International Journal of Differential Equations and Applications
Volume **23**, No. 1 (2024), pages: 83-95
ISSN (Print): 1311-2872; ISSN (Online): 1314-6084;
url: https://www.ijdea.eu

AUTOMATIC TRANSLATION OF SQUAD AND RACE QUESTION ANSWERING DATASETS IN BULGARIAN LANGUAGE

Simeon Monov[1], Detelinka Trifonova[2], Nikolay Pavlov[3],
Andrey Nikolov[4]

[1,2,3,4] Facaulty of Mathematics and Informatics
Paisii Hilendarski University of Plovdiv
236, Bulgaria Blvd., 4027 Plovdiv, BULGARIA

ABSTRACT:

There are many question-answering (QA) datasets, used in different natural language processing (NLP) tasks with SQuAD one of the most popular QA dataset around. RACE dataset is popular dataset for Multi Choice Question Answering (MCQA) task and used to evaluate and train MCQA models. These datasets are available in English language only.

We took these two datasets and translated them in Bulgarian language using automated translation techniques. After that we evaluated the new translated datasets on Extractive QA and MCQA tasks. Experimental results show, that our datasets can be effectively used to improve the performance of transformer models on QA tasks in Bulgarian language.

Key Words: Bulgarian dataset, question answering, LLM, transformer model.

Received: October 11, 2024 **Revised:** December 10, 2024
Published: December 19, 2024 **doi:** 10.12732/ijdea.v23i1.7

1. INTRODUCTION

With the advance of the transformer models and large language models (LLMs) the need for large train and validation datasets for specific tasks increased a lot. The LLMs are typically pre-trained on vast amount of text data in order for them to gain general

language understanding. During this process the models learn the language structure, grammar rules and gain reasoning abilities. Pre-trained models can perform limited amount of tasks such as text generation. They are also the foundation for fine-tuning, which is the process of adapting the general knowledge of the models to learn to perform more specific (downstream) tasks. Example downstream tasks are sentiment analysis, summarization, text classification, question answering, question generation, etc. Most modern LLMs such as GPT-4 [1], Llama [2], mT5 [3], Flan-T5 [4] are both pre-trained and fine-tuned on a lot of downstream tasks.

One very popular downstream task in NLP is Question Answering (QA). It is very important and challenging task. Recently very good results are achieved on this task mostly by fine-tuning pre-trained transformer models [5, 6, 7]. These models are extensively trained on this task and they are able to achieve state-of-the-art results. Very important part of these results is the availability of QA datasets for fine-tuning the models. Such large datasets exist for English language but unfortunately there are not many available for other languages such as Bulgarian language.

In this paper we take the approach on automatically translating two of the most popular QA datasets SquAD [8] and RACE [9] from English to Bulgarian language. We then fine-tune different models, which are pre-trained on Bulgarian data and evaluate their performance. We also use our dataset to evaluate the performance of different instruct LLMs using system and zero-shot prompting on the same task in Bulgarian language. In summary our contributions are:

- we created two Bulgarian QA datasets BGSQuAD and BGRACE by automatically translating them from English language and released the datasets;
- we trained several transformer models and evaluated their performance on Question Answer and Multi Choice Question Answer (MCQA) tasks;
- we additionally evaluated the performance of popular LLMs on our dataset using zero-shot learning.

2. QUESTION ANSWERING DATASETS

Question answering is one of the most researched areas in NLP. Majority of this research is in English language. Following is a selection of review and research articles with English benchmark datasets [10, 11, 12, 13].

2.1. THE STANFORD QUESTION ANSWERING DATASET

The most used QA validation dataset nowadays is the Stanford Question Answering Dataset (SQuAD). It is a crowd-sourced collection of question-answer pairs derived from Wikipedia articles, designed for answer diversity. SQuAD v1.1 includes 107, 785 question-answer pairs across 536 paragraphs. Each question refers to specific article/paragraph and the answer is also located inside the text of the article. The dataset is split into three sub-datasets: train, dev and test. Train and dev are used for training and validation purposes and test does not contain answers but only articles with questions.

2.2. RACE MULTI CHOICE QUESTION ANSWERING DATASET

The Large-scale ReAding Comprehension Dataset From Examinations (RACE) is a machine reading comprehension dataset with 27, 933 passages and 97, 867 questions from English exams designed for Chinese students aged 12–18. It includes two subsets, RACE-M for middle school and RACE-H for high school, with questions crafted by domain experts to test human reading skills rather than crowd-sourced or heuristic-based questions. Each question has four possible answers, with one correct choice.

