is being done using the cybersecurity tools.

Figure 6. Plans

4.4. Determination and analysis of input and output resources

In order to be able to do the analysis, we need certain input information. Based on it, we can determine the outputs or output information and results.(Fig.7). Output can also be defined as a management product.

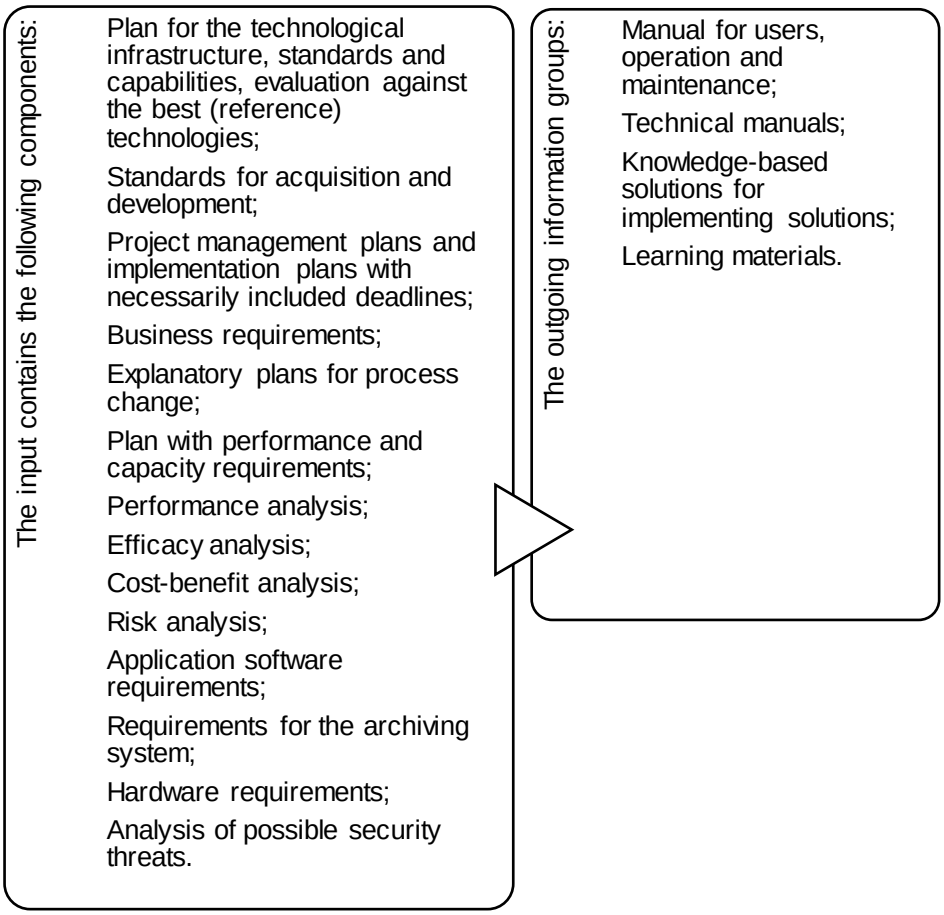


Figure 7. Input and output information.[8]

After acquiring sufficient quality output, we can move on to the activities of increasing the effectiveness of cybersecurity management and moving to the next level of maturity. The activities are as follows[8]:

- Development of a strategy for implementing the solution;
- Development of methodology implementation knowledge;
- Development of user manuals;
- Development of technical support documentation for system administrators;
- Developing and conducting trainings;
- Evaluation of test and training results, and formalization and implementation of the results.

5. Defining Tasks for the Implementation of Cybersecurity Objectives for Measuring the Effectiveness

The tasks of measuring cybersecurity performance can be divided in terms of protecting and enhancing enterprise assets [8].

Task 1. With regard to the company's assets (can also be defined as business goals) - finances, reputation and legality of the activity. The goal is to create resilience and attack recovery skills (disaster recovery, continuity of business processes).

Task.2. Regarding attack recovery (restoring the integrity and availability of information). The goals are to identify weaknesses in the system and to establish the act of the attack itself. This is the process for analysis in information security.

Task.3. Regarding the analysis process in information security, the goal is to establish the technological requirements for the system, weaknesses and threats.

For this type of analyses we need a standard metrics that can be found in any organization that exploit the digital infrastructure.(Fig.8)

We ca see the process of analyses in Fig. 9.

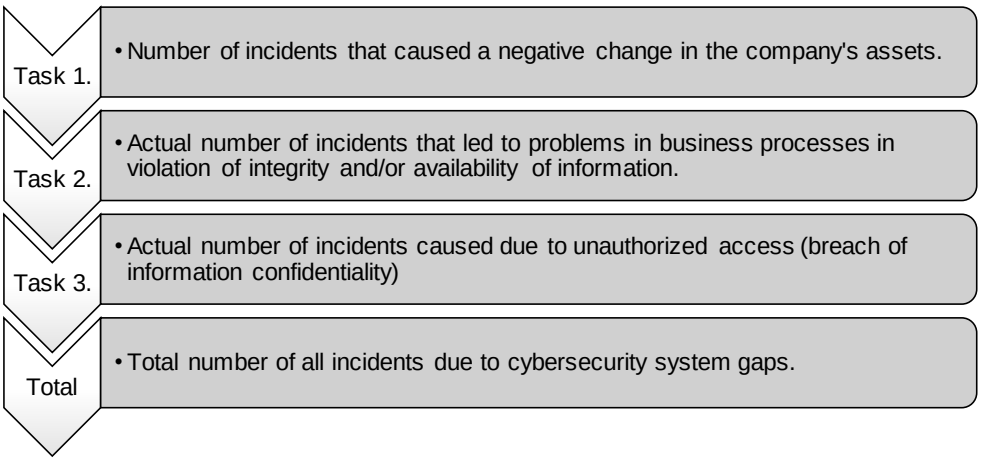


Figure 8. Metrics for the analyses of effectiveness

Example

As explained earlier, there is a different form and scope of the cyber security system depending on the object. For this reason, input and output information, digital infrastructure, and information asset owners affect M-MMM evaluation metrics differently. For this reason, we must consider the organization's context, structure, function, and goals.

One way of application is through case studies related to cyber security incidents. There are countless such cases. Of course, for a large part of them, certain details are not known, neither for the attack vectors nor for the technological details. This can be an obstacle only if there is a "zero day exploit". In most cases, it can be found that there are elements of the information system that are not updated or not supported. The reasons for this can be defined as the absence of process management.

In this way, by applying case studies, we can constantly test our system for vulnerabilities, based on incidents that have already occurred. This can be done through seven basic steps. We show this process in detail in Fig. 10.

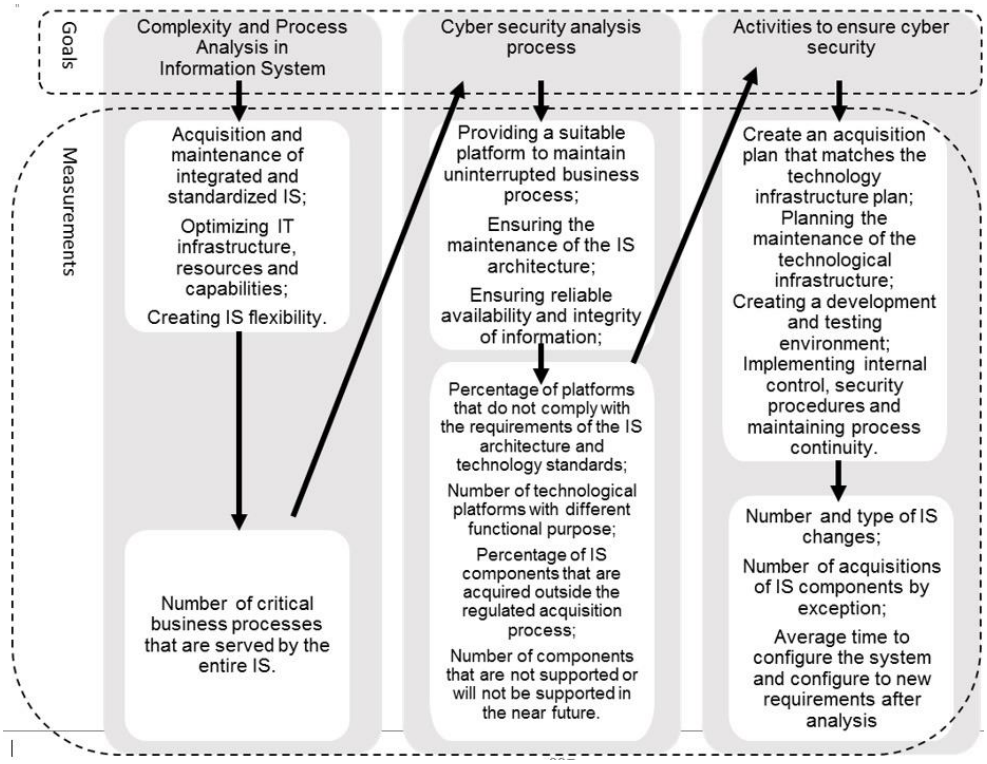


Figure 9. Analyses Process

Of course, the place of each stage and the details in which the analysis should be done are debatable, but these are the main elements that show effectiveness in practice. With a certain amount of experience, cybersecurity managers and employees at every stage should be able to make these analyzes intuitively.

6. Conclusion and future work

The idea of management system maturity assessment is an engineering approach to determine the effectiveness of processes in a particular area. It can effectively be applied to IT and cybersecurity process management as elements of IT and security. If we look at the cases in detail, when we exclude the technical problems (which, if we notice in the larger case of the cyber security problems, are not the decisive factor for the development of the attack into an incident), the problems are not related to the management of the processes related to cyber security - training of personnel, control and evaluation of the cybersecurity system.

Evaluating the maturity of a cybersecurity management system can greatly assist in preventing problems arising from known or foreseeable cybersecurity threats and issues. The MMM assessment can lead to increased trust in business partners. This should be done without further complicating processes and analyses. M-MMM offers such an approach, which is based on objective indicators.Една от следващите задачи ролята на управлението на активите и собствениците на информационните активи.

In the future work on the methodology, the distribution of roles and responsibilities in the management and analysis of the effectiveness of cyber security will be taken into account.

Detailed presentation of each element of the analysis through case studies. A similar representation of cybersecurity management can be seen in publicized incidents that lead to mass deployment of measures or updates to hardware management software elements.

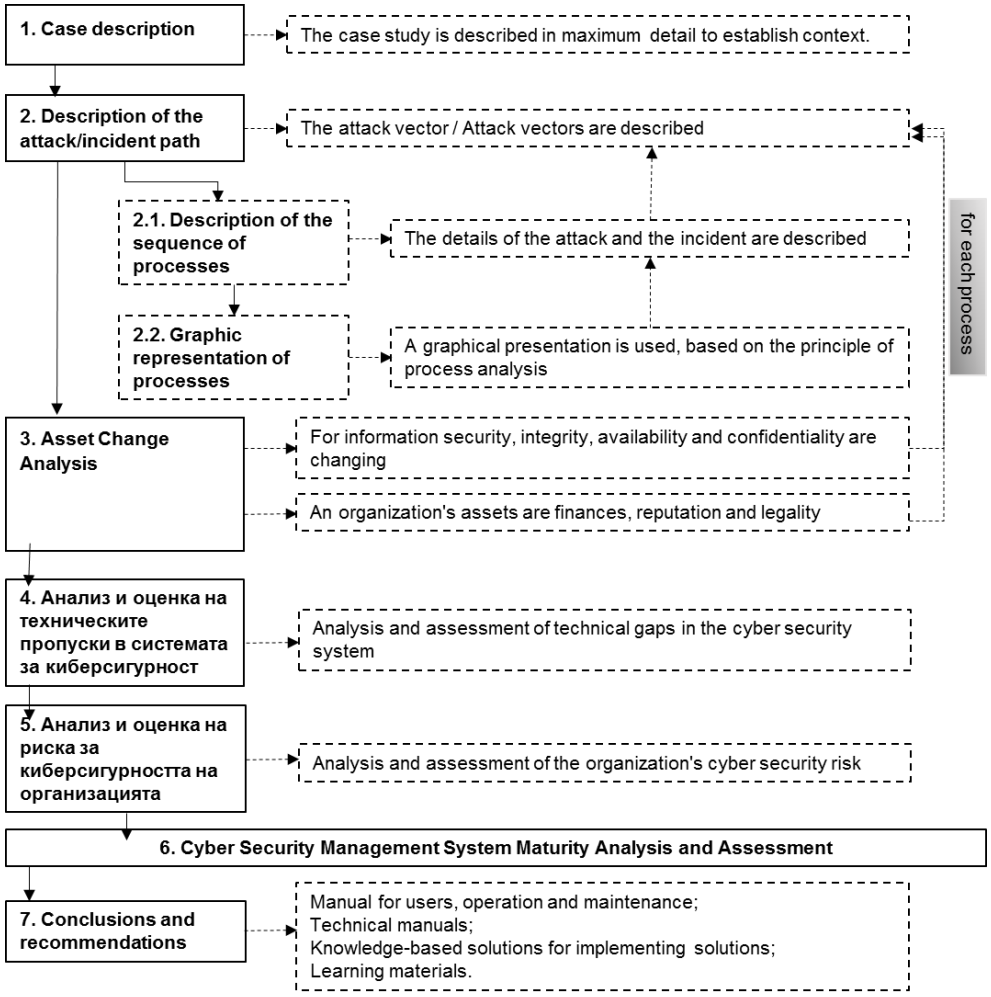


Figure 10. Process of case study analyses implementation in MMM

References

[1] T. COE, 'Where does the word cyber come from?', *OUPblog*, Mar. 28, 2015. https://blog.oup.com/2015/03/cyber-word-origins/ (accessed Jul. 12, 2022).

[2] 'Strategic Direction for Cybersecurity Maturity Model Certification (CMMC) Program', *U.S. Department of Defense*. https://www.defense.gov/News/Releases/Release/Article/2833006/strategic-direction-for-cybersecurity-maturity-model-certification-cmmc-program/(Jun. 20, 2022).

[3] J. Selinger, 'The Dod's Cybersecurity Maturity Model Certification and Resulting Tensions with U.S. Standard-Setting Policy and Established it Security Authorities'. Rochester, NY, May 16, 2021. doi: 10.2139/ssrn.3985000.

[4] S. Tzu, *The Art Of War*, First Thus edition. S.l: Filiquarian, 2007.

[5] S. Almuhammadi and M. Alsaleh, 'Information Security Maturity Model for Nist Cyber Security Framework', in *Computer Science & Information Technology (CS & IT)*, Feb. 2017, pp. 51–62. doi: 10.5121/csit.2017.70305.

[6] 'Cybersecurity Techniques: Reactive vs. Proactive | Advanced Business Equipment'. https://www.abecarolina.com/blog/cybersecurity-techniques-reactive-vs-proactive (Jun. 29, 2022).

[7] Н. Тагарев/Tagarev, 'VI. Инструментариум за оценка на информационната сигурност. Инструментариум за оценка на киберсигурността./ VI. Information Security Assessment Toolkit. Cybersecurity Assessment Toolkit.', in *Съвременен инструментариум за оценяване на киберсигурността/A modern security assessment toolkit*, Sofia: Издателски комплекс - УНСС/ Publishing Complex UNWE, 2017.

[8] Н. Тагарев/Tagarev, 'Част III. Интегрирана система за управление на киберсигурността на базата на ERP модели. Оценка на зрелостта на системата за управление на киберсигурността./Part III. Integrated cyber security management system based on ERP models. Cybersecurity Management System Maturity Assessment.', in *Съвременни подходи при управлението на киберсигурността. Част I. Теоретични и методически основи. / Modern approaches in cyber security management. Part I. Theoretical and methodological foundations.*, Sofia: Авангард Прима/Avangard Prima, 2021.

[9] A. M. Rea-Guaman, T. San Feliu, J. A. Calvo-Manzano, and I. D. Sanchez-Garcia, 'Comparative Study of Cybersecurity Capability Maturity Models', in *Software Process Improvement and Capability Determination*, Cham, 2017, pp. 100–113. doi: 10.1007/978-3-319-67383-7_8.

[10] R. Sabillon, J. Serra-Ruiz, V. Cavaller, and J. Cano, 'A Comprehensive Cybersecurity Audit Model to Improve Cybersecurity Assurance: The CyberSecurity Audit Model (CSAM)', in *2017 International Conference on Information Systems and Computer Science (INCISCOS)*, 2017, pp. 253–259. doi: 10.1109/INCISCOS.2017.20.

[11] 'Ibrahim, J., Saidin, A. Z., Dahlan, A. R. A., Aziz, N. A., Wahiddin, M. R., & Osman, R. A. H. (2015). A cybersecurity capability maturity model based on Maqasid Shari'ah (MS-C2M2). In International Conference on Maqasid Al-Shari'ah in Public Policy and Governance (IAIS Malaysia, 2015).' Accessed: Jun. 21, 2022. [Online]. Available: https://iais.org.my/attach/2015/maqasid2015/pdf/s6s2_paper.pdf

[12] S. Bilak and K. Brennan, 'Cybersecurity Capability Maturity Model (C2M2) - Cybersecurity Maturity Model Certification (CMMC) Supplemental Guidance (Draft)', CARNEGIE-MELLON UNIV PITTSBURGH PA, Jan. 2022. Accessed: Jun. 20, 2022. [Online]. Available: https://apps.dtic.mil/sti/citations/AD1168443

[13] 'CIRCULAR NO. A-130 TO THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES SUBJECT: Managing Information as a Strategic Resource'. Accessed: Jun. 21, 2022. [Online]. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf

[14] 'DOD Digital Modernization Strategy 2019', p. 72.

[15] B. Y. Ozkan and M. Spruit, 'Assessing and Improving Cybersecurity Maturity for SMEs: Standardization aspects'. arXiv, Jul. 03, 2020. doi: 10.48550/arXiv.2007.01751.

[16] 'Return On Security Investment (ROSI) A Practical Quantitative Model':, in *Proceedings of the 3rd International Workshop on Security in Information Systems*, Miami, United States, 2005, pp. 239–252. doi: 10.5220/0002580202390252.

[17] R. Bourgue, 'Introduction to Return on Security Investment', p. 18.

[18] N. M. Radziwill and M. C. Benton, 'Cybersecurity Cost of Quality: Managing the Costs of Cybersecurity Risk Management'. arXiv, Jul. 09, 2017. doi: 10.48550/arXiv.1707.02653.

[19] S. N. G. Gourisetti, M. Mylrea, and H. Patangia, 'Cybersecurity vulnerability mitigation framework through empirical paradigm: Enhanced prioritized gap analysis', *Future Gener. Comput. Syst.*, vol. 105, pp. 410–431, Apr. 2020, doi: 10.1016/j.future.2019.12.018.

[20] B. Yigit Ozkan and M. Spruit, 'A Questionnaire Model for Cybersecurity Maturity Assessment of Critical Infrastructures', in *Information and Operational Technology Security Systems*, Cham, 2019, pp. 49–60. doi: 10.1007/978-3-030-12085-6_5.

[21] N. Tagarev, 'Управление на промяната в информационната инфраструктура в обекти от критичната инфраструктура/ Change Management in Information Infrastructure for Objects of Critical Infrastructure', София, 2019, pp. 92–102.

[22] C. Glantz, S. Somasundaram, M. Mylrea, and R. Underhill, 'Evaluating the Maturity of Cybersecurity Programs for Building Control Systems', p. 12.
[23] O. O. Akinsanya, M. Papadaki, and L. Sun, 'Current Cybersecurity Maturity Models: How Effective in Healthcare Cloud?', p. 12.
[24] M. Spruit and M. Roeling, 'ISFAM: The information security focus area maturity model', *ECIS 2014 Proc. - 22nd Eur. Conf. Inf. Syst.*, Jan. 2014.
[25] 'SD2-C2M2 | PNNL'. https://www.pnnl.gov/copyright/sd2-c2m2 (accessed Jun. 21, 2022).
[26] S. N. G. Gourisetti, S. R. Mix, M. E. Mylrea, C. A. Bonebrake, and M. Touhiduzzaman, 'Secure Design and Development Cybersecurity Capability Maturity Model (SD2-C2M2): Next-Generation Cyber Resilience by Design', Pacific Northwest National Lab. (PNNL), Richland, WA (United States), PNNL-SA-141161, Apr. 2019. doi: 10.1145/3332448.3332461.
[27] 'Electricity Subsector Cybersecurity Capability Maturity Model v. 1.1. (February 2014)', *Energy.gov*. https://www.energy.gov/oe/downloads/electricity-subsector-cybersecurity-capability-maturity-model-v-11-february-2014 (accessed Jun. 21, 2022).
[28] K. Ferraiolo, 'The Systems Security Engineering Capability Maturity Model', p. 61.
[29] 'The CCSMM'. https://cias.utsa.edu/the-ccsmm.html (accessed Jun. 21, 2022).
[30] J. Payette, E. Anegbe, E. Caceres, and S. Muegge, 'Security by design: Cybersecurity extensions to project management maturity models for critical infrastructure projects', *Technol. Innov. Manag. Rev.*, vol. 5, no. 6, pp. 26–34, Jun. 2015, Accessed: Jun. 21, 2022. [Online]. Available: https://ir.library.carleton.ca/pub/12186
[31] R. Debreceny and G. Gray, 'IT Governance and Process Maturity: A Multinational Field Study', Feb. 2009, vol. 27, pp. 1–10. doi: 10.1109/HICSS.2009.278.
[32] W. Miron and K. Muita, 'Cybersecurity Capability Maturity Models for Providers of Critical Infrastructure', *Technol. Innov. Manag. Rev.*, vol. 4, no. 10, pp. 33–39, 2014.
[33] 'The P3M3 maturity model with an agile extension (P4M3)', *Henny Portman's Blog*, Aug. 27, 2018. https://hennyportman.wordpress.com/2018/08/27/the-p3m3-maturity-model-with-an-agile-extension-p4m3/ (accessed Jun. 21, 2022).
[34] B. Miller, 'The pathway to OPM3: a busy project manager's guide to advancing organizational maturity.', presented at the PMI® Global Congress 2004—North America, Anaheim, CA. Newtown Square, PA: Project Management Institute., 2004. Accessed: Jun. 21, 2022. [Online]. Available: https://www.pmi.org/learning/library/pathway-organizational-project-management-maturity-8221
[35] T. O. Group, *Open Information Security Management Maturity Model O-ISM3*. Van Haren, 2011.
[36] O. M. M. Al-Matari, I. M. A. Helal, S. A. Mazen, and S. Elhennawy, 'Adopting security maturity model to the organizations' capability model', *Egypt. Inform. J.*, vol. 22, no. 2, pp. 193–199, Jul. 2021, doi: 10.1016/j.eij.2020.08.001.
[37] I. Atoum and A. Otoom, 'A Classification Scheme for Cybersecurity Models', *Int. J. Secur. Its Appl.*, vol. 11, pp. 109–120, Jan. 2017, doi: 10.14257/ijsia.2017.11.1.10.
[38] U. M. Mbanaso, L. Abrahams, and O. Z. Apene, 'Conceptual Design of a Cybersecurity Resilience Maturity Measurement (CRMM) Framework', *Afr. J. Inf. Commun.*, vol. 23, pp. 1–26, 2019, doi: 10.23962/10539/27535.
[39] Isaca, *Cobit 4.1*. Rolling Meadows, IL: Isaca, 2007.
[40] E. Elue, CISA, and CDPSE, 'Effective Capability and Maturity Assessment Using COBIT 2019', *ISACA*, Jul. 27, 2020. https://www.isaca.org/resources/news-and-trends/industry-news/2020/effective-capability-and-maturity-assessment-using-cobit-2019 (accessed Jul. 11, 2022).

Practical Aspects of Using HDFS Data in SQL Server 2019 Big Data Clusters

Mitko Radoev

University of National and World Economy, Sofia, Bulgaria
mradoev@unwe.bg

Abstract. The modern database management systems provide users and developers of information systems with new and powerful tools for storing and processing large volumes of data. Microsoft SQL Server 2019 provides options for using and processing large volumes of heterogeneous data in Big Data Clusters. This document aims to examine different use-cases for processing data stored in Hadoop HDFS, along with relational data stored in Microsoft SQL Server.

Keywords. Microsoft SQL Server, Big Data Clusters, Hadoop HDFS.

1. Introduction

Microsoft SQL Server Big Data Clusters unite relational and big data. Big data clusters make it easy to integrate, manage, and analyze big data and relational data [1]. It provides new capabilities for accessing and processing data stored in many different environments, such as:

- Hadoop HDFS (read/write);
- Azure Blob Storage (read/write);
- Oracle (read only);
- Teradata (read only);
- MongoDB or CosmosDB (read only);
- Other sources via ODBC (read only).

These data sources can be accessed without installing any other software. SQL Server Big Data Clusters can query external data sources without moving or copying data.

The question arises, how to use data stored in Microsoft SQL Server together with data from other sources? This document will present some use-cases for jointly use of relational data and data stored in Hadoop HDFS.

2. Relational data

Before analyzing the possibilities for jointly use of relational and HDFS data, the characteristics of the relational data and the characteristics of the data stored in Hadoop HDFS will be considered separately.

Data stored in relational databases have many different properties. Their main characteristics are the following [2 - 4]:

- Structured data. The data stored in relational databases have a relatively constant structure
- Related data. Relational databases store data about real-world objects and the relationships between them.
- Metadata. Together with the data itself, the relational databases also store metadata or data about the data.

- Data sharing. The database is a common resource and it is used by many application programs.
- Restriction on data duplication. Relational databases are trying to eliminate or minimize data duplication.
- Data integrity. Relational databases usually ensure data integrity by imposing constraints.
- User views. Users are allowed to define different views to the same data.
- Data independence. Relational databases are based on a multi-level architecture that ensures the independence of the user views from the logical structure and the physical representation of the data.
- Ad hoc queries. Relational databases provide query languages that allow users to obtain the necessary information directly from the database.
- Transaction support. Relational databases provide a mechanism to ensure execution of transactions, i.e. sequences of actions that are logically related and should be implemented as one action.

Relational databases appear as a tool for organizing data at the enterprise level for the needs of the information system. Subsequently, the field of their application has expanded significantly and they are used in almost all areas. The emergence of big data poses serious challenges to relational databases, which they often find difficult to handle.

3. HDFS Data

Hadoop Distributed File System (HDFS) is a highly fault-tolerant distributed file system designed to run on low-cost hardware. HDFS provides high throughput access to application data and is suitable for applications that have large data sets. HDFS was originally built as infrastructure for the Apache Nutch web search engine project [5].

One of the first decisions that needs to be made when using Hadoop is to determine how the data will be stored. There is no such thing as a standard storage format in Hadoop. Just like classic file systems, Hadoop allows data storage in any format. Hadoop also provides built-in support for a number of formats optimized for Hadoop storage and processing.

This means that users have full control and a number of options for how data is stored in Hadoop. This applies not only to the raw data that are ingested, but also to the intermediate data generated during data processing and the extracted data that results from the data processing [6].

Each file format supported by Hadoop has its own specific features that make it more or less suitable for specific applications. These include plain text files or Hadoop-specific formats, such as Sequence file. There are also more complex, but richer file formats, such as Avro and Parquet. These different formats have different strengths that make them more or less relevant depending on the application and data types. It is also possible to create your own custom file format in Hadoop.

The types of data stored in Hadoop HDFS can be classified into 3 main categories:

- Text data. Very often Hadoop is used to store logs such as web logs and server logs. Text data is also found in many other forms: CSV files or unstructured data such as emails. The main consideration when storing text data in Hadoop is the organization of the files in the file system. In addition, a file compression format should be chosen, as text files can quickly take up significant space in the Hadoop cluster. In many cases, using a container format such as Sequence files or Avro instead of a plain text file provides a number of advantages.

- Structured text data. A more specialized form of text files are structured formats such as XML and JSON. Using these formats is a major challenge for Hadoop, as splitting XML and JSON files for processing is difficult, and Hadoop does not provide built-in tools. In this case, a container format such as Avro can be used. Avro's data conversion can provide a compact and efficient way to store and process data. Another option is to use libraries designed to process XML or JSON files.
- Binary data. Although text is typically the most common data type stored in Hadoop, it also can be used to process binary files such as images. For most cases of storing and processing binary files in Hadoop, using a container format such as Sequence file is preferred.

Applications that use data in Hadoop are based on single-write and multi-read file access model. Once created, saved and closed, the file must not be changed. This assumption simplifies data coherence issues and allows access to high-performance data.

In general, the data stored in Hadoop is very diverse both in terms of data type and structure, as well as the file formats used. Characteristic of all, however, is the large volume of data.

4. Combining relational and HDFS data: use-cases

SQL Server Big Data Clusters can query external data sources without moving or copying data. Back in SQL Server 2016, PolyBase was introduced as a tool for processing Transact-SQL queries that read data from external data sources. In SQL Server 2019, these capabilities have been extended to provide a variety of data source connectors, including Oracle, SQL Server, MongoDB (CosmosDB), and Teradata.

We can create external data sources and then we can query data from created external tables along with data from local tables. This option can be useful when we need to integrate data from different systems, for example to use data from different heterogeneous systems [7], to combine dynamic data with relational data [8], and in many other cases [9], [10].

The main idea is to combine and use together high-value relational data with high-volume big data stored in Hadoop HDFS. Of course, the possible applications of such combined data are many, but here an attempt will be made to describe some business-oriented use-cases in which the jointly use of relational and big data would be most useful.

Information systems in the banking and financial sector

Information systems in the banking sector are an example of successful use of both structured data stored in relational databases and unstructured and semi-structured data stored in other environments. On the one hand, the transaction processing systems used in customer service need a high degree of security and reliability of information, as well as transaction support. Therefore, these systems are often based on relational databases. On the other hand, the need to analyze a large volume of unstructured and semi-structured data requires their storage in an appropriate environment. Here are some specific examples of activities in the banking and financial sector that need to use relational data together with big data:

- Risk analysis for the bank's clients. Risk analysis serves banks to manage their financial security and offer personalized products and services to their customers. In addition to data on bank transactions, credit card transactions, loan payments and other financial details about each client, banks and financial institutions started to use other data sources - such as records of customer calls, chat and blogs, email correspondence and others. [11].

- Forecasting market behavior. This is another classic problem in the sector. Huge amounts of heterogeneous data from different sources (stock exchanges, commodity exchanges, currency trading, etc.) need to be analyzed together.

Information systems in the retail sector

Information systems in the retail sector include transaction processing of orders and payments from customers, which can use traditional databases. However, for the purposes of retail analysis, it is necessary to collect and store much more data about clients, their behavior and consumption - both online and in stores. New data used by retailers requires more complex processing, such as language processing, pattern recognition, sentiment analysis, and so on. The following use-cases can be indicated in the retail sector:

- Retail analytics for targeted and customized promotion and marketing.
- Retail analytics for dynamic pricing of products.
- Retail analytics for supply chain efficiency.

With ever-increasing interactions through social media and retail channels, customers are comparing products, services and prices to a number of online retailers. With this type of behavior, consumers can quickly move from one retailer to another, making it absolutely necessary for retailers to use this information [11].

Information systems in the telecommunication sector

Information systems of telecommunication companies also could be examples of using together relational and other data. Information about the contracts and billing information of the customers could be organized in a traditional database management system, but when they have to store billions of call record details, they need another way of storing data. In addition to structured data they have to store huge amount of semi-structured data. The telecommunication industry includes the local and long distance calls, wireless communications, text messaging, high speed internet, data communication, television streaming and many other activities. The companies target customers by catering specific products to stand out. Insights like location, interest, activities, and preferences of customers are being used to create marketing and promotional plans.

The following use cases in the telecom industry can be indicated:

- Call data records management.
- Network traffic analytics.
- Servicing of telecom data equipment.

Information systems in the healthcare sector

The healthcare sector is another example of the use of structured data stored in traditional databases, together with a huge amount of unstructured or semi-structured data. The big data generated in the healthcare sector is mainly due to patient record keeping and regulatory requirements. Data generated by electronic medical devices is difficult to process using traditional database management systems. The complexity and volume of health data is the reason why this data is stored in Hadoop or other similar environments. For millions of medical records, Hadoop provides storage for vast amounts of unstructured data. Here are some specific examples of activities in the healthcare sector:

- Health insurance. Health insurance companies need to assess the probability of getting a specific disease in different parts of the population. Such an assessment requires the processing of vast amounts of data, including patient records, demographic groups,

geographical regions, diseases, symptoms, etc. Hadoop can be useful in storing and processing such a huge volume of unstructured or semi-structured information.

- Processing of claims and payments on them. Storing and processing millions of claims every day involving unstructured information can be a major challenge for traditional databases.

5. Conclusion

There are many modern tools and technologies that must be combined wisely to achieve maximum effect. It is clear that there is almost no field of application of information technology, where data can be stored only in traditional databases, without using large volumes of unstructured or semi-structured data, i.e. big data. Also, there are still many cases in which relational database management systems have some advantages, such as transaction support, data integrity, relationships between data etc. In these cases, the jointly use of data stored in relational databases with big data stored in Hadoop or another environment can lead to a combination of the advantages of both technologies.

References

1. Microsoft Corporation, "Microsoft SQL Server 2019 and Big Data—a technical white paper," 11 2019. [Online]. Available: https://info.microsoft.com/ww-thankyou-SQL-Server-2019-Big-Data-WhitePaper.html?LCID=EN.
2. E. F. Codd, Relational database: A practical foundation for productivity, Communications of the ACM, Vol. 25, № 2, pp. 109-117, February 1982.
3. C. J. Date, An Introduction to Database Systems, 8th edition, Pearson/Addison Wesley, 2004.
4. T. Connolly, C. Begg, Database Systems: A Practical Approach to Design, Implementation, and Management, 4th edition, Pearson Education Ltd, 2005
5. Apache Software Foundation, "HDFS Architecture Guide," 22 8 2019. [Online]. Available: https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html.
6. Grover M., Malaska T., Seidman J., Shapira G. Hadoop Application Architectures, O'Reilly Media, 2015, ISBN: 9781491900086
7. M. Tsaneva, "A Practical Approach For Integrating Heterogeneous Systems," Business management, no. 2, p. 11, 2019.
8. P. Milev, "Technological Issues of Storing Dynamic Data in a Relational Database on Research Projects," Trakia Journal of Sciences, vol. 13, pp. 22-25, 2015.
9. V. Lazarova, "The Expansion of the E-business Information Systems in the Cloud- Problems and Solutions," Ikonomiceski i Sotsialni Alternativi, no. 4, pp. 34-47, December 2015.
10. V. Mihova, "Common Architecture Design of a Business Information System for Performance Management of the Business Applications," in 3rd International conference on application of information and communication technology and statistics in economy and education ICAICTSEE–2013, Sofia, Bulgaria, 2013.
11. Iconiq Inc., "Hadoop Use Cases" 11 1 2022. [Online]. Available: https://www.projectpro.io/article/hadoop-use-cases/232.

The Moonshot Phenomenon in the Business World

Kiril Dimitrov

University of National and World Economy, Sofia, Bulgaria
kiril.dimitrov@unwe.bg

Abstract. This paper reviews basic and emerging research trends on the moonshot phenomenon in the business world. It traces the emergence and development of the nuances for this construct, its relationships with the proclaimed corporate culture and some new attributes of management by objectives. Nevertheless the numerous realizations that the moonshot phenomenon has acquired – as goals, projects, tasks, specific leadership, skill and others, it seems subordinated to certain principles and approaches, constituting the so called moonshot thinking that is also succinctly described.

Keywords. Moonshot, corporate culture, proclaimed corporate culture, strategic management, exponential organizations.

1. Introduction

It is generally accepted that the term "moonshot" officially came into being at a joint session of Congress of the United States of America on 25th of May 1961 through a speech, delivered by the president John F. Kennedy, i.e. "this nation should commit itself to achieving the goal, before this decade is out, of landing a man on the Moon and returning him safely to the Earth" [1, 2, 15]. The principles of business administration imply that "the goal", formulated and communicated by Kennedy incarnates a "bold statement", grounded on "engineering expertise and insight", information about USSR's achievements and capabilities as the main competitor, and understanding NASA's team capabilities [3]. It is evident that the first application of the moonshot appears to be in the state administration. But management evolution has proved numerous times that many new management constructs, theories, methods, etc. are transposed later from the state administration (the military and other agencies) to the business world. That is why the purpose of this paper is to review and identify diverse realizations of the moonshot phenomenon in the business world. Since the moonshot phenomenon acquires new realizations in another societal sphere, it adds up new shades to its meaning, too. In such situation the academic publications (books and articles), describing this transition, turn out to be few. That is why blogs by opinion and business leaders and corporate websites are also reviewed to delve deeper into the specificity of the existence and evolution of the moonshot.

2. Transposing the moonshot as a goal for the business organization

The contemporary view to the theory of management by objectives permits defining the moonshot goal as an effective and efficient means of achieving the massive transformative purpose, formulated and pursued by the leaders and sometimes other constituencies of the business organization [4, 5, 6]. In particular the moonshot goal may be determined as an "ambitious", business goal that should be achieved within the next 5-10 years that categorizes it as a long-term one [5, 6]. It is strongly associated only with "the most game-changing

disruptive innovations" and related geographically with a global center for high technology and innovation – as the Silicon Valley in the United States of America [7]. Palao [5] implies its belonging to the aggregate of the proclaimed culture elements by emphasizing its inseparable connectivity and high congruence not only with the massive transformative purpose, but also with other official corporate documents as purpose, mission, vision, and organizational values. In its turn AQAL Group [17] preserve the hierarchical order between the massive transformative purpose and the moonshot, but present them as key inventories of the organizational manifesto – another proclaimed culture company document that may be created by the senior leadership and may attract diverse constituencies to the campaign of its realization. So, the moonshot goal receives triple realization within the set of proclaimed culture elements – first, an element, subordinated to the massive transformative purpose [4, 9], second – as an element, subordinated to the manifesto [17], and third - as an independent element with a rank, equal to the other items in the aforementioned set [12, 15].

It is worth considering the view of "Company X" – an exponential subsidiary of Alphabet that seems to be the initial driver of the transposition for the moonshot from the state administration to the business sector. This subsidiary uses the content approach to define the semantic area of the moonshot goal as "an intersection of three ingredients" [8], proposed by Eric "Astro" Teller – the first director of the subsidiary, also nicknamed as "Captain of Moonshots", as follows:

- The existence of a huge world problem, making an impact on the survival or the quality of life for millions or billions of people that should be identified by organizational leaders.

- The emergence of a radical, science-fiction based, potential solution to the aforementioned problem that even may even seem unrealizable at the present moment.

- The recent advent of a technological breakthrough, supporting leaders' hope that the realization of the aforementioned solution may become possible in the period of 5-10 years.

General approach of formulating a moonshot, consisting in five phases, is proposed by Diamandis [4]. Other variants of moonshot design approach exist, too – rendering the necessary phases to three [9] or introducing different values for chosen time measures [10] (see table 1).

Table 1. Approaches to design process of the moonshot

Steps	Diamandis's approach	SMExO community's approach	Johnson's approach	Minkowski's approach
1	Prepare a concise description of the organizational 5 year moonshot, using clear language, being impartial and specific in dates and numbers.	Formulate your organizational 5-year Moonshot.	Formulate your organizational moonshot (no end terms mentioned).	Choose a challenge for the company: (a) from the outside, i.e. (a Sustainable Development Goal) or (b) an internal goal for the company. Describe why it matters to the entity (motivation, exploration sphere, identified opportunities).

Steps	Diamandis's approach	SMExO community's approach	Johnson's approach	Minkowski's approach
2	Formulate a specific, measurable objective for the first year, embodying the first real action to achieve your moonshot.	Formulate a specific objective you can achieve this year that is intended to put your moonshot on schedule.	"What needs to happen in the next 3 months, for that to be true in 6 months?"	Select a technology in order to articulate the potential moonshot. Explore technology's second generation effect as a potential solution for the moonshot. This technology should trigger exponential growth for the entity.
3	Identify initiatives and actions that can be undertaken in the next 30 days to test and minimize the risk of your 1-5 year moonshot and objectives (ideas).	Formulate 3 specific sub-objectives you may achieve this month that are expected to put your one-year objective on schedule.	What needs to happen in the next month for that to be true in 3 months?	Develop a design question, e.g. "How might we solve world poverty by providing access to free, autonomous transportation to everyone?"
4	Search for evidence, supporting the achievability of your organizational moonshot.			Generate bold ideas to cope with the design question.
5	Identify a concrete action that may be undertaken at the moment to make immediate progress. It should be performed within the next one hour.			Formulate the organizational moonshot by means of discussion with peers.
Sources: [4, 9, 10, 16].				

The moonshot goal may also be described by means of its characteristics, identified by business leaders and researchers (table 2). Some of these characteristics are accompanied by succinct descriptions, while others are only presented by means of the respective labels.

Table 2. The characteristics of the moonshot

Author	Described characteristics
Diamandis [4]	- wonderful, beautiful, inspiring, poetic, beautiful, - occupying the role of "world unifier", - filled with "great technical challenges" and - related with numerous manifestations of true heroism

Author	Described characteristics
Jain and Schroeter [11]	- crazy - avoidance of competition, - associated with high risk - occupying the role of catalyst for frequent changes (pivots) in "all ground rules" - "smashing all existing paradigms" in the business and regarding the pursued strategies
Berman [12]	- wildly ambitious
Rockwell [13]	- mobilizing followers' efforts - engaging followers emotionally and rationally - build on the successes already achieved, but without repeating them - a catalyzer of numerous iterations and continuous adaptation - touching followers' lives
van der Pol [14]	- attracting and retaining great talent
X company [8]	- not always grounded on science or technology breakthrough
Anthony, Johnson [18]	- inspiring – stimulating people to perform at extraordinary high levels. - credible – grounded on detailed assessments of underlying technological trends that have to ensure to some extent its achievability. - imaginative – embedding „a meaningful break from the past".
Skibba [27]	- high reliability of the necessary resources (funding) - expressed public support (in and outside the company)

Most of the business organizations possess the ability (resources, time, and efforts) to pursue only a single moonshot goal with the exception of the mighty multinational companies (holding companies, conglomerates) that may contribute in numerous spheres through their diverse strategic units [14].

3. Other realizations for the moonshot

In addition to the moonshot goals a wide range of realizations for the moonshot phenomenon in the business sector exist. Some of them are directly targeted to the organizational setting, while others impact higher-rank systems (industries, regions, nations). Some of the main realizations for the moonshot phenomenon, based on the adopted approach to management, are, as follows:

- According to the process approach: (a) the term moonshot entered the business lexicon as shorthand for "a difficult or expensive task, the outcome of which is expected to have great significance" [18]; (b) taking into account the stance, occupied by Eric "Astro" Teller who emphasizes the act of moonshotting, describing it as "namely attacking a clearly defined problem through a radical solution supported by practical scientific data for huge upside potential" [22].

- Through the lens of the project-based management the moonshot represents: (a) an extraordinary project or proposal that fulfil three criteria, i.e. addressing a big problem, proposing a radical solution and utilizing innovative thinking and technology [6]. A six step plan is proposed to accomplish the moonshot project (table 3); (b) "risky, time-consuming

ventures" and/or "visionary projects", outlining the cherished company's development direction by the senior leaders, the ways in which the company may "get there", and constituting the workload of the most talented employees for long periods. The moonshot projects are required to align with the elements of the proclaimed culture as mission and values, as well as the pursued goals by the core company [19]; (c) "extremely ambitious project or mission" - i.e. an assignment, "undertaken to achieve a monumental goal" in order to solve the biggest societal issues and strengthen the resilience of business organizations, whole industries and the economy as a whole [23]; (d) a moonshot initiative, defined as "a novel, unexpected, ambitious, difficult, and risky mission" (e.g. an assignment), planned to exceed "the known limits" through disruptive innovations [25]. It is also depicted the image of the successful moonshot from this perspective - communicated imaginatively but simply, embedding the right mix of vision, framing, leadership, and team excellence.

Table 3. The plan for accomplishment of moonshot project

Steps	Description
1. Identify the huge problem	Generate great ideas
2. Assessing the high potential to overcome the huge problem	Mission impossible perspective
3. Create a project team of necessary experts.	Committed, motivated, collaborative people
4. Identify most difficult aspect of the project.	First, plan and perform the work activities, needed for the completion of the hardest part of the project.
5. Foster a growth mindset among the team members.	Failures should be viewed as the main source of learning and celebration events.
6. Actively searching for buy-in to the project at every level of your business organization.	
7. Fast implementation of the plan should follow.	
Source: [6].	

- The product management perspective creates another useful realization for the moonshot phenomenon within the organizational setting (a startup or an exponential subsidiary of big company). It is developed by means of creative adoption of agile and lean principles of management within prescribed sets of respective steps, i.e.: (a) the moonshot is described as a new product manager's skill [20] – design of "moonshot vision", defined as creating "an ambitious, groundbreaking"…"clear and compelling"… "product vision", associated with a specific set of steps to be followed by the product design teams and senior leaders in the companies, pursuing accelerated growth in these uncertain and disruptive times. (b) other set of steps is proposed for the achievement of "moonshot effect", specifying its unique deliverables as public beta or minimum viable product [21].

- The moonshot phenomenon may even accept the shape of great challenge to contemporary companies, not only struggling for their survival, but also striving for operational excellence and sustainable development, i.e.: (a) moonshot—a game-changing challenge that creates high performance teams, elevates leaders and catapults careers [21]; (b) the moonshot is perceived as "a huge, seemingly unachievable challenge" by Richard Branson – a globally famous entrepreneur [11, in the forward].

- The moonshot phenomenon manifestations may be assessed from leadership perspective, bringing to the forth cherished characteristics of the moonshot leaders and identification of common elements of moonshot leadership (table 4) [25].

Table 4. Leadership perspective of the moonshot phenomenon

Nuances	Description
1. Characteristics of moonshot leaders [25]	1.1. Future-oriented (Futurists), envisaging big ideas. 1.2. Entrepreneurial (Impact-drivers), opting for radical solutions with massive impact on companies and humans. 1.3. Adopting technologies (Technologists), experimenting with technology. 1.4. Oriented to innovations (Innovators), bringing ideas to life.
2. Elements of moonshot leadership [25]	2.1. Identifying a big issue to solve. 2.2. Searching for a radical solution. 2.3. Relying on disruptive technologies. 2.4. Oriented to intensive teamwork, extraordinary actions in daily work and unbelievable achievements. 2.5. Overcoming psychological barriers as (self-)doubts and lack of intent.
3. The traits of the person who pursues a moonshot [11]	3.1. "A mindset of abundance, sustained by possibility thinking, imagination and curiosity." 3.2. Strong desire to improve the lives of millions or even billions of people. 3.3. Sense of wonder. 3.4. "Courage to ask big, daring "what if" questions – even "crazy" questions".

- The change management perspective to the analyzed phenomenon has created the derivative construct of "moonshot transformation", described as an array of undertaken strategic moves by business organizations [26], as for example: (a) Designing not only new products or services, but also developing radical technologies with potentially great impact on company's mission; (b) The moonshot pursuit should be confined in a separate division away from the core activities of the company, doomed to pursuing of fast return on investments, in order to decrease the tensions and sharp contradictions between – on one hand, long-term versus short-term orientation in the completion of the undertakings, and non-profit driven versus for-profit oriented activities – on the other hand. (c) Providing the choice for the employees and other constituencies (e.g. clients) how to interact (work) with the company, i.e. by their physical presence in the office (on the site), entirely virtually or in a mixed way (hybrid), if and when it is possible. (d) Implementing mass customization through the use of exponential technologies (e.g. providing personalized academic education through artificial intelligence).

- The creativity perspective reveals the realization for the moonshot phenomenon as "a (crazy) idea" or "moonshot idea", supported by a specific design exercise [28, 29].

- The innovation perspective brings to the emergence of the term "moonshot innovation", equivocally described as "a wishful thinking" or business reality that both impose on the organizational leaders and employees the "painstaking task" of first searching for "the right questions" and after that starting to find clever answers for them [31].

- The practical side of revealing the essence of moonshot through provided specific examples constitutes another interesting realization for the moonshot phenomenon in business organizations. Two main streams may be identified here: (a) further specifying the massive transformative purpose by scope, depth and/or planning horizons (time) [4, 9, 15]; (b) using the same examples as those, proposed for the massive transformative purpose without taking into account the potential existence of the last one and any relationships with it [12, 30].

4. Principles and approaches underlying Moonshot thinking

The richness of realizations for the moonshot phenomenon in the business world is directed by stable and emerging principles of moonshot thinking, elaborated predominantly through the experience of Company X – a subsidiary of Alphabet, led successfully to exponential growth by its first director Eric "Astro" Teller. This entity officially proposes its own definition for "Moonshot thinking" as "pursuing things that sound undoable, but if done, could redefine humanity" [8]. Thus, the intersection sphere among the encountered and wisely chosen huge problems, affecting the lives of billions of people, the associated with them radical solutions and the breakthrough technologies attracts the attention of business leaders and diverse constituencies, striving to sustain business success and increase the organizational growth exponentially. The systematic efforts in this direction by the leaders of the aforementioned subsidiary brought to the creation of "a place where the processes and culture make it easier to make radical breakthroughs – repeatedly", i.e. a "moonshot factory" which performance at satisfactory levels is ensured by adhering to a set of guiding principles [8], as follows:

- The preference of aiming for 10X, not the incremental 10% in order to cope with exponentiating contemporary challenges and even anticipate them, create passion among the pursuers of the respective moonshot, gain freedom to change basic assumptions (i.e. all ground rules), search for and find out unexpected answers within the semantic area, locked between audacity, creativity and science fiction. This means deliberate and continuous prioritization of innovative behavior that may bring in the success not only in big and small companies, but also in the state administration. So, culture matters, especially the creation of innovation culture.

- Falling in love with the problem – identifying reasons and influencing factors, searching new approaches to solve the respective problem and assuming that technology is just a means to find the best solution.

- Communicating the real world (surveying public opinion or natural conditions) as early and as possible in order to identify disadvantages of the proposed solution and potential corrective actions.

- Increasing creativity in the company through structuring of diverse teams with members, originating from various communities, cultures, and professions in order to be able to present, challenge, defend and even partially accept mutually their unique stances and together generate better ideas through discussions and consensus.

- Prioritizing the work on the hardest, most important part of the problem and postponing, if possible, the completion of the necessary relatively simple tasks. Thus the speed of organizational (team) learning increases, regarding this problem, so decision-makers can sooner decide whether to persevere further or to redirect their efforts, resources and time to other bright ideas. Team culture development and long-term commitment constitute the two great managerial challenges here.

- Adopting of failure learning approach that requires the creation of specific organizational (team) culture, characterized by psychological safety for failing employees and managers, accepting failures as learning opportunities (lessons), and attributing them with the

same prestige as celebrating successes. Thus, the audacity, demonstrated among team members, increases and change resistance diminishes.

- Occupying the role of a chaos pilot who: (a) embraces the uncertainty, (b) perceives the unknown as an advantageous opportunity and a source of creative energy.
- Learning to love the first prototype ("version 0. crap" or"v0.crap") – the earliest, scrappiest version of the deliverable (product, service, process) that may provide the team with honest and open feedback in order to further elaborate and iterate their ideas by experimentation.
- Shifting perspective. The solution of big problems is not always complex and expensive, because changing held basic assumptions, regarding them and the creative use of multidisciplinary approach may generate results that greatly exceed the expectations.
- Adhering to the long-term view on the undertaken initiatives, ensuring intrepidity and tranquility, regarding necessary explorations, experiments, risks and audacious ideas.

Furthermore, the image of the "the moonshot thinking factory" that is expected to retain the success in the business organizations may be determined and monitored by multiple answering a number of questions [10], as follows:

- What are the involved constituencies? (departments, skills and expertise, insiders and outsiders)
- "What are they solving for?"
- "When and where do people meet?"
- "What is the reason people give for attending?"
- How is the momentum kept?
- How can it be detected that people's mindsets are changing?

The crisis management perspective also gives birth to another principle of moonshot thinking, describing it as a mixture of three complex ingredients - courage, resourcefulness and experience, contributing to better decision-making in business organizations in times of diverse crises (table 5) [24]. Even contributions of moonshot thinking to tackle some of the issues during the Covid-19 pandemic were identified, raising its status to an innovation driver, producing results in surprising ways [41].

Table 5. The three ingredients of moonshot thinking

Complex ingredient	**Description**
Courage	Reliance on intuition. Mutual trust among team members. Making decisions together and sharing the responsibility. Open communication environment and appreciation of diverse opinions. Believing in the positive outcome.
Resourcefulness	Building a plan on the decision when it needs to be made. Include all the team members in the potential solution of the problem. Relying on the best judgments, regarding the challenges. Communicating clearly, minimizing distractions. Focusing on what currently works (performs) well. Orientation to the efficient use of resources. Performing multi-perspective monitoring of the current situation. Initially working on multiple possible options and after that making a swift decision which one to follow, based on succinct discussion. Thinking all the time in options.

Complex ingredient	Description
Experience	Preliminary developing workarounds and procedures based on simulations of probable failures and finding appropriate solutions. All these serve as a sound base of experience to upgrade during a crisis. Performing reflection and learning session immediately after the end of the crisis.
Source: [24].	

The moonshot thinking is also associated with two other modern management approaches, as follows:

- Future-back planning that is used first to determine the future state of the respective business organization (or an entire industry) and after that directing the managers' attention backwards to the current moment in order to create actionable milestones. In this way the business organizations generate growth simultaneously from two sources – on one hand, the extension of the lifespan for the current core business activities for the company, and on the other hand, continuous identification and seizure of the new growth opportunities of tomorrow [32, 33].

- Infinite learning as a part of moonshot thinking brings forth organizational agility, resilience and exponential growth orientation through reassessing the impacts of the current exponential technologies, social and market changes, frequently going outside leader's and employee's comfort zone, regarding problems, challenges beliefs and basic assumptions in order to transform the company in a timely manner or even before the emergence of certain events in the business environment [34, 35, 36, 37].

5. The new attributes of Management by objectives and another related construct as reverberations of the moonshot phenomenon

The moonshot phenomenon is revealed by means of comparing several semantically related constructs. The first two of them represent the newness in the management by objectives as massive transformative purpose and big hairy audacious goals. The loonshot represents another related construct to the moonshot phenomenon, i.e. an idea (table 6).

Table 6. Semantically close constructs to the moonshot

Attributes	Description
Moonshot [38]	1. "An ambitious and expensive goal, widely expected to have great significance." 2. An effective and efficient means of achieving the massive transformative purpose. 3. Its formulation is based on future-back planning. 4. It should be achieved in the next 5-10 years. 5. Receiving the support by organizational leaders since its start until it is achieved or definitely abandoned in a timely manner.

Attributes	Description
Massive transformative purpose [39]	1. "The higher, aspirational purpose of the organization" or sincerely, confidently and officially expressed, shared, high generalized, corporate aspirations "to accomplish near-miracles" in order to win the "hearts and minds - imaginations and ambitions" of all constituencies, especially the external ones. 2. Strongly supported by organizational leaders.
Loonshots [38]	1. The most important ideas, projects or breakthroughs that do not realize seamlessly in terms of support, funding, recognition, etc. 2. These are accompanied by long periods of expressed skepticism, uncertainty, crush, ridicule or neglect and their champions are often considered crazy or even fired by their employers or abandoned by their sponsors. 3. But later on these recovered to transform into paradigm shifting innovations.
Big hairy audacious goals [40]	1. A long-term goal that is understood and supported by all personnel members in the business organizations. It engages, mobilizes and unites the people. 2. Present-future planning guides the formulation of the Big hairy audacious goals. 3. Sustainably supported by organizational leaders.

Thus, certain specifics of the moonshot phenomenon are identified and outlined: (a) the existence of potential hierarchical dependence among its realizations (i.e. between some of the constructs), (b) different attitudes of organizational leaders, regarding the provided support of its realizations, (c) simultaneous use of two directions of planning in the business organizations (i.e. present – future; future - back).

6. Conclusion

This review clearly shows the availability of great diversity in the stances on the moonshot phenomenon, occupied by company constituencies that justifies the identified and critically analyzed, numerous realizations. Corporate survival could not be sustained and organizational leadership positions could not be conquered and retained in the contemporary fast-changing, volatile, uncertain, complex and ambiguous business environment without realizing persistent pursuit of crazy ideas to solve the biggest problems of humanity, striving for a better world by creating the future and struggling for it through participation in deep transformations of societies and industries, creating extremely high value-added for the constituencies, attracting and engaging the best talent and solving a number of intermediate issues while iterating along the way of achieving your moonshot. That is why the design and implementation of moonshots in succeeding business organizations, no matter their embedded meanings (as specific goals, projects, tasks, specific leadership, skill and others) has become an indispensable part of the contemporary strategic management process and the official expressing of the intentions by the senior leaders.

Acknowledgements

This publication is funded by the University of National and World Economy (UNWE) Research Programme (Research Grant No. 02/2021).

References

1. Brinkley, D., American Moonshot. John F. Kennedy and the great space race, 1st edition, HarperCollins publishers, 2019.
2. VideoFromSpace, President Kennedy's Moonshot moment in front of Congress on 60th anniversary, 25th of May, 2021 (https://www.youtube.com/watch?v=aVX5j1SHfSo), accessed on: 08.11.2021.
3. Šandala, B., The Forgotten Truth About The Moonshot Goal, Product Coalition, 26th of December, 2020 (https://productcoalition.com/the-forgotten-truth-about-the-moonshot-goal-cf546b344c92), accessed on: 08.11.2021.
4. Diamandis, P., Find Your MTP & Moonshot Workshop (video), SU Global Summit 2019, December, 6th, 2019 (https://www.youtube.com/watch?v=vXJxRA_EfmQ), accessed on: 26.09.2021.
5. Palao, F., The Purpose Launchpad Guide, July, 2020 (https://www.purposelaunchpad.com/), accessed on: 08.07.2021.
6. Thackray, G., Moonshot thinking: disrupting the norm in business, Training Zone, 3d of November, 2017 (https://www.trainingzone.co.uk/develop/business/moonshot-thinking-disrupting-the-norm-in-business), accessed on: 12.11.2021.
7. Sculley, J., Moonshot! Game-Changing Strategies to Build Billion-Dollar Businesses, Rosetta books, 2016.
8. X company, website and sections: Moonshot blueprint; Moonshot thinking, 2021 (https://x.company/moonshot/), accessed on: 13.11.2021.
9. SMExO community, Massively Transformative Purpose (MTP), (n.d.). (https://usercontent.one/wp/www.smexo.dk/wp-content/uploads/2017/12/Abundance-digital-massive-transformative-purpose.pdf), accessed on: 31.09.2021.
10. Johnson, R., Breakthrough Innovation. Mindsets and Business Models, EDIE Sustainability Leaders Forum, a presentation, 2019 (https://event.edie.net/forum/wp-content/uploads/sites/32/2019/02/Leadership-Labs-Slides-Richard-Johnson.pdf), accessed on: 12.11.2021.
11. Jain, N., J. Schroeter, Moonshots. Creating a world of abundance. How entrepreneurs are enabling life without limits, Moonshots Press, 2018.
12. Berman, A., This Is How to Invent Radical Solutions to Huge Problems, Singularity Hub, section: InFocus, Nov. 15, 2016 (https://singularityhub.com/2016/11/15/this-is-how-to-invent-radical-solutions-to-huge-problems/), accessed on: 12.11.2021.
13. Rockwell, D., Moon Shot Goals, Leadership Freak blog, 17th of October, 2018. (https://leadershipfreak.blog/2018/10/17/moon-shot-goals/), accessed on: 120.11.2021.
14. van der Pol, H. J., Why your company needs a moonshot goal, Perdoo, Resources hub, July 1st, 2017 (https://www.perdoo.com/resources/moonshot-goal/), accessed on: 12.11.2021.
15. Minkowski, A thousand moonshots. Solving strategic company challenges with bold ideas, tab: Moonshots, section: Examples of moonshots, 2019 (https://1000moonshots.com/moonshots/), accessed on: 12.11.2021.
16. Minkowski, A thousand moonshots. Solving strategic company challenges with bold ideas, tab: Toolkit, 2019 (https://1000moonshots.com/toolkit/), accessed on: 12.11.2021.
17. AQAL Group, The Investment Turnaround is our Moonshot, October 16th, 2017 (https://aqalgroup.com/investment-turnaround/), accessed on: 12.11.2021.
18. Anthony, S., Johnson, M. What a Good Moonshot Is Really For, Harvard Business Review, section: Innovation, 14th of May, 2013 (https://hbr.org/2013/05/what-a-good-moonshot-is-really-2), accessed on: 12.11.2021.
19. Three Five Two, Moonshot Projects – Innovation or Wishful Thinking?, Three Five Two blog, section: Innovation, 2021 (https://www.threefivetwo.com/blog/moonshot-projects-innovation-or-wishful-thinking) accessed on: 12.11.2021.
20. Sapata, R. K., Moonshot thinking: How to Build Moonshot Vision as Product Managers, LinkedIn article, 22d of August, 2019 (https://www.linkedin.com/pulse/how-build-moonshot-vision-product-managers-ravi-kumar-sapata), accessed on: 12.11.2021.
21. Goldman, L., K. Purmal, A. Janze, The Moonshot Effect: Disrupting Business as Usual, Wynnefield Business Press, 2016.
22. Google Workplace, Moonshot Thinking from Astro Teller, YouTube video, 13th of November, 2013 (https://www.youtube.com/watch?v=cA_8IO3vbFs&t=17s), accessed on: 12.11.2021.

23. Stilley, W., Why We Need More 'Moonshot' Businesses, NASDAQ articles, JUL 28, 2021 (https://www.nasdaq.com/articles/why-we-need-more-moonshot-businesses-2021-07-28), accessed on: 12.11.2021.
24. Minkowski, How to: moonshot thinking in a crisis, 2021 (https://minkowski.org/moonshot-thinking/) accessed on: 12.11.2021.
25. Calvo, J., Journey of the Future Enterprise: How to Compete in the Age of Moonshot Leadership and Exponential Organizations, 1st edition, Libros de Cabecera, 2020.
26. Wee, S. C., How Google and Japan's top business school use 'moonshot transformation' to leap ahead, TechInAsia, section: Startups, 8 Nov 2018 (https://www.techinasia.com/google-japans-largest-business-school-moonshot-transformation), accessed on: 12.11.2021.Hhh
27. Skibba, R., What a modern moonshot would look like, ABC news, Inside Science, 21 July 2019 (https://abcnews.go.com/Technology/modern-moonshot/story?id=64444878), accessed on: 12.11.2021.
28. Genius Network, Tony Robbins & Peter Diamandis: Special Podcast, Live From The Genius Network, #306, Youtube video, 3d of May, 2016 (https://www.youtube.com/watch?v=eMAo6Hgb4bA), accessed on: 12.11.2021.
29. Villebaek, L., Moonshot Ideation Exercise, 1st of February 2021 (https://www.sprinthero.com/blogs/post/Moonshot-Ideation-Exercise), accessed on: 12.11.2021.
30. Bendix, A., 9 moonshot ideas that became real-life innovations, from driverless cars to birth-control pills, Business Insider, section: Science, August 14th, 2019 (https://www.businessinsider.com/moonshot-ideas-that-became-reality-2019-8#floating-cities-are-still-theoretical-but-floating-housing-complexes-are-already-standing-in-the-netherlands-1), accessed on: 12.11.2021.
31. Mention, A. L., J. Pinto Ferreira, M. Torkkeli, Moonshot innovations: Wishful Thinking or Business-As-Usual? Journal of Innovation Management (JIM) Vol.7, No.1, pp1-6, 2019 (DOI: https://doi.org/10.24840/2183-0606_007.001_0001), accessed on: 12.11.2021.
32. Johnson, M., J. Suskewicz, Lead from the future: how to turn visionary thinking into breakthrough growth, Harvard Business Review Press; Illustrated edition, 2020.
33. Johnson, M., R. Davis, A Future-Back Approach to Creating Your Growth Strategy, Innosight: Strategy and Innovation at Huron, section: Executive Briefing, February, 2014 (https://www.innosight.com/insight/a-future-back-approach-to-creating-your-growth-strategy/), accessed on: 12.11.2021.
34. Alayón, D., Understanding Moonshot Thinking, LinkedIn article, 31st of December, 2018 (https://www.linkedin.com/pulse/understanding-moonshot-thinking-david-alayón), accessed on: 12.11.2021.
35. Van Gasteren, W., What is a Growth Hacking Mindset? (+ 5 tips to develop it),Grow with Ward. Growth hacking coach, 23d of January, 2021 (https://growwithward.com/growth-hacking-mindset/), accessed on: 12.11.2021.
36. Klein, G., The Pre-Mortem Method: A practical way to do risk assessment, Psychology today, January 14, 2021 (https://www.psychologytoday.com/us/blog/seeing-what-others-dont/202101/the-pre-mortem-method), accessed on: 12.11.2021.
37. Hoffman, R., C. Yeh, B. Casnocha, Learn from People, Not Classes, Harvard Business Review, pp50-51, March–April, 2019.
38. Bahcall, S., Loonshot. How to nurture the crazy ideas that win wars, cure diseases, and transform industries, St. Martin's Press, 2019.
39. Ismail, S., M. Malone, Y. Van Geest, P. Diamandis, Exponential organizations. Why new organizations are ten times better, faster, and cheaper than yours (and what to do about it), New York: Diversion books, 2014.
40. Collins, J., J. Porras, Build to last: successful habits of visionary companies, New York: Harper Collins, 1994.
41. Cain, S., Moonshot thinking drove innovation during the pandemic in surprising ways, Fortune, June 25th, 2021 (https://fortune.com/2021/06/24/moonshot-thinking-drove-innovation-during-the-pandemic-in-surprising-ways/), accessed on: 29.11.2021.

Growth Rate of the Enterprise (Methods and Methodology for Analysis)

Rositsa Ivanova

University of National and World Economy, Sofia, Bulgaria
RIvanova@unwe.bg

Abstract. Development may be defined as evolution, as continuous transition from one stage to another, which is more perfect and at a higher level. Under the actual conditions of the economy, enterprises may develop in two opposite directions, each of which exists objectively and is formed under the impact of a system of factors – internal and external; direct and indirect; extensive and intensive, etc. One of the directions of enterprises' development is to growth or progress, and the other – to regression. The economic development may be considered in two aspects – macroeconomic and microeconomic aspect. The economic growth at enterprise level is the object of the study. It suggests improved methods for analysis of the enterprise's growth rate. The growth rate is analyzed and assessed as a stand-alone object of the analysis, i.e. as a resultative indicator. Thus, the power and direction of impact of direct factors affecting the enterprise's growth rate are identified.

Keywords. Economic growth, growth rate, impact, analysis, methodology

1. Introduction

The Covid-19 pandemic has had immediate impact on the development of the economy in global aspect. The occurrence of new virus mutations, the increased number of infected cases and the restrictive health and economic measures for overcoming the pandemic undertaken at the end of 2021 and the beginning of 2022 have slowed down the recovery and growth rate of big economies worldwide. In 2021, the economic development showed visible stagnation and at the same time, we witnessed inflation increase. To the analyzers' opinion "inflation at global level is due to the expanded money printing and to the increased credit with almost uncontrolled crediting." (12)

The measures undertaken in different countries for restricting and overcoming the consequences of the coronavirus crisis had direct impact on the economic growth both globally and at individual enterprises level. The undertaken restrictive measures and the closure of a number of businesses made the development of the economy stagnant. In practice, specific conflict between health and economic measures undertaken by the governments of different countries occurred. Currently, such conflict still exists. A typical example in this respect are the green certificates, which apply in Bulgaria since October 2021. Probably, this conflict should be resolved in political manner at macro-level.

The development of the Bulgarian economy is in unison with the development of the economies of other countries. The economic growth rate in Bulgaria slows down under the influence of various factors, the key ones being the increase of energy products and food prices; the inflation increase; the employment and unemployment rates; the reference interest rate and its dynamics; the fluctuations in the USD to BGN exchange rate, etc.

The slowdown of the economic growth at macro-level also affects the economic growth at enterprise level. To our opinion, there is two-sided relationship here. The economic growth

of the enterprises from different sectors and industries of the national economy influences the economic growth of the national economy, i.e. at macro-level.

A number of factors have impact on the acceleration or slowdown of the economic growth at enterprise level, i.e. at micro-level. In the context of health, economic, social and price crises, the economic growth of enterprises records a slowdown trend. This is due to the impact of the following more important factors: delayed payments among enterprises thus increasing their indebtedness; increase of international prices of a number of material resources, raw materials and energy sources; increase of consumer prices; difficulties and delays in some sectors and industries of the national economy; deterioration of the business conditions in the country due to the health crisis, uncertainty for the future, dependence of some sectors on the reallocation of budget funds, uncertainties with regard to the commencement of the National Recovery and Resilience Plan implementation, high inflation, which to a certain extent limits the actual increase of population's available income, thus having adverse effect on consumption, high international prices of energy sources, measures undertaken for keeping jobs and business of enterprises, unemployment rate, status and perspectives for development of the labour market, efficiency of labour and income levels, investments under the projects of the National Recovery and Resilience Plan and those funded under the multiannual financial framework of the European Union, etc.

All of the above reveals the relevance and importance of the issues of economic growth both at macro- and micro-level.

2. Economic Growth at Macro- and Micro-Level

The economic growth may be considered in two aspects – macroeconomic and microeconomic aspect.

In macroeconomic aspect, we can define it as an increase of the total volume and value of products and services manufactured by the relevant economy for a given period of time in comparison to the previous period, for example, the current towards the previous year. If the value of products and services manufactured and sold during the current year is higher than the vale for the previous year, the economy records economic growth, and vice versa – the economy is in recession.

The opinion that "the economic growth is a key objective of the macroeconomic policy and source of wealth of the nation" (Todorov, 2015, p. 86) is generally accepted. From this perspective, the economic growth in macroeconomic aspect may be considered from two points of view. The first one is to study the economic growth as a process and as an outcome formed under the influence of various factors, such as: geographic location of the country; status of and relationships among the institutions in the country; political context; digitalization of enterprises and economy; effectiveness of use of resources (natural and manufacturing resources – fixed tangible assets, material resources and labour force), etc. This perspective allows to differentiate between extensive and intensive type of growth. The extensive type of growth is due to the input of more resources, for example, increase of the number of workers, and in this case effectiveness is not high. The intensive type of economic growth is achieved by enhancing the positive impact of intensive, qualitative factors, such as the increase of the efficiency of labour, increase of load ratio and decrease of absorption ratio of fixed tangible assets, increase of material productivity ratio and decrease of materials-output ratio of production, acceleration of short-term assets' turnover, etc., thus achieving relatively higher effectiveness.

As far as the second aspect is concerned, the macroeconomic growth is considered as a factor that has direct impact on the social policy of the government, and therefore, on the development and condition of the social sphere of the country.

The relationship between the economic development (growth or recession) and the social sphere is especially relevant in the context of the Covid-19 pandemic. The impact of the Covid crisis on the Bulgarian economy can be seen in the negative trends in the changes of the macro-indicators' values. Since the beginning of 2020 such trends have been also witnessed in the economies of the other European Union member states and at global level in general. The measures undertaken by the Bulgarian government of Bulgaria were aimed at supporting the enterprises to overcome the crisis consequences and to keep the jobs of their staff. We share the opinion that "the economic uncertainty caused by the development and duration of the coronavirus epidemic and the unemployment increase due to entirely national regulatory reasons and in line with the global supply chains, would inevitably result in contraction of domestic consumption" (Yotsov et al, 2020, p. 37). This is exactly where the objective relationship between the economic development and the social sphere of the country is manifested. On the other hand, during the summer of 2020 the Bulgarian government undertook a number of measures to mitigate the tax burden for businesses that were most severely affected by the restrictions imposed for overcoming the Covid-19 related crisis. From our perspective, the opinion that "seeking more balanced indirect taxation (is reasonable – added by R.I.), however, the provision of tax preferences for sectors that are not fundamental for the economy and do not have the potential to be the drivers of economic development creates prerequisites and requests for overall restructuring of the tax system, which, if done under the pressure of economic drop and global recession, is not an appropriate decision in short-term aspect. Furthermore, the question of priority and social effect of taxation changes sought in the country has been raised, whereas the lack of differentiation of taxes for essential commodities and medicines remains difficult to understand and explain. Thus, founding the tax system upon indirect tax revenue with proportionate taxation of different income groups of the society is another factor for intensifying income inequality" (Yotsov et al, 2020, p. 38)

The individual enterprise is the basic unit in every country's economy. The achievement of economic growth at macro-level suggests and requires economic growth at micro-level – at enterprise level. We share the opinion that "in advanced European highly developed companies, growth is based on new knowledge and innovation technologies as a source of high added value" (Mitov et al, 2012, p. 126). It is underlined that "the overall economic growth depends on the growth of individual economic units, as there is an organic relationship between its two hierarchy variants (economic growth at macro-level and at micro-level – note, R.I.)" (Mitov et al, 2012, p. 128). We further share the opinion that "on one hand, the quantitative and qualitative changes of individual enterprises, being accumulated in complex combination, form the general economic dynamics. On the other hand, the qualitative and quantitative characteristics of the national economy define the general conditions and prerequisites for the economic growth achieved in business organizations (enterprises – note, R.I.)" (Mitov et al, 2012, page 128). This is how the dialectical relationship between the macro- and micro-level economic growth appears in practice. To this end, there is an opinion that "quite close relationships and dependencies between macro- and micro-level economic growth" (Nenov, Atanasov, 2000, p. 3-4) exist, which we share. It's worth noting that despite the objectively existing relationships and dependencies between the macro- and micro-level economic growth, "significant differences in their content, objectives, metrics, factors, realization conditions, development stages, etc. also exist. At macro-level, growth should be subject to targeted state policy, while the government creates the necessary prerequisites for its realization via "the invisible arm". At micro-level, growth is an in-house issue of managers and owners, it mainly depends on their in-house capacities and objectives, is realized in an individual and unique manner in every company, at individual rates, with alternating phases of evolutionary growth, drop or stagnation, is determined by the level of "managerial services"

and the increase of their quality in time, by the motivation and interaction between the interests of various groups in and outside the company, etc." (Nenov, Atanasov, 2000, p. 3-4).

We share the opinion that "both at national (macroeconomic level – note, R.I.), and at company level (business organization – at micro-level – note, R.I.), economic growth passes the following stages:

1st stage: resources-based growth (growth determined by factors));

2nd stage: effectiveness-based growth;

3rd stage: innovation-based growth" (Blagoev, 2019, p. 27).

We support the opinion that the economic growth "depends on and comes down to the interaction between the production process results, the potential of manufacturing resources with the effectiveness of (from – note, R.I.) their use. Their interaction determines the type or kind of growth and the effectiveness or the price of such growth" (Nenov, Atanasov, 2000, p. 39).

3. Enterprise's Growth Rate – Nature and Methods for Analysis

The enterprise's growth rate can be analyzed by means of two approaches – accounting and financial approach. The accounting approach is based on the use of profit as a key economic category, and the information for the analysis of the enterprise's growth rate is provided by the accounting department. The accounting approach is "an expression of the neoclassical economic theory according to which the main objective of every company is to maximize profit" (Kasarova, 2010, p. 227). The financial approach is founded on the "value and (this approach – addition, R.I.) is an expression of the concept for maximization of the company's value. Market information is used for its realization, in particular, information for the company's shares" (Kasarova, 2010, p. 227).

The enterprise's growth rate may be differentiated as internal growth rate and sustainable growth rate.

The internal growth rate expresses the maximum growth rate that the enterprise is able to achieve at the expense of equity increase rate, without the use of borrowings for its activity. This should be achieved by means of increase of retained earnings. The question about the method of allocation of the balance sheet profit – for reinvesting in the enterprise's business and for payment of dividends to owners (shareholders), is especially important.

The funding of enterprise's business with own capital sources only gives rise to the question about the formation of its capital structure. To our opinion, we should analyze and assess the values of indicators that characterize the overall capital structure. This means the following indicators:

1) Autonomy (independency) ratio. It determines the share of equity in the amount of the total capital resources. If the value of equity increases without changing the value of borrowings, the autonomy ratio value increases. This may have negative impact on the economic growth of the enterprise, provided the turnover of equity has slowed down. With view of the duration of the involvement of equity in the economic turnover of the enterprise, it is defined as fixed equity and working equity. The analysis has the important task to identify the amounts of fixed equity and working equity. Fixed equity is embodied in the fixed assets, and the working equity – in the short-term assets of the enterprise. The slowdown of both fixed and short-term assets' turnover rate has adverse impact on the enterprise's economic growth.

2) Financial dependency ratio. It shows the share of borrowings in the total capital resources. The increase of the share of equity results in decrease of the share of borrowings

in the total capital resources of the enterprise. The decrease of the interest-bearing borrowings has direct impact on the enterprise's financial performance by means of accrued interest expenses. The analysis has the task to identify the degree of financial pressure during the analyzed period, which is characterized through the ratio between short-term and long-term capital borrowings used for the enterprise's business. In the course of repayment of enterprise's short-term liabilities, financial pressure has direct effect on the liquidity and solvency of the enterprise and thus on the economic growth of the enterprise.

3) Financial autonomy ratio. It shows the amount of equity attributable to borrowings of BGN 1. The increase of the internal growth rate through the increase of equity results in increase of the financial autonomy ratio value.

4) Debt ratio. It shows the amount of borrowings attributable to equity of BGN 1. The increase of the amount of equity results in decrease of the enterprise's debt ratio. The risk for the enterprise's economic development and its internal growth rate here is hidden in the fact that the decrease of the debt ratio, all other things being equal, results in decrease of the positive effect and increase of the adverse effect of the financial leverage, respectively, thus decreasing the enterprise's equity return.

5) Fixed capital ratio. Fixed capital characterizes the permanently functioning funds in the enterprise's business. These are equity and long-term liabilities.

In the accounting model, the internal growth rate (IGR) is determined with the following formula:

$$IGR = \frac{ROA \times K^r}{1 - ROA \times K^r} \quad (1)$$

where:

K^r is the profit reinvestment ratio

ROA – return on assets.

Profit reinvestment ratio characterizes the enterprise's dividend and financial policy. It shows what part of the profit is kept in the enterprise. It is calculated with the following formula:

$$K^r = 1 - K^{id} \quad (2)$$

where: K^{id} is the dividend ratio.

Dividend ratio is calculated with the following formula:

$$K^{id} = \frac{D}{P^n} \quad (3)$$

where: - dividend per share;

P^n - balance sheet earning per share.

In terms of analysis, the earning per share is of particular interest. Its value shows the portion of earnings attributable to one share. It characterises that portion of earnings that is reserved as an internal source of funding of the enterprise's business.

To our opinion, the change rates (growth or decrease) of the balance sheet profit, of income and assets should be determined and compared with the internal growth rate for the analysed time period. If profit growth or decrease rates are higher than the internal growth rate

during the analysed time period, this means that the enterprise has used borrowings to fund its activity.

The methods for analysis of the internal growth rate will be illustrated with the following indicative data. (see Table 1).

Table 1. Enterprise's internal growth rate

Indicators	Previous year	Current year	Deviation	Change rate
I. Input data, BGN 000				
1. Balance sheet profit	2128	2473	345	16,2124
2. Average amount of assets	10140	9490	-650	-6,4103
3. Total amount of revenue	13520	13685	165	1,2204
4. Amount of charged dividends	398	495	97	24,3719
II. Additionally calculated indicators:				
5. Amount of charged dividends attributable to balance sheet profit of BGN 1, BGN (item 3 : item1)	0,1870	0,2002	0,0131	7,0212
6. Reinvested profit attributable to balance sheet profit of BGN 1, BGN (1 – item 5)	0,8130	0,7998	-0,0131	-1,6153
7. Asset-based return, % [(item 1 : item 2) x 100]	20,9862	26,0590	5,0728	24,1722
8. Internal growth rate, % [(item 7 x item 6) : (1 – item 76 x item 6)]	0,2057	0,2633	0,0576	28,0031

Source: own calculations

The data in table 1 show that the internal rate growth of the enterprise for the current year in comparison to the previous year has increased by 0,0576 % (0,2633 – 0,2057). The change rate of internal growth for the current year in comparison to the previous year records increase of 28,0031 %. It is higher than the increase rate of the balance sheet profit (16,2124 %) but is lower than the revenue growth rate (1,2204 %). At the same time, we can see a decrease of 6,4103 % in terms of the average amount of enterprise's assets for the current in comparison to the previous year. This means that in general there is an increase of enterprise's assets. To our opinion, the analysis should go further in order to identify and assess the turnover of fixed tangible assets and short-term assets separately. This is necessary due to the different economic nature of these assets and their different involvement in the enterprise's economic turnover. The acceleration of assets turnover, in particular, the turnover of short-term assets, which are characterised with faster turnover rate and liquidity, has positive impact on the acceleration of the enterprise's growth rate.

The sustainable growth rate characterises the maximum growth rate of the enterprise that can be achieved and maintained without changing the debt-to-equity ratio, i.e. without changing the value of debt ratio (financial leverage).

Sustainable growth suggests and requires good balance both between equity and borrowings and between short-term and long-term external capital sources used to finance the enterprise's activity.

We sustain our opinion highlighted in our previous studies that the enterprise's sustainable growth rate may be considered both in broad and narrow sense (Ivanova, 2014, 2019, 2020, 2022).

In broad sense, the sustainable growth rate expresses the increase rate of the total amount of revenue, which the enterprise can achieve without changing its financial policy. For the purposes of achieving comparability in the mathematical formula for calculation of the sustainable growth rate in broad sense, the balance sheet profit, the total revenue and the average amount of assets of the enterprise for the reporting period are considered reference indicators. The enterprise achieves the total revenue, which is recognised for accounting purposes and which is considered a direct factor for the formation and dynamics of the balance sheet profit for the same reporting period, namely from the use and the effectiveness of use of assets.

The sustainable growth rate in broad sense (SGR^{bs}) may be calculated with the following formula – model 1:

$$SGR^{bs} = K^{d} \times KN^{a} \times PT^{I} \times K^{r} \quad (4)$$

where:

K^{d} - debt ratio (debt-to-equity ratio) of the enterprise. It characterises the enterprise's financial policy;

KN^{a} - asset load ratio (total revenue to average amount of assets ratio). It characterises the effectiveness of use of assets;

PT^{I} - net revenue-based return (percentage of balance sheet profit to total revenue). It shows the amount of balance sheet profit attributable to revenue of BGN 100 gained by the enterprise;

K^{r} - profit reinvestment ratio.

According to the required information for the needs of financial management and for the governance of the enterprise in general, the formula for calculation of the net asset-based return may be translated through the method of determinant factor modeling:

$$ROA = \frac{Pf^{n}}{\bar{A}} \times 100 = \frac{Pf^{n}}{I} \times 100 \times \frac{I}{\bar{A}} = PT^{I} \times KN^{a} \quad (5)$$

where:

Pf^{n} - balance sheet profit;

$\bar{A}$ - average amount of assets.

- total revenue.

Therefore, the sustainable growth rate in broad sense may be presented with the following formula - model 2:

$$SGR^{bs} = K^{d} \times ROA \times K_{r} \quad (6)$$

This formula shows that the change of the enterprise's sustainable growth rate in broad sense is subject to the influence of the following factors:

1) Changes in the debt ratio.

2) Changes in the net asset-based return.

3) Changes in the profit reinvestment ratio.

Through the net asset-based return, the enterprise's sustainable growth rate in broad sense is also subject to the influence of the changes in the revenue-based return and the changes in the asset load ratio.

Profit from enterprise's principal activity, and in particular, profit from sales, has the biggest relative share in the balance sheet profit composition, and net sales has the biggest share in the enterprise's revenue. This gives us reason to analyze and assess the sustainable growth rate in narrow sense. Namely in narrow sense, the net sales increase rate that the enterprise could achieve while keeping its financial policy shows both the effectiveness of its overall governance and of its principal activity. The average amount of manufacturing assets (assets involved in the production process – fixed tangible assets and some short-term assets), net sales of products and profit from sale of products are considered reference indicators in this particular case.

The sustainable growth rate in narrow sense (SGR^{ns}) can be calculated with the following formula – model 3:

$$SGR^{ns} = K^{d} \times KN^{pa} \times PT^{NI} \times K^{r} \quad (7)$$

where:

K^{d} - debt ratio;

KN^{pa} - manufacturing asset load ratio (net sales of products to average amount of manufacturing assets ratio). It shows the effectiveness of use of manufacturing assets;

PT^{NI} - net sales-based return (percentage of profit from sale of products and net sales);

K^{r} - profit reinvestment ratio.

The formula of manufacturing assets-based return may be presented as a product of net sales-based return ($\frac{Pf^{s} \times 100}{NS}$) and the turnover of manufacturing capital ($\frac{NS}{\overline{A^{pa}}}$):

$$PT^{pa} = \frac{Pf^{s}}{\overline{A^{pa}}} \times 100 = \frac{Pf^{s} \times 100}{NS} \times \frac{NS}{\overline{A^{pa}}} \quad (8)$$

where: are net sales.

Therefore, the enterprise's sustainable growth rate in narrow sense may be also presented with the following formula – model 4:

$$SGR^{ns} = K^{d} \times PT^{pa} \times K^{r} \quad (9)$$

This formula shows that the change of the enterprise's sustainable growth rate in narrow sense is subject to the influence of the following factors:

1) Changes in the debt ratio.

2) Changes in the manufacturing assets-based return.

3) Changes in the profit reinvestment ratio.

Through the manufacturing assets-based return, the enterprise's sustainable growth rate in narrow sense is also subject to the influence of the changes in the net sales-based return and the changes in the manufacturing asset load ratio.

The methods for analysis of the sustainable growth rate is illustrated with the following indicative data. (see table 2)

Table 2. Enterprise's sustainable growth rate

Indicators	Previous year	Current year	Deviation	Change rate
I. Input data, BGN 000				
1. Balance sheet profit	2128	2473	345	16,2124
a) profit from sales	1968	2250	282	14,3293
2. Average amount of assets, including:	10140	9490	-650	-6,4103
A) average amount of manufacturing assets, including of:	9112	8680	-432	-4,7410
a) fixed tangible assets	6100	5900	-200	-3,2787
b) short-term assets	3012	2780	-232	-7,7025
3. Average amount of capital, including:	10140	9490	-650	-6,4103
a) equity	8000	7700	-300	-3,7500
b) borrowings, including:	2140	1790	-350	-16,3551
b1) long-term liabilities	890	840	-50	-5,6180
4. Total revenue, including:	13520	13685	165	1,2204
a) net sales	12160	11960	-200	-1,6447
5. Amount of charged dividends	398	495	97	24,3719
II. Additionally calculated indicators:				
6. Amount of charged dividends attributable to balance sheet profit of BGN 1, BGN (item 5 : item1)	0,1870	0,2002	0,0131	7,0212
7. Reinvested profit attributable to balance sheet profit of BGN 1, BGN (1 – item 6)	0,8130	0,7998	-0,0131	-1,6153
8. Autonomy ratio, BGN (item 3a : item 3)	0,7890	0,8114	0,0224	2,8425
9. Financial dependency ratio, BGN (item 3b : item 3)	0,2110	0,1886	-0,0224	-10,6260
10. Financial autonomy ratio, BGN (item 3a : item 3b)	3,7383	4,3017	0,5634	15,0698
11. Debt ratio, BGN (item 3b : item 3a)	0,2675	0,2325	-0,0350	-13,0962

12. Average amount of fixed capital, BGN 000 (item 3a + item 3b1)	8890	8540	-350	-3,9370
13.Fixed capital ratio, BGN (item 12 : item 3)	0,8767	0,8999	0,0232	2,6426
14. Asset load ratio, BGN (item 4 : item 2)	1,3333	1,4420	0,1087	8,1533
a) manufacturing assets load ratio, BGN (item 4a : item 2A)	1,3345	1,3779	0,0434	3,2504
15. Revenue-based return, % [(item 1 : item 4) x 100]	15,7396	18,0709	2,3312	14,8112
a) Net sales-based return [(item 1a : 4a) x 100]	16,1842	18,8127	2,6285	16,2411
16. Asset-based return, % [(item1 : item 2) x 100]	20,9862	26,0590	5,0728	24,1722
a) Manufacturing assets-based return, % [(item 1a : 2a) x 100]	21,5979	25,9217	4,3238	20,0194
17. Return on equity, % [(item 1 : item 3a) x 100]	26,6000	32,1169	5,5169	20,7402
18.Sustainable growth rate – broad sense, %				
a) model 1 (item 11 x item 14 x item 15 x item 7)	4,5639	4,8453	0,2815	6,1672
b) model 2 (item 11 x item 16 x item 7)	4,5639	4,8453	0,2815	6,1672
19. Sustainable growth rate – narrow sense, %				
a) model 3 (item 11 x item 14a x item 15a x item 7)	4,6969	4,8198	0,1229	2,6166
b) model 4 (item 11 x item 16a x item 7)	4,6969	4,8198	0,1229	2,6166

Source: own calculations

The data in table 2 show that the sustainable rate growth of the enterprise for the current year in comparison to the previous year in broad sense has increased by 0,2815 points (4,8453 % - 4,5639 %), and in narrow sense, the increase is in the amount of 0,1229 points (4,8198 % - 4,6969 %).

The impact of the direct factor on the deviation of the sustainable growth rate in broad sense may be determined according to the subsequent replacement method. To this end, the following system of values is constructed on the basis of the input analytical model (see model 1), and namely:

$$SGR_0^{bs} = K_0^d \times KN_0^a \times PT_0^I \times K_0^r = 0{,}2675 \times 1{,}3333 \times 15{,}7396 \times 0{,}8130 = 4{,}5639\ \%$$

$$SGR_{usl.1}^{bs} = K_1^d \times KN_0^a \times PT_0^I \times K_0^r = 0{,}2325 \times 1{,}3333 \times 15{,}7396 \times 0{,}8130 = 3{,}9668\ \%$$

$$SGR^{bs}_{usl.2} = K^{d}_{1} \times KN^{a}_{1} \times PT^{I}_{0} \times K^{r}_{0} = 0{,}2325 \times 1{,}4420 \times 15{,}7396 \times 0{,}8130 = 4{,}2901\ \%$$

$$SGR^{bs}_{usl.3} = K^{d}_{1} \times KN^{a}_{1} \times PT^{I}_{1} \times K^{r}_{0} = 0{,}2325 \times 1{,}4420 \times 18{,}0709 \times 0{,}8130 = 4{,}9256\ \%$$

$$SGR^{bs}_{1} = K^{d}_{1} \times KN^{a}_{1} \times PT^{I}_{1} \times K^{r}_{1} = 0{,}2325 \times 1{,}4420 \times 18{,}0709 \times 0{,}7998 = 4{,}8453\ \%$$

The difference between the first conditional value and the value of sustainable growth rate for the previous year shows the impact of changes in the enterprise's debt ratio:

$$SGR^{bs}_{usl.1} - SGR^{bs}_{0} = 3{,}9668\ \% - 4{,}5639\ \% = -0{,}5971 \text{ пункта.}$$

As a result of the debt ratio decrease by BGN 0,0350 (0,2325 – 0,2675), the enterprise's sustainable growth rate has dropped down by 0,5971 points. This is due to the greater decrease of the absolute amount of borrowings to equity, which has direct impact on the changes that occur in the debt ratio. All other things being equal, the decrease of the debt ratio results in decrease of return on equity.

The difference between the second and the first conditional value shows the impact of changes in the enterprise's asset load ratio:

$$SGR^{bs}_{usl.2} - SGR^{bs}_{usl.1} = 4{,}2901\ \% - 3{,}9668\ \% = +0{,}3233 \text{ пункта.}$$

Due to the effect of the asset load ratio increase by BGN 0,1087 (1,4420 – 1,3333), the enterprise's sustainable growth rate has increased by 0,3233 points. This is due to the increased effectiveness of use of assets, being expressed in the increase of revenue gained with average amount of assets of BGN 1 that the enterprise holds and controls.

The difference between the third and the second conditional values characterises the impact of changes in revenue-based return:

$$SGR^{bs}_{usl.3} - SGR^{bs}_{usl.2} = 4{,}9256\ \% - 4{,}2901\ \% = +0{,}6355 \text{ пункта.}$$

As a result of the revenue-based return increase by 2,3312 points (18,0709 – 15,7396), the enterprise's sustainable growth rate has increased by 2,3312 points.

The difference between the sustainable growth rate for the current year and the third conditional value of the sustainable growth rate shows the impact of changes in reinvested profit attributable to balance sheet profit of BGN 1:

$$SGR^{bs}_{1} - SGR^{bs}_{usl.3} = 4{,}8453\ \% - 4{,}9256\ \% = -0{,}0803 \text{ пункта.}$$

As a result of the decrease of profit reinvested in the enterprise's business, which is attributable to balance sheet profit of BGN 1, by BGN 0,0131 (0,7998 – 0,8130), the sustainable growth rate has decreased by BGN 0,0803.

The overall effect of the four direct factors described above has resulted in increase of the enterprise's sustainable growth rate for the current year in comparison to the previous year by 0,2815 points [(−0,5971) + (+0,3233) + (+0,6355) + (−0,0803)], i.e. as much as the change of sustainable growth rate for the current year in comparison to the previous year.

4. Conclusion

These results show that it is possible, expedient and appropriate to bind various aspects of the enterprise's activity and its effectiveness, which has direct influence both on the internal growth rate and on the sustainable growth rate of the enterprise. Objectively existing relationships and dependencies exist among individual factors. The complex impact of factors measures the dynamics of the enterprise's internal and sustainable growth rates, respectively. The analysis and assessment of such impact is an objective prerequisite for making appropriate decisions in order to achieve and maintain sustainable growth and financial stability of the enterprise.

In the context of digitalization of enterprises and the economy, the questions for the use of the advanced information technology capabilities for complex analysis of the financial and business activity of the enterprises from different sectors of the Bulgarian economy is extremely relevant. Linking the input information that is necessary for the analysis in a common electronic system, setting out the methods and methodology for analysis of specific objects, the definition of objectively existing relationships between them, being expressed with the specific formulas of the analysis and analytical models, and the generation of output analytical information in real time, are an objective prerequisite for making effective management decisions for financial and market stability and for achieving and maintaining financial equilibrium of the enterprise.

References

1. Blagoev, D., 2019. Growth of Innovative Companies. Real Estates Magazine, Sofia, volume III (I), page 23-30]
2. Ivanova, R., 2014. Analysis and Assessment of Enterprise's Growth Rate, MATTEH Magazine, volume 1, Economy and Management Section, Shumen, page 325-332
3. Ivanova, R., 2019. Aspects of Analysis and Assessment of Enterprise's Growth Rate. Compendium from the International Research and Practical Conference on Bulgaria of Regions' 2019. Plovdiv, University of Agribusiness and Rural Development
4. Ivanova, R., 2020. Analysis of Return of Enterprise's Assets /according to the example of industrial enterprises/. Sofia, UNWE Publishing House
5. Ivanova, R., 2022. Enterprise's Growth Rate (Methodology and Methods for Analysis). Theoretical Workshop "Man – Measure of All Things? Challenges of Post-Industrial Information Society". Sofia. Technical University. Faculty of Economy. Department of Legal and Humanitarian Sciences
6. Yotsov, V., D. Zlatinov, P. Lukanova, N. Nestorov, Y. Hristoskov, 2020. Annual Report 2020 "Economic Development and Policies in Bulgaria: Assessments and Expectations"
7. Kasarova, V., 2010. Growth Rate as Subject of Financial Analysis. International Research Conference on Advanced-guard Scientific Tools in Management. Ravda, page 223-235
8. Mitov, K., R. Koleva, M. Gergova, L. Doncheva, 2012. Effectiveness and Economic Growth of Industrial Business Today. Dialogue Magazine, Scientific Research Institute Svishtov
9. Nenov T., B. Atanasov, 2000. The New Model of Economic Growth – Objective Necessity for Successful Business. year IX, 2000, book I, Institute of Economic Studies (IES)
10. Nenov, T., B. Atanasov, 2000. Effectiveness and Growth of Industrial Companies Today. Varna, 2000
11. Todorov, I., 2015. Economic Growth of Bulgaria during the Period of Monetary Council. Economic and Social Alternatives Magazine, Sofia, UNWE, no. 3
12. Amendment and Supplement Act to the Value Added Tax Act, promulgated in State Gazette no. 55/19.06.2020
13. investor.bg/sviat/6/a/krasimir-petrov-ikonomikata-shte-se-zabavia-inflaciiata-shte-se-uskoriava-342821

Empirical Study of Digital Services in Municipalities

Katia Kirilova

University of National and World Economy, Sofia, Bulgaria
katia.kirilova@unwe.bg

Abstract. The paper focuses on the study of digital services offered by municipalities. Its concretization requires a study of the digital services offered by the municipalities in the Sofia region. These are a total of 22 municipalities with varying degrees of social and economic development. The survey is conducted based on official data from national registers. Finally, conclusions and recommendations are outlined.

Keywords. Digital services, municipalities, public administration.

1. Theoretical background and research methodology

The methodology used in the research suggests the performance of research and analysis. The most significant of these is the comparative analysis based on data from official sources. The methodology involves the division of actions and tasks into conditional stages. The phasing helps a lot to clearly define the goals of each of the stages and the correct reporting of the achieved results. In some studies, such a methodology can be modified and significantly parameterized. This is necessary in view of the specifics of the relevant subject area and the available data. The proposed methodology includes the following sequence of steps (Fig. 1):

- Determining the objects for analysis.
- Identification of official data sources.
- Study of the characteristics of each of the municipalities.
- Research of the offered digital services.
- Comparison of the number of offered services.
- Formulation of conclusions.

The first stage is related to determining the objects that will be analyzed. In this case, these are the municipalities of the Sofia region. They are 22 in number and have varying degrees of economic development. The second stage identifies the official data sources that can be used in this study. The most up-to-date data on the number of digital services offered by each of the municipalities are available on the website www.egov.bg [4]. The third stage is related to the proposal of the public profiles of the municipalities and the set of digital services offered in them. The study of these services is related to the analysis of the main directions in which the services are offered. The following is a comparative analysis of the digital services offered between the individual municipalities. Defining a set of characteristics is a complicated and complex problem [1]. This requires special attention to be paid to the individual characteristics and their importance for achieving the overall efficiency of the services. The last stage is related to the definition of conclusions and recommendations for future research.

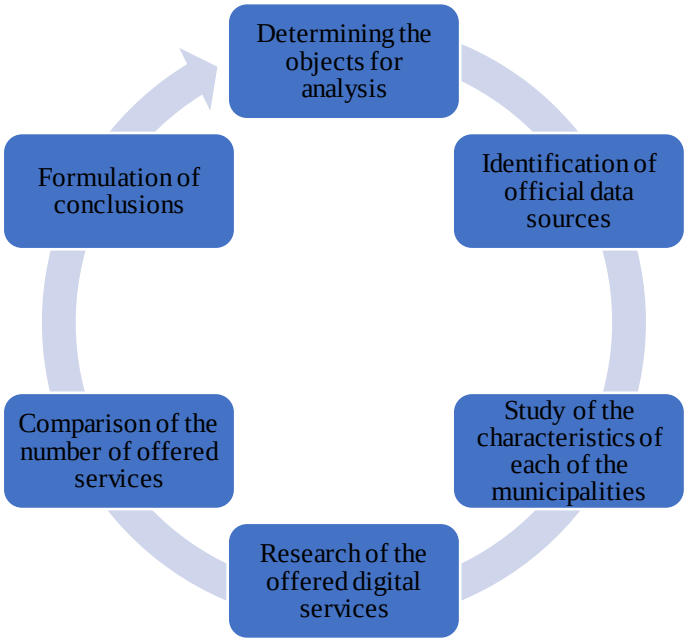


Fig. 1 Main stages of the proposed methodology

A very important element in the comparative analysis is the correct definition of the national data registers to be used. These issues are considered by several researchers, such as Kirilov[2]. Other authors emphasize that digital administrative services are a modern necessity for every society [3]. The level of development of e-government does not depend solely on the digital services offered. It should also consider the available technical infrastructure as well as the state of human capital.