2.3. NON-ENGLISH QUESTION ANSWERING DATASETS

Although there were numerous attempts, currently there are not so many free rich QA datasets in other language than English. In [14] the authors made a detailed survey on non-English Question Answering Datasets. In their work they reviewed existing datasets in 14 different languages, including Bulgarian language. Many of the reviewed datasets have been directly translated from English. For some of the more popular languages like Chinese more datasets exists but low resource languages such as Bulgarian there are no rich datasets which can be useful for fine-tuning task. In [15] the authors summarized the performance of 17 datasets in 14 languages, created from SQuAD with machine translation. Their findings can be seen in Table 1. Most of the data was translated using Google Translate and in some rare cases using different translation system.

Table 1: Overview of machine-translated SQuAD datasets [15]

Language	**Title**	**Translation**	**Best model**	**Evaluation metric [%]**	**Reference**
Arabic	Arabic-SQuAD	Google Translate	BERT	EM=34.10, F1=48.60	Mozannar et al., 2019 [16]
Bengali	Bengali-SQuAD	Google Cloud API	DistilBERT	EM=50.05, F1=51.18	Mayeesha et al., 2021 [17]
Czech	Czech SQUAD	LINDAT Translator	XLM-R large	EM=75.57, F1=79.19	Mackova & Straka, 2020 [18]
Danish	SQuAD-da	TAR Method/Moses			Carrino et al., 2020 [19]
French	SQuAD-fr	Google Translate	BERT base	EM=71.70, F1=86.70	Cattan et al., 2022 [20]
Hindi	Hindi SQUAD	Google Translate		EM=50.11, F1=53.77	Gupta et al., 2020 [21]
Italian	SQUAD-it	DeepL	BERT base	EM=75.00, F1=82.20	Croce et al., 2019 [22]
Korean	K-QUAD	Google Translate	BiDAF	EM=50.72, F1=71.50	Lee et al., 2018 [23]
Persian	ParSQuAD	Google Translate	mBERT	EM=67.73, P1=70.84	Abadani et al., 2021 [24]
Polish	PolAQ		mBERT	EM=72.56, F1=80.39	Jurkiewicz, 2020 [25]
Polish	SQUAD-pl	Google Translate	-	-	Brodzik, 2022 [26]
Portuguese	SQuAD_v1.1_pt	Google Cloud API	BERT	Acc=50	Carvalho, 2019 [27]
Portuguese	SQuAD_v2.0_pt	Google Cloud API	-	-	Janiake, 2020 [28]
Spanish	SQuAD-es-v1.1	TAR Method/Moses	mBERT	EM=48.30, F1=68.10	Carrino et al., 2020 [19]
Spanish	SQuAD-es-v2.0	TAR Method/Moses	mBERT	EM=76.50, F1=86.07	Carrino et al., 2020 [19]
Swedish	SQuAD-v2-sv	Google Translate	BERT base	EM=66.73, F1=70.11	Okazawa, 2021 [29]
Ukrainian	SQuAD-uk	Google Cloud API	mBERT	EM=56.10, F1=62.20	Tiutiunnyk & Dyomkin, 2019 [30]

3. THE BGSQUAD AND BGRACE DATASETS

We translated SQuAD v1.1 and RACE datasets using deep-translator library with google translation engine. The translation of both datasets took 5 days. During translation we kept the same source structure for the output data. For SQuAD we translated

train and dev subsets. For RACE, high and mid subsets of train, dev and test datasets were translated. In table 2 we can see examples of original English and translated questions from SQuAD. Table 3 shows examples of translation of questions and answers from RACE dataset.