2. Empirical analysis

The data on the digital services offered by the analyzed 22 municipalities are presented in Table 1 [4].

Table 1. Digital services by municipalities in Sofia region [4]

N	Name	Digital services (number)
1	Anton	135
2	Bozhurishte	132
3	Botevgrad	164
4	Godech	142
5	Gorna Malina	115
6	Dolna banya	132
7	Dragoman	136
8	Elin Pelin	122

9	Etropole	144
10	Zlatitsa	131
11	Koprivshtitsa	78
12	Kostenets	150
13	Kostinbrod	127
14	Mirkovo	121
15	Pirdop	149
16	Pravets	126
17	Samokov	138
18	Svoge	154
19	Slivnitsa	142
20	Stolichna obshtina	108
21	Chavdar	118
22	Chelopech	135

Fig. 2 also presents the graphic distribution of the offered digital services. As can be seen, some of the municipalities that offer the largest number of digital services are Botevgrad and Svoge.

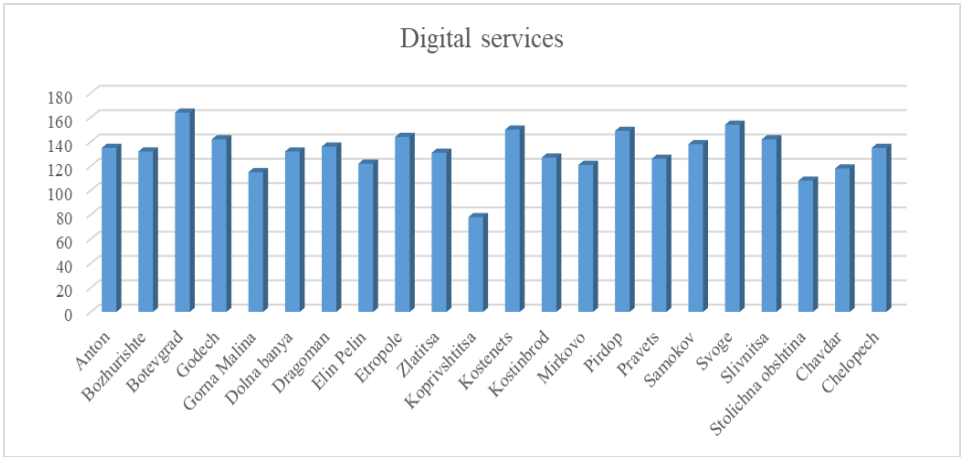


Fig. 2. Offered digital services

Of interest for the analysis is the top 5 municipalities in the Sofia region, which offer the largest number of digital services on the site www.egov.bg [4]. The graphic distribution of the municipalities in the top 5 is given in Fig. 3.

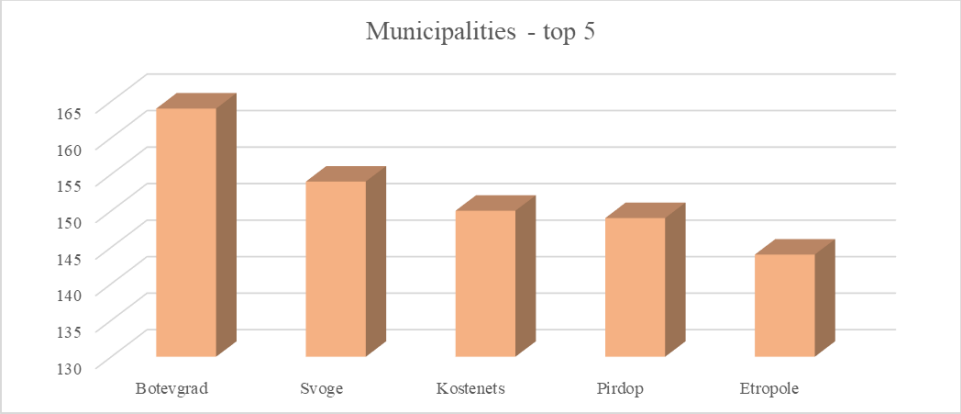


Fig. 3 Top 5 municipalities that offer the largest number of digital services

The study of the digital services provided by the municipalities shows that the three most common ways to get the finished results are:

- As an electronic document in the applicant's profile in the secure electronic service system.
- On paper at the counter in the administrative service center
- On paper through a licensed postal operator.
- These opportunities will be an extension in the future, which will undermine the improvement of services for citizens and businesses.

3. Conclusions

As a result of the research and empirical analysis, the following conclusions can be drawn:

- In recent years, the municipalities of the Sofia region have significantly developed the digital services they offer to citizens and businesses.
- The municipalities of Botevgrad, Svoge, Kostenets, Pirdop and Etropole offer the largest number of digital services.

Acknowledgement

This work was supported by the UNWE Research Programme (Research Grant No: NID NI-7/2020).

References

1. Kirilov, R., Features of application of software solutions in public projects risk management, Economic and Social Alternatives, Issue 2, 2017, ISSN 1314-6556.
2. Kirilov, R., Approaches for building information systems for monitoring the realization of students, Economic Alternatives, Issue 3, 2021, ISSN 1312-7462, https://doi.org/10.37075/EA.2021.3.09.
3. Milev, P., Application of data analytics for research of digital administrative services in Bulgaria, SHS Web of Conferences, Vol. 120, 2021, ISSN 2261-2424.
4. https://egov.bg/wps/portal/egov/nachalo

Characteristics and Preparation of Datasets for Machine Learning Algorithms

Genka Miteva, Alexandrina Murdjeva

University of National and World Economy, Sofia, Bulgaria
gmiteva@unwe.bg, amurdjeva@unwe.bg

Abstract. Data is an integral part of the prerequisites for applying machine learning algorithms correctly. These algorithms use data to learn, which makes the quality of datasets incredibly important for receiving accurate and useful results. Depending on the problem which needs to be solved, different algorithms can be used. After the appropriate dataset has been chosen, the data needs to be transformed into an acceptable input for the machine learning algorithm which has been selected. The paper provides a summary of the main characteristics of the datasets and offers their classification. The purpose of the proposed classification is to identify the main types of data and their relationship to ML algorithms. Finding the right match between the algorithms, the data required for them, and their basic characteristics will allow the search for opportunities to automate various tasks of data preparation and their potential interference in data storage and processing systems.

Key words. Data characteristics, datasets, data preparation, machine learning algorithms

1. Introduction

The process of learning is defined by acquiring new knowledge, skills or preferences or modifying existing ones. Learning comes naturally to humans, but that doesn't work for machines. Machine learning (ML) is a category of artificial intelligence that enables computers to learn on their own, or as the definition of Tom Mitchell states, machine learning is the study of computer algorithms that improve automatically through experience [1]. The idea behind it is learning through modifying actions to improve accuracy. Accuracy is measured in terms of the number of times the result from the action is correct.

The aim is to provide machines with the abilities of gathering data through senses similar to human ones and then processing the gathered data by using computational intelligence tools and machine learning methods to conduct predictions and make decisions at the same level as humans.

Data is incredibly important for the proper use of machine learning algorithms. Data must have a certain level of quality in order to receive accurate and useful results. Depending on the information problem which needs to be solved, different algorithms can be used. After the appropriate dataset has been chosen, the data needs to be transformed into an acceptable input for the machine learning algorithm which has been selected.

2. Data preparation

Algorithms use data to learn, and the quality of the data is very important for the final result. For machine learning, usually, data should be in the form of numbers, and some algorithms have more specific requirements. This is where data preparation comes into play. Usually, raw data cannot be used as-is. It needs to be cleared of errors and transformed into

proper input for the algorithm. Data preparation involves enhancing the data in an attempt to improve the performance of data analysis [2]. Data preparation comprises those techniques concerned with analyzing raw data so as to yield quality data, mainly including data collecting, data integration, data transformation, data cleaning, data reduction, and data discretization [3]. The tasks included in the process of data preparation usually differ from project to project, but some of the most frequently encountered ones are the following:

- Data Cleaning

Data cleaning deals with removing errors and inconsistencies from data to improve its quality. These errors can be related to misspellings, missing data, or invalid data. There are tools that provide functionalities for data cleaning, but a lot of the tasks have to be done manually.

- Feature Selection

A feature is an individual property of the object or process that we are observing which can be measured. ML can deal with data that has a very large number of features. Feature selection helps remove irrelevant features and improve efficiency while still providing good results. Feature selection uses methods such as filtering (a form of ranking the variables), wrapper methods (which uses a search algorithm to find a subset of variables that maximize the classification performance), embedded methods (feature selection as part of the training process), etc.

- Data Transformation

Usually, data transformation operations are organized in ETL environments. The main function of data transformation is applying a mathematical function to each value in a dataset to form a new value, according to the need of the ML algorithm. The process of data transformations is frequently integrated with other processes in a workflow. This task usually includes refinement, aggregation, or analysis.

- Feature Engineering

The main task of feature engineering is to construct suitable features to improve the performance of the ML algorithm. The difference between data transformation and feature engineering is that data transformation makes data understandable for the machine and does not create new features from existing ones. Feature engineering modifies features to create new ones which will make the ML algorithm more efficient.

- Dimensionality reduction

Dimensionality reduction transforms high-dimensional data into a meaningful representation of reduced dimensionality. The reduced representation still needs to account for the observed properties of the data.

The existence of a great variety of algorithms means that the proper algorithm needs to be selected according to the problem that needs to be solved. To select the most suitable algorithm, we need to know its logic and the cases in which it is used.

In practice, the selection process is based on comparing the result we are looking for with the result of the application of an algorithm. However, another important consideration when selecting an algorithm to work it is the data that we have available, what our capabilities are in terms of transforming it, etc. Based on this, we will try to make a classification of algorithms in terms of their data requirements. In order to do that, data characteristics should be defined.

3. Types of ML algorithms

There are different types of ML algorithms, but the most general classification is into 4 categories according to the applied approach:

- Supervised – the original definition of supervised learning can be interpreted as machine learning algorithms that use a set of examples that have the correct outputs [4]. In other words, supervised earning is a machine learning paradigm for acquiring the input-output relationship information of a system based on a given set of paired input-output training samples [5]. These examples are used as training sets. The algorithm learns to respond accurately by comparing outputs with those in the training set, where the data is labeled. This type of learning is mostly used when a prediction has to be made based on historical data. Classification, regression, and forecasting are supervised learning algorithms.
- Unsupervised – unsupervised learning aims to recognize existing patterns and structure which have not been discovered [6]. Based on these patterns, the algorithm creates rules. This type of ML algorithm is used when the categories of data are not known. The data is not labeled. Examples of such learning are Apriori algorithm, K-means.
- Semi-supervised – semi-supervised learning combines supervised and unsupervised learning. When some of the data is labeled, and some is not, semi-supervised learning is used.
- Reinforcement – reinforcement learning is used when the algorithm is only provided with a response that confirms if the output is correct or not. In other words, it uses observations gathered from the environment to take actions in order to maximize the reward and minimize error [7]. An example is Markov Decision Process algorithm.

According to the main topic of this paper – the data sets needed by ML algorithms, the key difference between supervised learning and unsupervised learning is supervised learning algorithms use a set of examples, which have the correct outputs. They use training sets that contain labeled data. Unsupervised learning is used when we don't know the patterns and structure of data – so there are no training sets used, and data is unlabeled.

4. Data characteristics

To correctly understand the relation between the data and ML algorithms, it is important to understand the nature of data and its main characteristics. Data for ML can be classified based on different criteria. Next are presented three possible classifications based on answers of the main questions: Can ML algorithms work only with numbers? Do we need to use only organized, structured data? And What should be the quality of the data?

Categorical and numerical data

- Numerical or quantitative data is a numerical measurement expressed in terms of numbers. It can be further broken down into two types.
  - Discrete data - discrete data usually represents items that can be counted. It can only take a certain number of values. These types of values are usually taken by counting.
  - Continuous data – continuous data has no restrictions on the values it can take. These types of values are usually taken by measuring.
- Categorical or qualitative data is a measurement expressed by means of a natural language description. It can sometimes be represented by numbers, but these numbers do not have a mathematical meaning.
  - Nominal data – it is used to name variables and is descriptive.

▪ Ordinal data is qualitative data that has some type of scale to it. It is usually categorized as qualitative data but exhibits both categorical and numerical data characteristics.

Traditional machine learning algorithms cannot work directly with categorical data – it should be transformed and somehow represented by numbers. That is why it is important to clearly define what numerical and categorical data are. With this in mind, the next important question is "How to 'translate' data into recognizable for ML format".

Structured, unstructured, and semi-structured data

Another type of classification is that data can be divided into structured, unstructured, or semi-structured.

- Structured data is typically categorized as quantitative data. It is highly organized and easily decipherable by ML algorithms. Usually, SQL (Structured Query Language) is used to manage structured data in a relational database. This type of data is easier to manipulate. It is typically easier to understand and interpret.
- Unstructured data is usually categorized as qualitative data and cannot be processed and analyzed via conventional data tools and methods. It does not have a predefined data model, so it is managed in non-relational databases. Another way to manage unstructured data is to use data lakes [8]. Such data includes text, mobile activity, sensor data. It can be collected quickly and easily but requires expertise and specialized tools.
- Semi-structured data can be described as the connection between structured and unstructured data. It does not have a predefined data model, just like unstructured data, but is easier to store. It uses metadata in the form of tags and semantic markers to scale the data into preset fields.

Traditional machine learning algorithms cannot use unstructured data directly. We need to find a way to represent the unstructured data using structured data.

Normalized and denormalized data

Data normalization is related to data cleaning. It makes data more cohesive and eliminates redundancy and inconsistencies so that the data can be stored in a database.

The main goal is to eliminate duplicate data within the data set. Another thing data normalization does is grouping data together logically. Sometimes, datasets have information conflicts and normalization helps resolve those as well. All in all, what normalizing data achieves is:

- Reduced data redundancy
- Organized sets of information
- Improved database flexibility
- More precise data analysis

Labelled and nonlabelled data

Another point of classification could be based on whether data is labelled or not. Data labelling refers to the process of adding tags or labels to raw data. These tags represent the class of objects that the data belongs to and help the machine learning model classify other, unlabeled objects.

Most practical ML models use supervised learning, which needs a labeled set of data, also called a training set, to learn from. The machine learning model uses labels provided by the human to train. After that, the model can be used on new data. Some common types of data labeling are:

• Computer vision – when building a computer vision system, images or pixels need to be labeled to generate the training set. It can be very basic, like labeling what the image contains, or it can be segmented at the pixel level.

• Natural language processing – natural language processing usually requires the text to be manually labeled to create a training dataset, for example, when we need to identify sentiment or intent

• Audio processing – audio processing converts different sounds into structured data, so it can be processed by machine learning algorithms. It often requires the sound to be manually transcribed into written text, to which we can add tags. Eventually, this becomes the training set.

5. Classifying ML algorithms according to data requirements

After going through the main characteristics of data for machine learning, we can try to classify ML algorithms in terms of their specific data requirements. This process can give us another point of view on choosing the most suitable algorithm.

The figure shows some commonly used algorithms from the fields of supervised and unsupervised learning, along with the problem it is usually used to solve, the type of algorithm and the type of learning it represents, and the type of data for this algorithm.

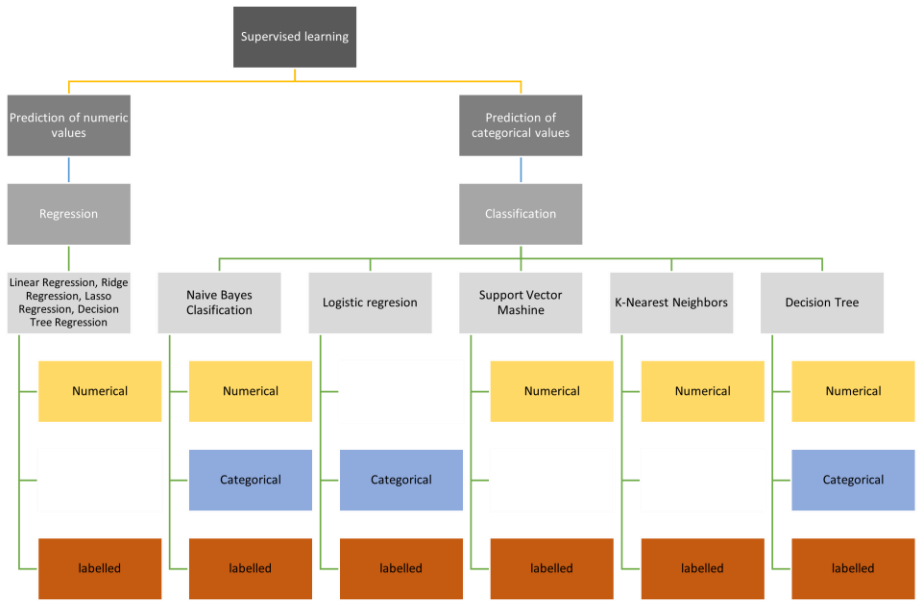


Fig. 1. Data in supervised learning

Regression models predict the output as a continuous variable. The most notable examples are:

• Linear regression, logistic regression – Linear regression is a classification algorithm for modeling the relationship between a continuous variable and one or multiple predictors. When the dependent variable is categorical, linear regression can be transformed into logistic regression. These are supervised learning algorithms, so it needs labelled data.

With classification algorithms, the output variable is a category.

• Linear SVM (Support vector machine), Kernel SVM – these are used to map nonlinearly separable functions into higher dimension linearly separable function. They can be used for both classification and regression problems. Linear SVM should be used when most dependent variables are numeric. It also needs labelled data as it is a supervised learning algorithm.

• K–means clustering aims to find groups in the data, with the number of groups represented by the variable K. It is an unsupervised learning algorithm, so it works with unlabeled data. It works with continuous numerical data.

• K-modes clustering is an unsupervised learning method that aims to divide the data points into a set number of groups. It is similar to K-means but uses categorical data.

There are algorithms that work with both numeric and categorical data. An example is the Decision tree algorithm, which is a supervised learning algorithm that is used for classification and regression. It predicts the value of the target variable by making decision based on simple rules.

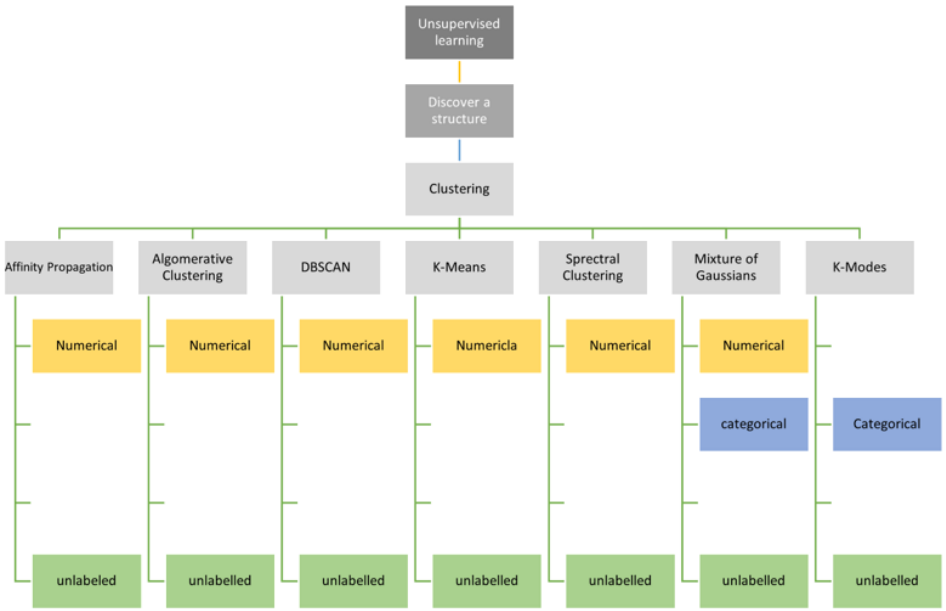


Fig.2. Data in unsupervised learning

6. Conclusion

Machine learning can provide us with valuable knowledge about the huge quantities of data available to us nowadays. However, in order to achieve the most accurate results possible, it is crucial to select the most appropriate machine learning algorithm for the problem at hand. This largely depends on the characteristics of the data we are working with. Furthermore, it is important to understand the specifics of each algorithm as they are greatly influenced by the datasets which are being used.

The proposed classifications show supervised and unsupervised learning algorithms, divided by the type of problem they solve and the type of algorithm, along with the type of

data they require to provide an accurate result. The proposed classification can be helpful when trying to automate the process of data preparation for machine learning.

References

1. Mitchell, Tom (1997). Machine Learning. New York: McGraw Hill. ISBN 0-07-042807-7. OCLC 36417892.
2. Lean Yu, Shouyang Wang, & Lai, K. (2006). An integrated data preparation scheme for neural network data analysis. IEEE Transactions On Knowledge And Data Engineering, 18(2), 217-230. doi: 10.1109/tkde.2006.22
3. Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. Applied Artificial Intelligence, 17(5-6), 375-381. doi: 10.1080/713827180
4. Stuart J. Russell, Peter Norvig (2010) Artificial Intelligence: A Modern Approach, Third Edition, Prentice Hall ISBN 9780136042594.
5. Liu, Q., & Wu, Y. (2012). Supervised Learning. Encyclopedia Of The Sciences Of Learning, 3243-3245. doi: 10.1007/978-1-4419-1428-6_451
6. Hinton, Geoffrey E. and Terrence J. Sejnowski. "Unsupervised learning : foundations of neural computation." (1999).
7. Kaelbling, Leslie P.; Littman, Michael L.; Moore, Andrew W. (1996). "Reinforcement Learning: A Survey". Journal of Artificial Intelligence Research. 4: 237–285 doi:10.1613/jair.301. S2CID 1708582. Archived from the original on 2001-11-20.
8. Structured vs. Unstructured Data: What's the Difference? | IBM. https://www.ibm.com/cloud/blog/structured-vs-unstructured-data

Challenges for the Electronic Management of Municipalities in the Southwestern Region

Katia Kirilova

University of National and World Economy, Sofia, Bulgaria
katia.kirilova@unwe.bg

Abstract. The study analyzes the degree of development of e-government in some municipalities in the Southwest region. The basis for these analyzes are the municipalities with a population of over 25,000 people. A study of the digital services is presented, as well as their characteristics. Finally, conclusions and recommendations are formulated.

Keywords. Digital services, municipalities, public administration.

1. Theoretical background

The Southwestern planning region includes municipalities from the districts of Sofia-city, Pernik, Kyustendil and Blagoevgrad. The region includes 48 towns and 902 villages.

Fig. 1 shows a map of the Southwestern region.

Fig. 1 Map of the municipalities in the Southwestern region [5]

Sofia-city has a determining place for the development of the whole region. The population of the region is 2,070,760 people. This is about 30.28% of the total population of the country [4]. This large percentage of the population also implies a greater interest in the development and use of electronic services and electronic management tools.

2. Research methodology

Performing comparative analysis is a complex and responsible task. This requires defining phases and stages in this activity. They cannot be used universally, and in each case the individual stages and tasks should be parameterized. Fig. 2 shows an exemplary sequence of stages and tasks comparing the degree of development of e-government in the municipalities of the Southwestern region. This is done based on the digital services offered by the individual administrations. The creation of an effective toolkit requires the prior establishment of design specifications for the individual software components of the services [1].

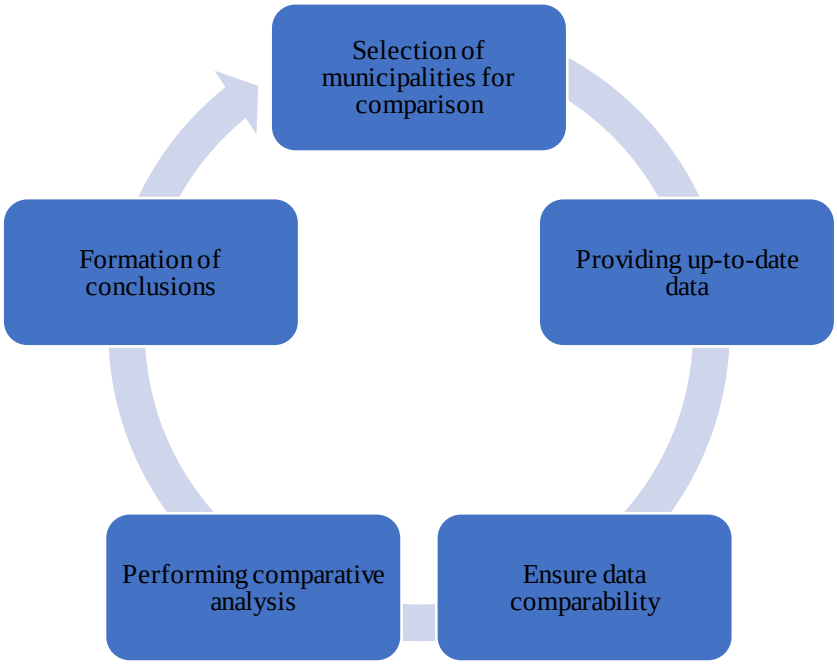


Fig. 2 Stages in performing the comparative analysis

The main stages are:

- Selection of municipalities for comparison.
- Providing up to date data.
- Ensure data comparability.
- Performing comparative analysis.
- Formation of conclusions.

The first stage is to determine the municipalities to be analyzed. In the specific study we have selected to analyze only the municipalities that have a population of over 25,000 people. These are the municipalities of Blagoevgrad, Petrich, Dupnitsa, Kyustendil, Pernik and Sofia. Some of these municipalities are in the forefront of another study in the subject area – a study of the media positioning of digital administrative services in Bulgaria [2]. The second stage is related to the provision of up-to-date data for the analysis. The main source of data is the government portal www.egov.bg [3]. Each of the municipalities is represented in it based

on the provided digital services. Ensuring the comparability of the data is very important for obtaining correct conclusions from the analyzes. Once the data is collected, a comparative analysis is performed. The last stage of the proposed methodology is the formulation of conclusions and recommendations.

3. Empirical study

The data for the selected 6 municipalities are shown in Table 1 [3], [4].

Table 1. Digital services by municipalities in Southwestern region [3], [4].

Municipality	Population	Digital services (number)
Blagoevgrad	67 810	142
Petrich	26 778	100
Dupnitsa	28 333	151
Kyustendil	37 799	164
Pernik	70 285	129
Sofia	1 221 172	107

As can be seen from the data in Table 1, Blagoevgrad municipality offers a total of 142 digital services. The services of Petrich municipality for 100. Both municipalities are in the same district of Blagoevgrad. The municipalities with a population of over 25,000 in Kyustendil district are Kyustendil municipality and Dupnitsa municipality. The digital services in the municipality of Dupnitsa are 151, and those in the municipality of Kyustendil are 164. It is also a leader in the provision of digital services by the analyzed municipalities in the Southwestern region. In Pernik district only Pernik municipality is analyzed, which offers 129 digital services. The digital services of Sofia municipality are 107. These data are presented graphically in Fig. 3.

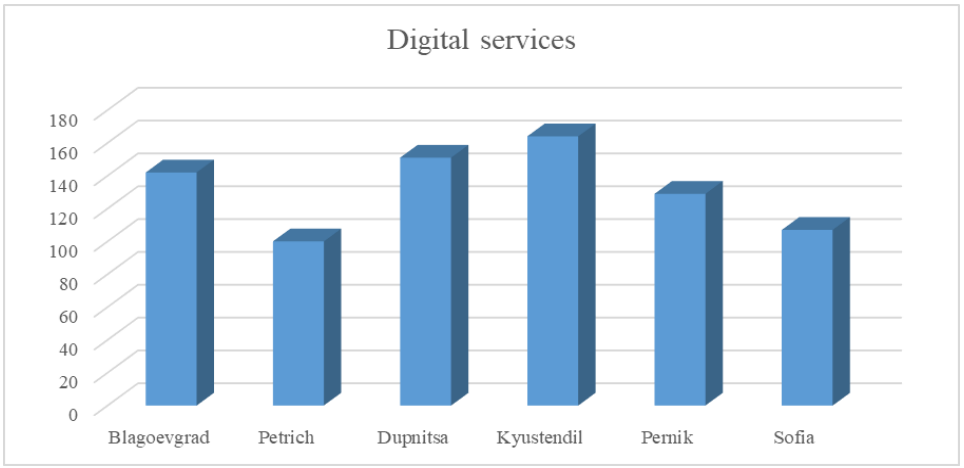


Fig. 3. Data for performing a comparative analysis

From the analyzed data the largest number of digital services on the site www.egov.bg [3] is 164 and is offered by the municipality of Kyustendil. The smallest number of digital services is 100 and is offered by the municipality of Petrich. This difference is significant given the population data. There are 37,799 people living in the municipality of Kyustendil and 26,778 in the municipality of Petrich. Similar differences are observed in the other municipalities. They are explained by the different usability of digital services by citizens and businesses, which determines their development.

4. Conclusions

As a result of the research and empirical analysis, the following conclusions can be drawn:

- There is no direct connection and dependence between the population of the analyzed municipalities and the number of digital services offered.
- Analyzes show that the municipality of Kyustendil offers the largest number of digital services.

Acknowledgement

This work was supported by the UNWE Research Programme (Research Grant No. NID NI-7/2020).

References

1. Kirilov, R., Approaches for building information systems for monitoring the realization of students, Economic Alternatives, Issue 3, 2021, ISSN 1312-7462, https://doi.org/10.37075/EA.2021.3.09.
2. Milev, P., Application of data analytics for research of digital administrative services in Bulgaria, SHS Web of Conferences, Vol. 120, 2021, ISSN 2261-2424.
3. https://egov.bg/wps/portal/egov/nachalo
4. https://nsi.bg/
5. https://bg.wikipedia.org/wiki/%D0%AE%D0%B3%D0%BE%D0%B7%D0%B0%D0%BF%D0%B0%D0%B4%D0%B5%D0%BD_%D1%80%D0%B0%D0%B9%D0%BE%D0%BD_%D0%B7%D0%B0_%D0%BF%D0%BB%D0%B0%D0%BD%D0%B8%D1%80%D0%B0%D0%BD%D0%B5

Organization of Data in Data Lake – Real-Life Practice

Eva Karkalikova, Alexandrina Murdjeva

University of National and World Economy, Sofia, Bulgaria
e.karkalikova@unwe.bg, amurdjeva@unwe.bg

Abstract: One of the biggest advantages of the data lake is its ability to store many different types of data - structured, unstructured, or semi-structured. It allows the flexibility of the hierarchical namespace, which means, that it represents to the users a nice possible way, in which they can organize their files. Because the data lake can support massive amounts of data, it is necessary for them to be very well classified, since the modern world, is collecting and storing data for almost everything. And therefore, for its correct ingestion, storage, and further processing, it needs to be standardized. In the following paper will be reviewed some of the most common questions, which appear in our minds, when talking about this kind of repository – like what it is used for, main conception and why it is better than the well-known Datawarehouse. Of course, the best practices will be mentioned also and a real-life example, representing the Azure Data Lake, as well.

Keywords: Data lake, organization, structure, files, data, best practice, real-life, Datawarehouse, conception, example.

1. Introduction

In the modern world almost, everything is generating data – from the smartphones trough IoT devices, different company systems etc. For curious human being, is very important to understand as much as possible the surrounding world and try to find hidden patterns in the data itself. By these new relationships between objects can be found, that are not directly connected one to each other, and for example if businesspeople understand the data within our company, they can make more reliable decisions and make the right choices.
The traditional Data Warehouse fits perfectly for the general idea of collecting information, but because of data's every day fast increasing volume from all different kind of sources, it is not capable for storing everything in different formats. For this reason, another target is needed, where the different formats of huge volume data can be stored. Here comes the role of the Data Lake. And beside which approach for data storage the company will choose, it needs to be sure, that in each moment the stored information is in secure and safe place, from where it can be used in the right way.

2. What Is a Data Lake?

A data lake[1] refers to a massively scalable storage repository that holds a vast amount of raw data in its native format until it is needed plus processing systems (engine) that can ingest data without compromising the data structure (Laskowski, 2016). The data lakes are typically built to handle large and quickly arriving volumes of unstructured data (in contrast to data warehouses' highly structured data) from which further insights are derived. Thus, the lakes use dynamic (not pre-build static like in data warehouses) analytical applications. The data in the lake becomes accessible as soon as it is created (again in contrast to data warehouses designed for slowly changing data).

It can store structured, unstructured data, or semi-structured [2], which means data can be kept in a more flexible format for future use.

A data lake is capable of store and analyzes petabyte-size files and trillions of objects and offers high availability, durability, and reliability.

In the modern world more and more data are coming straight from the IoT devices – from smartwatches and different kind of sensors, up to smart home voice controller and even light bulbs. The storing of the data becomes a challenge, because of its various formats, intensive streaming, and almost unlimited size.

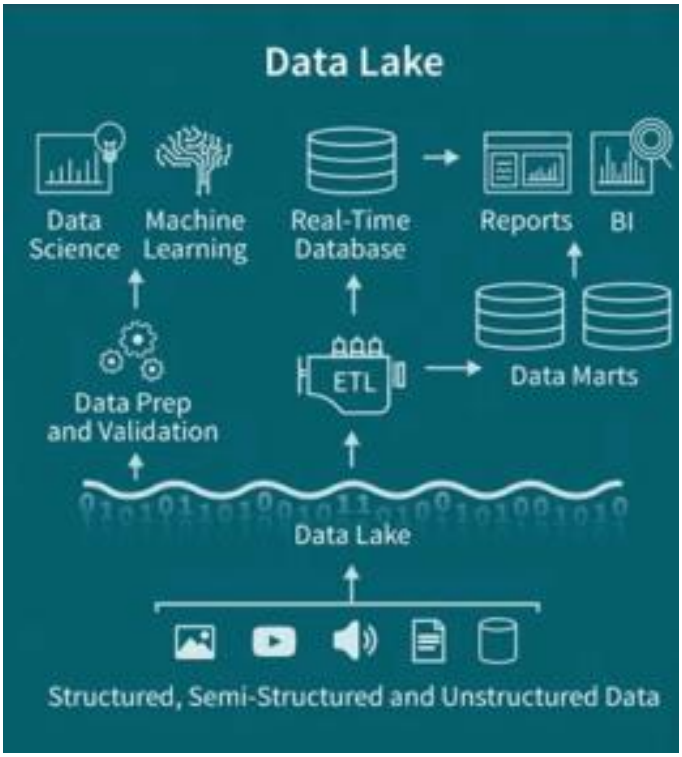


Fig. 1. Data Lake's place within Data workflow

In the above picture [3] is represented where is the data lake is positioned in the data workflow. Incoming data, landing in the data lake and from there it can be processed in different ways – for data science and machine learning, through data marts and the presentation of the data in different reports.

3. Data lake vs Data warehouse

The need of collecting data from different kinds of sources is not something new. Data warehouse solution has already taken place in the development and ETL processes. But since the volume of the data every single day gets bigger and bigger, the technology world is starting to talk about big data. Sadly, data warehouses are not so flexible and have limited sizes to offer such storage space, as needed. In general data, lakes and data warehouses are often confused, because as they have reconciliation points, they have some key differences, that are going to be listed below.

The distinction is important because they serve different purposes and while a data lake works for one company, a data warehouse will be a better fit for another.

3.1. Data structure: raw vs. processed

One of the big differences is that data lakes are capable to store unprocessed raw data (which can be in a variety of formats), while the data warehouse has some strict rules and must hold only clean data. One of the advantages of the data lake is that machine learning can be applied directly to the raw data because this kind of information fits best for such kind of algorithms. The risk of all that raw data, however, is that data lakes sometimes become data swamps without appropriate data quality and data governance measures in place. From financial perspective, the data warehouse has the benefit of storing only fully processed data, which is ready for creating meaningful reports and can be understood more easily. And by this criterion is cheaper, than the Data Lake, which, from the other side can store unprocessed data in raw format, on which can be applied Machine learning algorithms.

3.2. Purpose: undetermined vs in-use

As mentioned above, the data inside Datawarehouse has been processed, which means, that it had been trough some business rules and now is ready to be presented (with the right tools) to the target audience, which most of the times is the business. It does not need to have any additional information, that is not necessary, like inside the Data Lakes. In them raw data are stored and it's not sure when and if they will used at all, but it is a good idea to store it just in case.

3.3. Users: data scientists vs business professionals

If the data lake has a big volume of data and it is unorganized or are missing rules for logical separation of the data, it will be very difficult for all users to navigate within it. So, this makes the data usable only for data scientists and people, who are very close to the ingestion of the data and know its dataflow steps. By the other hand, data stored inside data warehouse is structured and easy to navigate and be used in presentation tools, charts, tables, etc. Because of this, it can be used by the business.

3.4. Accessibility: flexible vs secure

Data lake architecture has no structure and is therefore easy to access and easy to change. Plus, any changes that are made to the data can be done quickly since data lakes have very few limitations. Data warehouses are, by design, more structured and from this is coming the complicity of applying any changes. It is well known that their structure is inflexible, and a small improvement in its architecture can lead to a problem, because it may need many further modifications on other objects. This can be compared with house of cards – if one card is pulled in a wrong direction, the whole structure can fall apart.

4. How and why to organize Data Lake files

One of the benefits of the Data Lake is that it has very flexible structure – folders can be added and removed at any time. This often modifications do not mean, that the data lake can be a big mess – contrariwise this can be an interesting direction to think on how everything can be kept in the right place, so it is possible to navigate easier trough the huge amount of data in different formats and facilitate the further processing of the data. Data Lakes can be

considered as single repositories. However, data lake provides the flexibility to divide them into separate layers. According to Lingaro Group [4] can distinguished 3-5 layers that can be applied to most cases.

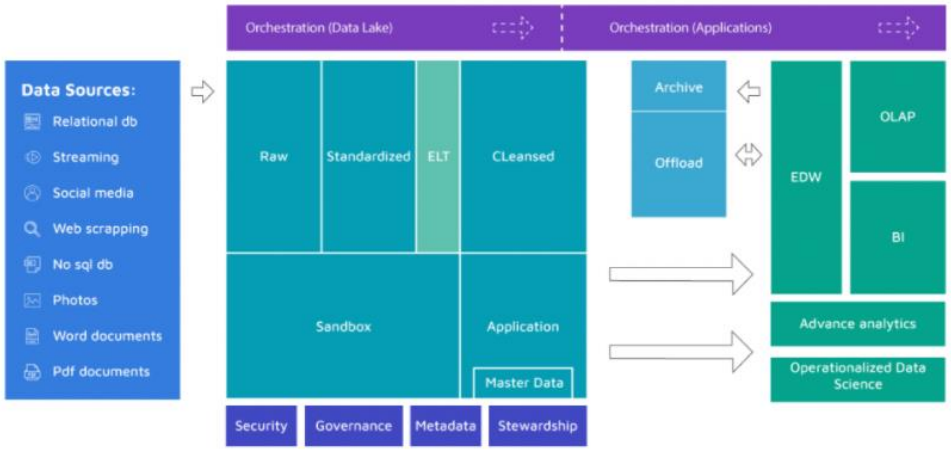


Fig. 2. Data orchestration inside data lake

These layers, represented in the above picture, are: Raw, Standardized, Cleansed, Application and Sandbox

- **Standardized and Sandbox** - In most of the cases Standardized and Sandbox are not necessary to be included in the development solution.
- **Raw data layer**[5] - Here is the landing zone of the incoming data. Its purpose is to store the information in its native format, without any cleansing of the data or any kind of enrichment. It should be optimized for fast and smooth data recording. Since many files from different sources are being collected, it is a good practice to have some folder structure organization inside. One of the best things, that can be implemented is to have a main folder of the area, followed by the data source/object/year/month. This layer can be compared with the Datawarehouse staging area.
- **Standardized data layer** - here the data can be pre-formatted, so it can best fit to the curation process logic. The structure must be the same as in raw. Even if there is a need, data can be driven in a lower stage of granularity. As mentioned earlier, it is optional.
- **Cleansed data layer** - mostly known as Curation layer. Here data is being enriched and saved in tables or files. In this layer the purpose of the data is well known,

which makes the logic and the whole processing logic to be very precise, to have quality output information, that will be used further. It can be said that this is the most complex part of the data lake.

- **Application data layer** - in some places can be named as Golden, Production, Trusted or Secure layer. It is used for Machine learning purposes. The data needs to be in the same structure, as in cleaned layer.

- **Sandbox data layer** - mainly data scientists are working with data from this stage. They are looking for some hidden patterns and correlations inside the data and 'play' with it with many different scenarios.

5. Best Practices for Data Lake Organization

In general, data lakes have the potential to significantly improve performance. But to make this happen, some simple [5] rules should be followed, offered by Upsolver team [6], to unlock the full power of the data lake.

5.1. Make several copies of the data

Another benefit of the data lake is that it offers the users almost unlimited size for storing data at a cheap price. This makes it a preferred choice when there is a need for holding big amounts of raw data. In the above section is mentioned, that the raw data is good for Machine Learning, but it has another benefit - it is like a backup for debug and recovery procedures when there is data loss further in the process.

5.2. Set a retention policy

It is good to store raw data, but this will be not forever. From time to time is necessary to revise all of it and delete those files, that are too old (the definition of "old" can be different for each organization, depending on its needs and purposes). By doing this the volume of the stored information and the price can be reduced.

5.3. Understand the data you're bringing in

During the data collection process there is no question like are all data needed, does it need to be stored in a specific location, or to create and organize the files in some hierarchical structure with easy navigation? These questions are simple at first sight, but if you start digging them you will see, that they are whole topics, to be carefully discussed. Answering these questions must be done before developing an ETL process, because after clearing the above points, you can save a lot of work, trying to save unnecessary information, that later, will be deleted and not used at all.

5.4. Partition your data

Partitioning the data can be compared to the "Merge small files" best practice. Its final goal is to make the data easy and fast to reach, find the needed information and further process it. It improves performance, by limiting the amount of query engines and sources, that are used for the runtime on the cluster. The most common way, that is applied to that when it comes to partition, is to do it by timestamp (e.g. year > month>day)

5.5. Readable file formats

The most optimized format, which is best for data lake reading and writing is the Apache Parquet. Many analytic tools have the ability to read such kind of data, because it is open-source.

5.6. Merge small files

Some files, like logs, are generated every day, or even they are coming each minute/second from different streaming sources. When storing the data, wise advice is to try

to merge these small files into one to allow the data lake to access, loop, and find the needed information much faster, rather if they were in thousands of separated mini-files.

5.7. Data governance and access control

When there is a massive amount of data, it must be guaranteed need, that it is accessed only by authorized users. Otherwise, not only the information is put in danger but also the whole company. Data Lakes provide a nice control model, named RBAC (Role-Based Access Control), which gives the ability to grant access to the data only on trusted accounts. This eliminates the risk of data getting corrupted or exposed to non-authorized persons.

6. A Real life Example

In this part an idea will be presented of organization of data lake, which was implemented in a working solution for a client based of data lake technology, provided by Azure. The goal was to provide structured and as much clear, as it can be, data for reporting purposes. To achieve this, data from nearly 30 sources must be collected. With this huge number of sources, in order not to mess up the data and to become from data lake a data swamp, it was crucial to think about a well-organized structure, in which the data will remain stable, reliable and for sure in the right place.

The main goal is the provide structure to data in Data Lakes that will not only organize them but also is well aligned to possible future use of this data – to make this data to be easily used by next transformation process which are responsible to provide quality data for business analysis.

The technological aspects of the projects were: data received as zip archives and each name was composed by the product category name and a timestamp., via SFTP servers, processed with the help of azure data factory,

The general approach is to store data in their original format and to store data prepared (in some scale) for further use.

The main arguments for applying this approach are:

1. To have separated levels for each stage of the data during its processing.
2. To navigate and easily find the needed information.
3. To perform some regular checks on the quality of the data (e.g. volume in the beginning in the process must match with the volume at the end point – there is no data loss during the whole operation).
4. To have cleaned, enriched and well-structured data, that will be understandable and meaningful for the business.

To achieve this whole idea, the following folder structure was developed (Fig. 3).

Figure 3. represents the concept structure, not the actual names of folders. There are 2 main levels – **sources** and **applications.** Since one and the same source (represented with "Source A" in Figure 3), which was providing data for most of these countries, decision was to have a structure from 3-4 levels and put as a next level from the source layer – **Country across** folder. And after this the **source's name** ("Source A"), followed by the Country name ("Country A"). Inside the country, there are the following layers:

- **Input** - was our landing zone – Every ingested archive is positioned here. Only data engineers had access to this layer.
- **Raw** - In the next step the archive is transferred, unzip and stored under raw in the respective product category folder, again, depending on the file name. A new folder, version folder, is created with name "Version "and the timestamp from the archives name. In this

folder the filed are stored. And because it is transactional data, there are fact data and some typical dimensions like market data, period, product dimension and etc. Only data engineers had access to this layer.

• **Curated** – from raw, next step is data to be enriched and saved in curated layer. The main processes, that are applied in this stage is to clean a little bit the data (remove strange symbols from the text values (if any), cast columns to the correct and needed data type etc.), compose separate columns for year and month, extracted from the full date one, add column with the timestamp, and at the end save it as a parquet file type, partitioned by the version, by following a naming convention of 2 parts – "version=" and the timestamp, which comes from the archive name (in raw). This is done in order to have some tracking and transparency of the processed data. Only data engineers had access to this layer.

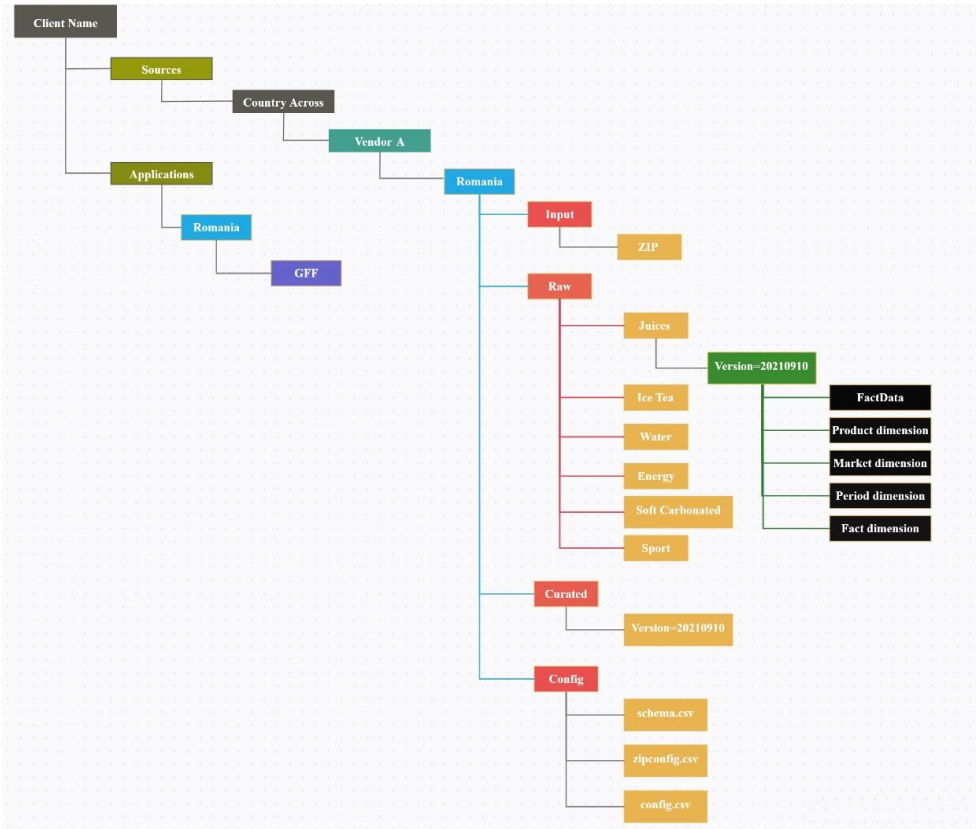


Fig. 3. The concept of folder structure, not the actual names of folders

• **Config** – this folder was used, to apply some rules and check the quality of the incoming data (without applying any modifications on it). This logic was implemented right after the raw was settled in its folder. So, for example – if in very important columns like year, or product volume there are missing values or even there are duplicates, or wrong file names or even missing files, the process stops, and the file did not get further to curated. By doing this some unnecessary usage compute power and cluster costs are reduced, because the datasets are very large sizes like 61 GB. And it does not make any sense if something is broken to

continue with the process. The results would be not correct. Which is also another best practice. Only data engineers had access to this layer.

- **GFF** (General Flat File) –Golden data, are data which was consumed by the power bi reports. They were embedded in one application. They are enriched the curated data by adding some columns with different calculations, or with additional information, aggregating or even disaggregating to lower level of granularity the data, etc. After Curated the parquet files were generated in the GFF folder. The business and data engineers had access to this layer.

7. Conclusion

In summary, in the practical example the general conception for the layers in Data Lakes is implemented and applied some of the best practices. And because of the very well-organized folders inside the data lake, there is no mismatching of the data, and the reports are depending on cleared, structured and reliable information. By this, the business can rule better and make the right decisions and remain competitive on the market.

References

1. Big Data, Fast Data and Data Lake Concepts (Natalia Miloslavskaya, Alexander Tolstoy; 2016) https://www.sciencedirect.com/science/article/pii/S1877050916316957
2. Data Lake vs Data Warehouse (n.d.; n.d.) https://www.talend.com/resources/data-lake-vs-data-warehouse/
3. What is a LakeHouse (Ben Lorica, Michael Armbrust, Ali Ghodsi, Reynold Xin, Matei Zaharia; 30.01.20), https://databricks.com/blog/2020/01/30/what-is-a-data-lakehouse.html
4. Data Lake Architecture: How to Create a Well Designed Data Lake (Pawel Mitrus; 05.06.2021) https://lingarogroup.com/blog/data-lake-architecture/
5. Data Lakes for Big Data [Report on Existing Data Lakes] (Abdelghny Orogat; March 2021), https://www.researchgate.net/publication/350382496_Data_Lakes_for_Big_Data_Report_on_Existing_Data_Lakes
6. 7 Best Practices for High-performance Data Lakes (Eran Levy; 19.12.2019) https://www.upsolver.com/blog/best-practices-high-performance-data-lake

Risks in Auditing Financial Statements Presented in the European Single Electronic Format

Iavor Bachev

University of National and World Economy, Sofia, Bulgaria
IavorBachev@gmail.com

Abstract. Public companies in the EU should, starting from 1 January 2020, mandatorily prepare their financial statements in accordance with the European Single Electronic Format (ESEF). As an EU Member State, we are obliged to comply with the European law which requires, effective from that date, listed companies to report their financial statements in the European Single Electronic Format (ESEF). This article will attempt to define risks in auditing the annual financial statements presented in electronic format.

Keywords. Format, audit, financial statements, ESEF.

1. Introduction

The introduction of the requirement for ESEF results from the desire to enhance accessibility, analysis and comparability of the information included in the annual financial reports of issuers of securities. The EU has provided for the regulatory technical standard on the European Single Electronic Format (ESEF) to be adopted in two stages. Until 2020, only the four single statements of the annual financial statements have to be converted in machine-readable language, and until 2022, this should also include the notes to the financial statements.

The adoption of Directive 2004/109/EC on the harmonisation of transparency requirements in relation to information about issuers whose securities are admitted to trading on a regulated market, amended by Directive 2013/50/EU,[1] marked the beginning of this process.

2. ESEF and impact on Bulgaria

In the Republic of Bulgaria, the requirement for use of ESEF is set out in §42 of the Transitional and Final Provisions of the Act to Amend and Supplement the Public Offering of Securities Act (Promulgated - SG No 42 of 2016). It specifies that after 1 January 2020, the issuer shall prepare the annual financial report under Article 100m in a unified electronic format of reporting, provided that the European Commission adopts the regulatory technical standards for specification of such format [4].

Supplementing Directive 2004/109/EC, in 2018, the Commission adopted Delegated Regulation (EU) 2019/815 on the European Single Electronic Format (ESEF). The Regulation sets out common rules on the format of the annual financial statements of issuers and more specific rules on marking up the single financial statements. These rules highlight more clearly the advantages of ESEF, as the combination of the XHTML format with the iXBRL mark-ups makes the annual financial reports to be both human-readable and machine-readable. This makes possible to process automatically larger volume of information which would facilitate the work of regulators and enhance exchange of financial information.

To assist users, in 2020, the Commission published Interpretative Communication on the preparation, audit and publication of the financial statements included in the annual financial reports drawn up in accordance with Commission Delegated Regulation (EU) 2019/815 on the European Single Electronic Format (ESEF) [6].

The Regulation [2] defines the following mandatory requirements for issuers:

- First: When drawing up and disclosing their annual financial reports, issuers shall use the eXtensible HyperText Markup Language (XHTML) format;

- Second: When drawing up their consolidated financial statements in accordance with International Financial Reporting Standards (IFRS), issuers shall mark-up those consolidated financial statements using inline eXtensible Business Reporting Language (iXBRL) and block-tag the notes to the financial statements.

- Third: iXBRL mark-ups shall comply with the ESEF Taxonomy, which is included in the Annexes to the ESEF Regulation and developed based on the IFRS Taxonomy published by the IASB. The Taxonomy does not define the content and presentation of the financial statements but defines the specific codes which may be used to identify (mark-up) the information included in the financial statements.

The Regulation requires public companies to use strictly specified computer languages and electronic forms. These are: HTML (Hyper Text Mark-up Language), XML (eXtensible Mark-up Language), XHTML, XBRL (eXtensible Business Reporting Language) and iXBRL (In Line XBRL). HTML is a standardized set of computer mark-up languages which are used for creating and structuring the elements of design of a website. XML is a computer language which codes a set of rules for creating documents in a way that makes them both human- and machine-readable. XHTML is a combination of XML and HTML which enables the HTML to be extended by using XML but preserves the functionalities and the backwards compatibility of HTML. XBRL is based on XML computer mark-up language used to report business information in a structure which is in line with the way in which companies usually structure and report business information, while iXBRL combines XBRL and XHTML so that reported data is machine-readable and also human-readable. The single report package comprises several electronic files: a file of the report itself (financial statements, management report and responsibility statements of relevant persons) in XHTML format and files with the taxonomy used. In its Manual, ESMA [7] encourages issuers to adopt naming convention which match {base} -{date} .zip. The {base} component of the filename should indicate the LEI of the issuer or the issuer's name (or an abbreviation of it); it should be of no more than 20 characters in length. The {date} component of the filename should indicate the ending date of the reporting period of reference. The {date} component should follow the YYYY-MM-DD format

The first risk in auditing the financial statements of public companies relates to the use of inappropriate computer languages and electronic forms. Issues are obliged to use strictly specified electronic forms which requires the use of specific software tools. The use of software tools is absolutely essential as the new requirements cannot be implemented without the use of technology. This involves the availability of specialized platforms which are licensed and can operate with these formats. An issue would arise if their number is limited or they lack the necessary authorization. Users would opt for those software solutions which allow quick and easy conversion of reports in the European Single Electronic Format, as well as ongoing update in line with changes in the taxonomy.

Another risk faced by auditors when using the ESEF relates to the provision of an audit opinion on whether the financial statements comply with the relevant statutory requirements provided for in the ESEF Regulation. The reason for the issue is that the ESEF Regulation does not contain specific requirements on the audit of financial statements prepared in ESEF. The predominant view is that statutory auditors should provide an audit opinion on whether the financial statements included in the annual financial reports comply with the relevant statutory requirements laid down in the Regulation [6]. The key argument is considered to be that in the audit report the statutory auditor expresses, where appropriate, an opinion on whether the financial statements comply with statutory requirements. As the requirements of the ESEF Regulation are legal requirements imposed at EU level, they are also statutory requirements. Therefore, statutory auditors are required to check the compliance of the financial statements with all applicable statutory requirements, including the relevant statutory requirements set out in the ESEF Regulation. The expression of such an opinion ensures that financial statements users will be equally protected, regardless of whether the financial statements are provided on paper or in ESEF.

In auditing, there is also a risk of inappropriate inclusion and disclosure of the audit opinion on whether the financial statements comply with the relevant statutory requirements provided for in the ESEF Regulation. The audit opinion on whether the financial statements comply with the relevant statutory requirements set out in the ESEF Regulation shall be included in the auditor's report. Under Article 28 of the Audit Directive, statutory auditors shall present the results of their audit in an audit report. Article 28(2)(c) specifies that the audit report shall include an audit opinion on whether the financial statements give a true and fair view in accordance with the relevant financial reporting framework, as well as an opinion on whether the financial statements comply with the applicable "statutory requirements" [5]. In this case, the "statutory requirements" correspond to the "relevant statutory requirements" laid down in the ESEF Regulation.

The ESEF Regulation does not specify whether the audit report should be included in the annual financial report or disclosed together with the annual financial report as a separate document. The Regulation allows for this issue to be resolved in accordance with the rules at national level, considering Art.4(4) of Directive 2004/109/EC which specifies that the audit report "shall be disclosed in full to the public together with the annual financial report" [1]. In the absence of specific rules in force at national level and taking into consideration the technical possibilities on a case-by-case basis, issuers may apply their preferred option.

There is also a risk of providing inappropriate audit opinion. To provide their opinion, auditors should check the compliance of issuers' financial statements with the provisions of the ESEF Regulation. The auditor must check if all the financial statements that are included in the annual financial report are prepared in a valid XHTML format and all consolidated financial statements that are drawn up in accordance with IFRS are marked-up using the XBRL mark-up language and the relevant taxonomy. The auditor shall obtain reasonable assurance that the financial statements, including relevant disclosures, are marked-up, where necessary, in all material respects, in accordance with the requirements of ESEF. To achieve this, the auditor would go through the usual stages of the audit methodology, such as engagement acceptance, planning and risk assessment, testing and obtaining evidence and reporting.

When accepting the engagement, the auditor must be aware of the software solutions used to convert the statements in the European Single Electronic Format and the taxonomy used. The auditor's approach would depend on whether the financial reporting forms in the ESEF format are available for the auditor still at an early stage of the audit or the company first provides to the auditor, for the purposes of financial statements audit, forms which have not yet been prepared in accordance with ESEF requirements and, later on, provides the

electronic report complying with the ESEF requirements. According to the guidelines of CEAOB, [8] in the first situation, the auditor can audit directly the content of the human-readable layer of the financial statements, since it needs to be the same as any other format of the financial statements (e.g., pdf, word). In this case, the auditor is not required to perform any additional audit procedures for the reconciliation between various formats of the financial statements. In the second situation, the auditor needs to perform reconciliation between the human-readable financial statements presented in the final electronic report and the initial information provided by the entity, on which the auditor performed the audit procedures.

When planning and determining the nature, timing and scope of the procedures to be performed and when evaluating whether the marked-up information is free from material misstatements, the auditor should use an appropriate materiality. According to the guidelines of CEAOB,[8] for the statutory audit purposes, the auditor should define the materiality in reference to the level of misstatement that could reasonably be expected to influence economic decisions taken by the users of the financial statements.

The auditor should identify and evaluate the risks of material misstatements in respect of marked-up information prepared by the issuer. Risk assessment will be based on the understanding of the process, put in place by the issuer to produce the marked-up information, and the understanding of the internal controls implemented by the issuer. Risks associated with the marked-up information may encompass completeness (omitted marked-up figures in the statements; omitted marked-up disclosures; required mark-ups relating to the identification of the entity are omitted) and accuracy (the marked-up information does not correspond with the financial statements; errors in extending taxonomy; misrepresentation of the accounting meaning of the number or disclosure as a result of marking-up by an inappropriate element of the core taxonomy).

After having identified risks, the auditor should define procedures in response to assessed risks. The guidelines of CEAOB recommend for the auditor to select one or more of the following procedures: using knowledge of the ESEF requirements, including the taxonomy, the auditor inspects the issuer's mark-ups, including related anchoring, and assesses whether the mark-ups are appropriate; develop an independent expectation of the appropriate mark-up and anchoring, if applicable, and compare the results to the issuer's marked-up financial statements as the basis for their substantive testing.

The conclusion of the auditor on whether the marked-up information complies with the ESEF requirements will be based on the results of the procedures performed. In cases where the mark-ups are materially misstated, the auditor should express a qualified or adverse opinion regarding this compliance. When the auditor is unable to obtain sufficient appropriate evidence in this regard, the auditor should express disclaimer of opinion on the compliance.

3. Conclusion

The advantage in using a single electronic format is that it allows to report and present business facts in a structured way so that they can be read both by users and by machines. The uniform structure of reports assists users in analysing data and making management decisions. Under the new rules, public companies will continue to prepare their financial statements complying with the provisions of IASs/IFRSs, but will have to present them in XHTML format. In case a company prepares consolidated annual financial statements, they will have to be converted and presented in XBRL format. Converting of the report in XBRL format is made by software tools which mark-up individual items in financial documents in accordance with the applicable taxonomy. Certain risks exist which will make auditors face challenges and possibilities of mistakes being made.

References

1. Directive 2004/109/EC (2004) available at: https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=celex%3A32004L0109. Last accessed on 14 September 2020.
2. Delegated Regulation (EU) 2019/815 (2019) available at: https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=CELEX:32019R0815. Last accessed on 14 September 2020.
3. Directive 2013/50/EU (2013) available at https://eur-lex.europa.eu/legal-content/bg/TXT/?uri=CELEX%3A32013L0050. Last accessed on 14 September 2020.
4. Act to Amend and Supplement the Public Offering of Securities Act (Promulgated -SG No 42). (2016).
5. The Audit Directive (2014) available at https://eur-lex.europa.eu/legal-content/BG/TXT/PDF/?uri=CELEX:32014R0537&from=EN. Last accessed on 15 September 2020.
6. Commission Interpretative Communication on the preparation, audit and publication of the financial statements included in the annual financial reports drawn up in accordance with Commission Delegated Regulation (EU) 2019/815 on the European Single Electronic Format (ESEF) (2019) Available at: https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=CELEX:52020XC1110. Last accessed on 15 September 2020.
7. ESMA ESEF Reporting Manual Preparation of Annual Financial Reports in Inline XBRL, 09 July 2020 | ESMA32-60-254rev mailbox: esef@esma.europa.eu.
8. CEAOB guidelines on the auditors' involvement on listed companies' financial statements in European Single Electronic Format. https://www.cposa.bg/attachments/article/349/Nasoki_KEONO_GFO_el_format_2020_03_25_final.pdf Last accessed on 15 September 2020.
9. Bachev,I. Edinniya evropeyski elektronen format na finansovite otcheti-predopredelenost I predizvikatelstvo za oditorite,Schetovodstvo I diggitalizatsya,sbornik s nauchni dokladi, Izdatelski kompleks-UNSS, 2021
10. Bachev,I. Dokladvane na nefinansova informatsiya-predlozhenia na EK za promyana I vazmozhni problem na transponiraneto na direktivata v Bulgariya,kolektivna monografiya na FSF,Ikonomikata na Bulgariya I ES, Izdatelski kompleks-UNSS, 2022

Reference Model for Cybersecurity in the Design of Online Services in Higher Education Institutions in Bulgaria

Elitsa Pavlova

University of National and World Economy, Sofia, Bulgaria
epavlova@unwe.bg

Abstract. The report examines the challenges of digital transformation and the creation of a reference model for cybersecurity in the design of online services in higher education institutions in Bulgaria. Cybersecurity has not yet been implemented in all IT processes and the main reason for this is the unique nature of the security systems of educational institutions. The reference model covers the standards COBIT 19, ITIL, ISO / IEC 27001, NIST CSF, CIS 20 and aims to minimize duplication of processes and security controls. The model has been tested in a web-based open source platform, which allows the choice between different standards. The platform can be used as a tool for a catalog of services, their audit and preparation of compliance reports.

Keywords. Reference model, Cybersecurity, Information security, Digital transformation, Access control, University, Bulgaria.

1. Introduction

The digital transformation in higher education institutions (HEIs) and research organizations is related to the use of new technologies, the organizational structure and the processes that ensure the connection between the goals of the university and its policy. Information systems and online services are an integral part of higher education, and incidents of disruption pose a serious threat.

Universities often take a reactive approach to cybersecurity, taking action only when a malfunction is detected or an attack occurs. A study on cybersecurity shows that in 2021 higher education institutions have modernized their existing IT systems by marking elements of compliance checklists without incorporating cybersecurity into their new systems and services [6].

The reliability of network and information systems is key to the smooth operation of all platforms and university sites, and therefore cybersecurity should be embedded in the early stages of planning, not as a process to mitigate a cyber-attack or recover from such.

The report examines the need for a cybersecurity model covering the design stages in order to improve the quality of education and identify key areas for cybersecurity. The reference model contains leading international approaches and standards, such as a framework for the control of information and related technologies (COBIT 19), a library of IT infrastructure (ITIL), a standard for information security management (ISO / IEC 27001), a framework for cybersecurity. National Institute of Standards and Technology (NIST CSF) and Controls for Effective Cyber Security (CIS 20). The methodology of the reference model considers the integration of both security solutions and security-related software solutions. The model identifies potential sustainability solutions based on the mission and descriptions of information systems and stakeholders, such as system administrators and cybersecurity experts. The model is limited to the design stage of electronic services, without going into details about their implementation and improvement. Security requirements include threat

modeling and analysis, which are the starting point for selecting appropriate countermeasures. Data and research related to cybersecurity in higher education in the European Union (EU) were used.

2. Description of the reference model

Directive 2016/1148 (NIS) is the first document adopted by the European Union for the protection of networks and information systems. It aims to address this need by presenting "measures to achieve a high common level of security of the Union's network and information systems in order to improve the functioning of the internal market". Article 14 (1) requires Member States to ensure that the electronic services offered have appropriate technical and organizational measures to manage the security risks of information systems. Paragraph 2 states that these measures must prevent and minimize the impact of security incidents and ensure the continuity of these services.

A report by the European Union Agency for Cybersecurity on minimum security measures for digital services compares security objectives with well-known standards, national frameworks and certification schemes [7]. It also publishes a set of guidelines that further describe the incident notification process imposed on digital service providers under Article 16 of the Directive.

In the context of cybersecurity, most research in higher education is done on risk management and data security, while little attention is paid to IT management and online service design. The importance of this problem in the education sector has increased with the growing online learning. Universities' online systems are the main target of attacks because their network includes confidential personal data about students and faculty, as well as rich scientific and intellectual property. Preventing cyber-attacks is not possible without special attention to administrative systems, as they require a secure configuration. The most common security risks include forgery of requests between sites, denial of service or falsification of credentials. A study on cyber risks shows that educational organizations have different priorities than other industries [8]. They focus more on the security of their endpoints (82%), pre-designed systems (79%) and virtual infrastructures (73%). Areas that are neglected include: the use of personal devices (49%), unstructured data stored in third-party data centers (33%), and employee activity (17%). The IT infrastructures of educational organizations are often in several buildings and the acquisition of centralized control over the operations and activities of the IT department is essential for maintaining security. One of the main problems is that each department or sector is responsible for the storage, processing and protection of their data, and employees often violate security measures. Managers treat cyber risks as a technical problem and delegate it to the IT department, and they are part of risk management. Some universities are introducing new cybersecurity protocols and checklists every day, but these efforts often lead to an unnecessary focus on formal compliance rather than real sustainability. The main methods of risk management within the standards and good practices are ISO / IEC 27001, COBIT 19, NIST.

The requirements to the reference model are based on a number of standards described in the article "Cybersecurity Design" [4]. These include minimizing the area of attack, default security, least privilege principle, in-depth protection, proven design models and protected components, security and privacy documentation. Mandatory parameters are a two-way link between stakeholders and the protection of identification data. Two technologies are key to introducing and enhancing cybersecurity: single sign-on, which reduces the number of passwords required, and multi-factor authentication, which adds additional security features. Inadequately protected internet network for access to university resources is a problem. In many cases, administrators do not apply basic "good IT hygiene", such as changing the default

passwords of network equipment. Some risk situations arise due to lack of rules, and in other cases, existing procedures are not applied.

Creating a cybersecurity model to meet the specific needs of a higher education institution is a long and complex process. Each university collects and processes different types of data, it is responsible for them and strives to protect their confidentiality. The security of information systems is becoming increasingly difficult to manage, so the principles of cybersecurity design must be at the heart of the life cycle of all online processes and services. Risk mitigation security controls can be configured to protect against certain types of threats. All controls must work together to provide critical and sensitive data to information systems.

According to the World Economic Forum, cybercrime costs the world economy $ 2.9 million every minute [13]. It is estimated that 80% of these attacks are targeted at passwords. The Authentication 2020 conference discusses growing cyber threats and the importance of access control in all learning and administrative processes [1]. Students, faculty, and staff need secure access to information resources and applications at all times, and from any device. A review of the websites of a number of leading universities (Auckland [10] and Cambridge [2]) shows that single sign-on is used successfully to access the most commonly used software applications and university services (eg teaching portals, staff self-service and shared disks). This creates a new set of security challenges that higher education institutions must meet.

3. Architecture of the reference model

The reference model covers the functions and controls established in COBIT 19, ITIL, ISO / IEC 27001, NIST CSF and CIS 20, focusing on the integration of cybersecurity in the stages of designing online services in higher education. It is based on the COBIT methodology, because this approach includes the way the organization organizes and controls the processes related to information technology. The framework identifies which information criteria are important for IT processes, as well as which resources (people, applications, information and infrastructure) are included in them.

The researched reference model consists of categories, objects, profiles, processes, harmonized with the leading standards, which allow the university to achieve its goals in the field of cybersecurity.

The categories are performed depending on the risk profile of the information system and are applicable to both the design of online services and risk management. They are part of the NIST CSF framework and to each of them there are referenced references to documents from other standards and approaches, for specific controls and processes for cybersecurity. The model has five areas for cybersecurity:

1. Access control
2. Awareness and training
3. Data security
4. Information protection
5. Technology Protection

Sites are information systems that are related to information security policies, user authentication profiles and methods, asset management and control. The main university information systems are shown in the table below.

Table 1. Basic university information systems, subsystems / services

IT systems	Subsystems / Services
Administrative systems	Profiles and authentication Archiving and storage Information security Communication and cooperation
IT systems	Networks and infrastructure Servers and data Web development and hosting Hardware, software and applications
Learning activity	Awareness and training
Science	Research

Accounts are a unique alignment of administrative roles with access control and can be used to improve cybersecurity by comparing a "current" to a "target" profile.

The processes are related to the design of online services - input and output documents, service requirements, information flow diagrams, structures and databases. Implementing a model for designing online services or modifying existing ones in a consistent and uniform manner requires the establishment and maintenance of common system-specific procedures, inventory, base configuration, and common secure configurations.

In order to unify the services, it is proposed to devide them into groups so that in one group there should be services with similar basic configurations for design and requirements for cybersecurity. Application development lifecycle procedures describe how basic configurations are used to manage and change within the organizational process and how to identify unauthorized change within the system. Having a clearly defined process for evaluating and approving change requests ensures that each proposed and accepted change is evaluated in a consistent and repetitive way from a security perspective.

4. Areas of cybersecurity

The functioning of the university is based on the collection, processing and sharing of information, and information security ensures the optimization of security measures of all systems and administrative processes at the university. The effectiveness of the mechanisms for protection of information assets in the administrative units depends on a systematic approach, which includes management of all elements of administration. The University Risk Management study says that identifying information threats and managing risk are important elements of cybersecurity [9].

The principles of the EU Cyber Security Systems Design Framework include a good understanding of the system's functions, identifying its shortcomings and analyzing the architecture during design, reviewing code during programming and testing penetration testing before launching the system. On the other hand, the NIST framework focuses on security-related processes that can be integrated into the existing system development lifecycle. The reference model has five areas of cybersecurity: access control, awareness and training, data security, information protection and technology protection.

Access Control

Access control covers all credentials that are issued, managed, verified, revoked and audited for authorized devices, users and processes. The controls related to access control affect all information systems and have the highest level of priority. This area has four sub-

categories - Credentials, Physical Access to Assets, Remote Access and Access Permissions. The table below shows each sub-category and its reference documents.

Table 2. Access control, by categories and reference documents

Sub category	Reference documents
Identification data	CIS CSC 1
	COBIT 19 DSS05.04, DSS06.03
	ISO/IEC 27001:2013 A.9
Physical access to assets	CIS CSC 12
	COBIT 19 DSS01.04, DSS05.05
	ISO/IEC 27001:2013 A.11
Remote access	COBIT 19 APO13.01, DSS01.04, DSS05.03
	ISO/IEC 27001:2013 A.6, A.13
	NIST CSF PR.AC-3
Access permissions	CIS CSC 3, 5, 12, 14, 15, 16, 18
	ISO/IEC 27001:2013 A.6, A.9
	NIST CSF PR.AC-4

Awareness and training

Awareness and training is an area that determines whether university professors and staff receive cybersecurity training and can perform their duties and responsibilities in accordance with relevant policies, procedures and agreements. This area has three sub-categories - Security Training, Eligible Users and Cyber Security Officers and Third Party Stakeholders. The table below shows under the Awareness and Training category and the reference documents to it.

Table 3. Awareness and training, by categories and reference documents

Sub category	Reference documents
Security training	CIS CSC 17, 18
	COBIT 19 APO07.03, BAI05.07
	ISO/IEC 27001:2013 A.7
Privileged users and employees	CIS CSC 5, 17, 18
	COBIT 19 APO07.02, APO07.03, DSS06.03
	ISO/IEC 27001:2013 A.6, A.7
	NIST CSF PR.AT-2
Stakeholders from third countries	CIS CSC 5, 17, 18
	COBIT 19 APO07.03, APO10.04, APO10.05
	ISO/IEC 27001:2013 A.6, A.7
	NIST CSF PR.AT-3

Data security

The data is managed in accordance with the university's risk strategy for protecting the confidentiality, integrity and availability of information. This area has five sub-categories - data security, asset management, data protection, software and information integrity testing

and development and testing environment. The table below shows all sub-categories and reference documents.

Table 4. Data security, by sub-categories and reference documents

Sub category	Reference documents
Data security	CIS CSC 13, 14
	COBIT 19 APO01.06, BAI02.01, BAI06.01, DSS06.06
	ISO/IEC 27001:2013 A.8
Asset management	CIS CSC 1
	COBIT 19 BAI09.03
	ISO/IEC 27001:2013 A.8, A.11
	NIST CSF PR.DS-3
Data protection	CIS CSC 1, 2, 13
	COBIT 19 APO01.06
	ISO/IEC 27001:2013 A.6, A.7, A.8, A.9, A.13, A.14
Check the integrity of software and information	CIS CSC 2, 3
	ISO/IEC 27001:2013 A.12, A.14
	NIST CSF PR.DS-6
Development and testing environment	CIS CSC 18, 20
	COBIT 19 BAI07.04
	ISO/IEC 27001:2013 A.12

Information protection

Security policies are maintained and used to protect information systems and assets. This area has seven sub-categories: basic information systems configuration, systems management lifecycle, configuration change control processes, information archiving, physical work environment provisions, security and cybersecurity processes in human resource practices. The table below shows each sub-category and its reference documents.

Table 5. Information protection, by categories and reference documents

Sub category	Reference documents
Basic configuration of information systems	CIS CSC 3, 9, 11
	COBIT 19 BAI10.01, BAI10.02, BAI10.03, BAI10.05
	ISO/IEC 27001:2013 A.12, A.14
Systems management life cycle	CIS CSC 18
	COBIT 19 APO13.01
	ISO/IEC 27001:2013 A.6, A.14
	NIST CSF PR.IP-2
	ITIL V3 2011 SDP1.1, SDP1.2, SDP1.3, SDP1.4, SDP1.5, SDP1.6, SDP1.7, SDP-2.1-11
Configuration change control processes	CIS CSC 3, 11
	COBIT 19 BAI06.01, BAI01.06
	ISO/IEC 27001:2013 A.12, A.14
	ITIL V3 2011 SDP-1.1-7 SDP-2.1-11
Archiving information	CIS CSC 10

	COBIT 19 APO13.01
	ISO/IEC 27001:2013 A.12, A.17, A.18
Physical work environment	COBIT 19 DSS01.04, DSS05.05
	ISO/IEC 27001:2013 A.11
Protection processes	COBIT 19 APO11.06, DSS04.05
	NIST CSF PR.IP-7
Cybersecurity for human resources	COBIT 19 APO07.01, APO07.02, APO07.03, APO07.04, APO07.05
	ISO/IEC 27001:2013 A.7, A.8

Security technology

Technical security solutions are managed to ensure the security and resilience of systems and assets, in accordance with relevant policies, procedures and agreements. This area has four sub-categories. The table below shows each sub-category and its reference documents.

Table 6. Security technology, by categories and reference documents

Sub-category	Reference documents
Portable devices	ISO/IEC 27001:2013 A.8, A.11
	NIST CSF PR.PT-2, PR.PT-3, PR.PT-4
	CIS CSC 3, 11, 14
Configuring systems	COBIT 19 DSS05.02
	ISO/IEC 27001:2013 A.9
	CIS CSC 8, 12, 15
Communication networks	COBIT 19 DSS05.02, APO13.01
	ISO/IEC 27001:2013 A.13
	CIS CSC 9, 14, 15, 18
Network integrity	NIST CSF PR.AC-5

5. Approbation of the reference model

When choosing a platform to use to verify and validate the reference model, the most popular open source content management systems are considered. The literature review includes software solutions for creating different types of websites - WordPress, TYPO3, Joomla!, Drupal, Contao, Neos, WooCommerce, OpenCart, AbanteCart, PrestaShop [3]. The specific functionalities in the areas of installation and configuration, user management, services, content creation, data filtering, possibilities for changes, improvements and extensions of the source code and adaptation of the system to the individual requirements for design and security are compared. WordPress is the most popular system, but software updates related to significant administrative efforts are mentioned as a disadvantage. WooCommerce is a universal set of tools that can be integrated into any WordPress site. The downside is that it relies heavily on paid plug-ins for advanced features and customization of service attributes. TYPO3 is a great solution for large e-commerce portals, but its implementation and administration requires expertise. Drupal offers a wide range of extensions for building complex domain structures [11]. An article on the Hongqiat website ranks OpenCart first among online trading platforms. A comparative analysis between WordPress and OpenCart shows that each solution presented is based on software that has a decisive influence on the orientation of the project [5].

Based on the functional requirements for the system, the OpenCart platform was chosen. It is designed as a modular content management system that allows easy management of services. It is possible to create multiple administrator roles and different access rights so that an administrator can access only the sections of the platform that he needs. Add an unlimited number of services, create links to a selected category or subcategory (eg information systems, cybersecurity), service status, related services, etc. Add groups of attributes (eg standards) and attributes (eg cybersecurity controls). Ability to filter and sort services by specific functions (eg service criticality). Adding information pages (eg general information on the reference model, standards and approaches, etc.). The platform has different interfaces - Template management, module management, automatic image sizing, environment and reporting, editing services, reports and statistics, SSL security. The admin panel has an overview of all services, users, the latest edits, data saving and recovery tools, open source editing code and extensive documentation.

The reference model is published at: https://csservices.online. The main buttons in the site menu are: Home, Reference Model, Standards, Information Systems, Cybersecurity and Web Services. Information pages, including a description of the reference model, the process of working with it, as well as information about the referenced standards.

Designing an online service includes:

• Identification and determination of the subject of the service;

• Clarification of the type of source document and regulatory requirements for the service, as well as users;

• Construction of information flows;

• Determining which information systems the future service will work with and which system components should be included.

• The main areas of cybersecurity for the specific type of service are reviewed and the security controls from the reference documents are included.

• Determining the status of the service (active, inactive, in the process of updating, in the process of testing), which changes during development.

In the verification of the reference model, software expertise, dynamic and formal method were used. Software expertise is applied to any property of the web platform, as well as to any stage of its development. The main advantage of this type of verification is the high efficiency, which depends on the experience and motivation of the verifiers. The dynamic method analyzes and evaluates the properties of the web platform based on the results of real work. The main disadvantages of this method are that it allows to find only errors that occur during operation, and its implementation requires the creation of testing systems. Formal verification proves the correctness of the software. Compared to software expertise and dynamic methods, it is the most effective and reliable method of verification, but its application requires highly qualified specialists to implement it.

The platform has been tested in all modern browsers for desktops Firefox, Chrome, Opera, IE, Edge and Safari (Mac, Windows and Linux), as well as for phones and tablets (eg iOS Safari on iPhone / iPad, Chrome and Firefox on iPhone / iPad / Android) [12]. Limitations in validation are related to the use of various operating systems and protocols that are installed by customers.

6. Conclusions

There are many international standards, but none of them provides the full set of mechanisms and procedures for IT management and cybersecurity in universities. Universities are working on quantitative economic analysis in the field of cybersecurity, but are hampered by a lack of data on the number and impact of cyber-attacks. They invest in security controls

and processes that overlap with basic controls, and this often leads to increased complexity and security gaps. A major problem is that security controls are applied equally to all data and functions, without prioritizing the protection of critical data. Combining several standards for IT management and cybersecurity can help find effective and secure security controls to be used in developing a reference model for service design.

The proposed reference model allows for a choice between individual standards and related security controls. The model can be used as an audit tool to provide guidance on the implementation of specific security policies and the preparation of compliance reports. Its main characteristics are improving the communication and exchange of information between the various structural units, creating conditions for the rapid introduction of new online services, standardized criteria for analysis and increasing cybersecurity in universities in Bulgaria.

References

1. Authenticatecon. "Authenticate 2020," 2020. https://authenticatecon.com/.
2. Cambridge College. "Single-Sign-On (SSO)." *Cambridge College Edu*, 2020. https://www.cambridgecollege.edu/single-sign-sso.
3. comments, 30 Jun 2014 Nitish Tiwari Feed 634up 25. "3 Open Source Content Management Systems Compared." *Opensource.Com*. Accessed February 24, 2022. https://opensource.com/business/14/6/open-source-cms-joomla-wordpress-drupal.
4. Eaton. "Cybersecurity Design Principles," n.d. https://www.eaton.com/us/en-us/company/news-insights/cybersecurity/cybersecurity-design-principles.html.
5. D, Waqas. "OpenCart Vs. WordPress - Who Is the Winner?" *Fullstop*, May 14, 2020. https://fullstop360.com/blog/opencart-vs-wordpress-who-is-the-winner/.
6. Dave Burg, Mike Mason, and Richard Watson. "Cybersecurity: How Do You Rise above the Waves of a Perfect Storm?" *EY*, 2021. https://www.ey.com/en_bg/cybersecurity/cybersecurity-how-do-you-rise-above-the-waves-of-a-perfect-storm.
7. Markopoulou, Dimitra, Vagelis Papakonstantinou, and Paul de Hert. "The New EU Cybersecurity Framework: The NIS Directive, ENISA's Role and the General Data Protection Regulation." *Computer Law & Security Review* 35, no. 6 (November 1, 2019): 105336. doi:10.1016/j.clsr.2019.06.007.
8. Netwrix. "Top Cybersecurity Risks in Education." *Netwrix*, 2021. https://blog.netwrix.com/2017/09/05/infographics-top-cybersecurity-risks-in-education/.
9. Research Gate. "Risk Management in Universities," 2020. https://www.researchgate.net/publication/321746840_Risk_Management_in_Universities.
10. The University of Auckland. "What Is Single-Sign-On?" *Www.Auckland.Ac.Nz*, 2021. https://www.auckland.ac.nz/en/about-us/about-the-university/identity-and-access-management/single-sign-on/details.html.
11. Hongkiat. "Top 10 Open-source e-Commerce Shopping Carts," September 22, 2021. https://www.hongkiat.com/blog/opensource-ecommerce-shopping-carts/.
12. "Web Testing," n.d. https://app.eu-central-1.saucelabs.com/live/web-testing.
13. World Economic Forum. "Forgotten Your Password?" *World Economic Forum*, 2021. https://www.weforum.org/press/2020/01/forgotten-your-password-not-having-one-will-make-you-safer-says-world-economic-forum/.

Analysis of Serverless Cloud Data Warehouse Solutions

Geno Stefanov

University of National and World Economy, Sofia, Bulgaria
genostefanov@unwe.bg

Abstract. The traditional Data Warehouse solutions and architectures are well known. Building and running a traditional Data Warehouse, central repository of data coming from one or more data sources, has always been complicated and expensive task. Scaling is another tough activity in the Big data era. It is hard to leverage between slow performance and investing in costly upgrade processes. On the other hand, the cloud (serverless) Data Warehousing solutions are getting more and more market share and popularity. This is no accident, the serverless Data Warehouse handles most complexities, such as reliability, security, efficiency, and costs optimizations, and provide a consumption-based billing model for their usage. The paper aims at analysing the benefits and drawbacks of a serverless Data Warehouse solution and architecture.

Keywords. Serverless DWH, Cloud computing, Big Data, Cloud ETL.

1. Introduction

A data warehouse can be described as an integrated central repository for all the data that an organization consolidates from various sources – data which can then be accessed and analysed to make adequate decisions. This data could be from multiple data streams, the IoT (Internet of Things), transactional databases, and internet. More specifically, "a Data Warehouse is a subject oriented, integrated, non-volatile and time variant collection of data in support of management's decisions" [1]. Thus, a DW represents an integrated collection of historical business data, supporting business analysts in the decision-making process. In that sense, the DW represents the "corporate memory" of all business data that is relevant for strategic decisions [2].

The traditional data warehouse architectures usually consist of three layers - data source layer, data warehouse layer and analytical layer (see fig. 1). Data warehouses are constantly growing in size and complexity as a business evolves. This growth means that the system resources required to effectively run the data warehouse increase over time. The requirements for a data warehouse is storage, processing, memory and input-output capabilities. These requirements are grouped into the following categories: software requirements, security requirements and hardware requirements.

As the traditional data warehouses are prone to inflexible storage capacity, technical difficulties, and high operational overhead due to hardware maintenance needs, many businesses are moving their data warehousing to the cloud.

The goal of this paper is to analyse the cloud based serverless Data Warehouse solutions in terms of their pros and cons compared to the traditional on premise architecture.

The rest of the paper is organized as follows: Section 2 introduces the cloud based Data Warehouses, as well as presenting and analysing the main vendors of cloud Data Warehouses. In Section 3 an analysis for cloud based vs traditional on premise Data Warehouse solution is presented and explained. Finally, in section 4, some conclusions and future work are inferred.

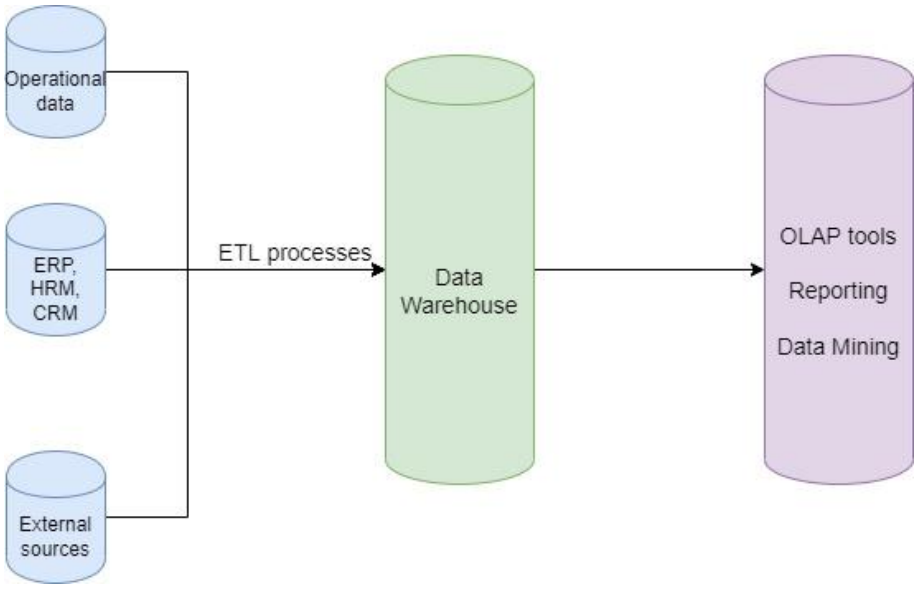


Fig.1 Data warehouse architecture vision of creating web applications

2. Data warehousing in the cloud

A cloud data warehouse uses the space and compute power allocated by a cloud provider to integrate and store data from disparate data sources for analytical querying and reporting. Cloud data warehouse provides SAS (storage as a service). The traditional data warehouse software, hardware and infrastructure configuration is taking a lot of time and requires skilled professionals. It also takes a lot of time to optimize and manage the system. Data warehouse scaling up is not an easy and is resource exhaustive task. In the cloud based Data warehousing the scaling can be achieved effectively and easily. Using cloud, data can be scaled up or down instantly without any hassle on the other hand.

Cloud data warehouse are efficient for analytics because they use columnar storage and massively parallel processing (MPP) which allows better performance for running complex queries. Cloud data warehouse allows enterprises to move their focus from system management to analysis whereas traditional data warehouse focuses on data management. Cloud data warehouse provides cost benefits by eradicating hefty upfront costs. Using cloud data warehouse there are no hardware, server rooms, IT-related staffing issues, or operational expenses to maintain your DWH. Cloud data warehouse reduces the cost and complexity of managing on-premises systems so customers can focus on extracting value from their data rather than on maintaining hardware and software infrastructure

Amazon Redshift

In November 2012, Amazon Web Services (AWS) launched Redshift - "Amazon Redshift is a fast, fully managed, petabyte-scale data warehouse service that makes it simple and cost-effective to efficiently analyze all your data using your existing business intelligence tools"[3]. This can be considered as the beginning of the cloud data warehousing.

Amazon Redshift enables SQL-querying of exabytes of structured, semi-structured, and unstructured data across the data warehouse, operational data stores, and a data lake with the possibility to further aggregate data with big data analytics and ML(Machine Learning) services.

The Redshift architecture is presented in fig. 2 and is built around two types of nodes: a single Leader Node and a cluster of Compute Nodes. This solution runs on a SQL platform, based on PostgreSQL and is built to integrate with Amazon Web Services as well as Amazon S3 (Redshift). It is available in the PaaS model and is based on Massively Parallel Processing.

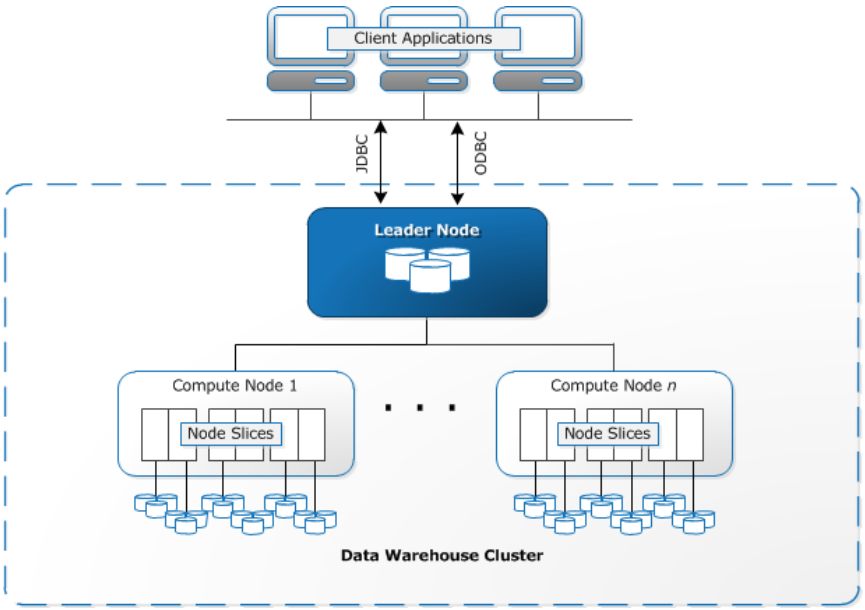


Fig. 2. Amazon Redshift data warehouse architecture

Google BigQuery

BigQuery allows for cost-effective exabyte-scale storage with tables having up to 10,000 columns. It's most effective when main analytical queries either filter data according to partitioning or clustering or require scanning the entire dataset.

BigQuery's serverless architecture decouples storage and compute and allows them to scale independently on demand. This structure offers both immense flexibility and cost controls for customers because they don't need to keep their expensive compute resources up and running all the time. This is very different from traditional node-based cloud data warehouse solutions or on-premise massively parallel processing (MPP) systems.

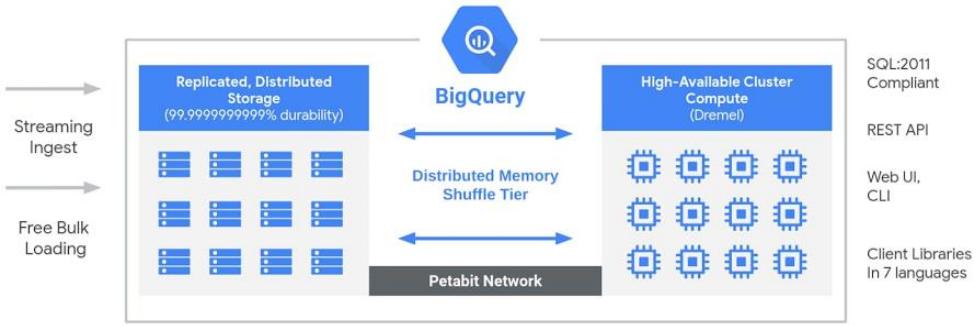


Fig. 3. BigQuery architecture

Azure Synapse Analytics

Azure Synapse Analytics "is a limitless analytics service that brings together data integration, enterprise data warehousing, and big data analytics"[4].

As enterprise data warehousing, Azure Synapse Analytics is good for integrating data from hundreds of data sources across the company's divisions, subsidiaries, etc. for analytical querying to be performed in seconds. Reporting on all management levels, from C-suite to directors, managers and supervisors, is protected with a fine-grained data access control.

Synapse SQL uses a node-based architecture. Applications connect and issue T-SQL commands to a Control node, which is the single point of entry for Synapse SQL.

The Azure Synapse SQL Control node utilizes a distributed query engine to optimize queries for parallel processing, and then passes operations to Compute nodes to do their work in parallel.

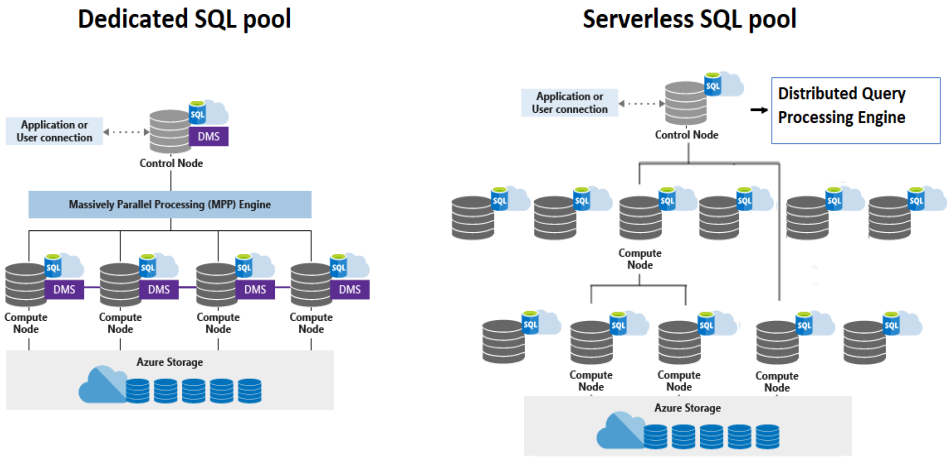


Fig. 4. Azure Synapse Analytics architecture

3. Cloud vs On-Premises Data Warehouse solutions

Cloud data warehouses provide the same benefits that drive organizations to migrate other applications to the cloud. The cloud data warehouse maintenance, scaling or upgrading is easier as compared to traditional data warehouse because in cloud data warehouse you do not need to maintain indexes, clean files or update Meta data etc. The cloud provides cheap storage and on-demand compute. The cloud data warehouse has lowered the cost of complexity and lengthy time-to-value, which has limited the adoption and successful use of traditional data warehouse technology.

Although, the on-premise traditional Data Warehouse has some benefits compared to the cloud version. First and main is that the organization has complete control over the data, software and hardware used to implement the solution. If there is failure of the system, the IT profecionals within the company are having physical access to the hardware and access to every layer of software to facilitate troubleshooting.

The key differences between traditional data warehouse and cloud data warehouse are explained in Table 1 bellow.

Table 1. Cloud vs traditional Data Warehousing

Characteristic	Cloud data warehouses	On-premises data warehouses
Scalability	Instant up or down scaling	Scalability depends on the solution hardware and software infrastructure and may be time and resource consuming due to reconfiguring/purchasing hardware, software, etc.
Availability	Less than 0.01% of downtime with leading cloud providers (Amazon, Microsoft, Google).	Depends on the quality of available hardware, software, and the competence of the IT professionals implementing the solution.
Security	The cloud provider is responsible for the infrastructure and data safety.	Depends on the competence of the IT professionals implementing the solution.
Performance	Serve multiple geographic locations and has great query performance (measured in seconds).	If the scalability challenge is solved, show excellent query performance (measured in milliseconds).
Cost-effectiveness	No hardware related costs. Pay only for used storage and computing resources (the pay-as-you-go cost model).	The solution requires significant initial investments (hardware, IT team, training, etc.)
Flexibility	A Cloud based data warehouse is flexible and automatically grows when its need increases.	It may result in over provisioning and overpaying

4. Conclusion

The modern serverless Data Warehouse is changing the game in the field of big data and business intelligence by providing an easy, effective, and powerful way to achieve the set goals. It's obvious to feel pleased by current speed and storage capacity at which the data can be retrieved and processed. The hardware and the performance are getting better with the passage of time.

In this paper the main characteristics of cloud-based ETL tools for IoT and Big Data era were briefly explained. A sample ETL process was proposed to manage data generated by IoT devices.

References

1. Amazon,2022, https://docs.aws.amazon.com/redshift/index.html
2. Belev, I., Alternatives for Storing and Validating XBRL Data, ASRJETS, 2019, Vol. 60, No. 1, pp.191 - 201, ISSN (Print) 2313-4410, ISSN (Online) 2313-4402.
3. Belev, I., Relational Database Model for Defining Web Automation Integration Rules, Proceedings of International conference on application of information and communication technology and statistics in economy and education (ICAICTSEE – 2016), UNWE -Sofia, 2016, pp. 533-537, ISSN 2367-7635 (PRINT), ISSN 2367-7643 (ONLINE)
4. Delchev, D., Lazarova, V., Big Data Analysis Architecture, Economic Alternatives, 2021, Issue 2, pp. 315-328
5. Inmon, W. (2005). Building the Data Warehouse. John Wiley & Sons, New York, 4th edition.
6. Kimball, R. (2002). The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Datawarehouses. John Wiley & Sons, New York, 2nd edition.
7. Marzovanova, M., Creating annotations and indexes in system for intelligent tagging and search in unstructured data, Proceedings of International conference on application of information and communication technology and statistics in economy and education (ICAICTSEE – 2013), UNWE - Sofia, 2014, ISBN 978-954-644-586-5.
8. Microsoft,2022, https://azure.microsoft.com/en-us/services/synapse-analytics/#overview
9. Mihova V., Murdjeva A. Metadata for generating a specific data warehouse. International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE-2012), Sofia, Bulgaria, 2012. (http://icaictsee.unwe.bg/pastconferences/ICAICTSEE-2012.pdf)
10. Milev, P., Conceptual Approach for Application of Data Science in Web Analytics, Proceedings of International conference on application of information and communication technology and statistics in economy and education (ICAICTSEE – 2019), UNWE -Sofia, 2019, pp. 223-228, ISSN 2367-7635 (PRINT), ISSN 2367-7643 (ONLINE).
11. Minchev, Z., Boyanov, L., Future Digital Society Resilience in the Informational Age, Sofia: SoftTrade & Institute of ICT, Bulgarian Academy of Sciences, 2019
12. Murdjeva, A., Stefanov, G., 2012. Analysis of Conflicts during the Integration of Heterogeneous Multi-Dimensional Data Marts. In: International Conference on Application of Information and Communication Technology and Statistics in Economy and Education ICAICTSEE, 5-6 October 2012, UNWE, Sofia, Bulgaria
13. Stefanov, G., Analysis of Cloud based ETL in the Era of IoT and Big Data, Proceedings of the 9th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2019), UNWE - Sofia, 2020, pp. 198 - 202, ISSN 2367-7635.
14. Tsaneva, M., A practical approach for integrating heterogeneous systems. Business Management 2/2019 (http://bm.uni-svishtov.bg/title.asp?title=1383).
15. Tsaneva, M., Integration of Business applications via Data-driven File Import, Proceedings of International conference on application of information and communication technology and statistics in economy and education (ICAICTSEE – 2020), UNWE -Sofia, 2020, pp. 429-436, ISSN 2367-7635 (PRINT), ISSN 2367-7643 (ONLINE).