Table 2: Original and translated SQuAD context and question example

SQuAD v1.1	BGSQuAD

{
"context": "In December 1878, Tesla left Graz and severed a ll relations with his family to hide the fact that he dropped out of school. His friends thought that he had drowned in the Mur River. Tes la went to Maribor (now in Slovenia), where he worked as a draftsman for 60 florins a month. He spent his spare time playing cards with lo cal men on the streets. In March 1879, Milutin Tesla went to Maribor to beg his son to return home, but Nikola refused. Nikola suffered a nervous breakdown at around the same time.",
"qas": [
{
"answers": [
{
"answer_start": 24,
"text": "left Graz"
},
{
"answer_start": 24,
"text": "left Graz"
},
{
"answer_start": 24,
"text": "left Graz and severed all relations with h is family"
}
],
"question": "What did Tesla do in December 1878?",
"id": "56dfa7887aa994140058dfa9"
},
	{
	"context": "През декември 1878 г. Тесла напуска Грац и прекъс ва всички отношения със семейството си, за да скрие факта, че е напус нал училище. Приятелите му помислили, че се е удавил в река Мур. Тесл а отива в Марибор (сега в Словения), където работи като чертожник за 60 флорина на месец. Той прекарваше свободното си време в игра на кар ти с местни мъже по улиците. През март 1879 г. Милутин Тесла отива в Марибор, за да моли сина си да се върне у дома, но Никола отказва. Ни кола получава нервна криза горе-долу по същото време.",
	"qas": [
	{
	"answers": [
	{
	"answer_start": 24,
	"text": "напусна Грац"
	},
	{
	"answer_start": 24,
	"text": "напусна Грац"
	},
	{
	"answer_start": 24,
	"text": "напуска Грац и прекъсва всички отношения със семейството си"
	}
	],
	"question": "Какво направи Тесла през декември 1878 г.?",
	"id": "56dfa7887aa994140058dfa9"
	},

Table 3: Original and translated RACE article and questions example

RACE	BGRACE
{ "id": "high1717.txt", "answers": ["B", "C", "B", "D"], "options": [["his parents were too careless", "his parents thought he had watched too much TV", "Santa Claus was not satisfied with Jack's be havior", "Santa gave the TV to another child as a present"], ["Santa would not know where the tree was.", "Santa would be a ngry and would not give her any gifts.", "Her big brother might lau gh at her.", "Santa might think she was a \"bad\" child."], ["Lydia's mother was very Strict with her", "Lydia believed in Santa when she was young", "Lydia was naughty when she was young", "Lydia liked taking pictures with Santa"],	{ "id": "high1717.txt", "answers": ["B", "C", "B", "D"], "options": [["родителите му бяха твърде небрежни", "родителите му смятаха, че е гледал твърде много телевизия", "Дядо Коледа не беше доволен от поведението на Джак", "Дядо Коледа подари телевизора на друго де те"], ["Дядо Коледа нямаше да знае къде е дървото.", "Дядо Коледа ще ше да бъде ядосан и нямаше да й даде никакви подаръци.", "Големият й брат може да й се смее.", "Дядо Коледа може да я помисли за „лошо “ дете."], ["Майката на Лидия беше много строга с нея", "Лидия вярваше в Дядо Коледа, когато беше малка", "Лидия беше палава като млада", "Л идия обичаше да се снима с Дядо Коледа"],

```
    [
      "She was afraid that Santa would get mad with her.", "She was
 afraid that Santa would get too tired and hurt himself.", "She dis
liked the idea that Santa would keep a copy of her picture.", "She
feared that she would appear in the Santa's naughty list."
    ]
  ],
  "questions": [
    "The real reason why Jack's TV was taken away is that ___ .",
"Which of the following is NOT one of the reasons why Lucy didn't w
ant the tree to be moved?", "We can learn from the third story that
 _ .", "Why didn't Lydia want to take pictures with Santa?"
  ],
  "article": "One year, my school report made my parents angry. On
Christmas Eve, all the presents were stolen, along with our TV. My
parents told me that there were no presents because Santa was very
angry with my behavior over the past year. The next year on Christm
as Eve I slept downstairs with a plastic sword waiting for Santa to
 make sure that he didn't steal the new TV. The next morning, when
I woke up, I saw Santa standing there. As soon as I saw that there
were no presents, I grabbed my plastic sword and ran at him, shouti
ng angrily: \"THIEF! THIEF!\"\nJack\nWhen I was young, we always ha
d a specific room for the Christmas tree. My mom never really liked
 the location, so one year she moved the tree into another room. I
```

```
    [
      "Тя се страхуваше, че Дядо Коледа ще й се разсърди.", "Тя се
страхуваше, че Дядо Коледа ще се умори и ще се нарани.", "Тя не хар
есваше идеята, че Дядо Коледа ще запази копие от нейната снимка.",
"Страхуваше се, че ще се появи в списъка на палавите на Дядо Коледа
."
    ]
  ],
  "questions": [
    "Истинската причина, поради която телевизорът на Джак беше отне
т е, че _ .", "Кое от следните НЕ е една от причините Луси да не ис
ка дървото да бъде преместено?", "От третата история можем да научи
м, че _ .", "Защо Лидия не искаше да се снима с Дядо Коледа?"
  ],
  "article": "Една година докладът ми от училище ядоса родителите м
и. На Бъдни вечер всички подаръци бяха откраднати, заедно с телевиз
ора ни. Родителите ми казаха, че няма подаръци, защото Дядо Коледа
беше много ядосан от поведението ми през изминалата година. На след
ващата година на Бъдни вечер спах долу с пластмасов меч и чаках Дяд
о Коледа да се увери, че няма да открадне новия телевизор. На следв
ащата сутрин, когато се събудих, видях Дядо Коледа да стои там. Щом
 видях, че няма подаръци, грабнах пластмасовия си меч и се втурнах
към него, викайки ядосано: „КРАДЕЦ!\nДжак\nКогато бях млад, винаги
имахме определена стая за коледната елха. Майка ми никога не е харе
свала мястото, така че една година премести дървото в друга стая. Б
```