Empirical Study of Leadership Training

Tanya Pancheva

Sofia University St. Kliment Ohridski, Sofia, Bulgaria
t.pancheva@abv.bg

Abstract. The research is focused on leadership training issues. A process for developing empirical research on leadership training is presented. On this basis, a survey model was developed. The application of the survey provides information on the current state of this type of training in Bulgaria. Finally, conclusions and recommendations are outlined.

Keywords. Empirical study, leadership, training.

1. Research methodology

The development of research methodology is important to obtain effective results. In part of the scientific research, this is obtained through the creation of a survey model. It defines certain groups of questions. They should be directed according to the individual characteristics of the studied object. In the specific case, the following questions are defined:

- What is the total number of employees in your organization?
- In what type of settlement does your organization carry out its main activity?
- How long has your organization been operating?
- How long ago did you organize the first leadership training?
- How often do you organize leadership trainings?
- Who are the leadership lecturers who run the training courses?
- Leadership training is aimed at: acquisition of knowledge and skills
- What is the main purpose of the leadership training offered?
- If leadership training is aimed at teaching skills, how is this achieved?
- What leadership training methodology do you use?
- Do you vary your leadership training methodology for different audiences?
- Do the trainers have a specific toolkit to check the result obtained?
- If you have a result checker, please share what it is.
- In your opinion, do you have a good practice from the trainings to share with other trainers?
- Do you have a specific concept or definition of leadership that you follow as a basis for the structure and content of the training?
- If a certain understanding of leadership is used in the training, could you share it:
- The reason your organization conducts leadership training.

The proposed methodology includes the following sequence of steps (Fig. 1):

- Defining the sections in the survey card;
- Development of questions;
- Creating a questionnaire;
- Programming of questionnaire for online survey;
- Conducting the survey.

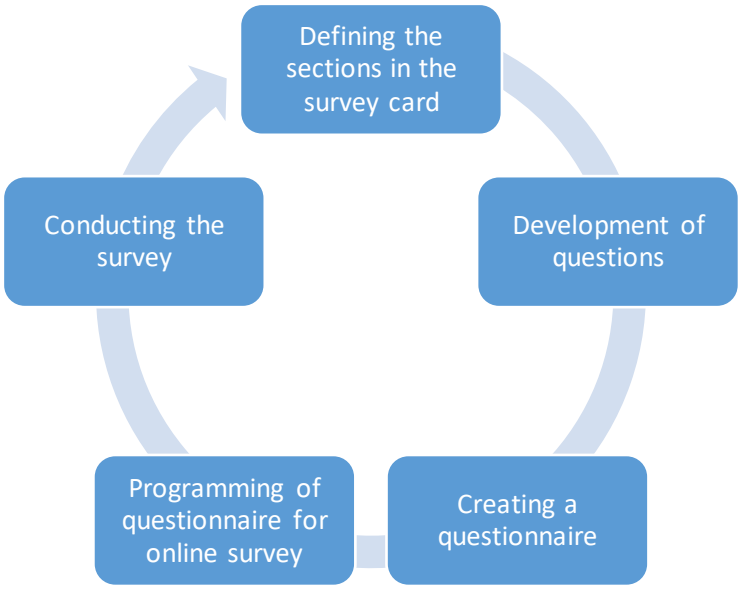


Fig. 1. Main stages of the empirical study

During the first stage, the groups of questions are defined. This is very important for the proper structuring of the survey card. Good practice shows that shorter questionnaires are more effective. For these reasons, it is optimal to have 3-5 groups of questions. Next is the development of the questions from the survey card. Possible answers are defined for each of them. For some of the questions, there is a possibility of 1 answer. For other questions, there is a possibility of 2 or more answers. In this way, the development of the final questionnaire is carried out. Next is the programming of the questionnaire. This gives opportunities for online research. The last stage is the conducting of the survey among the respondents.

2. Empirical analysis

The empirical study was conducted among 22 respondents. In fig. 2 is a screenshot of the survey card. It shows that 40.1% of respondents have a staff of up to 9 people. 26.4% of the organizations have a staff of 10 to 49 people. 13.6% of respondents have a staff of up to 249 people, and 9.1% have more than 250.

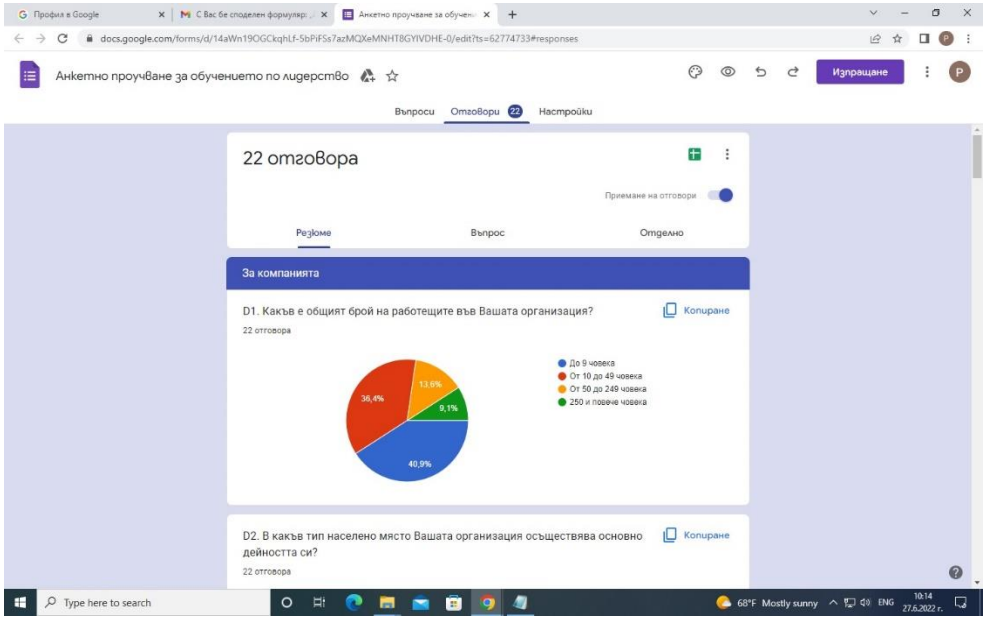


Fig. 2 Questionnaire - Part 1

In fig. 3 shows the territorial distribution of the respondents.

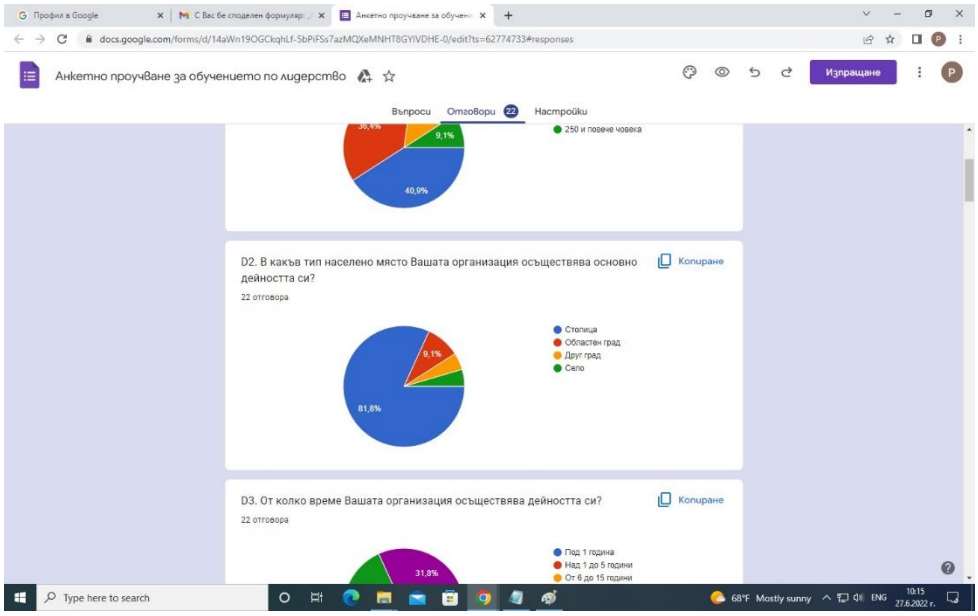


Fig. 3 Questionnaire - Part 2

Along with the size of the organization, the second question is aimed at determining the place for carrying out the main activity. 81.8% of the respondents carry out their main

activity in the city of Sofia, and 9.1% in a regional city. The number of organizations that carry out their main activity in another populated place is small.

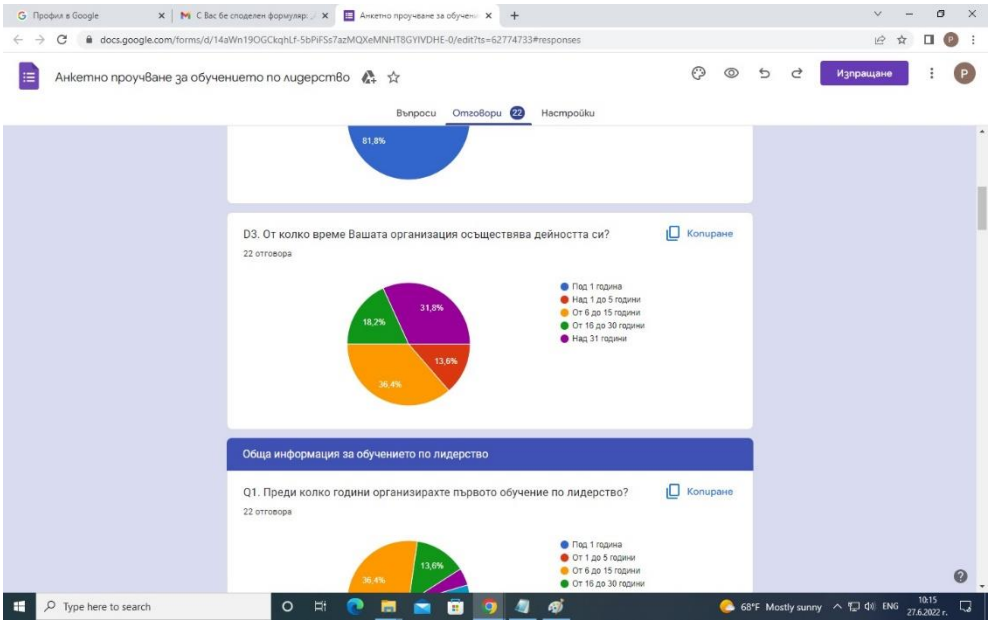


Fig. 4 Questionnaire - Part 3

Another important characteristic of the organizations, which has a direct relation to the object of research, is the period from which it carries out its activity. The longer the organization has been in business, the more likely it is that it has established itself in the market and that the leadership training offered has proven its effectiveness. The following results were observed in the conducted survey: the largest number of organizations carrying out similar activity between 6 and 15 years, namely 36.4%. Followed by the organizations with 31.8% that have been conducting similar trainings for more than 31 years. From 16 to 30 years are 18.2% of the organizations, and 13.6% have been performing similar activity relatively recently - over 1 to 5 years. Of the respondents who answered, there are no representatives who conduct leadership training for less than 1 year (fig. 3).

3. Conclusions

As a result of the research and empirical analysis, the following conclusions can be drawn:

- A questionnaire has been created to study leadership training in Bulgaria;
- A survey was conducted among 22 respondents.

References

1. Internet site https://forms.gle/rWUHfvVL6Qcqyora8 .

Methodology for Training in the Module "Databases" with the Help of MS Access

James Osondu

Faculty of Economics and Business Administration, Sofia University "St. Kliment Ohridski", Sofia, Bulgaria
josondu@yahoo.com

Abstract. Microsoft Access is very useful and helpful in teaching and demonstrating database topics in universities. One of the people that often needs to learn about databases is university students based on their given curriculum for a required degree program. That is right because there are many economic activities that makes use of a database. For example, a manager, customers, and even employees of a business entity need to understand about a database. That is why Microsoft Access can help you to teach and demonstrate economic activities of an economic entity so one can understand databases. So, we will discuss about how Microsoft Access can be used to teach topics in databases in universities. Since Microsoft Access is a Database Management System (DBMS), it is ideal to teach the basic principles of relational databases. its collaboration with other software like PowerPoint, Excel, Word makes it ideal for Demonstrating which is a teaching method employed by my university. Access can be used for Presentation, Exercises, Tutorials, Examples, and Modeling for relational database models. An example used here is the teaching of the Topic of Tables, Keys, Relationships. This topic is taught through Presentation, Exercises, Tutorials, Examples, and Modeling for relational database models.

Keywords. Database management system, Microsoft Access, Database curriculum, Relational database models, Training.

1. Introduction

This article is broken into 4 parts. The first section describes Microsoft Access and its different features. The second section talks about the Outline of how Microsoft Access might be used in supporting the core topics of a Database curriculum. The third section discusses the Outline of how Microsoft Access might be used to support specifically the database core topics. The fourth part is the conclusion of the article.

Description of Microsoft Access and its different features:

- Microsoft Access is a part of the Office Suite, which allows users to create databases through a database management system. It is a file server-based database. It allows users to store or import data or directly link to data stored in other applications and databases, such as Microsoft Excel or ASCII' (Wikipedia, 2021).
- It allows users to create reports, forms, queries and macros, design tables and build relationships.
- Microsoft Access has a 'navigation pane' where data is displayed which lists tables, forms, reports, and objects in the opened database. It is a simple, quick, and effective way to open a form, to run a report or to enter data.
- It also has a 'ribbon' that is on the top of the screen which shows all the commands

for working with a database. Each ribbon tab contains its own group, and each group contains the commands for a particular activity such as 'sort and filter'. The ribbon makes it a lot easier for the users to find specific icons. For example, if the user wanted to create a form, they click onto create in the ribbon then in the forms group they would click on form wizard.

- One of the benefits of Microsoft Access is its compatibility with structured query language (SQL) as queries can be viewed graphically or edited as SQL statements where these statements can be used directly with macros.

2. Outline of how Microsoft Access might be used in supporting the Database curriculum core topics

The Database course is an integral component of any degree program offered by the faculty of Economics in Sofia University. The programs are the 3 bachelor's degrees in English, French, and German - Economics and Finance, Business Administration and Accounting, Finance and Digital Applications.

The introductory database course for undergraduates by Prof. Shvetner is assembled around three main areas:

- Primary focus: Data Modeling
- Primary focus: Database Programming using Structured Query Language (SQL)
- Secondary focus: Database System Architecture and Administration

Another feature of Access is that it is also possible to write SQL statements and run them against the MS Access database. This buttresses the second main goal of the course: To teach Structured Query Language (SQL).

Microsoft Access could be used as the program for the Design, produce and evaluate core topic.

Microsoft Access could be used as the problem for the Design, produce and evaluate core topic; as the problem could be creating a database for a video club (example), the solution would be creating the database to suit the requirements of their staff, movies, clients and/or costs, then they evaluate the outcome.

Microsoft Access could be used as a real-life application in Data handling and examine how data is formed, coded, sourced, categorized, stored, compressed, and secured in this program. Students will be able to visualize and undertake some of these topics on the actual software program.

Ethical and social issues could be explored using Microsoft Access in real life applications and how these issues can be correctly addressed to client's information stays confidential and protected.

Microsoft Access also helps people in respect to the skills users need to operate Microsoft Access efficiently and correctly.

As Microsoft Access is software it can be utilized as a prime example of the core topic software to explore its purpose, functions, and features (Board of Studies NSW, 2003).

These features above, support most of the topics covered in the introduction to databases course. When used for project work, where a complete database application is created, MS Access can provide a real sense of accomplishment when students see the entire system running together.

3. Database curriculum core topics

Microsoft Access can be a practical way of teaching students how a database is designed and how it works which may help students to easily identify key factors through several tasks such as creating a table, as they have already completed the task. An example of this can be through the lesson plan and assignments provided by the instructor.

It can be used for project based learning as a series of activities can be implemented for the students to design, produce and evaluate solutions for an identified need or problem which could be a new video club company has launched, thus students have to create a database with new clients, their personal details and documentation relating to work that needs to be completed and what has been completed for specific clients. This project scenario or other scenarios can be developed to focus on problem-solving, generating ideas, modelling, managing, communicating, collaborating, and evaluating solutions. The project can be linked to their interests and everyday lifestyles by choosing a line of work their interested in or create a Self Organized Learning Environment (SOLE)where they can choose a scenario of their choice however, following specific criteria set out by the instructor. An example of this can be through the assignment and notes provided by the instructor however, students can create their own scenario, following specific guidelines and need to focus on creating functions stated on the curriculum (file, record, field and data, character, data dictionary, validation and verification checks, input and store data, query searches and sorts, mail merge, design and perform calculations) over the duration of this subject area (Board of Studies NSW, 2003).

*Using a **Table** Object* in Access helps because a database table is similar in appearance to a spreadsheet, in that data is stored in rows and columns. As a result, it is usually quite easy to import a spreadsheet into a database table. The main difference between storing your data in a spreadsheet and storing it in a database is in how the data is organized.

To get the most flexibility out of a database, the data needs to be organized into tables so that redundancies do not occur. For example, if you are storing information about staff, each staff should only need to be entered once in a table that is set up just to hold staff data. Data about movies will be stored in its own table, and data about branch offices will be stored in another table. This process is called normalization.

Each row in a table is referred to as a record. Records are where the individual pieces of information are stored. Each record consists of one or more fields. Fields correspond to the columns in the table. For example, you might have a table named "Staff" where each record (row) contains information about a different staff, and each field (column) contains a different type of information, such as first name, last name, address, and so on. Fields must be designated as a certain data type, whether it is text, date or time, number, or some other type.

Another way to describe records and fields is to visualize a library's old-style card catalog. Each card in the cabinet corresponds to a record in the database. Each piece of information on an individual card (author, title, and so on) corresponds to a field in the database.

Tables are essential objects in a database because they hold all the information or data. For example, a database for a business can have a Contacts table that stores the names of their suppliers, e-mail addresses, and telephone numbers. Because other database objects depend so heavily on tables, you should always start your design of a database by creating all its tables and then creating any other objects. Before you create tables, consider your requirements, and determine all the tables that you might need.

Other important database topics which can be explained by access are Table and field properties; in Access, Tables and fields also have properties that you can set to control their characteristics or behavior.

In an Access database, table properties are attributes of a table that affect the appearance or behavior of the table as a whole. Table properties are set in the table's property sheet, in Design view. For example, you can set a table's Default View property to specify how the table is displayed by default.

A field property applies to a particular field in a table and defines one of the field's characteristics or an aspect of the field's behavior. You can set some field properties in Datasheet view. You can also set any field property in Design view by using the Field Properties pane.

Another core topic is about Data types:

- Every field has a data type. A field's data type indicates the kind of data that the field stores, such as large amounts of text or attached files.
- A data type is a field property, but it differs from other field properties as follows:
  - You set a field's data type in the table design grid, not in the Field Properties pane.
  - A field's data type determines what other properties the field has.
  - You must set a field's data type when you create the field.

You can create a new field in Access by entering data in a new column in Datasheet view. When you create a field by entering data in Datasheet view, Access automatically assigns a data type for the field, based on the value that you enter. If no other data type is implied by your input, Access sets the data type to Text. If needed, you can change the data type by using the Ribbon.

*Another core topic is about **Table relationships.***

Although each table stores data about a different subject, tables in an Access database usually store data about subjects that are related to each other.

For example, a database might contain:

- A customers table that lists your company's customers and their addresses.
- A products table that lists the products that you sell, including prices and pictures for each item.
- An orders table that tracks customer orders.

Because you store data about different subjects in separate tables, you need some way to tie the data together so that you can easily combine related data from those separate tables. To connect the data stored in different tables, you create relationships. A relationship is a logical connection between two tables that specifies fields that the tables have in common.

*Another core topic is about **Keys***

Fields that are part of a table relationship are called keys. A key usually consists of one field but may consist of more than one field. There are two kinds of keys:

- Primary key: A table can have only one primary key. A primary key consists of one or more fields that uniquely identify each record that you store in the table. Often, there is a unique identification number, such as an ID number, a serial number, or a code, that serves as a primary key. For example, you might have a Customers table where each customer has a unique customer ID number. The customer ID field is the primary key of the Customers table. When a primary key contains more than one field, it is usually composed of pre-existing fields that, taken together, provide unique values. For example, you might use a combination of last name, first name, and birth date as the primary key for a table about people. For more information, see adding or changing a table's primary key.
- Foreign key: A table can also have one or more foreign keys. A foreign key contains values that correspond to values in the primary key of another table. For example, you might have a Movies table in which each movie has a director ID number that corresponds to a record in a Directors table. The director field is a foreign key of the Movies table.

The correspondence of values between key fields forms the basis of a table relationship. You use a table relationship to combine data from related tables. For example, suppose that you have a Directors table and a Movie table. In your Directors table, each record is identified by the primary key field, ID.

To associate each movie with a director, you add a foreign key field to the Movie table that corresponds to the ID field of the Directors table, and then create a relationship between the two keys. When you add a record to the Movie table, you use a value for directors ID that comes from the Directors table. Whenever you want to view any information about a movie's director, you use the relationship to identify which data from the Directors table corresponds to which records in the Movie table.

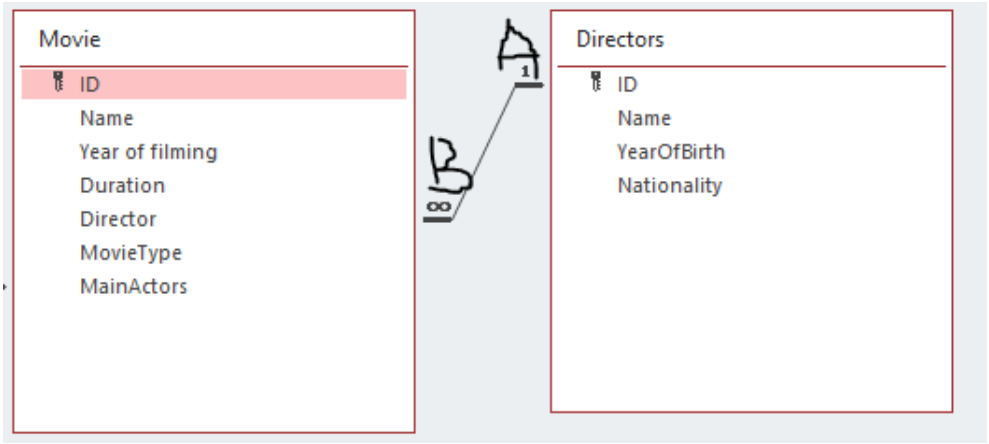


Fig. 1. Primary &Foreign Keys

A. A primary key, identified by the key icon next to the field name.

B. A foreign key — note the absence of the key icon.

Do not add a field if you expect that each unique entity represented in the table might require more than value for the field. Continuing the preceding example, if you want to start tracking orders placed by your customers, you do not add a field to the table, because each director will have more than one movie. Instead, you create a new table to store movies, and then create a relationship between the two tables.

Some of the Benefits of using relationships are.

Keeping data separated in related tables produces the following benefits:

- **Consistency -** Because each item of data is recorded only once, in one table, there is less opportunity for ambiguity or inconsistency. For example, you store a customer's name only once, in a table about customers, rather than storing it repeatedly (and potentially inconsistently) in a table that contains order data.
- **Efficiency -** Recording data in only one place means you use less disk space. Moreover, smaller tables tend to provide data more quickly than larger tables. Finally, if you do not use separate tables for separate subjects, you will introduce null values (the absence of data) and redundancy into your tables, both of which can waste space and impede performance.

- **Comprehensibility** - The design of a database is easier to understand if the subjects are properly separated into tables.

Plan your tables with relationships in mind. In Access, you can use the Lookup Wizard to create a foreign key field if the table that contains the corresponding primary key already exists. The Lookup Wizard creates the relationship for you.

The Microsoft Access also offers collaboration with traditional Office desktop applications such as Word, Excel, Power point and can be used for Lecture Notes, Exercises, Examples, Textbooks, and Presentations. The creation of lecture notes, exercises, power point presentations can be put in an Access container or database. Creating a separate database for the course content that would include attachments for lecture notes, exercises, PowerPoint presentation and textbooks and links to external internet resources is required. Figure 2 bellow shows the logical table called Topics that contains fields for the main database topics, links, attachments, and Comments.

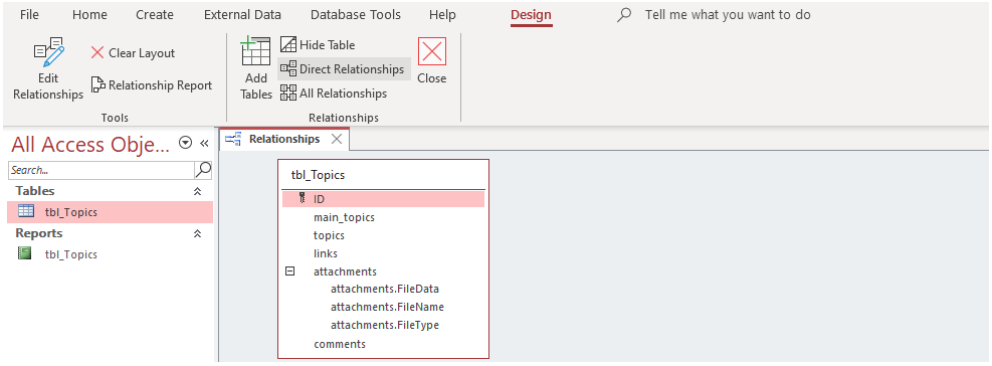


Fig. 2. Course content

4. Conclusion

The first section explains Microsoft Access and its different features. The second section explains how Microsoft Access might be used in supporting the Database curriculum core topics. The third section describes how Microsoft Access might be used to support specific core topics like Tables, relationships between tables, primary key, foreign key, data types. In addition, the creation of lecture notes, exercises, power point presentations can also be put in an Access container or database. Creating a separate database for the course content that would include attachments for lecture notes, exercises, PowerPoint presentation and textbooks and links to external internet resources is required. This methodology is suitable for a Demonstrator style teaching.

References

1. Wikipedia (2021, June 20). Microsoft Access. Retrieved June 20, 2021, from Wikipedia: http://en.wikipedia.org/wiki/Microsoft_Access
2. Why I teach intro to databases using Microsoft Access, article posted on Posted on January 25, 2017 by holowczak ; accessed on June 20, 2021, http://holowczak.com/why-i-teach-intro-to-databases-using-microsoft-access/?doing_wp_cron=1624049062.2942008972167968750000
3. Board of Studies NSW (2003). Information and Software Technology Years 7-10: Syllabus. Sydney: BOS

Simulation Modelling in the Teaching of Mathematics

Svetlana Vasileva

NMHS, Dobrich, Bulgaria
svetlanaeli@abv.bg

Abstract. The paper discusses some aspects of the application of queuing theory (and queuing systems) and the simulation modeling environment GPSS Studio to create demonstration models and use them in profiling training in Mathematics in high school. The classic use of simulation modeling is to study the variety of systems. But "observation windows" of simulations in GPSS World and GPSS Studio reveal new advanced capabilities of GPSS for integrating simulation modeling in the learning process. The paper gives some examples of the use of imitation models and demonstration capabilities of GPSS World and GPSS Studio in teaching materials in Practical mathematics (Empirical distributions).

Keywords. Queuing systems, simulation models, Probability distributions, GPSS.

1. Introduction

One of the most important tasks of modern education systems is the formation of universal learning activities that develop in learners not only the ability to learn, but also the ability for self-development and self-improvement. [1] Some of the cognitive universal learning activities are the sign-symbolic universal activities, which involve the acquisition of model-building techniques. Therefore, it is the teachers' main goal to use a variety of graphic tools widely during classes. This is due to the fact that the verbal form for presenting information in the learning process is not optimal and is not universal. [1]

On the other hand, modeling is an action that begins in elementary school and passes into further human activity and reaches a new level of its development. Modeling helps to make the studied object more accessible. Thus, through modeling "can reduce the study of the unknown - to the known, from the simple - to the complex" [1].

Educational modeling is one of the forms of education that contributes to the development of interest in the subject and increases the motivation for learning [2], because working with the model, students "have the opportunity to study the process in dynamics, to clarify the interrelationship of the components of the studied system"[2].

Simulation modeling (SM) allows to study the analyzed or designed system "according to the scheme" of the operational research, which consists of interconnected stages [3]: • content statement of the task; • development of a conceptual model; • development and software implementation of a simulation model; • check for correctness and credibility of the model; • assessment of the accuracy of the modeling results; • planning and conducting experiments; • making decisions.

This allows to use simulation modeling as a universal approach to decision-making in the conditions of uncertainty and accounting in the models of factors that are difficult to formalize. But perhaps the most attractive and most important side of SM for integration not only in management and economics, but also in education, is the application of the basic principles of the systems approach in solving practical problems.

The purpose of the report is to show some new opportunities in secondary school (profiling preparation in mathematics and informatics [4]), opening up with the application of modern information technologies, which can be used to create effective curricula and courses, especially with the inclusion of simulation modeling tools.

2. Queuing Systems and Simulation Modelling

Simulation modeling is one of the most popular means of modeling systems. SM allows queuing systems (QS) to be studied under different types of input flows and request intensities at system inputs and to determine their main characteristics. It can be said that SM can serve as a link between learning mathematics and learning informatics (programming).

The system for simulation modelling - GPSS World provides users with software tools for interacting with the system in the process of simulation research and software tools that directly implement the modeling process (parser, model interpreter, model clock, management of event lists, etc.). [5] The GPSS World Extended editor of provides the means and ways to lower the qualification requirements for both the simulation modeling expert and the end user using the created models. In the extended editor, most operations and actions during the entire cycle of simulation studies are automated. [5] and [6]

GPSS World. The GPSS simulation modelling language is mainly used to build event discrete simulation models and run experiments on a computer [7], [8], [9], etc. This implies the use of the environment of GPSS in the educational process as well, both for demonstrating the operation of studied systems and for developing by the students themselves models for the management and operation of different systems, as well as for demonstrating different types of probability distributions.

The main and effective application of SM is in the study of various QS, but simulation modeling can also be successfully used as a teaching tool, especially in disciplines in the content of which the topics of concurrency control of processes (transactions) and the like are studied., described in QS terms as service requests arriving at the respective system according to a probability distribution law.

The main and effective application of SM is in the study of various QS, but simulation modeling can also be successfully used as a teaching tool, especially in disciplines in the content of which the topics of concurrency control of processes (transactions) and the like are studied., described in QS terms as service requests arriving at the respective system according to a probability distribution law.

The main and effective application of SM is in the study of various QS, but simulation modeling can also be successfully used as a teaching tool, especially in disciplines in the content of which the topics of concurrency control of processes (transactions) and the like are studied., described in QS terms as service requests arriving at the respective system according to a probability distribution law. From the point of view of the many available examples of models of various systems where the transactions enter for servicing, according to different distribution laws, GPSS could be used in profiled mathematics training as well.

3. Basic Elements of the Simulation Modelling Language

Blocks and transactions. The GPSS language is a declarative language built on the principle of object-oriented programming languages. Basic elements of the language are transactions and blocks, which represent respectively dynamic and static objects of the model [7], [8], [9], [10], etc. The operation of a GPSS-model under the control of the simulator consists in moving the transactions from some blocks to others analogously to the moving of the requests (corresponding to the GPSS transactions) in the modeled QS.

Blocks are routines containing a set of parameters (operands) for addressing them. The transfer of the control from block to block is realized through the movement of transactions in model time. The call of the subprograms of the blocks is done through the movement of transactions.

Objects of type "resources" - analogues of service devices in real systems. Objects of this type include devices, multi-channel devices (storage, memories) and logical keys. As in any object-oriented language, in GPSS each object has properties and methods that modify those properties (called in the GPSS standard Numeric Attributes - SNA).

Variables. Arithmetic variables allow to calculate arithmetic expressions that consist of operations on the system numeric attributes of the objects..

Функции. A function can specify a continuous or discrete functional relationship between the function's argument and its value. In the two-phase locking models shown in distributed databases, described in [11], for example, functions are used to set the probabilities with which a transaction will "read" or "update" an element.

Cells and matrices, save values - used to store user numeric information. Writing to these objects is performed by the transactions. In the models shown, cells are used for counters of generated and fixed transactions.

Queues. In any system, the flow of transactions can be held up due to resource unavailability. In this case, the transactions queue up. Points in the model can be specifically defined at which to collect queue statistics (queue loggers). Then the interpreter will automatically collect queue statistics (length, average queue time, etc.).

Tables. The tables serve to collect statistics for random variables set by the user. They consist of frequency classes in which the number of hits in a given interval of a specific quantity is recorded, such as the system numerical attributes [10], such as: total number of transactions entered into a given block; total number of entries in a storage, etc. For example, to construct the transaction service diagram from the Centralized Two-Phase Locking model [11], shown in fig. 1, the DaTable table is defined in the GPSS program [11], in the manner described in table 1 and demonstrated in formula (1).

Table 1 Structure of the command Table

Table name	**TABLE**	Value of the standard numeric attribute	First boundary value	Width of intermediate intervals	Total number of intervals, including the left and the right
DaTable	**TABLE**	***M1***	***300***	***200***	***30***

The table DaTable fixes the value of the system standard numeric attribute M1. M1 is the residence time of the transaction in the model [9] and [10].

DaTable TABLE M1,300,200,30 (1)

Before the transaction leaves the model, with the TABULATE DaTable command, the transaction's residence time is recorded (fixed) in the table.

4. Observation Windows

Blocks Window – different mode of operation - for brief or detailed reference during simulation, containing information about labels and names of blocks, number of enters in each block, log with chronological tracking of actions by blocks in model time [9];

Table Window – frequency distribution diagram of the tabulated transactions in a table defined in the model. The window of the selected table presents a dynamic histogram, which

is extremely useful for displaying served/received transactions, frequency distribution and other characteristics [9]. Fig. 1 shows the diagram formed by GPSS on the table *DaTable* described in table 1 and formula (1).

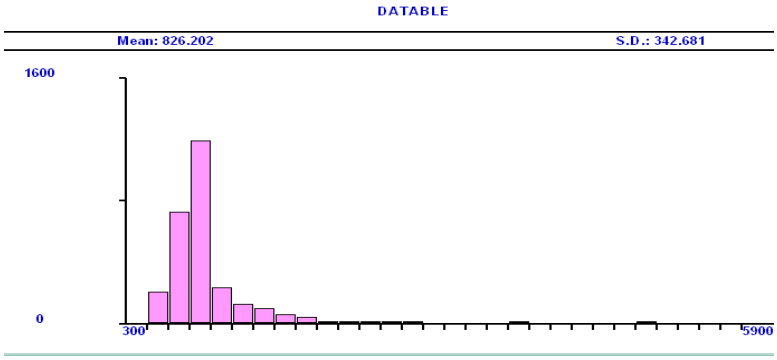


Fig. 1 Frequency Distribution of the Transactions Response Time in the Centralized Two Version Two-Phased Locking [11]

Facilities Window – the Facility entities window (fig. 2, bottom left) – information on: Number/name of the facility; Number of entered transactions; Utilization factor; etc. [9];

Matrix Window – It is shown on fig. 2 (the two windows on the right) – describes the results in values for the common resource for the transactions and their comparison at specified units of model time [9];

Plot Window – Function or expression graph window [9]. It can be used to trace the values of arbitrarily composed expressions;

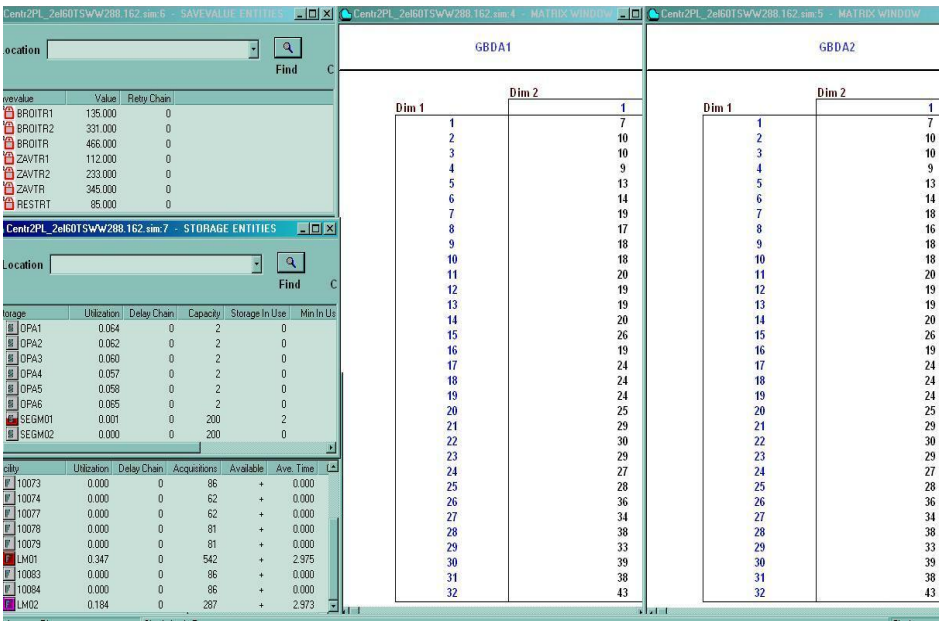


Fig. 2 Combined window for demonstration of the execution of a modeling algorithm [12]

Queues Window – Displays information about: Queue Number/Name; Maximum tail length; Current queue length; Total number of entries; Average queue length; Average waiting time of the transactions in the queue; etc. [9] and [10];

Savevalues Window - Fig. 2 the upper left panel. Any change to the variables (usually counters) defined in the model can be monitored [9];

Storages Window – Window of the multichannel devices (Fig. 2, middle left panel). During a simulation, it can be seen any change on: the number of free units of the capacity of each of the multi-channel devices; the minimum and maximum used capacity of the device; the number of entries in the corresponding multi-channel device, etc. [9] and [10].

Fig. 2 shows how several different windows can be opened and arranged on the screen in the GPSS World environment and in them the operation of the modelled system can be observed. The example shows a combined window for observing a simulation of Centralized Two-phased Locking with embedded timestamping [12] (such algorithms are studied in the topics Transaction Management Algorithms and Distributed Database Management Systems by disciplines such as Databases).

The modelling algorithm, a snapshot of which is demonstrated in fig. 2, is described as a protocol in [13]. In the example of fig. 2 the "combined window" makes it possible to monitor:

- The values of the elements of the matrices *GBDA1* and *GBDA2*. Arrays model the local databases in which the two copies of the data items are stored. In the matrix windows *GBDA1* and *GBDA2* can see the synchronous update of the values of the local copies of each data element. The first column of the matrices *GBDA1* and *GBDA2* are the numbers of the data elements, and the second column – the values of the replicas of the corresponding data elements;
- The SaveValues window at the top left of fig. 2 makes it possible to monitor the counters: *BROITR1* and *BROITR2* – number of transactions generated, processing 1 and 2 elements respectively, *BROITR* – total number of transactions generated, *ZAVTR1* and *ZAVTR2* – number of transactions leaving the model, processing 1 and 2 data elements respectively, *ZAVTR* – total number of transactions that left the model, *RESTRT* – number of restarted transactions.

5. Probability distributions

In the curriculum for profile preparation in mathematics [4] in Module 3 called "Practical Mathematics" for the objectives of the module it is written: "... building skills for modeling real or theoretical situations with the studied tools, skills for interpreting the obtained results and for considering the problems posed as a set of separate problems". Bearing in mind the syllabuses of Module 1 "Geometry" and Module 2 "Elements of Mathematical Analysis", which involve only mathematical techniques, it can be said that the GPSS World environment with over twenty probability distributions from the built-in library of procedures (Beta; Binomial; Weibula; Discrete uniform; Gamma; Geometric; Laplace; Logistic; LogLaplace; LogLogistic; LogNormal; Normal; Inverse Weibull; Inverse Gaussian; Negative binomial; Pareto; Pearson Type V; Pearson Type VI; Poisson; Uniform; Triangular; Exponential, etc. [9]), to describe incoming transaction flows, can be successfully integrated both in lessons for new knowledge and exercises, and also as the development of student projects for participation in student scientific conferences or interdisciplinary student developments.

From queuing theory, it is known that the outgoing QS flow with m devices waiting for the simplest input flow with parameter R and exponential service time distribution with parameter μ is the simplest flow with parameter $\lambda = min\{\lambda, m\mu\}$ [10]. Therefore, complex QSs

can be constructed where the output of one service system is input to other systems. In this way, multi-phase systems and networks of queuing systems are obtained.

Therefore, complex QSs can be constructed where the output of one service system is input to other systems. In this way, multi-phase systems and networks of queuing systems are obtained.

Therefore, complex QSs can be constructed where the output of one service system is input to other systems. In this way, multi-phase systems and networks of queuing systems are obtained. The distribution of QS outflows has a complex probabilistic nature and can only be studied by observing the operation of the corresponding QS through modelling" [10]. In other words, observing the execution of GPSS transactions modeling processes in a real system modeled with QS is a way for students to understand the meaning of studying the topic of Empirical Distributions and also to absorb the learning material successfully. Last but not least, they will learn to think systematically and apply this approach, both in their future professional activities and in their civil and personal lives.

Let's look at an example shown in [10, p. 78], modeling the customer service of a bank:

The arrival time of clients is given by the distribution function ARRIVE. Seven cashier counters are open, a queue can form at each of them. When a client arrives at the bank, if at least one cashier is free, the client immediately approaches this cashier. Otherwise, it joins any current shortest queue. The order of servicing clients from the queue is FIFO.

Service at the cash register is divided into five types of operations, the frequency of which and the average service time are set by the MEAN function.

In models with several queues, simultaneous events may occur at the customer's arrival and at the end of the cashier's service. If the "arrival" event is considered before the "end of service" event, then the device will be considered busy at the time of checking, and the newly arrived client will not be able to occupy the cashier. Therefore, it is necessary to ensure priority in the processing of the event "release of the cashier". This can be done by setting the priority level (PRIORITY 1 block) between the GENERATE and ADVANCE blocks – fig. 3.

To implement the search for a free cashier, the TEST E block, labelled NEXT, is used, which checks whether the cashier is free. The maximum cashier number is stored in the transaction P5 parameter. By the same parameter, a loop is formed to check all devices (LOOP block). The number of the first free cashier is stored in the transaction P3 parameter, and the transaction is submitted for servicing. If all devices are busy, then the transaction moves to the next block after the LOOP, that is, to the TEST block marked with the SIT label. This block compares the queue lengths for all devices. Initially, for comparison, the number 1000 is placed in the P2 transaction parameter. Similarly, a loop through all queues is organized using the P1 of the transaction parameter. Initially, the maximum queue number is placed there. If the length of the next scanned queue is less than the previous one, then the queue number in the P3 parameter and its length in the P2 of the transaction parameter are stored. After viewing all the queues (going to the next block after LOOP), these parameters will contain the number of the minimum queue and its length, respectively. In FIG. 3 e shows the text on the GPSS software.

```
* A function of the service time
020   MEAN FUNCTION RN1,D5
   0.0,450/.29,750/.61,1000/.85,1500/1.0,3000
* A function of the entry time in the model
030   ARRIVEL FUNCTION RN1,D6
   0.0,50/.25,100/.60,150/.80,200/.90,250/1.0,300
```

```
040   T_1    TABLE  M1,200,600,10     ; Arrival time data
collection

050  ***************************************************
055          GENERATE  FN$ARRIVEL     ; Customer login
060          ASSIGN    1,7            ; Set the MAX queue number
065          ASSIGN    2,1000         ; Memorizing a large number
070          ASSIGN    4,FN$MEAN     ; Save in P4 the time for
service

075          ASSIGN    5,7            ; Set MAX number of the
cashier
080   NEXT   TEST E    F*5.0,FAC      ; Is there a cashier
available?
085          ASSIGN    3,P5  ;"Yes", we remember the cashier's
number
090          TRANSFER  ,QUI           ; The customer goes to the
                                      ;checkout for service
095   FAC    LOOP      5,NEXT         ; Cycle across all cashiers
100  ***************************************************
105   SIT    TEST L    Q*1,P2,SIS     ; MIN queue search
110          ASSIGN    3,P1           ; We remember the queue
number
115          ASSIGN    2,Q*1    ; We remember the length of the
queue
120   SIS    LOOP      1,SIT          ; Loop through all queues
125   QUI    QUEUE     P3             ; Queuing
130          SEIZE     P3             ; Occupancy of the cashier
135          DEPART    P3             ; Leaving the queue
140          PRIORITY  1
145          ADVANCE   P4             ; Customer service
150          RELEASE   P3             ; Releasing the cashier
155          TABULATE  T_1            ; Time of residence+
+165          TERMINATE
170  ***************************************************
175          GENERATE  14400          ; Timer for working hours
180          TERMINATE 1
```

Fig. 3. GPSS model of the servicing the clients in a bank [10]

Fig. 4 shows the simulation report, and a histogram of the time spent in the system is shown in fig. 5.

Students can compare graphs of the frequency distribution of the response time of different algorithms operating a system, compare them to published standard specifications. They can verify that the modelled system is working correctly at all and conduct all kinds of experiments [10] and [14]. And in terms of probability distributions describing incoming transaction flows, one can change the parameters and observe how the histograms change during model execution.

FACILITY	ENTRIES	UTIL.	AVE. TIME	AVAIL.	OWNER	PEND	INTER	RETRY	DELAY
1	8	0.878	1581.250	1	64	0	0	0	2
2	11	0.906	1186.364	1	77	0	0	0	2
3	9	0.920	1472.222	1	76	0	0	0	2
4	11	0.941	1231.818	1	74	0	0	0	1
5	11	0.955	1250.000	1	73	0	0	0	2
6	10	0.969	1395.000	1	54	0	0	0	2
7	12	0.990	1187.500	1	71	0	0	0	2

QUEUE	MAX	CONT.	ENTRY	ENTRY(0)	AVE.CONT.	AVE.TIME	AVE.(-0)	RETRY
1	2	2	10	4	0.500	720.000	1200.000	0
2	2	2	13	6	0.549	607.692	1128.571	0
3	2	2	11	3	0.708	927.273	1275.000	0
4	2	1	12	3	0.837	1004.167	1338.889	0
5	2	2	13	3	0.951	1053.846	1370.000	0
6	2	2	12	2	1.097	1316.667	1580.000	0
7	2	2	14	1	1.153	1185.714	1276.923	0

TABLE	MEAN	STD.DEV.	RANGE			RETRY	FREQUENCY	CUM.%
T_1	2162.308	948.108				0		
			200.000	-	800.000		5	7.69
			800.000	-	1400.000		11	24.62
			1400.000	-	2000.000		15	47.69
			2000.000	-	2600.000		12	66.15
			2600.000	-	3200.000		13	86.15
			3200.000	-	3800.000		6	95.38
			3800.000	-	4400.000		3	100.00

Fig. 4. Simulation report

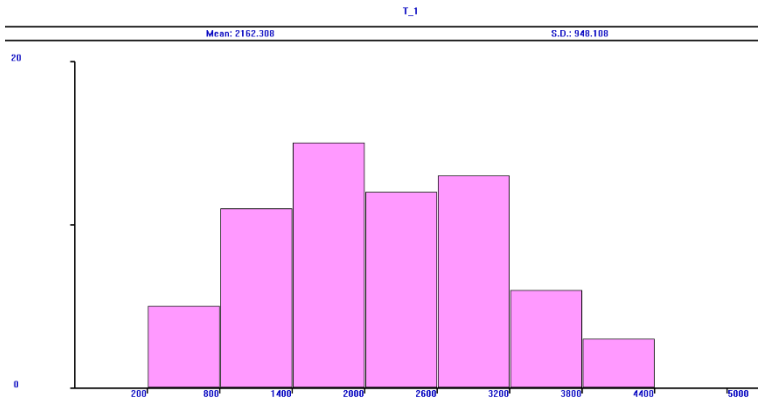


Fig. 5. Frequency distribution of the Response time of the transactions

6. Conclusion

The capabilities to set up and conduct experiments on a modelled system, as well as the automated creation of simulation models in the GPSS World Advanced Editor, the "standard combo window" for monitoring model performance, and the user-friendly graphical interface of the simulation study automation system GPSS Studio, make it make it extremely suitable for educational demonstrations of the operation of various systems. The tabulation of functions and the automatic construction of histograms (polygons) of the frequency distribution of the

service time are a very suitable tool for considering examples from the work of real systems, and not just abstract mathematical examples, which are very often difficult for students to perceive and are ”boring” to them.

By the forms that can be created by the means of the Universal Form Editor [5] and [6] the demonstration materials could be even more attractive and informative for the students. On the other hand, the simulations recorded on film represent the operation of the modelled systems in different modes, almost realistically "in action". And the embedding in electronic textbooks (in online courses or in electronic teaching aids, etc.) provides opportunities for their asynchronous distribution.

Last but not least, through the creation of complex (requiring the integration of different activities and correspondingly different participants) learning materials in different disciplines, students can be attracted to real participation in projects and develop them themselves for Olympiads, competitions and student conferences in mathematics and informatics. Projects in which they solve real problems (usually beneficial for the school), with similar activities to prepare for their future realization as specialists and, not least, to learn to think systematically and to organize and work in a team.

References

1. Gorsheneva M.V., 14 Feb 2021, Technology of simulation modeling in elementary school (https://urok.1sept.ru/articles/685435)
2. Makarova, S.V., 2000, Improving the methodology of teaching information technology in a pedagogical university based on the inclusion in the course content of questions of modeling real situations in human information activity. 2000. (http://www.dslib.net/teoria-vospitania/sovershenstvovanie-metodiki-obuchenija-informacionnym-tehnologijam-v-pedagogicheskom.html)
3. Elina Computer, Simulation Applications for Mass Use Technology (http://elina-computer.ru/static/proekty.html), 2021
4. Republic of Bulgaria, Ministry of Education and Science, 2018, Curricula for profiled training, https://www.mon.bg/bg/100598
5. Vlasov, S., V. Devyatkov, F. Isaev, M. Fedotov, Simulation studies using GPSS WORLD - new opportunities. Automation in industry, 2012, № 7. p. 3–8.
6. Fedotov, M., V. Devyatkov, SASS for Professionals - Extending the Capabilities of the Advanced GPSS World Editor. Proceedings of the 6th All-Russian scientific-practical conference „Simulation modeling. Theory and practice" (SMMOD-2013). Vol. 1,“FEN” Press, Academy of Sciences of the Republic of Tatarstan, Kazan, p. 356-360.
7. Boev, V., *Systems Modeling. GPSS World Toolkit. St Petersburg*, BHV - Petersburg, 2004.
8. Kudryavcev, E., *GPSS World Fundamentals of simulation modeling of various systems*, Moscow, DMK Press, 2004.
9. Minuteman Software. *GPSS World Tutorial Manual*. (http://www.minutemansoftware.com/tutorial). 2010
10. Tomashevskii, V., E. Zhdanova, *Simulation modeling in the GPSS environment*. Bestseller, Moscow, 2003
11. Milev, A., S. Vasileva. Modeling of Two-version Algorithms for Two-Phase Locking in Distributed Databases. International Conference Automatics and Informatics’09, 3-4 October 2009, Sofia, Bulgaria, pp.I-13 – II-16
12. Vasileva, S., D. Vasilev. Some approaches to integrating simulation modeling into learning. Fifth International Conference on Engineering, Technologies and Systems TECHSYS 2016, Technical University – Sofia, Plovdiv branch, 26 – 28 May 2016, Plovdiv, Bulgaria, pp. II-245 – II-250, 2016.
13. Kaur, M., Transaction processing in Distributed Databases, Computer Supported Cooperative Work 12, 2003.
14. Elina computer, Advanced simulation editor for GPSS World language. User's manual. Kazan, OOO „Elina computer”, (www.elina-computer.ru), 2011.

Challenges to the Accounting and Supervision of Insurers' Activity

Maia Natchkova

University of National and World Economy, Sofia, Bulgaria
mnachkova@unwe.bg

Abstract. The Directive "Solvency II" of the European Parliament introduces a new regulatory framework with the objective of optimization of the financial stability of insurance companies. It constitutes a new system of insurance supervision in Europe, which shall lead to improved comparability, and controlling of the operation of insurance companies and to establishment of a common supervisory body at Pan-European level. The new supervision model allows intermediate location, recognition and signalling of illiquidity, overindeptedness, and insolvency and liquidation risks of insurance companies. The regulations, principles and concepts of the Directive establish new requirements regarding the accounting and reporting of insurance companies with the objective of unification and standardization of the accounting principles and practices on the field of insurance operation in the European Union. The intent is to establish common harmonized accounting and reporting in order to improve the comparability, the true and accurate recognition of accounting information in financial reports of insurance companies, to improve the management of their assets, liabilities and capital, and to build up precise systems for risk management. Special attention is paid to the specific operation of insurers in a high-risk environment, which leads to the formation of specific objects of financial reporting and audit, preparation of specific so called Supervisory Financial Reports.

Key words: Insurers, insured persons, insurances, fair value, risk management.

1. Introduction

The Directive "Solvency II" of the European Parliament introduces a new regulatory framework with the objective of optimization of the financial stability of insurance companies. It constitutes a new system of insurance supervision in Europe, which shall lead to improved comparability, and controlling of the operation of insurance companies and to establishment of a common supervisory body at Pan-European level. The new supervision model allows intermediate location, recognition and signalling of illiquidity, overindeptedness, and insolvency and liquidation risks of insurance companies. The regulations, principles and concepts of the Directive establish new requirements regarding the accounting and reporting of insurance companies with the objective of unification and standardization of the accounting principles and practices on the field of insurance operation in the European Union. The intent is to establish common harmonized accounting and reporting in order to improve the comparability, the true and accurate recognition of accounting information in financial reports of insurance companies, to improve the management of their assets, liabilities and capital, and to build up precise systems for risk management.

Special attention is paid to the specific operation of insurers in a high-risk environment, which leads to the formation of specific objects of financial reporting and audit, preparation of specific so called Supervisory Financial Reports.

2. Challenges to the Insurers' Accounting

Insurers' specific activity determines the organizational and methodological features of their accounting and the specific nature of their reporting items – assets, capitals, revenue, expenses, contractual relationship. As an information system that registers, processes, controls and analyses the information generated as a result of the insurers' business, accounting is characterised both with its internal organization and with specific external regulation implemented by the state supervisory authority. The Insurance Supervision Division at the Financial Supervision Commission supervises the accounting activity and organization of insurers in order to guarantee the safety of insurance policyholders that they will get insurance protection with regard to their insurance claims upon occurrence of insured events. Irrespective of the specific features of insurance business, its accounting model should adhere to the requirements of the national accounting legislation and of the International Accounting Standards (IAS)/ International Finical Reporting Standards (IFRS). The organization, methodology and methods of insurers' accounting are legally governed by the Accounting Act [1], IFRS 4 Insurance Contracts and the new IFRS 17 Insurance Contracts [2], which will replace IFRS 4, however it is still subject to postponement until its official enforcement, the Code of Insurance [3] and ordinances of the regulatory authority [4]. Insurers' accounting ensures ex-ante, ongoing and ex-post control of the fulfilment of insurers' contractual commitments to insurance policyholders under the insurance contracts they have entered into with, and implements the following procedures:

A. Accounting assessment of short-term (in non-life insurance) and long-term (in life insurance) obligations undertaken by the insurers to the insured persons and to the beneficiaries under the insurance policies they have entered into with;

B. Timely checks for identifying the necessary amount of insurance (technical) provisions determined by the actuaries of the insurance companies;

C. Inspection of the quality of financial assets and financial instruments in which the allocated insurance (technical) provisions are invested and which are used to cover such provisions;

D. Control of solvency, liquidity and return of insurers, which enables them to meet possible future losses as a result of occurrence of assumed insurance risk.

The organization of accounting in accordance with the accounting, investment and insurance policy implemented by the insurers is a complex and continuous process. Such process objectively requires to develop specific rules, principles and procedures in the organization of insurers' accounting that would make it one of a kind and different from the accounting in other sectors and industries of the economy.

The main challenges that determine the future development and improvement of the insurers' accounting information system are as follows:

1) Defining the degree to which legal regulations will perform regulatory functions with regard to the insurers' accounting system, including determining the level of impact of government decisions and tax legislation. Standardisation is considered as a tool for globalization of insurance business, and not as an opportunity to achieve total unification and standardization of its accounting information system at national and world level. Total unification and standardization of accounting are impossible due to the various legal regulations and insurance business regulators; the different fiscal regime in individual countries; the different status and specificities of individual insurers; the different provision of staff that have wide knowledge in the field of accounting, finance, insurance, mathematic modelling and economics;

2) Observance of approved national and international classifications, norms and standards while pursuing insurance business and while organising its accounting;

3) Determining the level of impact of economic theories in the field of insurance on the insurers' accounting system;

4) Obtaining reliable information on the chronology and consistence of the insurance business and its results;

5) Maintaining open information system with view of providing the opportunity to add new accounting items that have occurred in the course of pursuing the insurance business;

6) Identifying the opportunities for carrying out effective ex-ante, ongoing and ex-post control at the accounting system's input and output point;

7) Proper choice of evaluations for assets and insurance liabilities (technical provisions) of the insurers. Such choice is made on the basis of performed analysis of the structure of insurers' assets and capitals;

8) Determining and limiting the amount of losses upon occurrence of unforeseen obligations for payment of large compensations and claims to insurance policyholders due to the large-scale occurrence of the insurance risk;

9) Determining the amount of allocated and free insurance (technical) provisions of the insurers during the current and the previous reporting year;

10) Determining the practices used for reduction of risks – identifying the options for the insurance risk not to occur and for the insurers not to pay insurance compensations or claims to insured persons or beneficiaries;

11) Ensuring the fulfilment of obligations under concluded insurance contracts. The complex set of various factors that determine the development trends of the insurance business have strong impact on the opportunities for investment of insurance provisions and insurers' own funds in financial assets and financial instruments. The covers of insurance (technical) provisions (financial assets and financial instruments) are used as a guarantee to the insured persons that the insurers will fulfil their commitments to them under the insurance contracts they have entered into. Last but not least, it's worth noting the specific requirements to the insurers' accounting imposed by the state supervisory authority in terms of the ways and forms of investing their equity and their insurance (technical) provisions. The legislator defines the types of investments and their amount, as they are of material importance in terms of macroeconomics and should be in compliance with the common interest for the economic development of the society;

12) Making differentiation between the transactions and other events that have occurred after the date of preparation of the insurers' financial statements, which result in recognition of assets and liabilities in the Statement of Financial Position or give rise to contingent receivables and payables, which are only disclosed in their financial statements;

13) Improvement of the organization of analytical reporting of individual accounting items in various breakdowns and options for their reliable comparison and carrying this information forward to the relevant synthetic level. Insurance business serves significant number of entities that often make more than one insurance of different type and characteristics. This defines the need of the high number of analytical methods and procedures used by the insurers in order to obtain as full and as detailed information as possible. The specific conditions for work with insured persons suggest continuous monitoring of their analytical files under separate insurance contracts. The direct accounting control of sub-accounts of clients under different types of insurances is also control over plenty of economic processes carried outside the geographic and economic sector of the insurer's business;

14) Organising off-balance reporting of possible future receivables and future liabilities for payment of claims reported by the insurance policyholders with regard to the probability

of occurrence of insurance risk assumed under concluded insurance contracts. Accounting information for such future receivables may be recorded off-balance as an increase of contingent receivables against increase of total liabilities, and upon possible future payment of reported claims it may be recorded as an increase of contingent liabilities against increase of off-balance total assets. Upon occurrence of actual receipt of such receivable and payment of liability, off-balance contingent commitments should be derecognized and the occurred actual assets and liabilities should be recorded in the Statement of Financial Position (Balance Sheet) of the insurers. By the time of completion of the insurance relationship, which does not usually happen within one financial calendar year, all relationships between the parties are hypothetical ("pending") and therefore no final financial result /profit or loss/ could be formed under the insurance transaction;

15) Disclosure in the financial statements of implemented policies and undertaken actions that are logically bound with the changes occurred in the financial position and the financial results, in the cash flows and the items of equity, in the values of key financial indicators for the insurers' business. Specific set of financial reports and schedules are made on monthly basis in contrast to sectors and reporting units whose activity falls within the calendar annual cycle;

16) Maintaining sufficient capital provision of the insurance company;

17) Obtaining objective and full, timely and reliable, true and fair, authentic and logical information about the financial position of insurers, which can be used for making effective management decisions with view of reducing the relevant adverse effects in their business. "True and fair presentation require reliable disclosure of the effects of transactions, other events and conditions in compliance with the definitions and criteria for recognition of assets, liabilities, revenue and expenses, as set out in the applicable accounting standards." [5];

18) Protection of the insurers' property;

19) Improving the competitive power and return of insurers and reinforcing their position on the insurance market;

20) Conservative application of prudential principle when reporting the insurers' business for the purposes of avoiding collapse of their business as a result of the world crisis due to the COVID-19 pandemic. Under the conditions of crisis, insurers must foresee additional options for financial support and corporate funding from the shareholders, from corporate groups that control the insurer, from the reserve funds allocated during the years, from the government via subsidies and governmental grants, from reinsurers by entering into reinsurance contracts, or from banks via loans for their business, etc.

While developing their accounting policy, insurers should take into account the following key rules for reporting of their specific business:

1. Proper classification of their own funds as fixed and additional;

2. Verification of the availability of own funds and their ability to cover losses in case of adverse development of the insurance risk;

3. Verification of the classification of their own funds by tiers – capital of first, second and third tier, and whether their entry to the respective tiers corresponds to the solvency and capital adequacy requirements as set out in the Directive "Solvency II" of the European Parliament, and namely:

a. the eligible amount of tier 1 should be higher than one third of the total amount of eligible own funds,

b. the eligible amount of tier 3 should be lower than one third of the total amount of eligible own funds,

c. the eligible amount of own funds for covering the solvency capital requirement should be equal to the sum of the total amount of capital of tier 1, the eligible amount of capital of tier 2 and the eligible amount of capital of tier 3. [6];

4. Verification of the excess of funds – whether they are reported as accumulated profits that have not been allocated for distribution to the insured persons or third parties (policyholders and beneficiaries);

5. Verification of the amount of eligible own funds – whether it has dropped under the minimum capital requirement and whether there is a risk of such drop during the next three months, in order to avoid the insurer becoming insolvent and over-indebted and being declared bankrupt or liquidated;

6. Testing the adequacy of liabilities for the purposes of verifying the availability of required insurance (technical) provisions that are necessary to cover their liabilities to the insurance policyholders;

7. Calculation of insurance (technical) provisions based on the present value of insurance liabilities while using the risk-free rate of return. Specific risk margin should be added to it and the cash flows should be discounted with the resulting amount. The fair value is calculated on the basis of the applicable market value. Due to the lack of active market of insurance (technical) provisions, they should be measured at fair value by means of the discounting methods;

8. Verification of the methods for insurance (technical) provisioning applied by the actuaries. Such verification refers to the consideration of all acquisition costs incurred during the administration of insurance liabilities. These are costs for claim settlement of insured persons, all payments to them, including future voluntary bonuses that the insurers expect to pay to insured persons, whether such payments are guaranteed by contract or not;

9. Segmentation of insurance liabilities in homogeneous risk groups by types of activities, by geographic, organizational and sectoral features for the purposes of their proper evaluation;

10. Taking into account the effect of inflation risk when evaluating insurance liabilities;

11. Observance of the principle of market coordination when calculating insurance liabilities. To this end, the information from the financial and insurance markets is taken into account for:

- the changes of the market prices of financial assets and financial instruments,
- the changes of market interest rates,
- the changes of the market prices of real estates,
- the changes of the exchange rates and credit spreads over the reference interest rate,
- the adverse changes of the frequency of occurrence of insurance risks,
- the unusual accumulation of risk due to emergency situations,
- the increase of the level of risks related to the change of insurance administration costs,
- the increase of the risks as a result of the world crisis due to the announced COVID-19 pandemic, etc.;

12. Taking into account all unrealized profits that occur on the basis of changes in the fair value of insurance (technical) provisions, directly in current profit or loss. Verifying the possibility of these unrealized profits to influence the insurers' equity in the form of hidden capital (revaluation) reserves of their insurance liabilities formed as a result of their fair value measurement. Otherwise, if such unrealized profits from fair value measurement of insurance liabilities are stated as capital (revaluation) reserves, they could result in outflow of insurers' funds in the form of payment of dividends to investors and could cause drop in their liquidity and deterioration of their solvency.

3. Challenges to the Supervision of the Insurers' Business

The state supervisory authority – Insurance Supervision Division at the Financial Supervision Commission carries out continuous ex-post control of:

A. the observance of contractual commitments by the insurers to the insurance policyholders,

B. the sufficiency of allocated insurance provisions,

C. the fair market assessment of obligations undertaken under concluded insurance contracts,

D. the qualities of assets the insurance provisions are invested in,

E. the insurers' solvency limit, which allows them to meet probable future losses due to the occurrence of assumed insurance risk.

The state supervision allows early identification, recognition and alert for the risks of insurers becoming insolvent, over-indebted, bankrupt and liquidated, for the purposes of protecting both the insureds' and insurers' interests. On the other hand, the state supervision over the insurers' business guarantees the stability of insurance and financial markets. The insurance supervisory authority needs to be flexible in its external control thus being able to react quickly to the changing conditions of the insurance market due to the global pandemic crisis. The terms and conditions of such control are set out in the Directive "Solvency II" of the European Parliament. The Directive is aimed at creating common basis for accounting and supervision of insurers' business in the European Union, while taking into account the specificities of the national legislations. The International Association of Insurance Supervisors /IAIS/ requires additional, careful, prudent and transparent assessment of insurers' business. To this end, Directive "Solvency II" of the European Parliament:

- defines the principles for assessment of insurance liabilities (allocated insurance provisions) in the reserve fund;
- gives guidance for the covers of these provisions (assets in which insurance provisions are invested);
- determines the key requirements to insurers' equity.

There are three main areas (pillars) in which the rules set out in Directive "Solvency II" of the European Parliament are grouped:

- **Pillar 1** – it comprises quantitative requirements for capital adequacy, for assessment of assets and liabilities, for insurers' own funds. This pillar defines the Solvency Capital Requirement, Minimum Capital Requirement and technical provisions;
- **Pillar 2** – it comprises qualitative requirements and supervision. It defines the rules for establishment of internal control and risk management system – requirements to the persons charged with the governance of the insurance company; to the risk management to the internal control; to the independent financial audit; to the actuarial function. This pillar sets out that the quantitative requirements for the amount of insurance (technical) provisions, their covers (the assets they are invested in) and equity should be transformed in qualitative changes, i.e. to develop risk protection models that take in consideration the individual situation of common risks of every insurer. Such situation would result in determination of new requirements to the insurers' guarantee capital. Such qualitative requirements are firmly bound with the individual risk profile of the respective insurance company. This pillar provides for the condition that insurers should make own risk and solvency assessment for the purposes of achieving stable and reasonable management, as well as assessment of the general solvency requirements with view of their specific risk profile. The development of specific internal risk management model that is compliant with the relevant insurer precisely determines its own

one of a kind risk profile. The implementation of a single standard risk management model may not be applied to each individual risk profile. An internal model reflects the actual risk situation, as by rule, it corresponds to the respective insurer. The new supervisory system is characterised by shifting the burden from a quantitative to a qualitative supervision, more strongly oriented to the individual risk profile of the insurer and with the requirement for an active risk management. The entire system of risk management and control processes in the companies is called ORSA – own risk and solvency assessment. The provisions of this pillar refer to insurers with regard to their abilities to have internal control methodology, which should be defined more precisely. It gives main guidelines for its establishment with regard to risk management, insurance contract administration management, claim handling management and allocations to the reserve fund (allocation of insurance provisions); business management (liquidity, solvency and return requirements, taking into consideration of the relationship between the asset and the liability risk profile) and the reinsurance management (having adequate reinsurance program corresponding to the reinsurer's solvency);

- **Pillar 3** – it comprises accounting and public disclosure of information from the insurers' financial statements. It further provides for the requirements and guidelines to reporting for supervision purposes and obliges the insurers to publicly disclose annual reports that contain information about their solvency, performance and financial position. Thus, it ensures transparency on the insurance market, achieves the desired market discipline and implements a reliable risk management. Insured persons and financial market should be informed about the assets of an insurer and the related risks. Depending on their amount, assessment and status, insured persons should be able to make informed decisions whether to enter into contracts with this insurer. Such transparency is aimed at improving the market disciplines of supervised entities.

The objective of Directive "Solvency II" of the European Parliament is to achieve:

A. Financial stability and risk management – to ensure financial stability of insurers by introducing specific solvency requirements that take into account the insurance, market, credit and operational risks;

B. New approach for supervision practices – to change the way supervisory authorities work. It introduces new requirements for moving from simple compliance assessment of capital requirements based on a set of rules to proactive supervision of insurers' risk management based on a set of principles.

C. *Group supervision* – to introduce a new group supervision system for the insurers' business aimed at improving supervision and risk management.

The main challenges to the supervisory authority under the conditions of the world COVID-19 pandemic refer to the stronger need for it to audit:

1) Insurers' material violations of the legal, regulatory or administrative provisions, which determine the conditions for issuing and observance of the license for pursuing insurance business;

2) Breach of the Solvency Capital Requirements;

3) Breach of the Minimum Capital Requirements;

4) Breach of the insurer's continuous business, i.e. observance of the "going concern" principle of accounting;

5) Lack of ORSA – Own Risk and Solvency Assessment;

6) Strict monitoring of the financial position and determining the financial performance of the insurer and its solvency limit, with view of ensuring timely reaction upon occurrence of adverse trends. The significant decrease of the insurer's key financial stability and solvency indicators may be an indicator to this effect,

7) Improper determination of the solvency limit due to the high levels of risk of insolvency, overindebtedness, bankruptcy and liquidation of insurers;

8) Breaches of the internal model adopted by the insurer to bind its minimum required capital with its insurance (technical) provisions and to calculate its guarantee capital, resulting in decrease below the minimum required fixed capital and in loss of insurance license;

9) Audit of the insurer-specific insurance aggregate diversification in comparison to the general situation in the sector, branch or industry (e.g. existence of low capitalization in the sector).

10) Audit of the relationship between the fair value and the asset and liability management. Non-conformities occur upon stating items in the insurers' balance sheet according to the present value method, where the two sides of the balance sheet are not well balanced between each other (i.e. mismatching). On the other hand, stating of items in the balance sheet according to the present value method identifies the mismatches between assets and liabilities at an early stage, which enables the insurance supervisory authority to obtain timely information about the weaknesses of the audited insurer's asset and liability management and to order the respective remedial measures before declaring the insurer insolvent. The establishment of working portfolio of quickly returnable assets facilitates the insurer's asset and liability management. Thus, the insurer will be obliged to implement high-quality value added asset and liability management not only by the supervision according to the second pillar of Directive "Solvency II" of the European Parliament, but by the mechanisms of the market of assets as well.

11) Audit of the special Solvency and Financial Position Statement made by the insurers. The audit defines whether the insurers have provided in this financial report accessible and complete in all material aspects, comparable and consistent in long-term aspect, appropriate, reliable and understandable information for:

- the nature, volume and complexity of the business and its results;
- description of the management system and assessment of its adequacy in terms of the insurer's risk profile;
- separate description of the risk exposure, concentration, decrease and sensitivity of each category of risk;
- separate description by assets, insurance-technical provisions and other liabilities of the legal grounds and methods used for their calculation and the mismatches of their use;
- description of capital management, including description of the structure and amount of own funds; their quality; the solvency capital requirement and the Minimum Capital Requirements); the amount of any breach of the Minimum Capital Requirements and of any significant breach of the Solvency Capital Requirements during the reporting period, notwithstanding the fact that such breach could be remedied afterwards; explanations about the origin of the breach, the consequences of its occurrence and the remedial actions undertaken by the insurer.

Directive "Solvency II" of the European Parliament allows the insurers not to disclose in the Solvency and Financial Position Statement specific confidential information relevant to the obtaining of unjustified competitive advantages by their competitor insurers. To this end, they should obtain special permit from the insurance supervisory authority. The very importance of insurance at macro-level is another prerequisite for the need of ongoing control on the financial position of the respective insurer. Therefore, the results of this regular monthly reporting should be strictly monitored and controlled not only by the respective insurer but by the state authorities as well, and namely, Insurance Supervision Division at the Financial Supervision Commission.

The world crisis that has occurred as a result of the announced COVID-19 pandemic proved the effectiveness of Directive "Solvency II" of the European Parliament for the protection of the insurance sector against market shocks. While supervising the business of the insurers, the insurance supervision authority should to sufficient extent take into account the nature of the long-term insurance business and whether there is a need of provoking changes in terms of illiquid liabilities, capital risk or risk margin, which would improve the risk sensitivity, facilitate the projection of long-term illiquid liabilities and encourage long-term investments. It should require to provision of adequate guarantees for protection of the insured persons' interests.

4. Conclusion

Every insurer is exposed to the effect of different risks. Under the conditions of crises, the need of use of risk management systems by the insurers become more than great. Insolvency cannot be avoided despite the sufficient capital adequacy of the insurer if its risk management is insufficient or non-compliant with its internal control and the external control carried out by independent financial auditors and by the insurance supervisory authority. The synchronized application of IFRS 17 Insurance Contracts and the Directive "Solvency II" of the European Parliament will in any case result in radical restructuring of accounting, independent financial audit and public supervision of the insurers' business.

Good performance may be achieved through skilful planning, development of reliable business plan that is fully compliant with the needs of the insurance market, prices, trends and integration of the national and international insurance business, to enable insurers to survive and flourish under the market economy conditions, be financially sound, solvent, and have good financial results.

References

1. Accounting Act, promulgated in State Gazette (SG) no. 95 of 08.12.2015, effective from 1.01.2016, last amendment and supplement, no. 19 of 05.03.2021;
2. International Financial Reporting Standards (IFRS), EU Official Journal, Commission Regulation (EU) 2016/2067 of 22 November 2016, L323; Commission Regulation *(EC) No 1126/2008* of 3 November 2008 adopting certain international accounting standards in accordance with Regulation (EC) No 1606/2002 of the European Parliament and of the Council, as last amended, IFRS 4 *Insurance Contracts*, paragraphs 15 – 19; Bases for conclusions, paragraphs OZ94 to OZ104;
3. Code of Insurance, promulgated in SG no. 102 of 29.12.2015, effective from 1.01.2016, last amendment and supplement, no. 21 of 12.03.2021;
4. Ordinance No 53 of the Financial Supervision Commission of 23.12.2016 on the requirements to reporting, assessment of assets and liabilities and formation of technical provisions of insurers, reinsurers and the Guarantee Fund, promulgated in SG no. 6 of 19.01.2017, effective from 19.01.2017, last amendment and supplement no. 2 of 04.01.2019;
5. Ivanova, R., Chukov, Kr., Ivanova, R., *Financial and Business Analysis*, UNWE Publishing House, Sofia, 2019, page 30
6. Solvency II Directive of the European Union, 22.04.2009, article 100

Classification, Comparison and Criteria for Choosing NoSQL Databases

Mariana Kovacheva

University of National and World Economy, Sofia, Bulgaria
mkovacheva@unwe.bg

Abstract. This paper presents the classification types of NoSQL databases and outlines the main similarities and differences between them. In the paper are included basic features of Document-oriented databases, Key-value databases, and Column-oriented databases. Development of technology and digital transformation generates great amounts of data of all types and because of that more storage is needed, which NoSQL databases can provide. Based on the classification of NoSQL databases, some of them are examined more thoroughly. In the paper a comparison between MongoDB, Apache HBase, Neo4j and Amazon DynamoDB is presented. Criteria for choosing NoSQL databases is introduced.

Keywords. NoSQL, Document-oriented databases, Key-value databases, Column-oriented databases

1. Introduction

Technologies are developing with fast rates and every device nowadays generates data every second that is working. Digitalization covers new processes of peoples lives everyday with which the data that is collected increases and needs new ways for storing. NoSQL databases are able to adapt to the needs of the new technologies to store different types of data as not only structured, but semi-structured and unstructured.

2. Classification types of NoSQL databases

NoSQL database is a non-relational database, which uses unique keys to store data instead of storing it in traditional way in rows and columns. The NoSQL databases provide more flexibility because they do not have structured schema and do not define each table and column. Non-relational databases and especially NoSQL, are used for large volumes of data and when flexible models are used. Operations in them are faster, because the data is stored in nodes, and it's processed there. [1] [2]

According to the classification types, NoSQL databases can be Key-Value pair, Column-oriented, Document-oriented and Graph based. [3] NoSQL databases have a lot of similarities as some differences between the different types. This type of databases can be scaled horizontally if needed and depending on the volumes of data need to be stored. All NoSQL databases have a flexible schema and can process all types of data – structured, semi-structured and unstructured. Depending on the problem that needs to be solved, a type of NoSQL database can be chosen. For this aim in the paper will be presented the main characteristics of the four types of NoSQL databases.

The Key-Value (KV) stores is the simplest type of NoSQL database as data is stored in key/value pairs, and this type of database is optimized for reading and writing data. The key serves as a unique identifier. The key-value stores maintain the data as pair which consists of an index key and value. The KV stores has some resemblance with the relational databases, but each table has only 2 columns. Examples of KV stores are Amazon DynamoDB, Aerospire, Couchbase, Redis, etc. [4]

Column – oriented databases organize information into columns. This type of NoSQL database functions the same way as relational one, but the data model is much more flexible and scalable. The storage by rows makes this type of database a lot faster and more efficient. Column – oriented databases are used in businesses that send massive amounts of data from multiple sources. Examples of column-oriented databases are Apache HBase, MariaDB, Apache Kudu, Apache Parquet, etc. [5]

Document-oriented databases store data as documents. Instead of storing data into tables/rows and columns/ as in the relational databases, document-oriented use flexible documents. Document is the record in this type of NoSQL database and usually stores information about an object and its related metadata. The data is stored into field-value pairs and these values can be of different types and structures. Documents can be stored into XML, JSON, BSON, etc. Examples of document-oriented databases are MongoDB, DynamoDB, PostgreSQL, etc. [6]

Graph databases are built to store and navigate relationships. This type of NoSQL databases uses nodes and edges. The nodes are used to store data entities. The edges are used for storing relationships between the entities. Examples of Graph databases are Neo4j, Amazon Neptune, IBM Graph, etc. [7]

3. Features of the four types of NoSQL databases

3.1. Features of Key-value database

- **Scalability**

Key-value databases are highly partitionable, allow horizontal scaling at scales that other types of databases cannot achieve. These types of databases are more flexible and have faster read and write process. The database searches for a single key and returns its value.

- **Consistency**

For the key-value database to be consistent it's important the data through all nodes to be updated. The KV databases that use replication must ensure that all clones are consistent, especially when retrieving information from a particular key.

- **Transactions**

The number of transactions that can be performed for a given period.

- **Availability**

The data stored in the database must be always available for the readers and writers.

- **Partition tolerance**

The database should continue functioning even when there are some failures and because of that the data should be stored not only in one place, but in multiple. This will help the system to continue functioning.

3.2. Features of Column-oriented database

- **Scalability**

Column-oriented databases have excellent scalability, as they can be expanded numerous times over clusters.

- **Flexibility**

The columns in the column-oriented database do not have to look like one another. New and different columns can be added without distressing the whole database.

- **Load time**

Data is easily loaded, which means that big amounts of data could be loaded very fast.

3.3. Features of Document-oriented database

- **Model**

In this type of database, the data is stored into documents and this type of database is with flexible schema. This allows that not all documents will have the same schema, but of needed document-oriented databases allow schema validation.

- **Flexibility**

Document-oriented databases have flexible schema, which means that there is no need of predefining it and it is adaptable to change. Any changes can be made to the design without any problem for the structure.

- **Distribution**

Document-oriented databases allow horizontal scaling and data distribution. Horizontal scaling is cheaper than the vertical one.

3.4. Features of Graph database

- **Flexibility**

Graph databases have flexible schema and you do not need to predefine it. This is one of the best advantages of the Graph database as additional graphs can be added without interrupting the current structure.

- **Performance**

The performance of the graph databases stays steady even when the loads of data is growing.

- **Agility**

This type of database develops with the needs of the business and the applications that uses it. Graph databases can be easily developed for the aims of the business, and they are easy for maintenance.

4. Comparison between MongoDB, Apache HBase, Amazon DynamoDB and Neo4j

4.1. MongoDB

MongoDB is document-oriented database, which uses JSON like documents. MongoDB works with the concepts of collection and documents. The documents are sets of key-value pairs and do not need any predefined schema, which allows for them to have different structure even when they are in the same collection. [10]

4.2. Apache HBase

Apache HBase is column-oriented NoSQL database. The data is stored into individual columns, which are indexed by a unique row key. HBase runs on top of Hadoop. Apache HBase is very effective for managing large datasets, and because of the work based on HDFS, both the data and requests are distributed across HBase clusters, which gives an opportunity for receiving results from queries in seconds. [11]

4.3. Amazon DynamoDB

Amazon DynamoDB is a NoSQL database which supports strong key-value and document structures. DynamoDB uses hashing and B-trees for managing the data. When data is entering, firstly it's distributed into different partitions by hashing on the partition key. Amazon DynamoDB is designed to run high performance apps at any scale. This NoSQL database has reliable performance, even when scaling is needed. [12] [13]

4.4. Neo4j

Neo4j is a graph NoSQL database in which the data is stored instead in rows and columns, it has nodes, edges, and attributes. Each node and edge can have any number of attributes. The nodes are the entities in the graph that can be tagged with labels. Relationships provide connections between two node entities. They always have direction, type, start and end node. As the nodes can have any number or type of relationships without making the performance worse. [14]

Table 1. Comparison of NoSQL databases [15]

	MongoDB	Apache HBase	Amazon DynamoDB	Neo4j
Database model	Document - oriented	Column -oriented	Key-value store	Graph based
Description	Uses JSON like documents and works with the concepts of collection and documents. The documents are sets of key-value pairs and do not need any predefined schema, which allows for them to have different structure even when they are in the same collection.	The data is stored into individual columns, which are indexed by a unique row key. HBase runs on top of Hadoop and is very effective for managing large datasets, and because of the work based on HDFS, both the data and requests are distributed across HBase clusters, which gives an opportunity for receiving results from queries in seconds.	DynamoDB uses hashing and B-trees for managing the data. When data is entering, firstly it's distributed into different partitions by hashing on the partition key. DynamoDB is designed to run high performance apps at any scale. It has reliable performance, even when scaling is needed.	Instead in rows and columns, this type of database has nodes, edges, and attributes. Each node and edge can have any number of attributes. The nodes are the entities in the graph that can be tagged with labels. Relationships provide connections between two node entities.
Schema	Schema - free	Schema - free + Schema possible	Schema - free	Schema - free + Schema possible
Transaction concepts	Multi-document ACID	Single-row ACID	ACID	ACID
Partitioning	Sharding	Sharding	Sharding	when using Neo4j Fabric
Replication	Multi-source deployment Source-replica	Multi-source Source-replica	Yes	Causal Clustering only on enterprise version
MapReduce	Yes	Yes	No	No
Companies using the DB	Toyota, Forbes, Humana, Trade Ledger, Expedia, Bosch, Royal Bank of Scotland, CISCO, SAP, Adobe, etc.	Adobe, Astro, Facebook, Twitter, Yahoo	Amazon, Zoom, Disney, Snap, Dropbox, Netflix, Samsung, Airbnb	eBay, Adobe, Department of Education UK, IBM, Microsoft, Volvo, Telenor, NASA, Cisco, Swiss International airlines, HP

5. Criteria for choosing NoSQL databases

Data storage is becoming more and more important and because of all the amounts of data that is generated daily, even every second of the day, companies and clients need good criterions based on which to choose the correct NoSQL database for their needs. [8] [9]

Different NoSQL data is stored in different types of NoSQL databases, i.e., they differ by the model they have. If the wrong or inappropriate model of data is chosen, the application using it might not be successful in its work afterwards.

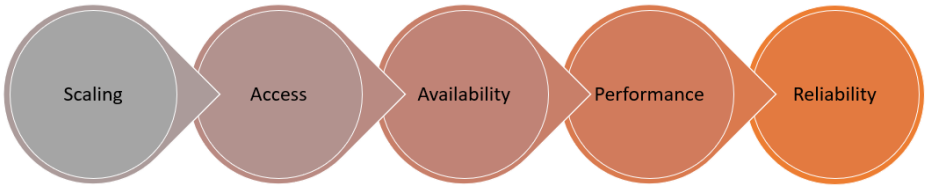


Fig. 1. Criteria for choosing NoSQL databases

5.1. Data Scaling

Other thing that needs to be considered is the expected growth and scale of the application. A NoSQL database should be capable of efficient and easy data scaling.

5.2. Access

One of the main requirements when choosing a NoSQL database is how the data will be accessed and what will be the requirements of the application.

5.3. Availability

The most important points to be covered when talking about data availability and NoSQL databases are the architecture of the database, consistency, and replication. There shouldn't beany shared resources and points of failure.

5.4. Performance

The performance of the NoSQL databases should meet the requirements of the users, such as fast reading and writing process, efficient use of memory. Performance is important for the cost. When a NoSQL database is high performing, it needs less hardware and nodes, which means that the support is not so expensive.

5.5. Reliability

Reliability shows the probability of a system to work without any problem/failure for a period as we talk about reliability both in terms of robustness and response times. Still NoSQL databases lack to provide reliability and consistency, even though they provide scalability, fast access, and high performance.

6. Conclusion

One of the most important decisions to be taken nowadays it's not why to choose between SQL and NoSQL, but which NoSQL database to chose for the needs of the application that will be supported by it and for the organization. Non-relational databases are the best

option for storing semi-structured and unstructured data, which are the types of data that are mostly generated from the newer systems. The NoSQL databases are better for using when there is no clear predefined structure for the database of the system. Information can be always added without any problem because there is no need for the data to be in the same structure as the other. NoSQL databases gives us the opportunity not only to store unstructured data but to work with dynamic schema with all the types – column-oriented, key-value, document-oriented and graph-based.

References

1. "NoSQL," NoSQL, Jun. 30, 2020. https://www.techopedia.com/definition/27689/nosql-database
2. "NoSQL Definition," Aug. 27, 2013. https://techterms.com/definition/nosql
3. S. Sayyad, "Types of NoSQL Databases," Types of NoSQL Databases. https://www.educba.com/types-of-nosql-databases/
4. A. Willaims, "Key-value store databases," Key-Value Databases, Explained, Apr. 08, 2021. https://www.kdnuggets.com/2021/04/nosql-explained-understanding-key-value-databases.html#:~:text=Examples%20of%20Popular%20Key%2DValue%20Databases&text=Amazon%20DynamoDB%3A%20Probably%20the%20most,optimized%20for%20in%2Dmemory%20storage.
5. A. Williams, "Column-oriented databases," Column-Oriented Databases, Explained, Feb. 12, 2021. https://www.kdnuggets.com/2021/02/understanding-nosql-database-types-column-oriented-databases.html
6. "Document-oriented database," What is a Document Database? https://www.mongodb.com/document-databases#:~:text=A%20document%20database%20(also%20known,for%20developers%20to%20work%20with.
7. "Graph databases," What is a Graph database? https://aws.amazon.com/nosql/graph/
8. "NoSQL Database Evaluation Guide." Couchbase. [Online]. Available: https://info.couchbase.com/rs/302-GJY-034/images/NoSQL_Evaluation_Guide.pdf
9. Navdeep Singh Gill, "NoSQL Databases Overview, Types and Selection Criteria," Feb. 12, 2019. https://www.xenonstack.com/blog/nosql-databases
10. "MongoDB - Overview," *MongoDB - Overview*. https://www.tutorialspoint.com/mongodb/mongodb_overview.htm
11. "What is Apache HBase?" https://aws.amazon.com/big-data/what-is-hbase/
12. "What is DynamoDB?" https://www.dynamodbguide.com/what-is-dynamo-db/
13. "Amazon DynamoDB." https://aws.amazon.com/dynamodb/
14. "What is a Graph Database?" https://neo4j.com/developer/graph-database/
15. "System Properties Comparison." https://db-engines.com/en/system/Amazon+DynamoDB%3BHBase%3BMongoDB%3BNeo4j

Trends in Leadership Training in Bulgaria

Tanya Pancheva

Sofia University St. Kliment Ohridski, Sofia, Bulgaria
t.pancheva@abv.bg

Abstract. Conducting an empirical study is extremely important to achieve the objectives of the present study. It can answer a number of questions about trends in leadership training. In accordance with the object of research - leadership training programs, the research proposes the development of a specialized survey card. Finally, conclusions and recommendations are outlined.

Keywords. Empirical study, leadership training, training.

1. Empirical analysis and trends

The respondents included in the present study are organizations that conduct leadership training. The current list of respondents participating in this study is available in the records of the Ministry of Education and Science at the following Internet address: https://rq.mon.bg/programs/approved-programs [1].

After the respondents were identified, an empirical study was conducted. It covers issues of leadership training. The results obtained are:

As stated above, leadership training can cover a variety of topics, as well as the needs of the trainees are different. This provides a number of explanations regarding the answers to question 6. As can be seen from Fig. 1, 10 organizations, which represents 45.5% of the respondents, answered that the lecturers during the leadership trainings are invited, according to the request of the users. At 7 of the organizations, the lecturers are permanent collaborators of the organization. This represents 31.8% of the respondents, and these responses indicate that organizations that conduct leadership training apply flexible approaches in the interest of the trainees.

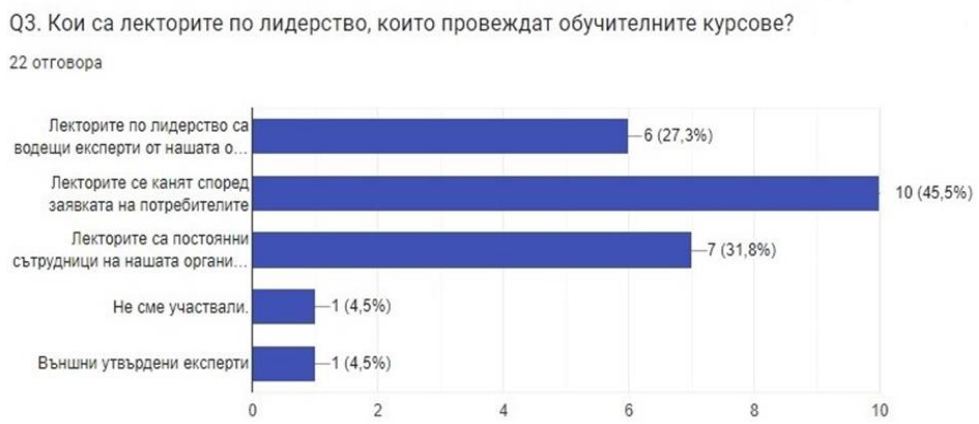


Fig. 1 Responses received to "Who are the leadership lecturers who run the training courses?"

Evidence in this direction is the convincingly given answers to question 7 about the focus of leadership training. For 86.4% of respondents, it is aimed at acquiring both knowledge and skills. This is of great importance for achieving high efficiency of the training process (Fig. 2).

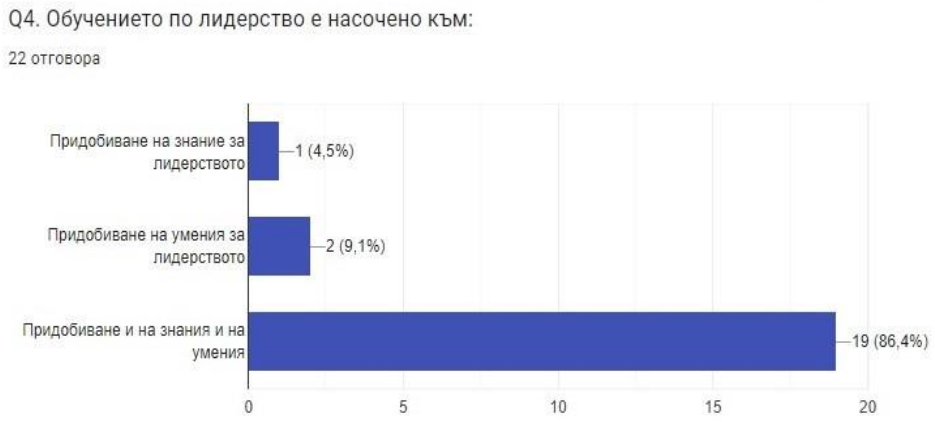


Fig. 2 Responses received to the question "Leadership training is aimed at"

Regarding the main goal of the training process, the respondents are divided in the following two directions: for 81.8% of the respondents, the main goal of the offered leadership training is to improve the individual performance of the course participants in view of the specifics of the team, the goals and the environment. In 18.2% of the organizations, the main goal is to improve the individual performance of each participant by himself (Fig. 3).

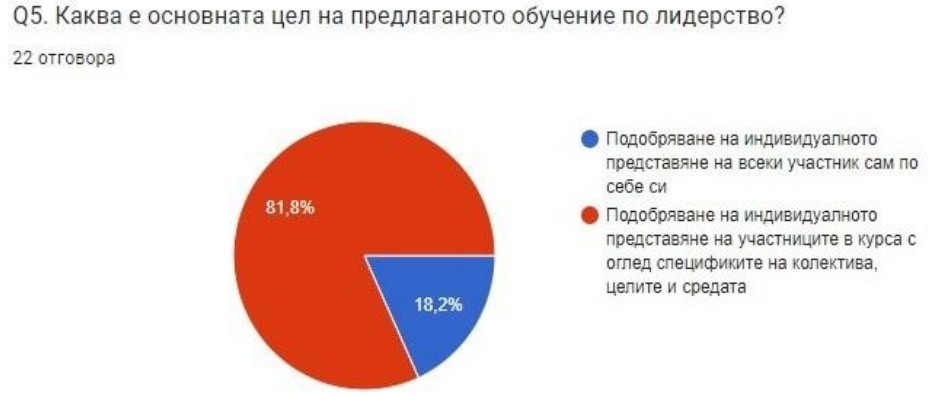


Fig. 3 Responses received to the question "What is the main purpose of the leadership training offered?"

It is obvious that such leading organizations, which have had the opportunity to conduct multiple leadership trainings over the years, use a certain methodology for this training. In view of the goals set in the dissertation work, it is important to establish the origin of this

methodology, as well as its adaptation and parameterization over time. When conducting the empirical study, more than half of the organizations (54.5%) indicated that the methodology used was an adapted form of known methodologies. 36.4% of the respondents use a specific own methodology, and 9.1% apply a known classical methodology in the process of leadership training. These results can be seen from the diagram presented in Fig. 4.

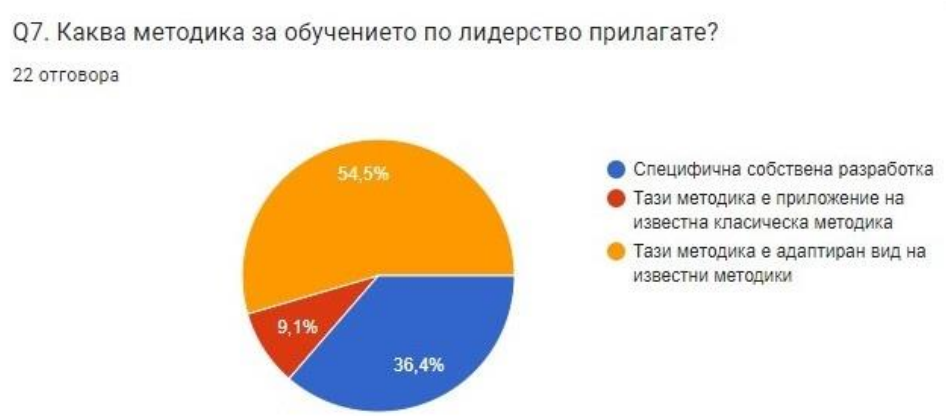


Fig. 4 Answers received to the question "What leadership training methodology do you use?"

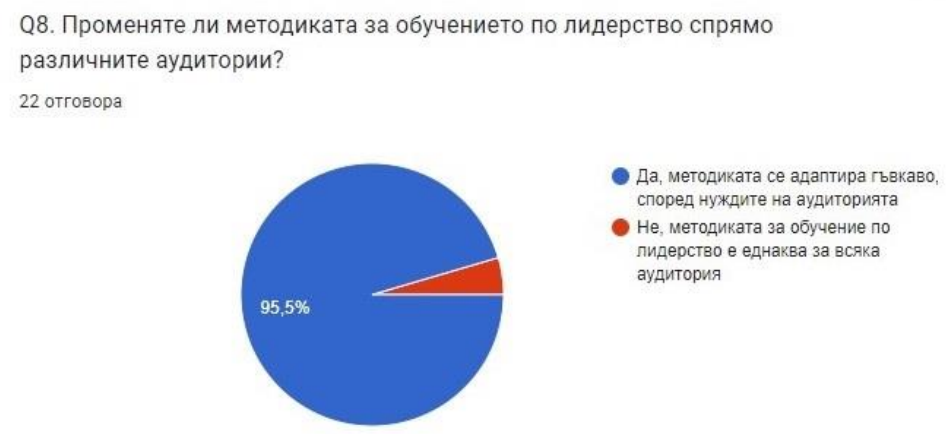


Fig. 5 Responses received to the question "Do you change the methodology of the leadership training according to the different audiences?"

In Fg. 5 presents the received answers to the question "Do you change the methodology for leadership training in relation to different audiences?" As can be seen from the presented data, in 95.5% of the organizations it is constantly necessary to change and parameterize the methodology for training in relation to different audiences. Only 4.5% indicate that such a change is not necessary. Applying innovative and flexible approaches to the target audience is a very important tool for the success of any training. This determines the next question in the questionnaire, namely: "Do the trainers have a specific toolkit for checking the obtained result?". The results are presented graphically in Fig. 6.

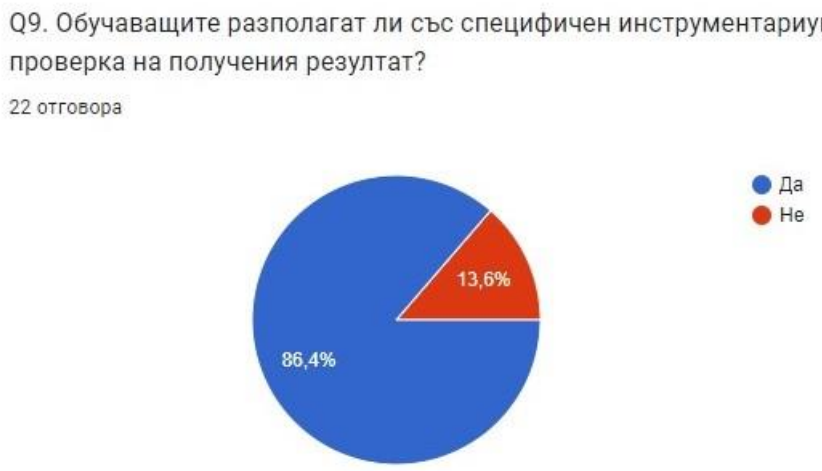


Fig. 6 Received answers to the question "Do trainers have a specific toolkit for checking the result obtained?"

The answers to the questions also outline the main trends. They are in the following directions:

- Development of leadership training methodologies;
- Offering attractive training programs;
- Attracting quality specialists for lecturers.

2. Conclusions

As a result of the research and empirical analysis, the following conclusions can be drawn:

- Systematic research on leadership training is important;
- Broader leadership study models offer greater learning opportunities.

References

1. Internet site https://rq.mon.bg/programs/approved-programs.

New Requirements that Insurers Face in a Situation of Global Health Crisis

Simeon Grigorov

University of National and World Economy, Sofia, Bulgaria
simeongrigorov@abv.bg

Abstract. The business of insurance companies is crucial for the society all over the world. They accumulate and manage a serious cash in with the care of good trader and cool investor, while should keep the principles of reliability, liquidity and diversification of their portfolios in interest of all the insured. In a result of the announced COVID-19 pandemic, the insurers are facing a serious problem related to the undertaken by them insurance risks. Based on the developed and always changing accounting principles, rules, specific practical procedures that they have adopted and that determine their accounting, investment and insurance policies, insurers may adequately address the changes that have occurred in their insurance business at global level in a situation of COVID-19 pandemic. The aim is to minimize the occurred risk, to limit the events risk possibilities, and to keep adequate liquidity.

Key words. Insurers, insured persons, financial statements, crisis, COVID-19 pandemic.

1. Introduction

This publication focuses on the challenges facing the insurance business as a whole, as well as on the new requirements for the specific accounting of insurers in the context of the global health crisis caused by the COVID-19 virus. Special attention is paid specifically to the new requirements that insurers face in a situation of global health crisis, on the one hand, and the high-risk environment in which they conduct their business, on the other. These specific conditions require the emergence of new requirements and significant changes in existing concepts, rules and principles of accounting, analysis and management of the insurance business.

The business of insurance companies is crucial to societies all around the world. They accumulate and manage serious funds with the care of a good trader and investor, while respecting the principles of reliability, liquidity and diversification of their portfolios in the interests of all policyholders..

The crisis caused by the COVID-19 pandemic continues to have a significant impact on people, society, business and the world economy. The insurance business is certainly no exception. Due to the importance of the insurance business for society, insurers must maintain both its continuity and the smooth service of insured persons and third parties benefiting from insurance during the crisis. The insurance sector cannot stay away from all the problems caused by the pandemic, but unlike other sectors and sectors of the country's economy, it responds quickly to changing economic conditions following the crisis.

The process of recovery of the global economy from the pandemic faces insurers with a number of challenges, but also offers them many new opportunities to develop their business in the medium and long term. As a result of the declared pandemic by COVID-19, insurers are facing a serious problem related to their insurance risks. However, their main goal remains

unchanged - offering insurance protection, assuming insurance risk and settling the claims of insurance policy holders in the event of the insurance event at the expense of the formed insurance money fund from the collected insurance premiums. They can achieve this goal through the following levers:

A) Reduction of the degree of manifestation of the insurance risk and prevention of its mass realization.

B) Limitation of damages from insurance events.

C) Achieving high profitability and good liquidity so that they can easily meet their obligations to insured persons and third parties benefiting from insurance.

2. Impact of the COVID-19 Pandemic to the Insurance Business

The impact of the COVID-19 pandemic on the insurance business is significant. It varies depending on the products and types of coverage offered by insurers and has a significant impact on the amount of insurance premiums on some insurance lines such as travel insurance, cultural and entertainment events, tourism and trade credit due to expectations of significant losses from these businesses. . Other insurance lines, such as motor insurance and property insurance, remain relatively stable in times of crisis. The amount of personal / personal / insurance claims / claims / has decreased significantly due to the imposed lockdowns, although there is a new trend of frequent unscrupulous claims, to which insurers should pay special attention.

The COVID-19 pandemic has led policyholders and insurers to question the need for or usefulness of certain insurance products, such as travel insurance. It is no longer needed by many insured persons or it cannot cover all insurance risks in the new reality. For this reason, insurance companies should evaluate their products and, if they see signs of unfair treatment as a result of the crisis, then take neceserry corrective action.

Undoubtedly, the most affected sector of the pandemic is the one of health insurance. „They face a serious threat under the conditions of a global health crisis due to the announced COVID 19 pandemic. As a result of this pandemic, they assume additional liabilities under the concluded health insurances that cover health insurance risks related to diseases caused by the COVID 19 virus. There is a large-scale occurrence of the assumed health insurance risk under the conditions of global health crisis, which gives rise to an urgent need of sufficient liquid funds for payment of large-scale compensations for the healthcare activities, medical goods and services and subscription healthcare service used by the health-insured persons and by the beneficiaries to the health insurances.” [1]

On the other hand, health insurance in Bulgaria meets certain requirements set out in the Directives of the European Commission [2], whose main goal is to create a single and competitive market for health insurance products. The COVID-19 pandemic has forced people to take out more and more health insurance to cover the risks associated with their health and, as an option, to seek better quality medical services.

From an investment point of view, the brief instability in the financial markets failed to significantly affect insurance companies. This fact is observed in particular in the stock exchange trading of government bonds, which remained relatively unaffected by the crisis. The COVID-19 pandemic has shown that society's resilience to severe shocks in general must be strengthened, whether related to human health, natural disasters and catastrophes - such as earthquakes or climate change, economic crises of large-scale and complex cyberattacks.

The role of insurance is to accept these new challenges by applying adequate measures to deal with emerging new types of insurance risks. The aim is to minimize the risk, to limit the risk of events and to maintain adequate liquidity.

3. Key Changes in the Insurance Business Following the COVID-19 Pandemic

Measures taken around the world to reduce the spread of COVID-19 require insurance companies in most jurisdictions to activate their plans to ensure the continuity of insurance services during a public emergency. At the same time, the Insurance Supervisors are closely monitoring the ability of policyholders to access insurance coverage and file online claims, despite the exceptional circumstances imposed by the global health crisis following the COVID-19 pandemic.

Many jurisdictions have asked insurers to complete questionnaires or questionnaires on the implementation of their business continuity plans, thus monitoring their ongoing development. For these reasons, the key changes that are taking place in the insurance business as a result of the COVID-19 pandemic concern the following main areas:

1) Health and safety - insurers must ensure the protection of the health and safety of all their clients, employees, regulators, insurance brokers and agents. Insurers can achieve this new requirement by enabling alternative employment arrangements for their employees, by adapting to possible restrictions on access to offices, by digital delivery of services and at the same time - ensuring the continuity of the insurance business;

2) Remote work with clients - a significant increase in the investments of insurance companies in outsourced workstations, servers and cloud space. Digitalization of insurance services based on cloud technologies will accelerate automation and become a standard tool for all insurers;

3) Use of mass platforms - especially when offering insurance products for health prevention. The increasingly popular "telehealth" platform poses a new challenge to health insurers, who need to expand the range of health insurance services they offer in order to successfully adapt to this innovation;

4) Secure payments and refunds - a new trend in the insurance business is refunds, bonuses or discounts on insurance premiums due to reduced use of motor vehicles due to announced lockouts and forced people to stay at home. It is applied by insurance companies mainly in the USA, thus contributing to the reduction of harmful gas emissions and adverse climate changes on the planet;

5) Regulation against digitalisation - with the process of digitalisation of the insurance business, conflicts are expected to arise regarding the regulations of insurance in different countries. This requires urgent measures to find practical solutions and changes in the legislation of individual Member States of the European Union regarding digitalization processes;

6) Concentration risk mitigation strategy - the intentions of offshore centers in India, Eastern Europe, Ireland, China to concentrate insurance services need to be reviewed. Insurers must implement a strategy to mitigate the risk of concentration by expanding their insurance operations to more locations;

7) Rethinking/reconsidering the management of exposure to financial instruments - a better understanding and management of potential exposures in the context of the COVID-19 pandemic will help insurers to better plan future claims against them;

8) Providing new opportunities for policyholders to file claims and receive insurance coverage - the experience of pandemics in most countries is extremely limited and insurance coverage is not specifically designed for the resulting costs and losses. Uncertainty among policyholders about the possibility of their insurance to cover the occurrence of a possible insurance event also remains significant. In response to this problem, a number of insurance supervisors and individual insurance companies and associations have published statements

and guidelines on the types of costs and losses that a particular insurance company will or will not cover.

9) Review and further analyses of brokerage and distribution models - models for the use of insurance brokers and agents for the distribution of insurance products should be reviewed in certain markets within reasonable limits in order to connect the insurer more directly with the end customer using digital insurance offerings products;

10) Transfer of capital to the insurance sector - "The intersectoral transfer of capital to the insurance sector is a serious factor in strengthening insurance competition. The different rate of profit in individual industries, including the insurance sector, is the main prerequisite for the transfer of capital to the insurance sector and an important factor in developing the competitiveness of insurers.."[3]

4. Key Changes in Insurers' Accounting Requirements Following the COVID-19 pandemic

The accounting work of insurance companies monitors, measures, evaluates, registers, analyzes and controls all the economic assets, their sources of formation (capital) and their movement in the cycle of insurance activity, the specific type of business relations, which differ from other types of business by content, deadlines, manner and technology of implementation.

In accordance with the requirements of the Accounting Act, insurers develop their own accounting policies that take into account the essential features of their insurance business. "Accounting policy is a universal method that combines accounting principles, rules, instrumentation and specific practical procedures, which form the organization and technology of accounting reports. It takes into account the impact of the economic, political and social environment, and the factors that influence the insurers 'activity. Through their accounting policy, insurers present the economic information about the specific processes and the underlying operations within their insurance business in a true and fair, authentic and comparable manner. "[4] Based on already developed and constantly changing accounting principles, rules, specific practical procedures that they have adopted and that determine their accounting, investment and insurance policies, insurers can adequately manage with the changes that have occurred in their insurance business globally in a COVID-19 pandemic situation.

The key changes that insurers make in their accounting job concern the following specific objects of their activity:

1) True, accurate and reliable determination of the amount of allocated insurance (technical) reserves, which are the main guarantee of insurers for their ability to cover all claims of policyholders in the global health crisis. "These are specific obligations of insurers to the insured persons and are intended to cover the eventual realization of the insurance risk. The technical provisions represent a reduction in the resources of the insurer in connection with the settlement of its existing liability to the insured persons as of the date of the Statement of Financial Position. They represent specific liabilities that do not always become due due to the probability that the insurance risk will not materialize, resp. the insured event does not occur. This fact is associated with the process of their release, which leads to their recognition as current insurance income for insurers. They represent the main part of the liabilities of an insurance company, the largest balance sheet item in the Liabilities of its Statement of Financial Position (Balance Sheet). " [5];

2) Correct choice of the estimates of the coverage of the insurance reserves (financial assets and financial instruments) and of the technical reserves of the insurers in the conditions of global health crisis;

3) Ensuring the formation of a specific technical reserve (Reserve Fund), which one to cover the threat of occuracy of mass insurance risk in a result of the global spread of the COVID virus 19;

4) Accurate determination of the adequacy of the insurance premiums on the concluded insurances in the conditions of global health crisis in order to cover the insurance risks, especially in the field of health insurance;

5) Accurate determination of the sufficiency of the insurred amounts and indemnities under the concluded insurances in order to cover the arisen insurance risks in the conditions of global health crisis, especially in the field of health insurance;

6) Stronger preliminary, ongoing and ex-post control of insurers' accounting systems in the context of the global health crisis;

7) Timely and effective recalculation of solvency capital requirements and profitability indicators, assumptions on technical provisions and other techniques for reducing insurance risk and managing this risk by insurers in the context of the global health crisis caused by Covid-19 pandemy;

8) Ensuring sufficient liquidity through applying constant control over the availability and spending of funds, their investment in assets and instruments in the context of the global health crisis;

9) Specific disclosures in the financial statements of insurers concerning the changes that have occurred in their activities in the context of the global health crisis. The management bodies of these companies are required to prepare and present specific financial statements in accordance with the International Accounting Standards. [6] Their form and content are provided for in IAS 1 *Presentation of Financial Statements* and in IAS 7 *Statement of Cash Flows*. [7] „Financial statements are prepared on the basis of the assumption that the entity is going concern and will continue as going concern in foreseeable future. It is assumed that the entity neither intends nor needs to liquidate or significantly shrink its operations."[8] „ In these conditions, there is a need to improve, expand and adapt the information from the financial statements of enterprises in line with the strategy for smart, sustainable and inclusive growth of their business. The financial statements must accurately and fairly present the property and financial condition of enterprises, financial results, changes in cash flows and equity." [9]

Insurers also prepare an Annual Activity Report, which contains the most important information for consumers and investors about their activities. In the context of the global health crisis, insurers are required to disclose in their financial statements also the following information:

- the new levels of insurance risks and the probabilities of the mass occurrence of insurance events;
- the adverse effects that have occurred and the respective significant risks that could occure in a reault of COVID-19;
- the measures taken for a quickly resolve the problems that have arisen in their activities as a result of the declared pandemic by COVID-19;
- the plans adopted by them for rapid adaptation of their business, for further strengthening of their position on the insurance market, increase of their competitiveness and profitability, etc..

5. Conclusion

Insurance business is from a crucial importance for any society in the world. The cash resource that it accumulates and disposes of, should be managed with the care of a good steward. So far, the sector is mamging well with the situation caused by the COVID-19 pandemic, but problems are constantly arising. In connection with this pandemic, insurers can understandably assume additional responsibilities under the insurance policies (especially health insurance), covering the new insurance risks arising from diseases, complications and damages generally caused by the COVID-19 virus. There is a mass manifestation of the assumed insurance, incl. of health insurance risk in the context of the global health crisis. There is a need for sufficient liquidity to cover the indemnities for the services used by the insured persons for health activities, medical goods and services and subscription health care. In this regard, insurers must make efforts to reduce the degree of manifestation of insurance risks, including and the health insurance risks posed by the global health crisis. In this way, they will limit the occurrence of insurance events and will be able to meet their obligations to the insured persons without any problems.

References

1. Natchkova, M., Challenges faced during the independent financial audit of financial statements of health insurers, 10TH INTERNATIONAL CONFERENCE ON APPLICATION OF INFORMATION AND COMMUNICATION TECHNOLOGY AND STATISTICS IN ECONOMY AND EDUCATION (ICAICTSEE – 2020), NOVEMBER 27-28, 2020, UNWE, SOFIA, BULGARIA;
2. Directives 73/239/EEC, 88/357/EEC, 92/49/EEC, 98/78/EC, 2001/17/EC, 2009/138/EC;
3. Natchkova, M., Analysis of the competitiveness of insurers in Bulgaria in the introduction of new insurance products, collective monograph "The economy of Bulgaria and the European Union in the digital world ", UNWE, SOFIA, 2019, p. 626, ISBN 978-619-232-208-3;
4. Natchkova, M., Accounting policy of insurers in crisis conditions, International journal "KNOWLEDGE", Institute or Knowledge Management, Skopje, Scientific Papers, Vol.41.1, 08/2020, p.151, Global Impact and Quality Factor 1.822 (2018), ISSN 2545-4439, ISSN 1857-923X
5. Milanova-Tsoncheva, E., Natchkova, M., Mavrudiev, H., Accounting of Insurers and Social Insurers, S., Publishing House - UNWE, ISBN 978-619-232-116-1, 2018, p.310;
6. Accounting Act, promulgated in State Gazette (SG) no. 95 of 08.12.2015, effective from 1.01.2016, last amendment and supplement, no. 105 of 11.12.2020;
7. International Financial Reporting Standards (IFRS), 2019;
8. Ivanova, R., Optimisation of the Content of the Entities' Management Report for the Needs of the Financial and Business Analysis of the Smart, Sustainable and Inclusive Growth, Management and Sustainable Development Magazine, LTU, Sofia, no. 3/ 2011;
9. Chukov, K., Ivanova, R, „Financial and economic analysis", UNWE, SOFIA, 2019;
10. Code of Insurance, promulgated in SG no. 102 of 29.12.2015, effective from 1.01.2016, last amendment and supplement, no. 64 of 18.07.2020;
11. https://www.carat.fiserv.com/en-us/resources/insurance-industry-adapts-to-the-new-normal/
12. https://www.oecd.org/finance/Insurance-sector-responses-to-COVID-19-by-governments-supervisors-and-industry.pdf
13. https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_4764

Data Quality Scoring Model for Increasing the Organization Efficiency

Daniel Delchev

University of National and World Economy, Sofia, Bulgaria
ddelchev@unwe.bg

Abstract. The largest organizations invested tens and hundreds of millions of euro/dollars in analytical systems and data warehouses and failed to achieve their goals. This results into new approach of information management. This is the essence and in the scope of current paper, which presenting general Data governance model and deep dive into Data Quality Scoring Model in order to foster set up of strong and prudent data quality management process. Process which is essential also for organizations working and aiming monetization of Big Data. With the help and applications of model those organizations could improve their efficiency and fruitful usage of Big data. In the paper are covered all main 5 pillars of new Data Governance model – Processes; Governance model; Organization; Standards; Technology and corresponding value for organization reflected into revenues; operating margin and assets. Also in the scope is the Big data governance role in organization with definition of the role and reflection on three layers in each organization: Strategic; Tactical and Operational layer with respective tasks and responsibilities. As part of overall Data Governance current paper provides opportunity to deep dive in specific component – Big Data quality scoring model – as vital element ensuring connection between business users with examples what happen when Data Quality framework is not defined and what are the benefits when it is defined. Below is presented an introduction of Data quality Scoring Model as management tool that characterizes level of organizational refinement in addressing Data quality processes, solutions, issues resolution. The model classify maturity in organizations into 4 buckets: Chaotic; Defined; Controlled and Continuing Improvement. The lack of data quality management of company data and Big data leads to mistrust in models, reports and undermine data driven-decision process with respective decrease in overall organization efficiency.

Keywords. Data Governance, Data Quality management, five pillars of Data Governance, Data Quality Scoring Model.

1. Introduction

Data Governance enables the organization to manage and utilize data as the elementary key asset and to provide compliance with regulatory requirements.

Strong data quality is key for every organization as it facilitates the strengthening of our customer base by further developing organization capabilities to offer the right products, via the right channels, at the right point of time, to the right customers.

Achieving the transformation to a data-driven organization is only possible with a stringent Data Governance. Data must be pro-actively managed, from the source to the final report, across departments and countries, and the quality of data must be ensured along every step of the data flows by defining efficient, automated processes combined with data quality measurement, controlling and reporting.

1.1 Definition of Data Governance

We could pay attention initially on the definition of Data Governance introduced by the Data Governance Institute presented below: "Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods." [1]

When we use the aforementioned definition and look at it through the prism of digital transformation of financial institution to data-driven organization we could fine tune Data Governance definition to the following:

Data Governance are the methodological, conceptual, organizational and technical rules, responsibilities, standards and procedures related to data, aligned with the organization's Mission & Vision, Strategy, Transformation Map and values. The goal is to use company data with their maximum potential in every aspect considering ethical principles within the organizational business processes.

When we talk for Data Governance in Financial institutions, especially in Europe we should consider the backbone of latest regulations in the area and especially BCBS239 - Basel Committee on Banking Supervision paper Principles for effective risk data aggregation and risk reporting. The BCBS239 officializes one of the most important lessons learned from the global financial crisis in 2007. The lesson learned was that banks' information technology (IT) and data architectures were not capable to respond to the need of broad management of financial risks. It appears that banks couldn't aggregate risk exposures and identify concentrations quickly and accurately at the bank group level. Same issue was observed across business lines and between legal entities. The weak risk data aggregation capabilities and risk reporting practices lead some banks to unable to manage their risks properly. This impacted badly the banks themselves and the stability of the financial system was not guaranteed.

To resolve that, the Basel Committee issued supplemental Pillar 2 (supervisory review process) guidance to helps banks to enhance their ability in identifying and managing bank-wide risks. In particular, the Committee emphasized that a good risk management system should have good management information systems (MIS) applied at business and bank-wide level. References to data aggregation were included by The Basel Committee as part of its guidance on corporate governance.

The benefits of improving the risk data aggregation capabilities were widely recognized in the banking industry. The improvements are in terms of strengthening the capability of the risk function to make judgements. Which helps to gain efficiency, less probability of losses and better strategic decision-making, and increased profitability. [2]

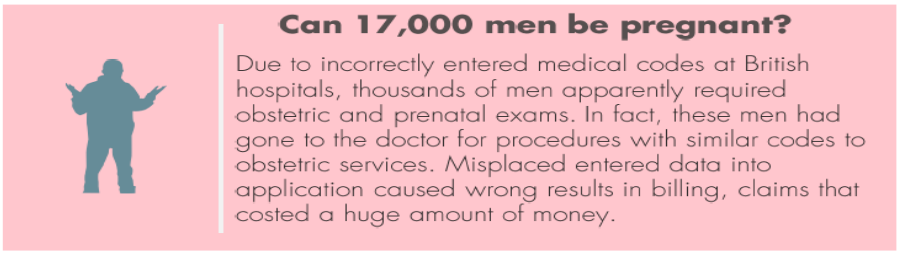
1.2 Definition of Data Quality management

Data is vital asset of every organization in the contemporary world. Therefore, managing the quality of the data becomes more important each day because organizations are continue collecting data every day. Data Quality Management can be defined as a set of practices to maintain high quality information. These set of practices go throughout the process of handling data; from acquiring it, implementation, distribution, and analysis.

In essence why we need data quality management in each organization. The proliferation of data in the digital age has presented a real challenge – data crisis. The data crisis leads to low quality data in its volumes that makes it hard for the businesses to process it out, and in some cases, the data is unusable. Data Quality management has become an important process used to make sense out of data. It helps organizations by pointing out errors in their data which need to be fixed. It also helps in assessing if the data in organization's systems is accurate enough to fulfil their purpose [3, 4].

Based on the reasons described above we could briefly describe Data Quality as an essential part of Data Governance. The definition focuses over Data Quality as the planning, implementation, and control of activities that apply quality management techniques to data to assure the data are fit for consumption and meets the needs of data consumers.

2. Data Governance Matters



It appeared that seventeen thousands of British men turned up at hospitals to be treated for many pregnancy-related services in 2009 and 2010, things like obstetric exams and midwife services - obviously a mistake. Researchers studying the data think they're the result of something way more boring: medical coding errors.

The research was published in the British Medical Journal, with the intention to draw attention to how much data gets entered incorrectly in the country's medical system. In the reality, these guys weren't turning up at the doctor for pregnancy-related services. They were at their doctor for procedures that had medical codes similar to those of midwifery and obstetric services. Reading summary of the mistaken data we may come up to the wrong conclusion that the health system needs more midwife and less physicians. [5]

The largest organizations invested tens and hundreds of millions of dollars in analytical systems and data warehouses and failed to achieve their goals such as:

- Long time of data warehouse development – too long time-2-market;
- Insufficient attention to data warehouse data model and the large costs of fixing errors in the model, the absence of a business data model;
- Unclear and immeasurable data quality, lack of trust of the information from the data warehouse;
- Scattered data-marts for different departments and analytical tasks, the inability to reconcile data and reports from different departments;
- The inability to compare information due to chaos with reference data;
- There is no clearly defined responsibility for data and data quality in the organization.

3. The Five Pillars of Data Governance and Topic Areas and Responsibilities

In order to achieve accurate, timely and secure information the organizations should develop and implement overall Data Governance system – based on the following 5 pillars (Figure 1 includes the 5 pillars of Data governance and their link to accurate, interoperable, timely and secured information)

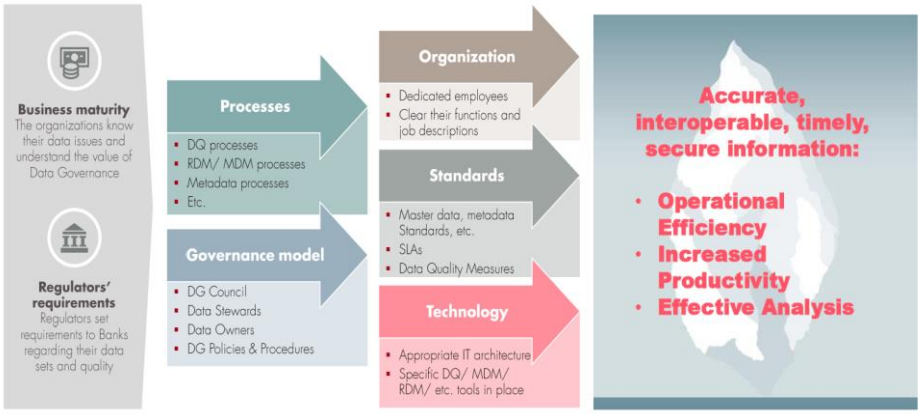


Fig. 1. Five pillars of the Data Governance

The five pillars of Data governance are transformed into specific topic areas and responsibilities for a concreate financial institution presented in Figure 2.

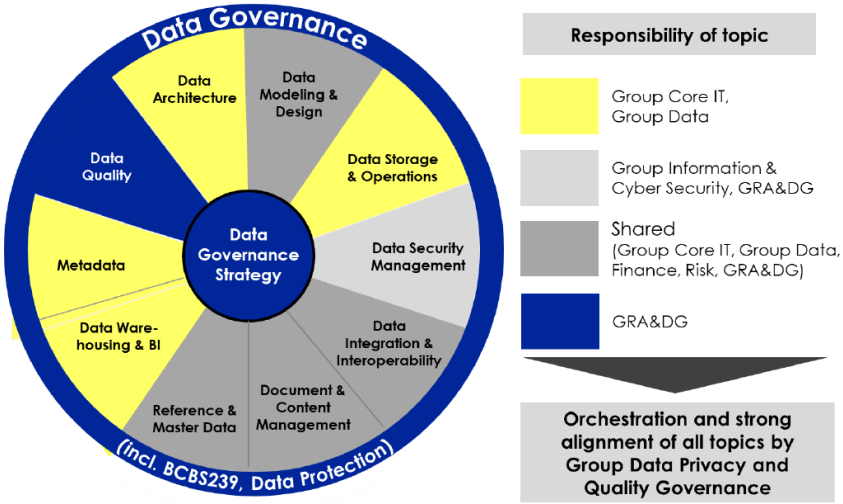


Fig. Data Governance topic areas and responsibilities

The different colors of the building blocks represent proposed distribution of the responsibilities within one financial institution and are adjusted from the author of current paper taking into consideration organizational structure of the specific financial institution classified as one of 125 significant banking group in Europe:

• The blue topic areas Data Governance (incl. BCBS 239, Data Protection), Data Quality Management and Data Governance strategy are defined and pro-actively driven by Group Regulatory Affairs and Data Governance division (GRA&DG) in alignment with Group Core IT, Group Data, Finance and Risk.

• The dark grey topic areas Data Modelling and Design, Data Integration and Interoperability, Document & Content Management and Reference Data Management are covered by Group Core IT, Group Data, Finance, Risk and GRA&DG in shared responsibility.

• The yellow topic areas Data Architecture, Data Storage & Operations, Data Warehousing & BI and Metadata are in the responsibility of Group Core IT and Group Data, whereas Group Core IT and Group Data define and implement solutions taking into consideration the defined Data Governance.

• Only one Data Governance topic area – Data Security Management (light grey) is in the full responsibility of Group Information and Cyber Security, which aligns with GRA&DG especially regarding defined Data Governance roles and responsibilities and data privacy requirements.

• Data Governance Operational Steering Committee (DG OSC) and Data Governance Board (DGB) are the responsible decision bodies as outlined in the Data Governance bylaws.

4. What Influences Data Quality

Many influence factors can affect Data Quality during the life cycle of data.

Table 1 describes main activities which contain possible triggers with potential negative impact on Data Quality. Strict awareness and consideration of these areas during the daily work improves the level of Data Quality (Table 1).

Table 1. Activities and potential cause for missing Data Quality

Activity		Potential cause for missing Data Quality
(I)	**Data Entry**	Incorrect or incomplete information stored on data entry and/or data conversion in the source.
(II)	**Processing**	Insufficiently defined process or discontinuity/gaps in data processing, e.g.: • Fast propagation/transformation of data without time to perform Data Quality Checks sufficiently • Data delivered 'out of context' (e.g. foreign key references without access to the referenced data source to perform DQ-checks) • 'Wrong timing' or asynchronous processing • Undocumented or too many manual intervention
(III)	**Architecture**	Systems, applications and data architecture that are inconsistent or allow data redundancies. Lack of data integration and interoperability (e.g. Data Linage Level 3).
(IV)	**Definitions**	Inconsistent or redundant definitions of business terms and data elements (Synonyms, Homonyms).
(V)	**Data Consumptions**	Using or transforming data different as specified and required (the semantics used for data entry are different then for subsequent usage). Use of unregistered/personally maintained data (e.g. locally stored list of values for report creation).
(VI)	**Data Decay**	Data can become outdated or obsolete over time if not actively maintained. Unused legacy systems are not decommissioned.

5. Data Quality Dimensions

Understanding the Key Data Quality Dimensions is the first step in assessing and continuously improving Data Quality. To be processed and interpreted effectively and efficiently, data must meet a certain set of quality criteria. The following dimensions also refer to BCBS 239 principles and are defined to measure and evaluate Data Quality (Table 2).

Table 2. Data Quality Dimentions and definitions

Data Quality Dimensions		Definition
(I)	**Completeness / Unambiguous**	Data are not missing and is of sufficient breadth and depth of the task at hand. Values are present in the attribute that requires them.
(II)	**Accuracy / Free of Error**	Data are correct and reliable ('error free') from the entry at the source to the reporting at the end.
(III)	**Consistency**	A given set of data must be consistent across different IT systems of the Group.
(IV)	**Timeliness**	Data are available as expected and sufficiently up-to-date for the task at hand.
(V)	**Uniqueness**	Aggregate data are free from any duplication from filters or other transformations of source data.
(VI)	**Validity**	Data are founded on an adequate and rigorous classification system.
(VII)	**Availability/Accessibility**	Data are easily accessible and usable for relevant parties when they are required based on the need to know principle.
(VIII)	**Traceability**	History, processing, location and naming of any piece of data under consideration can be easily traced to the origin.
(IX)	**Punctuality**	Data are submitted in time as expected and defined.
(X)	**Reliability**	Data are regarded and clearly identifiable as true and credible.
(XI)	**Recoverability**	Organisation can recover data in case of force majeure loss of data.

6. Poor Data quality has its price

Costs of poor data quality could be listed as the following:

- Fines/legal/liability exposure costs;
- Lost opportunity (e.g. missed prospects for cross-sale);
- Non-recoverable costs (e.g. fraud);
- Hurt customer confidence, shareholder confidence;
- Wrong decisions/actions;
- Data verification/cleanup/correction costs;
- Redundant data handling and support costs;
- Scrap and rework.

In Figure 3 is illustrated common data architecture of a financial organization and costs of poor data quality.

Through the whole Data flow are marked weak points resulting into poor data quality issues. In principle presented architecture consists of 4 layers from Source systems, via Data Warehouse and some Dara Marts to reports. Here there is no deferiation between on-premise or cloud solutions and example is applicable for both with clarification that on cloud environment is more appropriate for usage of big data and data enrichment.

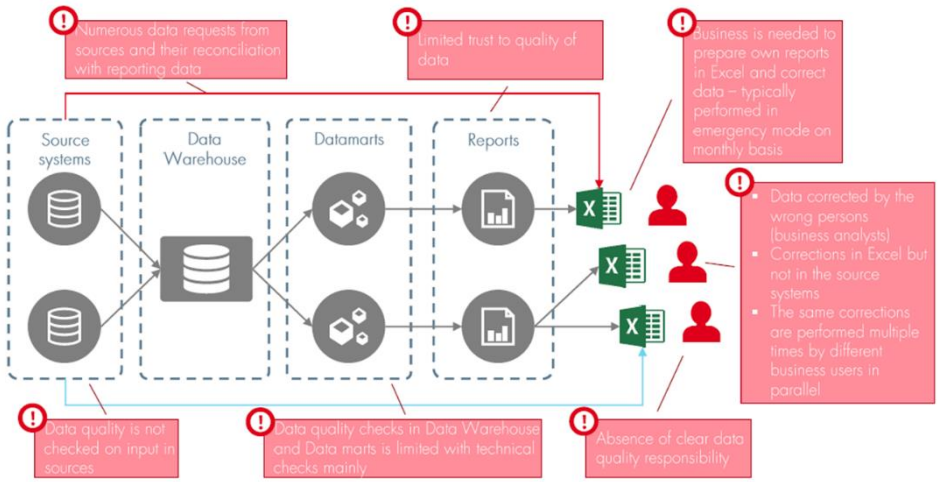


Fig. 3. Illustration of common data architecture of a financial organization and costs of poor data quality

In next Figure 4 is presented the same basic data architecture of a financial organization which is enhanced through implemented Data Quality Framework and respective enhancements and elimination of weak points which were leading to poor data quality.

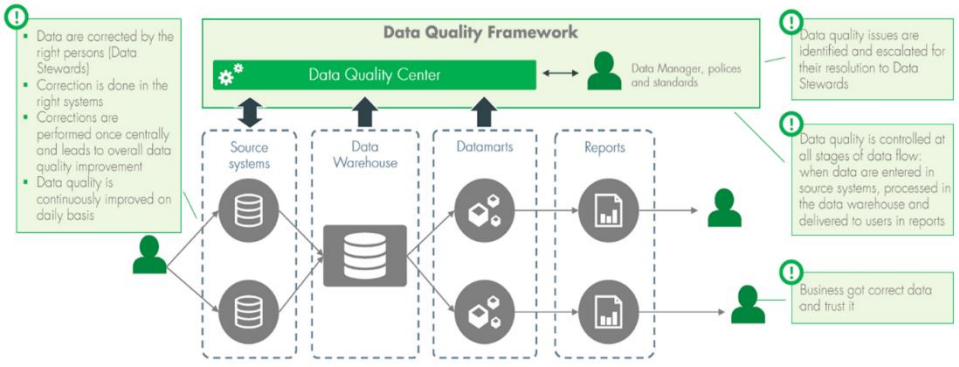


Fig. 4. Illustration of common data architecture of a financial organization enhanced through implemented Data Quality Framework.

The comparison on high level what happens when Data quality is defined and not defined or implemented is presented on Figure 5 where we could find visualized direct connection between Data Quality Framework and respective organization business users.

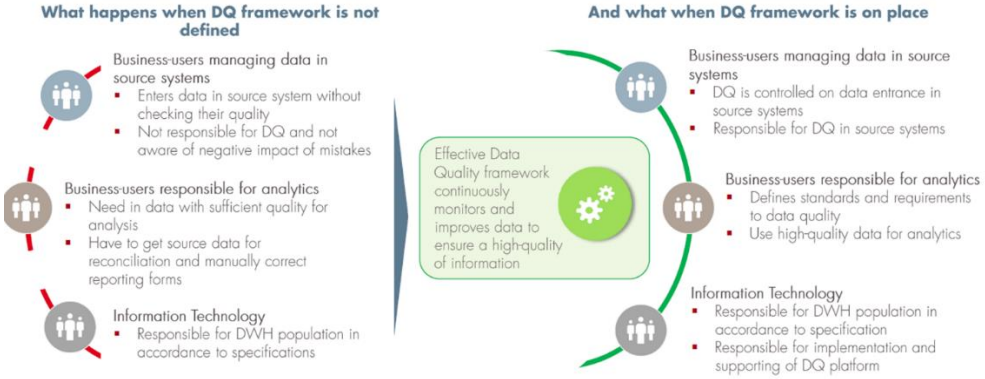


Fig. 5. Connection between Data Quality Framework and respective organization business users.

According organization DQ Lab - bad data is dangerous to your business. It is even more detrimental to your business than a lack of data at all. Data-related regulations are evolving very fast, almost at the speed of evolving the digital world itself. Therefore, it is becoming crucial for organizations to apply proper data management strategies. Businesses working with improperly maintained or organized data may struggle to demonstrate compliance. Regulations such as GDPR, CCPA, FR Y-14Q lead organizations to have stronger trust around the data used internally as well as the date they send externally to the company. Organizations that find themselves in possession of sensitive personal and financial data such as banks face up more heavy data management requirements.

Good quality data enables:

1. Better decision making

- More accurate and realistic decision making

- Confidence in decision making and eliminates the need to guesstimate and saves the unnecessary costs of trials and errors.

2. Better Targeting

It's critical to know who your prospects are – something that an organization can only manage analyzing and understanding data. Using high-quality data from your current customer base, you can create user personas and anticipate the needs of the new opportunities and target markets.

3. Effective Marketing

Assessing the results we may identify what's working and what's not in marketing activities. This is possible if only we have data with good data quality

4. Better Customer Relations

Good customer relations are vital for any successful business. Creating bond with your customers starts with understanding what they need. By analyzing the information you already have in your database you may predict their interests and preferences. Good quality data is verry important in process of preparation of proper marketing content and helps to avoid spamming of you customers ant targeted audience.

5. Competitive Advantage

Good quality data provides you with a clearer image of your industry and its dynamics. Your marketing decisions will be more confident as well as your long terms plan will be more precise and accurate. You will be ble to respond to market dynamics in a more proper manner. It will also be easier for you to understand the needs of your customers, which will help you sale more. [6]

In order to have appropriate tool which assess financial organization maturity in Data quality here is presented developed model, based on series of interviews with a financial institution mid-level management in areas of Finance, Risk Management, Data management, IT, Compliance and Business areas.

The models defines financial institution maturity on 4 different levels – Chaotic, Defined, Controlled or Continuing Improvement. The expert based model is presented in Table 3 below and has 5 variables assessed with score from 1-4 (1-lowest score and 4-highest score). The variables covered the following: Process; Governance; Resources; Standards and Tools.

The limitation of costs of poor data quality could be achieved with grades classifying the organization close to border between Defined and Controlled. The ranking of organization steadily in Controlled bucket is enough for most of financial institutions to be compliant with the respective data management regulations in Europe.

Ranking of organization in bucket Continuing Improvement could be excellent base for data-centric and digitally transformed organizations.

Table 3. Data Quality Dimentions and definitions

	Variable 1	1p.	2p.	3p.	4p.	Level
Process	*No DQ processes*	☆ 1				CHAOTIC - Reactive(Level 1)
	Some not standartized processes	☆ 1	☆ 2			DEFINED - Controlled(Level 2)
	DQ processes are defined to greater	☆ 1	☆ 2	☆ 3		CONTROLLED - Proactive(Level 3)
	DQ processes are defined and fully implemented	☆ 1	☆ 2	☆ 3	☆ 4	Continuing Improvement - Predictive(Level 4)
	Variable 2	**1p.**	**2p.**	**3p.**	**4p.**	**Level**
Governance	*No governance& roles or Informal roles of Data Owners*	☆ 1				CHAOTIC - Reactive(Level 1)
	Define Data Owners &Stewards	☆ 1	☆ 2			DEFINED - Controlled(Level 2)
	Governance processes and roles are	☆ 1	☆ 2	☆ 3		CONTROLLED - Proactive(Level 3)
	Top level executive involvement in data quality initiatives	☆ 1	☆ 2	☆ 3	☆ 4	Continuing Improvement - Predictive(Level 4)
	Variable 3	**1p.**	**2p.**	**3p.**	**4p.**	**Level**
Resources	*No dedicated employees and lack of plans to ensure HR resources*	☆ 1				CHAOTIC - Reactive(Level 1)
	No dedicated employees, but budgeted/included in Business plan	☆ 1	☆ 2			DEFINED - Controlled(Level 2)
	Dedicated employees	☆ 1	☆ 2	☆ 3		CONTROLLED - Proactive(Level 3)
	Dedicated employees, knowledge transfer	☆ 1	☆ 2	☆ 3	☆ 4	Continuing Improvement - Predictive(Level 4)
	Variable 4	**1p.**	**2p.**	**3p.**	**4p.**	**Level**
Standards	*No standards*	☆ 1				CHAOTIC - Reactive(Level 1)
	Limited DQ standards	☆ 1	☆ 2			DEFINED - Controlled(Level 2)
	Standards are implemented across data domains	☆ 1	☆ 2	☆ 3		CONTROLLED - Proactive(Level 3)
	Standards are implemented and continuously reviewed	☆ 1	☆ 2	☆ 3	☆ 4	Continuing Improvement - Predictive(Level 4)
	Variable 5	**1p.**	**2p.**	**3p.**	**4p.**	**Level**
Tools	*No special tools*	☆ 1				CHAOTIC - Reactive(Level 1)
	Excel, some customs tools	☆ 1	☆ 2			DEFINED - Controlled(Level 2)
	Separate DQ tools	☆ 1	☆ 2	☆ 3		CONTROLLED - Proactive(Level 3)
	Centralized and integrated solution	☆ 1	☆ 2	☆ 3	☆ 4	Continuing Improvement - Predictive(Level 4)

Linking data owners, data scientists, business analysts and employees who use the data in their day-to-day job, in a cultural framework to collaborate and to put the data at the center of decision-making process is one of the primary goals here. This comes up with new data-driven use-cases, discovering patterns in data and experimenting with analytics solutions to understand what works in your organization.

There are several prerequisites for establishing a data-driven culture - access to data, governance of usage and quality of data, and appropriate technologies to prepare and analyze data. [7]

7. Conclusion

We could easily conclude that in general, not only for financial institutions, but also for customer and data driven organizations the topic for data quality is essential. In a nutshell it is essential if the organization wants to give customers the best experience and simultaneously making decisions using accurate data.

Fair conclusion is that great customer experience leads to happy customers, brand loyalty, and higher revenue. In case the organization is using poor-quality data, than it mostly guessing at what its customers want. There is even worst case that the organization might be actively doing things its customers dislike.

Current model for Data Quality Organization Maturity could be used as management tool for organizations aiming at increasing efficiency and monetization of Big data through various business models and data initiatives.

Ensuring data-based decision taking, relying on properly governed and easily accessible data could be competitive advantage of each organization and opportunity to increase share of its happy customer and employees.

References

1. The Data Governance institute Article Defining Data Governance, https://datagovernance.com/defining-data-governance/
2. Basel Committee on Banking Supervision paper Principles for effective risk data aggregation and risk reporting, https://www.bis.org/publ/bcbs239.pdf
3. What is data quality management?, https://www.dqlabs.ai/blog/what-is-data-quality-management/
4. The Data Governance institute Article Focus Areas for Data Governance, https://datagovernance.com/focus-areas-for-data-governance/
5. Britain's 17,000 Pregnant Men Aren't Really Pregnant, Sam Ro Apr 8, 2012, https://www.businessinsider.com/britain-17000-pregnant-men-2012-4
6. Top 5 benefits of data quality, https://www.dqlabs.ai/blog/top-5-benefits-of-data-quality/
7. Data driven culture, https://bi-survey.com/data-driven-culture

Framing a Consumption Ecosystem by Developing Courses Designed for the Digital Age of Higher Education

Vanya Slantcheva

Management, Trade and Marketing College, Sofia, Bulgaria
v.slantcheva@mtmcollege.bg

Abstract. The paper reviews the hybrid model adoption in higher education (HE). In accelerating the blending of in-person and online learning, the hybrid shift incorporates digital tools to help educators assess student engagement, thus providing them with a roadmap for how to refine and improve their courses and teaching methods to empower students' learning. The general idea considered here is that designing courses for the digital age is to reveal the servuction (service production) and the consumption ecosystems within HE service delivery. It addresses the issue of digital roles assumed by the teacher and the student in order to exchange knowledge for skills' pick-up and competence proof. Using a design thinking technique, a mirrored service blueprint of course design for undergraduates is unfolded for the servuction and the consumption ecosystems to be planned out.

Keywords. Higher education, hybrid classes, ecosystem, course design, digital roles, undergraduates.

1. Introduction

The purpose of the report is to review the contemporary concept of digital transformation in order to outline the crucial evolution from a value chain's theory that uses data for operational efficiency to a digital ecosystem's framework that provides data-driven services. The main goal is to address the framework of higher education (HE) service delivery in terms of its service production (servuction) ecosystem – that connects parties involved in producing and selling, – and consumption ecosystem of conjoining traditional and non-traditional complementors. Becoming aware of the HE institution's disposition to profound disruptive change, digitally powered exchanges between the students and the complementors are to be revealed.

The most significant consideration is that framing a consumption ecosystem for the HE is grounded by induction in means of academic course design that empower students to own their learning. It is important to understand empowerment as a privilege which allows students to replicate, up-plicate or duplicate their learning as continuous interactions within the unfolding HE consumption ecosystem.

2. Engaging in Digital Transformation and Understanding Consumption Ecosystems in HE Industry

Tiers of Digital Transformation in Terms of HE Industry. Shifting from Traditional HE Practices

The business world is shifting from an industry-based framework and an emphasis on products/services to an emphasis on data and digital ecosystems as networks of data generators and recipients. It is important to understand the framework—including the tiers, the idea of

interactive data, and the different types of ecosystems—and understand which tier the organization is in, which type of ecosystem it operates in, and the organization's strategic intent. [14] This modern data includes a shift from episodic to interactive data. [16] Episodic data is collected intermittently after discrete events. It is aggregated and analyzed for insights after the fact. Episodic data and its after-the-fact insights are closely guarded and not usually shared outside the firm. Interactive data, on the other hand, is a continuous stream of data in digital platforms. The unique characteristics of interactive data are that it comes in large volumes; it has a real-time component that can be shared; and it can be used to create new digital experiences. Interactive data can be generated from assets, from sensors, from users, and from other sources. With interactive data, there is a reversal of roles where products support data, signifying the value of data over products. [14, 16].

Part of digital transformation is an evolution from value chains to digital ecosystems. Digital transformation involves using modern digital technology to enrich and augment a value chain by creating "digital ecosystems". There are two types of digital ecosystems: 1) production ecosystems and 2) consumption ecosystems (Fig. 1). Establishing a framework for digital transformation, M. Subramaniam (2021) identifies four tiers. Tier 1 to tier 3 are production ecosystems. These are essentially value chains enriched and augmented by modern digital technologies. From that perspective, transformation starts by using data for operational efficiencies and proceeds to offering revenue-generating, data-driven services from digital platforms.

The framework involves thinking beyond traditional "value chains" to ecosystems [16].

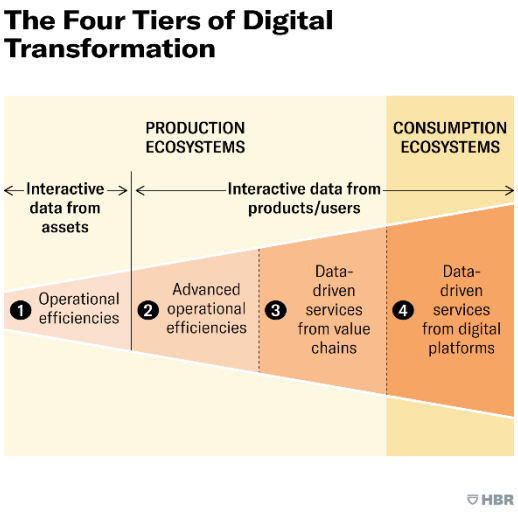


Fig. 1. The Four Tiers of Digital Transformation
(Source: Subramaniam, M. 'The 4 Tiers of Digital Transformation', *HBR Webinar*, Dec. 1st, 2021.)

In a traditional value chain, the complementors are separate from the value chain. With modern digital technologies, it becomes possible to connect the complementors to a company's digital ecosystem. Thus, conjoining traditional complementors creates a "consumption ecosystem." To operate in consumption ecosystems, companies must facilitate exchanges between customers and external entities, extending the value chain into a digital platform. (Fig.2).

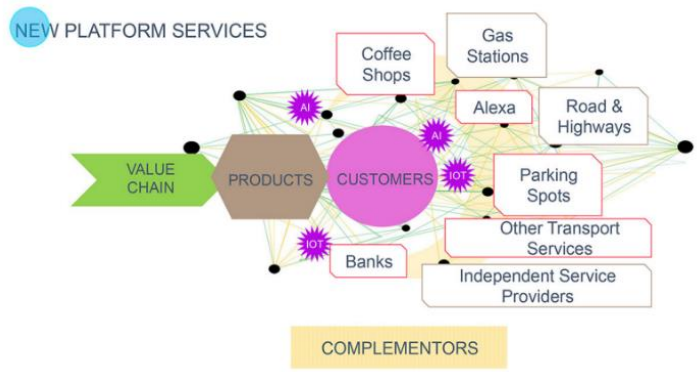


Fig. 2. New Platform Services Embracing Complementors into Consumer Ecosystem.
(Source: Subramaniam, M. 'The 4 Tiers of Digital Transformation', HBR Webinar, Dec. 1st, 2021)

Consumption ecosystems emerge after a product is sold through linkages enabled by sensor data generated as the product is being used. [12] Consumption ecosystems differ from production ecosystems which are built on linkages within value chains — as those between suppliers, R&D, manufacturing, assembly, and distribution channels — that connect the parties involved in producing and selling. Many companies use smart connected products within their value chains to improve their assembly processes, customer relationships, and after-sales service. But consumption ecosystems, which involve tracking products after they are sold and sharing sensor data with parties outside the value chain, are new. And they require new capabilities. The second shift occurs within the new consumption ecosystem when companies go beyond its produce-and-sell role, becoming a platform creator and orchestrator [12].

In terms of Ch. Lovelock's conceptual framework for servuction system (service production system) that refers to service operations, delivery, the other contact points with the service, and the customer in person [10], it helps to configure the HE production ecosystem. It integrates well with the overall educational process from within, and performs a series of service encounters as "moments of truth" a student experiences with the service.

Activating HE Consumption Ecosystem and Third Parties Outside Its Value Chain

As service consumption is process consumption, and the consumer perceives the service production process as part of its consumption – and since the consumption of the service process is a critical part of the service experience [3] – a consumption ecosystem of a higher education (HE) service comes with the service encounter. Service encounter is 'a period of time during which a customer directly interacts with a service' [12], i.e. when student interacts with faculty members, administrative and non-administrative staff (service employees), in peer-to-peer relations (fellow customers), with LMS (technology, equipment, platform). Hence, a concern related to moving to remote and hybrid service encounter was how to ensure that in-person interactivity would cross over into the virtual environment through interaction with peers and faculty outside of class. For remote students, interaction and opportunities for learning and social experiences should advance after they exit the virtual classroom [9]. And not to lose that whole social and learning structure of collective interactions, as well as those with third parties, post-pandemic ways to facilitate learning in and outside the classroom is to be embraced for further implementation.

Furthermore, paraphrasing Subramaniam and Piskorski's (2020) definition, a HE consumption ecosystem also emerges after the educational service is sold through linkages enabled by (sensor) data generated as the service is being perceived and experienced. A shift from traditional practices is that HE institutions need to manage exchanges between their students and third-party providers to offer value-added services for them, going beyond the encounter-and-post-encounter stages [10] and the sell-produce-deliver[-consume] role to run 'tethered digital platforms' [12] outside its value chain. There the HE service initiates digitally powered exchanges among students and third-party providers inside or outside the university campuses (Fig. 3).]

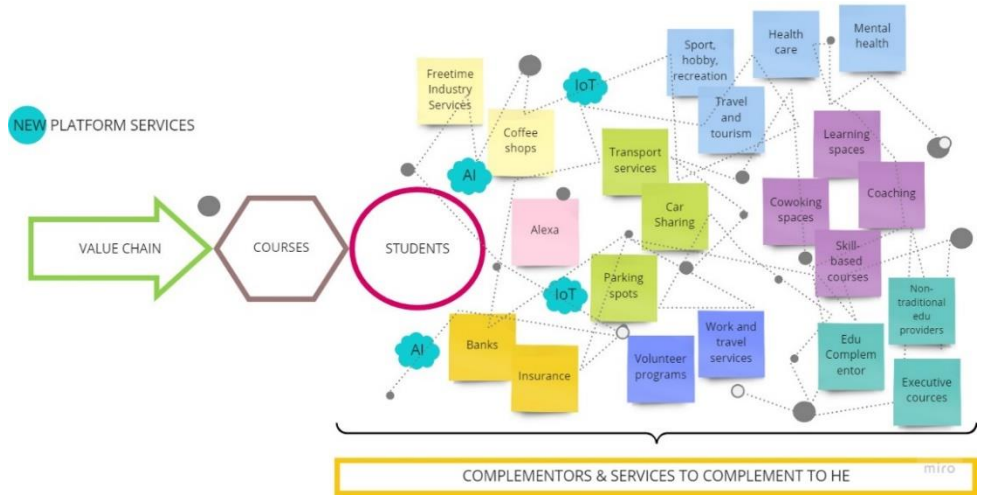


Fig. 3. Digitally powered exchanges among the students and the complementors. (Adapted from: Subramaniam, M. 'The 4 Tiers of Digital Transformation', HBR Webinar, Dec. 1st, 2021)

A tethered digital platform has four essential components [12]:

• *a sensor* tied to the service – HE organization could use an internet smartphone app as its sensors;

• *sensor data* generated when the service is perceived and experienced – it could analyze the app's data to yield insights on a student's synchronous and asynchronous engagements, learning style, preparation, assignment take-on and completion, sharing, working solo or as a team, personal desires and lifestyle, skills requirements to facilitate exchanges between that student and merchants;

• *platform users*, for whom exchanges are orchestrated; merchants or non-traditional edu service providers could use the platform to compete for the student's spending; and

• *a platform service* that is generated through those exchanges among service and third-party entities – the platform would expand the HE institution's traditional services into broader purchase experiences for students.

Hence, the HE institution needs to craft a strategy that optimizes the scope, uniqueness, and controllability of its services' sensor data. Subramaniam and Piskorski establish the following framework that offers four potential tethered digital platform strategies: full, enabled, collaborative, and hybrid [12]. And growth opportunities within depend on the scope and uniqueness of a product's sensor data and how much control the company maintains over that data (Fig. 4).

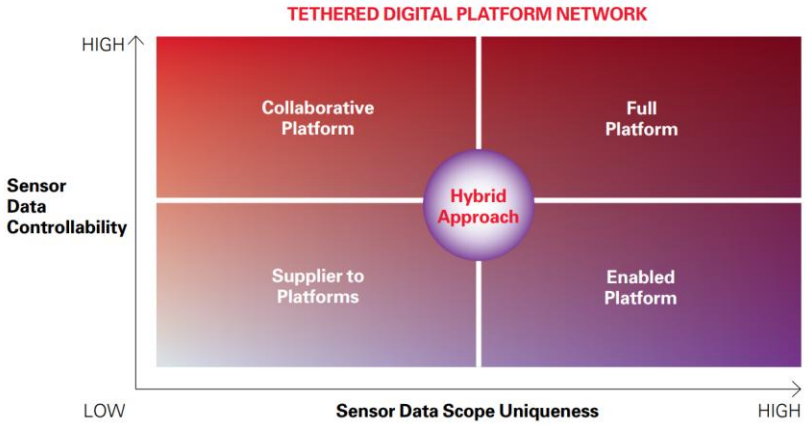


Fig. 4. Tethered Digital Platform Network.
Source: Subramaniam, M. and M. J. Piskorski, 'How Legacy Businesses Can Compete in the Sharing Economy', MIT Sloan Management Review, Volume 61, No. 4, pp. 31-37, Summer 2020, p. 34

A full platform involves orchestrating data exchanges to provide services. It is best when the organization has full control over service data that is unique and faces few sharing restrictions. If data is not unique or the organization can't share data freely with others, an enabled platform (which allows others to exchange data to provide services) or a collaborative platform (which involves running an own platform and providing services as part of another organization's or a merchant's platform) could be a better platform strategy choice. [12]

HE institutions orchestrate unique sensor data to provide their educational, administrative and social services but they could vary by their scope according to organization size, structure, geography, cycle degree covered, profile of the scientific field, extent of specialization, value-added services included. That reflects on the strength of data controllability the institution has got and its disposition to exchange them with third parties.

3. Designing Courses that Empower Students to Own Their Learning

Building Mature Digital Learning Ecosystems

Evolution in the higher education ecosystem happens through "punctuated equilibrium:" long periods of relatively slow change interspersed with occasional moments of rapid adaptation. The pandemic period is a punctuation moment. Educators, faced with unprecedented urgency, have been working hard to restore teaching and learning using technology, innovation, and collaboration. While being on varying stages of digital transformation, universities have developed their own digital competencies. DeVaney et al. (2020) have suggested a framework to help universities identify where digital learning fits into their education ecosystem and transform their teaching and learning. They consider that building mature digital learning ecosystems is about assembling digital newcomers, inviting emerging adopters and demanding advanced HE institutions.

For *digital newcomers* it is easier to take quick actions through the current state of technology and platform choices – extensive broadband access, reliable communications tools, user-friendly video conferencing, and widespread smartphone adoption. Institutions easily and cost-effectively secured licenses for students, and faculty immediately started engaging online. The faculty and students get comfortable with the medium, seeking help from peer institutions,

consultants, and companies to train the group on what it means to teach effectively in an online setting.

Emerging adopters used early adopters among departments, faculty, and staff as mentors and key architects of their strategy, empowering them with authority, resources, and decision-making latitude to adopt turnkey solutions. They accelerate the production of online courses, supplement with widely available open content from other institutions. They can minimize human curation by using machine-learning solutions to map the most relevant courses to their curriculum; they explore virtual and take-home labs for courses that require hands-on problem solving.

Advanced institutions drive their digital strategy scaling the infrastructure across all programs and using online courseware. They accelerate pedagogical innovations to serve diverse online communities with varying socio-economic backgrounds and increase the commitment to creating an inclusive environment for learning experiences. Outside of the virtual classroom, community engagement has been strengthened through crowd-sourced notes, study groups, virtual live-streamed events. Advanced institutions are best positioned to explore immersive technologies like augmented reality (AR) and virtual reality (VR). These universities are well positioned to become leading contributors to the global HE ecosystem with their expertise and content. [5]

"Designing Experiences and Letting Go"

Educators' experiences over the pandemic period have stretched their thinking around *what* and *how* students have been engaged. In that respect R. Austin discusses a 4-S framework of a post-pandemic course design which should reflect that work and which 'optimizes student discovery and empowers students to own their learning' [1]. Even physical classrooms have gradually been reopened and a hybrid model of blending in-person and online learning has been applied. Educators can use the techniques they have experimented online with to develop better course perception and learning experiences. In other words, the online and hybrid shift encompasses digital tools to help teachers creating and assessing student engagement, thus providing instructors with a roadmap for how to refine and improve their courses and teaching methods. According to R. Austin the most meaningful student experiences are those in which they: a) are actively engaged; b) feel responsible for their own learning; and 3) have significant control over their experience. [1, 2]

R. Austin applies Kolb's experiential learning model, for the 4-S process – struggle, structure, systemize, and synthesize – can be repeated as often as necessary and refine the findings until they meet instructor's learning objectives. As Figure 5 shows, different forms of both digital and in-person learning can be inserted where they would best guide students through the process. [1]

The learning process of an autonomous course, or the series of service encounters within that course, begin by "presenting a challenge for the students to grapple with" [1]. To have students *struggling* – being on their own, they need to ask themselves what is happening, what the problem is, what information is available and if whether it is relevant or additional – is needed. This step can be done in person, but there is also a wealth of digital tools that provide students with simulated cases or digital workspaces. After students have struggled on their own, they are encouraged to *structure* and categorize the information they already know to determine if and how it applies to the problem. They have to recognize how to categorize information as to relate on it, and also what to analyze to generate information needed, using virtual or in-person group work, online discussions, and other activities that center on student-driven exploration and dialogue. The teacher eventually guides students toward systemizing the structured information into overarching frameworks. To start seeing general principles emerging students need to be guided through connecting structured information. This is

facilitated through "live discussions that build off points raised and emerging themes from earlier online asynchronous discussion, or by inviting students to reflect on conclusions reached while doing group work." Once students have reflected on the matter at hand, the instructor steps in more directly to help them *synthesize* their new frameworks, establishing group consensus and drawing conclusions. This might involve presenting a synthesized model of the key information through live instruction, an edited video lesson, or other instructor-driven 'rescue frameworks' in the form of instructor-led analysis to help fill in any gaps in student understanding [1, 2].

Unquestionably, graduate students may be perfectly happy working at their own pace with minimal oversight. But undergraduates rarely operate in the same way [12]. There are plenty of approaches educators can take to guide and support undergraduate students and help them work at a steady pace *outside* of class. Adopting strategies for incentivizing pre-class preparation, making it available for coaching and supporting group projects can provide better student engagement and a structure they need to thrive, and different from "creating a narrative and the space to discover on their own" [12]. Doing so educators can meet their students where they are rather than expecting them to perform at a level they haven't quite reached.

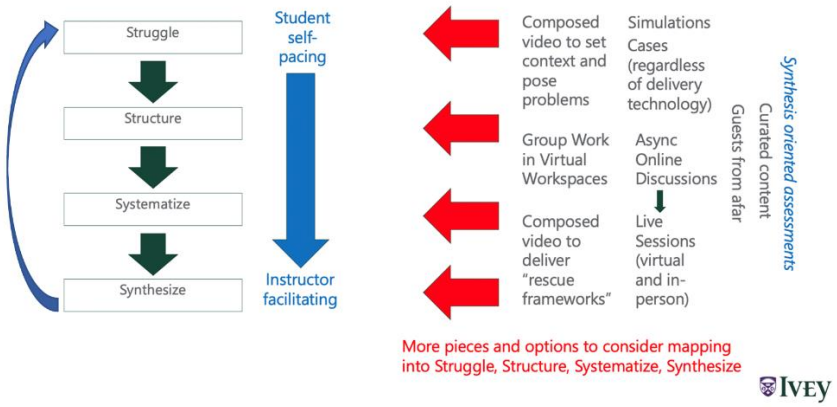


Fig. 5. As students move through the four phases of the 4-S framework, control of the unfolding experience moves from the student to the instructor.
(Source: Robert D. Austin, "Designing Better Courses: Blending the Best of Pre- and Post-P andemic Pedagogy," Harvard Business Publishing Education, July 21, 2021)

4. Framing a Student Consumption Ecosystem for Continuous Learning Interactions

To design an educational course for the digital age we should inductively accept it as an autonomous service design. It embraces a substantial process of learning as series of service encounters and assumed predetermined roles for student involvement. Such a service blueprint is used as a premise for activating a sequential HE consumption ecosystem, for third parties outside HE servuction ecosystem to be tethered in a digital platform. Moreover, if the course design purpose is to have students empowered to own their learning, they could share it with third parties by means of their experience and compatibility with the service, considering the HE institution controllability over the data.

Here role and script theories can help us understand, design and manage student behavior during HE service encounters both inside and outside the servuction ecosystem and sequentially revealed to the HE consumption ecosystem. Hence, the Service Blueprint tool is used to map those that students experience along an academic course delivery. See Table 1.

Both educator and student roles are sets of behavior patterns learned through experience and communication, to be performed by each in a certain learning or social interaction in order to attain engagement, learning and productivity in educational goal accomplishment. The representation of Prof. Austin's 4-step framework [1] in terms of a course delivery script below specifies the sequence of educator's and students' behaviors to teach and to learn. Students learn their own scripts through experience, education and communication on the occasion of the course within the HE servuction ecosystem. Those scripts are more flexible because it is about a highly customized educational service and may vary by course, situation and student. Pandemic reality harnessed HE institutions to produce a digital conversion of the traditional service script, using digital technology to transform it from a high-contact service into a low-contact one [10]. The post-pandemic present has taken the baton of the digital transformation of virtual learning, inviting HE institutions to modify the remote encounter to a hybrid one where "digital technology could serve to substitute, augment and complement face-to-face" [6] encounter. Moreover, the exposure to engaging in both in-person and virtual learning activities has been changing student attitudes toward digital courses, and introduces new players in the higher education market. Instead of relying on traditional academic institutions, learners gravitate toward alternative, non-traditional education providers due to their more user-friendly, appealing, and often cheaper online offerings. [8]

From that perspective, Table 1 blueprints a mirrored HE consumption ecosystem as a counterpart of the HE servuction ecosystem. There students engage being familiar with their course scripts of episodic interactions within the education servuction process, but reflectively they can replicate, up-plicate or duplicate their own academic learning in a complementary learning with scrips of continuous interactions within a consumption journey they can build themselves.

In the rehearsed service blueprint teacher's script presents the pedagogy that can transform episodic teacher-student interactions to continuous ones. Such a teaching plot is in conjunction with the post-pandemic demands for continuous education. [8] It serves as an invitation to them to expand their thinking around their latest research considering its teaching importance and media potential as a research topic. They need to lean into networking to stay in contact with lifelong learners and to increase their visibility. Teachers also need to embrace the importance of their pedagogy, balancing their approach to teach and research.

Their visibility, closeness to companies and organizations, and personal brand effectiveness will be of growing value to continued education programs.

In such a blueprinted framework HE institutions will need to foster communities of lifelong learners, protecting themselves from the disruptive tendencies envisioned. Building strong communities among all university stakeholders is a protective measure. Universities need to attract alumni as learners for life offering learners more choice and flexibility of participation. In that respect they will also need to set them apart from non-traditional educational service providers. HE institutions are advised to craft collaboration with edtech and big tech to ensure their continuous education offerings.

Table 1. Mirrored HE consumption ecosystem embracing tethered digital platform network of non-traditional edu service providers

Opportunities & Insights	HE consumption ecosystem mirrored	*Student Competitiveness with Non-traditional Edu Service Providers*				
Support Processes			Virtual Workspaces		Virtual Workplaces	**Line of Internal Interaction**
Backstage Actions		**Tethered Digital Platform Compatibility**				**Line of Visibility**
Frontstage Actions of Non-traditional edu providers		Complementor – a pedagogical partner for educators and learners worldwide	Complementary teaching and learning	Attending Professional classes (executive edu courses)	Complementor - working, Freelance, Influencer	**Line of Interaction**
Student Actions			External engagement, experience, group & individual interactions	Metaversal learning/ training; Remote expertise	Remote expertise	
Phase Name		**Struggle**	**Structure**	**Systemize**	**Synthesize**	
Evidence		simulations, cases digital tools that provide students with simulated cases composed video to set context and pose problems	group work		live sessions (virtual & in-person composed video to deliver "resque frameworks" edited video lessons	
Environment		in person; digital workspaces	asynchronized online discussions; in-person or group work in digital workspaces	group work in digital workspaces		
Student Actions	HE servuction ecosystem	asking questions: what is happening, what the problem is, what information is available and if it is relevant or additional is needed on their own	categorizing info to relate on; analyzing to generate info needed	seeing general principles emerging; reflect on conclusions	establishing group consensus and drawing conclusions	**Line of Interaction**
Frontstage Teacher Actions/ Facilitation		presenting a challenge to students to grapple with; setting context and posing problems regardless of delivery technology	encouraging students to categorize what they know; online asynchronous discussions, other activities that center on student-driven exploration and dialogue. live discussions	guiding students through connecting structured information virtual & in-person	delivering "rescue frameworks"	**Line of Visibility**
Backstage Actions		**Digital Platform Compatibility**				
Support Processes		On campus Augmented, residential model Faculty and Program Departments	LMS, Digital Workspaces IT	Collaborative teaching and learning Physical and Digital Workplace Library and Publishing Subsidiary	Virtual Workplaces Administrative Infrastructure	**Line of Internal Interaction**
Opportunities & Insights		*Student Competitiveness with the HE Service*				

(Adapted from: Robert D. Austin (2021); Nielsen Norman Group Service Blueprint Map Template© by Sarah Gibbons)

5. Conclusion

The review clearly shows that for HE institutions, conceiving of digital ecosystems is to be a combination of servuction and consumption ecosystems. Each university's digital transformation depends on how it uses its data within its production and consumption ecosystems, advancing through the four tiers of M. Subramaniam's framework. Which tier a HE institution lands on depends on the nature of its service delivery scope and its strategic intent to market consolidation. All they depend on strong investments in infrastructure to support servuction processes and to set digital platforms to overcome current episodic student-

teacher and student-service patterns of interactions and to foster lifelong learning. Further extensive research is to be undertaken. The first area of research would be to analyze student competitiveness with the service within its HE servuction ecosystem and to reveal it as a learner competitiveness within HE consumption ecosystem by its comparison with the traditional HE institutions vs. non-traditional education service providers. A second area of a research might be to explore the HE institutional practices in crafting strategies that optimize the scope, uniqueness, and controllability of its services' sensor data for orchestrating student exchanges among the HE service and third-party entities.

References

1. Austin, R. D. 'A 4-Step Framework for Post-Pandemic Course Design: To Boost Engagement, Let Students Take Charge of Their Own Learning', HPB Education, August 5th, 2021. (https://hbsp.harvard.edu/inspiring-minds/a-4-step-framework-for-post-pandemic-course-design)
2. Austin, R. D., 'Designing Better Courses: Blending the Best of Pre- and Post-Pandemic Pedagogy', HBP Education Webinar, July 21st, 2021. (Designing Better Courses: Blending the Best of Pre- and Post-Pandemic Pedagogy Harvard Business Publishing Education)
3. Baron, S. and K. Harris. Services Marketing" Texts and Cases. 2nd ed., Palgrave: New York, 2003.
4. Carioca, M. F., Stahl, W. and N. Khalil, "How the Pandemic Has Changed Us—and Higher Education: Reflections from a Dean, a Professor, and a Student", HBP Education – Inspiring Minds, 2021. (How the Pandemic Has Changed Us and Higher Education Harvard Business Publ. Education)
5. DeVaney, J., Shimshon, G., Rascoff, M. and J. Maggioncalda, "Higher Ed Needs a Long-Term Plan for Virtual Learning", HBP Education – Inspiring Minds, May 6, 2020. (Higher Ed Needs a Long-Term Plan for Virtual Learning | Harvard Business Publishing Education)
6. Govindarajan, V. "COVID-19's Impact on the Future of Higher Education: What University Leaders Should Be Thinking About Now", HBP Education Webinar, May 20, 2020. (COVID-19's Impact on the Future of Higher Education: What University Leaders Should Be Thinking About Now Harvard Business Publishing Education)
7. Grönroos, Ch. Service Management and Marketing: A Customer Relationship Management Approach. 2nd ed., John Willey & Sons: Chichester, 2000.
8. Kaplan, Andreas. "5 Ways Universities Can Stay Relevant for Lifelong Learners", Inspiring Minds, HBP Education, 2021. Professionals Need to Keep Their Skills Fresh. Will They Turn to Higher Ed? | Harvard Business Publishing Education
9. Lakhani, K. R. and L. M. Viceira, "5 Key Lessons from HBS's Pandemic Teaching Transformation: And What Learnings Will Stick Moving Forward", HBP Education – Inspiring Minds, 2021. 5 Key Lessons from HBS's Pandemic Teaching Transformation Harvard Business Publishing Education.
10. Lovelock, Ch. And J. Wirtz, Services Marketing: People, Technology, Strategy.6th ed., Pearson Prentice Hall, 2007.
11. Ringland, G. and L. Young. Scenarios in Marketing: From Vision to Decision. John Wiley & Sons Ltd., Chichester, 2006.
12. Roberto, M.,"How to Help Undergraduates Work Smarter Outside of Class: Strategies to Incent Better Class Prep and Support Group Projects", HBP Education – Inspiring Minds, October 15, 2021. (hbsp.harvard.edu/inspiring-minds/how-to-help-undergraduates-work-smarter-outside-of-class)
13. Shostack, G., 'Planning the Service Encounter', in J. S. Czepiel, M. R. Solomon and C. F. Surprenant (eds), The Service Encounter, Lexington Books, MA, 1985.
14. Subramaniam, M. and M. J. Piskorski, 'How Legacy Businesses Can Compete in the Sharing Economy', MIT Sloan Management Review, Volume 61, No. 4, pp. 31-37, Summer 2020. (How Legacy Businesses Can Compete in the Sharing Economy (mit.edu))
15. Subramaniam, M. 'The 4 Tiers of Digital Transformation', at Featured Article - Professor Mohan, also in 'The 4 Tiers of Digital Transformation', HBR Webinar, with Julie Devoll, Moderator, (HBR Webinar Summary) 2021 (https://hbr.org/webinar/2021/11/the-4-tiers-of-digital-transformation).
16. Subramaniam, M. 'Are You Using the Right Data to Power Your Digital Transformation?', HBR, Dec. 3rd, 2021. (Are You Using the Right Data to Power Your Digital Transformation? (hbr.org))

Opportunities for Development of an Experimental Prototype for Structuring of Data in Web-based Information Systems

Yavor Tabov

University of National and World Economy, Sofia, Bulgaria
jtabov@unwe.bg

Abstract. The paper addresses issues related to the presentation of text data from web-based information systems in a structured form. The main concept of the research is focused on the possibilities for analytical processing of the set of data in the online environment. The specifics of the data in the web-based information systems are presented. A model prototype for data structuring in this type of systems is proposed. Promising directions for future research are outlined.

Keywords. Internet, web-based information system, data structuring.

1. Introduction

Software solutions for extracting data from web-based information systems and presenting the extracted text data in a structured form apply different approaches to achieve their goals. From a technological point of view, the respective approaches have many common features. The study of the characteristics of the most popular software solutions in this field can be used as a basis for summarizing the characteristics of this type of software solutions in general. The preliminary expectation of the study is that this type of software solutions will have a minimalist user interface for configuring the extraction of data from the relevant web-based information systems. The purpose of using specialized tools is to extract from web-based information systems only some parts of the relevant web pages that are of interest from the point of view of the data they contain.

Today, much of the data of interest to organizations is in an unstructured form within various web-based information systems. In this context, extracting data from web-based information systems is a key activity that is crucial for organizations. According to some subject matter studies, more than 80% of useful business data is in unstructured form [3]. Research on this topic in the public sector shows that increasing amounts of data are being processed in the field of administrative services [8]. The ability to analyze data from a variety of sources is at the heart of an essential part of contemporary research in any subject area [9]. Web space is a prime example of storing huge data sets.

According to several studies in the subject area, unstructured data is constantly increasing in volume, due to the data that is created daily in modern world-famous platforms with rich user content, such as YouTube, Facebook, Twitter, LinkedIn, and others [4, 6]. This content is of interest to the management of the organizations given the possibilities for analysis of the relevant data [10]. The presented content management systems are essentially web-based information systems. Web-based information systems can be described in the form of different types, in terms of their purpose and the services they provide:

- Websites.
- Online stores.
- Social media.

- Blogs and forums.
- Content management systems.
- Content sharing systems.
- Content publishing systems, etc.

In the present study, we will adhere to the understanding that all systems that provide custom views, described in text form through the HTML hypertext markup language, are web-based information systems. In the context that all systems operate in some environment [5], web-based information systems operate within the web concept. Web-based information systems are made up of several web pages that are accessed using a web browser via the HTTP protocol and using a request-response communication [15]. In essence, web-based information systems have the usual characteristics for each system [1]. Web-based information systems exist within the web concept and in this sense the Internet is their environment. From a technological point of view, web-based information systems are software solutions built with the help of web-based technologies. The satisfaction of Internet users with this type of software solutions has become one of the most important indicators of their success since the beginning of the digital age.

2. Architecture of an experimental prototype

Unlike traditional information systems, which are specifically designed for professional users in organizations, web-based information systems are designed for a wide audience on the Internet with a variety of computer skills. This includes more direct consumption and interaction with data in web-based information systems compared to traditional information systems [2]. A characteristic feature of web-based information systems is their multi-layered logical architecture, usually consisting of three layers [14]:

- User interface.
- Business logic.
- Data.

The user interface provides access to the content and functionality of the web-based information system in a user-friendly form. Various management reports and information are usually displayed within this layer [7]. Business logic includes a set of software modules that perform the transformation of the relevant data of the web-based information system into useful information according to the rules of the subject area. The data represents the relevant content of the web-based information system.

According to other sources in the subject area, web-based information systems are systems based on web technologies that share different types of Internet protocols and play an important role in collecting, storing, and transmitting data from different sources within the web concept [16]. This type of information systems shares a high degree of community formation and user-level content generation [11]. Business organizations use web-based information systems to improve the level of management processes in organizations. The development of many business areas is directly related to the application of web-based information systems, including for the purposes of managing communication with other organizations. In this sense, web-based information systems and related technologies for their construction play a significant role in business development [16].

Regarding the creation of an experimental prototype for presenting text data from web-based information systems in a structured form, an architectural scheme of the software implementation of the experimental prototype should be drawn up, which includes:

- Choice of software technologies.
- Design of building components and their interactions.

- Drawing up a scheme of application in a real environment.

Fig. 1 presents an architectural scheme of the experimental prototype for presenting text data from web-based information systems in a structured form, which includes the main components of the prototype and the interactions between them.

In terms of its way of working, the experimental prototype for presenting text data from web-based information systems in a structured form is of the client-server type. It includes a web server that supports the PHP programming language [13] and various file formats, such as HTML, as well as a database management system. In the field of web server is located the software implementation of the experimental prototype, which is divided into two separate parts, namely:

- Logical part.
- User interface.

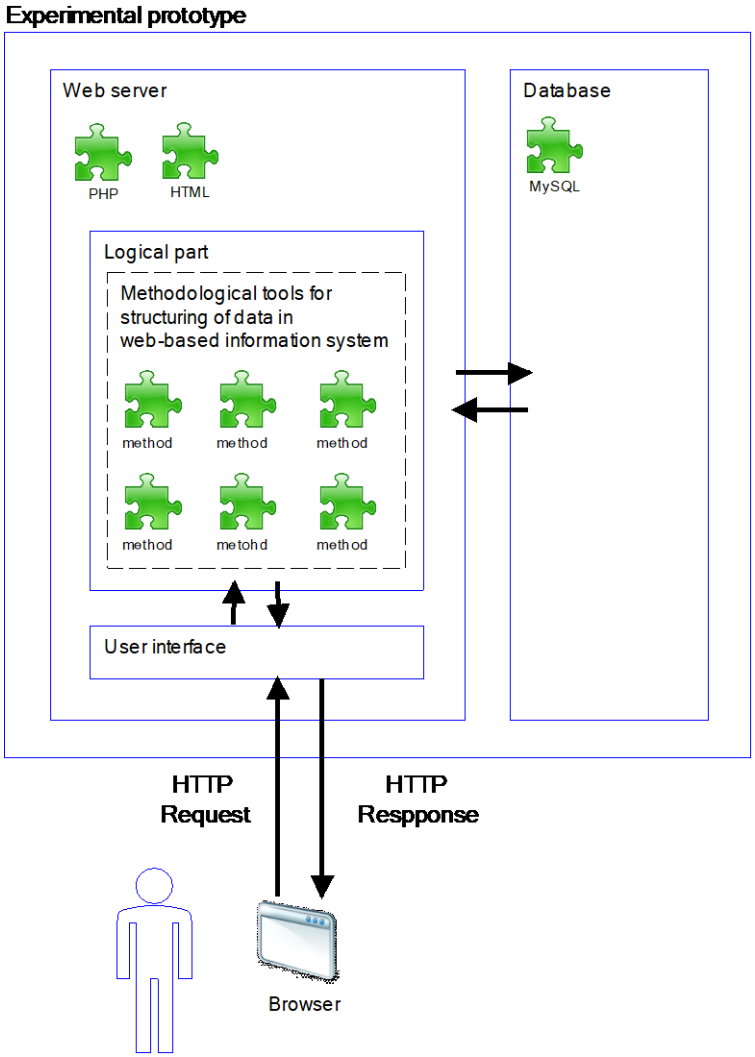


Fig. 1. Experimental prototype - components and interactions

3. Features of an experimental prototype

The logical part of the experimental prototype includes the software tools used to implement the methodological tools (components, classes, program libraries, etc.) and all computational procedures and algorithms that build the overall functionality of the experimental prototype for presenting text data from web-based information systems in a structured form.

The user interface acts as a bridge for management between the user and the logical part and is an organization of visual program controls (buttons, labels, tables), which manages and monitors the work of the experimental prototype for presenting text data from web-based information systems in structured form.

The database management system is a software solution for managing and storing collections of logically related data, which in the context of the experimental prototype for presenting text data from web-based information systems in a structured form can represent settings and practical results. In the current implementation, the database management system is MySQL [12].

Through a web browser, the users of the experimental prototype for presenting text data from web-based information systems in a structured form access the user interface and perform certain operations (commands are given in the form of HTTP requests). This leads to the interaction of the user interface with the logical part and thus establishes the control of the experimental prototype – certain input parameters are set when applying the methodological tools. The work of the logical part is related to interaction with the database, which includes operations of loading and writing the relevant data, which are created when working with the experimental prototype for presenting text data from web-based information systems in structured form.

Fig. 2 presents a scheme of application of the experimental prototype on web-based information systems.

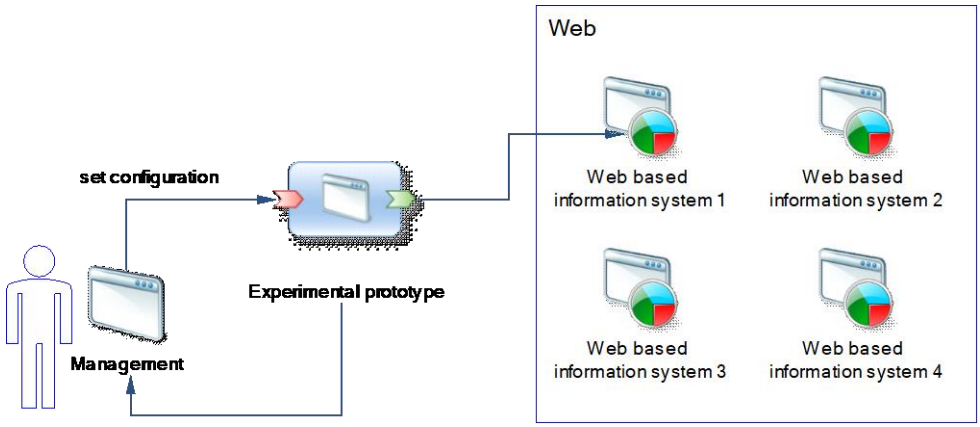


Fig. 2. Application of the experimental prototype on web-based information systems

When conducting an experiment, the user should set the appropriate settings. The experimental prototype for presenting text data from web-based information systems in a structured form receives text data (based on established settings) from a web-based information system. The processing of this data determines the life cycle of the experiment, starting with a comparison with several templates for presenting text data from web-based

information systems in a structured form and ending with visualizing the text data presented in a structured form. The implementation of each method leads to obtaining a corresponding result to the user.

Going through the different stages of the life cycle of the experiment in practice is an application of methodological tools for presenting text data from web-based information systems in a structured form. This methodological toolkit includes the use of templates and fragments, which realizes the presentation of text data in a structured form. These templates and fragments are presented in an abstract way, but for the needs of the experimental prototype it is necessary to prepare a way to describe them. This description is performed according to certain rules, which are defined by the extensible XML markup language [17]. Templates and fragments consist of one or more descriptions of grouped elements. To conduct the experiment, two types of grouped elements have been defined – reference type and detail page type. The term reference type refers to areas of a web page that are usually more than one, have similar or identical characteristics, and are at the same level in the HTML tree structure. These areas can be hyperlinks or other HTML elements that store hyperlinks. They represent a variety of web addresses that link to categories of lists or to detailed pages of descriptive content. This type of area is described in templates and fragments by the defined grouped reference element. There may also be areas in a web page that store a lot of interconnected text data. A typical example of such an area is a web publication, which consists of individual text fragments, such as title, date, author, description, etc. This type of area in templates and snippets is described by the defined grouped detail page element. The two grouped reference and detail page elements are a reference type element and a detail page element, respectively. These elements in XML format are defined by tags and represent descriptions of tree structures, which consist of many hierarchically organized sub-elements. Tags are characterized by names and attributes.

4. Conclusion

Based on the presented study the following conclusions can be outlined:

- Data structuring is a very important research issue given the possibilities for further analytical processing of structured data.
- The data in the web-based information systems are of interest to organizations for structuring and subsequent analytical processing.
- The development of an experimental prototype in the subject area is a challenging task that requires the use of specific methodological tools, given the peculiarities of data in web-based information systems.

For these reasons, the described experimental prototype is a promising area for future research.

References

1. Bagad, V., Management Information Systems, Technical Publications, 2008, ISBN 9788184313673.
2. Cheung, C., Lee, M., The Structure of Web-Based Information Systems Satisfaction: An Application of Confirmatory Factor Analysis, Springer, Berlin, Heidelberg, 2008, ISBN 978-3-540-68257-8.
3. Das, T., Kumar, P., BIG Data Analytics: A Framework for Unstructured Data Analysis, International Journal of Engineering and Technology (IJET), Vol. 5, No. 1, 2013, ISSN 0975-4024.
4. Eberendu, A., Unstructured Data: an overview of the data of Big Data, International Journal of Computer Trends and Technology (IJCTT), Vol. 38, No. 1, 2016, ISSN 2231-2803.
5. Gupta, H., Management Information System, International Book House PVT. LTD., 2011, ISBN 978-93-81335-05-5.

6. Khan, N., Yaqoob, I., Hashem, I., Inayat, Z., Ali, W., Alam, M., Shiraz, M., Gani, A., Big Data: Survey, Technologies, Opportunities, and Challenges, The Scientific World Journal, 2014.
7. Kirilov, R., The challenge – systems with open business logic, Proceedings of 5th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE), 2016, ISSN 2367-7635.
8. Kirilova, K., Methodological aspects of administrative service in the public sector, Economic and Social Alternatives, Issue 2, 2018, ISSN 1314–6556.
9. Milev, P., Application of data analytics for research of digital administrative services in Bulgaria, SHS Web of Conferences, Vol. 120, 2021, ISSN 2261-2424.
10. Stefanova, K., Kabakchieva, D., Business intelligent systems, Publishing complex UNWE, Sofia, 2015.
11. https://udspace.udel.edu/handle/19716/28502
12. https://www.mysql.com/
13. https://www.php.net/docs.php
14. https://www.researchgate.net/publication/4056556_A_three-tier_system_architecture_design_and_development_for_hurricane_occurrence_simulation
15. https://www.sciencedirect.com/science/article/pii/B9780128040188000073
16. https://www.ukessays.com/essays/information-systems/web-based-information-systems.php
17. https://www.w3.org/TR/xml11/

Approaches to Higher Security Level for Hadoop Environment

Ivona Velkova

University of National and World Economy, Sofia, Bulgaria
ivonavelkova@unwe.bg

Abstract. In the age of digitalization and digital transformation, data protection has proved to be one of the most sensitive areas on which everyone has turned their attention. The current paper looks at another aspect of security - providing secure access to the data itself in the Hadoop system as a high-risk area. The paper presents compare and evaluates the principles of operation of several major Apache projects in the face of Knox and Ranger, whose activities focus on identifying, authorizing, auditing, and managing each user with access to the system with included security policies. Current and future trends in their usage are considered.

Keywords. Hadoop, security, tools, Knox, Ranger, encryption, big data, 3A.

1. Introduction

Nowadays, data is the most critical and protected resource for any business. Extracting knowledge from them brings the greatest value to the existence of business. For this reason, most hacker attacks are very often aimed at company data - stealing, falsifying, using, etc. for a profit of third parties.

With the increasing size of data worldwide, various software and cloud solutions are being developed to make data storage, retrieval, and processing better and easier in real-time. One of the most successful solutions is that of Apache - Hadoop. Its main goal is to manage large amounts of data in a single and reliable environment. Its two components: Hadoop Distributed File System - HDFS and Map Reduce are the core of its operation. HDFS is the main storage system as it supports very fast data transfer speeds between nodes and allows the system to continue to operate without interruption in times of node failure. Map Reduce, in turn, is a software framework for analyzing and transforming a very large set of data into parallel programming for distributed processing in clusters.

With the adoption of Hadoop in almost every sector such as business, finance, healthcare, military, education, government, etc., its security features and data protection become a major concern [1].

2. Hadoop security

The Hadoop system was developed for local use with features, including security, which was not very important at the time. The reason for this is the distribution of sensitive data in individual clusters and very limited access to them.

With the advancement of digital transformation and the transition from various web services to the cloud, Hadoop does not compromise - it becomes surrounded by ever-increasing tools and applications that become available to various users throughout the organization. In this way, sensitive data is compromised. The innovations that are needed create new security

concerns for organizations struggling with Hadoop's technological advances and the rapidly changing rules of online environments [1], [2].

Access to data is improving over time. Hadoop HDFS does not store unencrypted data and does not have access to unencrypted data encryption keys that satisfy two types of encryption: encryption at rest and encryption during transportation. Encryption at rest refers to the encryption of data when the data is on a permanent carrier (disk). Encryption during transportation encryption means encrypting data when it travels on the network [2].

Another type of protection is the authentication of a user and the authorization of the access he has to his role in the organization. When using identity and access management, attention is paid to the concept of "3A": authentication, authorization, and audit. These three concepts are related and form the verification process when it comes to managing access to a resource in a system. Authentication is based on the idea that each user has unique information that distinguishes him from other users to provide proof of identity when they identify themselves. Authentication of a user or service request does not automatically give the user unrestricted access to all data in the Hadoop cluster. Access rights can be set for parts of HDFS and even for specific files and data types. Hadoop, by its very nature, does not perform authentication, which can have severe effects on all business data. To overcome this challenge, when talking about the authentication of users in a Hadoop cluster, Kerberos, which is an authentication protocol, is used. Hadoop uses the Kerberos protocol to ensure that the person making the request is whom they say they are. In security mode, all Hadoop nodes use Kerberos to perform cross-authentication. This means that when two nodes talk to each other, each makes sure that the other node is the one that says it is. Additional protection that Kerberos uses is secret-key cryptography to provide authentication to client-server applications. It is important to know that Kerberos is not a mandatory cluster requirement for Hadoop, which makes it possible to run entire clusters without deploying or implementing any protection [3], [4].

To authorize the user, Hadoop HDFS checks the files and directory permission after user authentication. In HDFS, all files and different directories have a group and an owner. In almost all cases, the owner and members of the group have different permissions on the files and directories. Each client process that will have access to HDFS has a two-part identity, which is a username and a list of groups. HDFS checks the permissions for the file or directory accessed by the client, and if the username matches the name of the file/directory owner, HDFS performs a test for owner permissions, and if the file/directory group matches one of the members of the file. the list of groups in the client access process, then HDFS performs the group permission test. If there is no compliance, the operation is considered unsuccessful [3].

The core technology of Hadoop HFDS has directories called encryption areas. When data is saved to Hadoop, it is automatically encrypted and assigned to a specific encryption area. Encryption is specific to the file, not the area itself. This means that each file in the zone is encrypted with its own unique Data Encryption Key /DEK/. Clients decrypt data from HDFS using a DEK, then use DEK to read and write data. Encryption areas and DEK encryption occur between file system levels and database architecture. Encryption keys must be managed, which is the work of the Hadoop Key Management Server /KMS/. KMS generates encryption keys, manages access to stored keys, and manages the encryption and decryption of HDFS clients. KMS is a Java web application with client and server components that communicate with each other via HTTP and REST API. KMS security includes HTTPS secure transport communication and support for HTTP SPNEGO Kerberos authentication [5], [6].

3. Apache Projects

The Hadoop system has resources to help improve security. Knox and Ranger are Apache open-source projects. Knox provides a security management framework and supports security implementations in Hadoop clusters. This project support monitoring, authorization management, auditing, and policy implementation in Hadoop clusters. It provides initial access to all interactions with Hadoop clusters. Knox also maintains improved security, as it can integrate with enterprise identity management solutions and is compatible with Kerberos [7], [8].

The Ranger project focuses on developing tools and techniques to help users implement and standardize security in Hadoop clusters. It provides a centralized framework that can be used to manage security policies at the resource level using user authorization roles. Ranger helps administrators apply group access rules, data type, and more [7].

- Knox

Apache Knox Gateway creates a security perimeter between Hadoop and the rest of the world by providing a REST API gateway to interact with Hadoop clusters. This allows Knox to both protect multiple clusters and present the REST API user with a single endpoint [8]. In this way, anyone who tries to break into the system, but without the necessary data or with the wrong credentials, will be rejected. Knox redirects external requests to an internal Hadoop service using the service name and URL of the service definition [10].

Apache Knox can be seen as a firewall for the Hadoop system, which is responsible for providing authentication to users. It authenticates user credentials using LDAP and Active Directory. Only successfully authenticated users have access to the Hadoop cluster. Knox is responsible for setting who is privileged to access your data through HTTP username and password authentication. In addition to user access audit [3].

All communication with Hadoop is done through Knox Gateway, which controls and moderates it. The process simplifies access to Hadoop: instead of connecting to different Hadoop clusters, all of which have different security policies, Knox becomes the single entry point for all Hadoop clusters in the organization with the idea of hiding Hadoop's structure from the outside world [7], [8].

- Ranger

Apache Ranger formerly known as Apache Argus is a framework for monitoring and managing the overall data security of the Hadoop platform. This is an authorization system that provides or denies access to Hadoop cluster resources such as HDFS files, Hive tables, etc. based on pre-defined security policies and then enforced consistently across the HDP stack. Apache Ranger uses Kerberos to authenticate the user and collaborates with Apache Knox for authorization [9]. Apache Ranger provides an additional program for synchronizing users and downloading users and groups from LDAP or Active Directory. Information about the found user or group is stored in the Ranger portal and is used to define a policy [11]. The user's request assumes that it is already authenticated when it comes to Ranger. It has different authorization functionality for different Hadoop components such as Hive, HBase, etc. To achieve this goal, Ranger uses plugins. Plugins are small programs written in the Java programming language that is embedded in the processes of each component of the cluster. When a user resource access request is available, the plugin intercepts the request, collects the data, and compares it against the user data on the master server. They then download the policies that correspond to the roles that the user has and return access to the resource if one is available. For example, the Apache Ranger plugin for Apache Hive is built into HiveServer2.

This project provides a security approach for the Hadoop cluster, providing the user with centralized control of the RBAC-based File Policy Management (Role-based access

control) user interface [4]. In terms of communication, the Ranger project provides the user with the REST API to provide a communication platform and manage the overall data security of the Hadoop platform [9].

It is important to know that it's working as the sole authorizer. This means that only the security policies and rules implemented in Ranger apply. System administrators can create a security policy for a specific set of resources for example folders and files and assign a specific set of permissions /read, write, execute/ to a specific set of users and groups. Security policies are stored in a single place in the policy manager [12].

4. Work process comparison of Knox and Ranger

The two Apache projects are interconnected and can operate in collaboration mode. They create a higher level of security and ensure proper access under certain restrictions.

Knox is the gateway or can be viewed as such. The first step, which decides whether to allow the user access to the Hadoop cluster or not based on user data. Ranger, in turn, determines who can access certain resources in a Hadoop cluster using policies and predefined roles. The working process is presented in Figure 1.

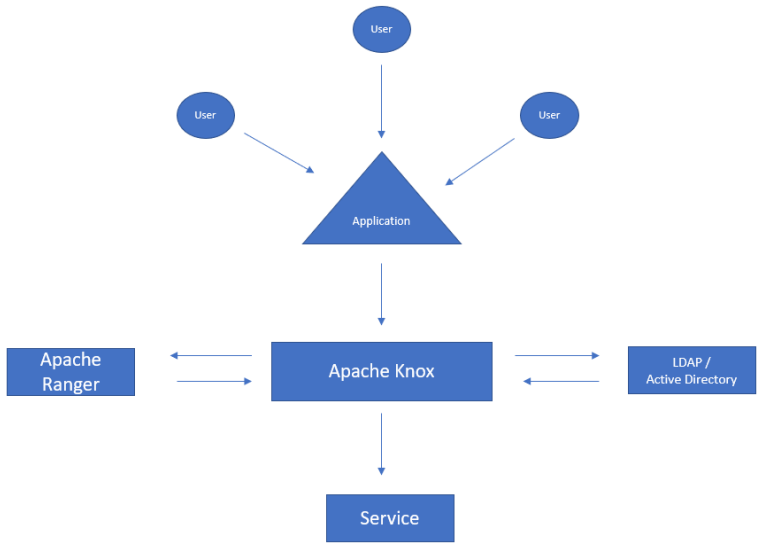


Fig. 1. Workflow of Apache Knox and Apache Ranger.

As shown in Figure 1, some users want to access a resource in the Hadoop system. When entering their credentials to log in to the authentication system with the request and this data is sent to the following instance Apache Knox.

Knox, for its part, verifies user credentials by querying an LDAP server and, based on the result of the request, decides whether to grant or deny access to that user to the interface it is trying to access. After granting user access, Knox will receive an authentication token from Kerberos to be able to pass the request to the target Hadoop component. All communication between the user and Knox is encrypted via the SSL (HTTPS) protocol.

In the next step, Knox connects to the LDAP protocol from the Active Directory to confirm the data entered by the client. If there is a confirmation, Knox starts checking the stored services to find the requested service by the user by requesting it from a Service. If the

requested service is found, Knox starts a follow-up check on the type of authentication to know exactly what the required parameters are for the requested service and to compare them with those sent by the customer.

After verifying access to one of the Hadoop components and Knox has made a connection to the Apache Ranger for authorization, the user is expected to take the next action. When the latter is inside the system and tries to read/write in HDFS or one of the components of Hadoop, Apache Ranger sends the user's request information to Ranger to confirm the rights and roles of the user to provide and modify the requested data and audit what the user wanted to do.

5. Conclusion

Despite the large size of the data and its unstructured appearance, Hadoop's system manages to store, process, and extract knowledge from the results. However, for data to be protected, the work must be transferred to secure environments. Even though the Hadoop system was not originally designed as the most secure environment, with the development of technology and the entry into the IoT environment, developers are creating various tools to help secure the system. Some of the best known are Kerberos, Apache Knox, and Apache Ranger, which are the subject of this paper. With the capabilities they offer and the functionality to work in collaboration, Hadoop and data access are becoming a more secure work environment.

References

1. Dr. M Praveen Kumar, Sampurnima Pattem, Security Issues in Hadoop Associated With Big Data
2. Shefali Aggarwal, Keeping Big Data Safe: Common Hadoop Security Issues and Best Practices, 2015, https://www.qubole.com/blog/hadoop-security-issues/
3. Introduction to Hadoop Security – How to secure a Hadoop Cluster?, https://data-flair.training/blogs/hadoop-security/
4. Aisha Javed, 3 basic a's of identity and access management authentication, authorization, and accounting, 2019, https://www.xorlogics.com/2019/04/15/3-basic-as-of-identity-and-access-management-authentication-authorization-and-accounting/
5. Priya P. Sharma, Chandrakant P. Navdeti, Securing Big Data Hadoop: A review of security Issues, Threats and Solutions, 2014, https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.637.2627&rep=rep1&type=pdf
6. Cloudera, Configuring the Key Management Server (KMS), https://docs.cloudera.com/documentation/enterprise/6/6.1/topics/cdh_sg_kms.html
7. Walker Rowe, Introduction to Hadoop Security, 2016, https://www.bmc.com/blogs/hadoop-security/#
8. Jitendra Gautam, Securing the Hadoop environment, 2020, https://www.linkedin.com/pulse/machine-learning-big-data-introduction-hadoop-security-gautam
9. 5 Hadoop Security Projects, 2020, https://www.integrate.io/blog/5-hadoop-security-projects/#knox
10. Yinyi Soo, Discussion and Comparison of Several Hadoop Security Tools, 2020, https://ysoo23.medium.com/discussion-and-comparison-of-several-hadoop-security-tools-b4532a8c67f9
11. Knox Implementation with AD/LDAP with HDP, 2020, https://bd-practice.medium.com/knox-implementation-with-ad-ldap-with-hdp-7e84c6305b60
12. Cloudera, Apache Ranger, https://www.cloudera.com/products/open-source/apache-hadoop/apache-ranger.html