4. TESTING AND EVALUATION OF THE DATASETS

We conducted two main types of tests:

1. Fine tune transformer models on the train subset of both BGSQuAD and BGRACE datasets. Evaluation was done on dev dataset from BGSQuAD and test dataset from BGRACE.

2. Use zero-shot and system prompting to test different instruct models. These tests were performed on the full datasets.

The following transformer models were used for our tests:

- **bert-base-multilingual-uncased** — This Bert [31] is a multilingual version that supports 104 languages, designed without case sensitivity.
- **roberta-base-bulgarian** — A Bulgarian-language adaptation of RoBERTa [32], pre-trained to support various NLP tasks.
- **mT5-base** — The multilingual variant of the T5 [33] model, pre-trained on text across 101 languages. It's designed for tasks like translation, summarization, and question answering.
- **mT5-small** — A smaller variant of the multilingual T5 model, designed to handle multilingual tasks efficiently with lower computational resources.
- **mT5-large** — A larger variant of the multilingual T5 model, offering higher capacity and better performance on complex multilingual tasks due to its increased model size.

- **GPT-4o-mini** — A compact version of GPT-4o optimized for efficiency, ideal for tasks requiring reduced computational resources.
- **Llama-3.1-8B-Instruct** — A larger version of the LlaMA model, designed for high-performance NLP tasks.
- **Llama-3.2-3B-Instruct** — A smaller variant of the LLaMA model, optimized for efficiency, suitable for medium-scale NLP tasks.
- **BgGPT-7B-Instruct-v0.2** — This model is adapted and fine-tuned to understand and generate natural language specifically for Bulgarian.

All tests were performed on a single NVIDIA Tesla V100 GPU with 32GB VRAM.

4.1. BGSQUAD RESULTS

BGSQuAD dataset was tested by fine-tuning Llama-3.2-3B-Instruct (for 1 epoch) and mT5-base, small and large models (for 3 epochs). Also, Llama-3.1-8B-Instruct, Llama-3.2-3B-Instruct and GPT-4o-mini were tested using zero-shot and system prompting. Both EM and F1 metrics were measured. The results are shown in Table 4.

Table 4: BGSQuAD dev evaluation results

Model Name	**Metrics**	
	EM	**F1**
mT5-large - 3 epochs	34.3%	**69.1**%
mT5-base - 3 epochs	32%	66.4%
mT5-small - 3 epochs	26.2%	59.2%
Llama-3.2-3B (fine-tuned) - 1 epoch	**36.9%**	59.3%
Llama-3.2-3B	10.8%	30.6%
Llama-3.1-8B	16.6%	38.5%
GPT-4o-mini	12.1%	41.8%

4.2. BGRACE RESULTS

BGRACE dataset was tested by fine-tuning Llama-3.2-3B-Instruct (for 1 epoch), roberta-base-bulgarian and bert-base-multilingual-uncased (for 3 epochs). Zero-shot prompting tests were performed on Llama-3.1-8B-Instruct, Llama-3.2-3B-Instruct, gpt-4o-mini

and BgGPT-7B-Instruct-v0.2. Accuracy percent was measured for all the tests. The results are shown in Table 5.

Table 5: BGRACE test evaluation results

Model Name	**Accuracy**
roberta-base-bulgarian – 1 epoch	39%
roberta-base-bulgarian – 3 epochs	42%
Llama-3.1-8B	26%
Llama-3.2-3B	16%
Llama-3.2-3B (fine-tuned) – 1 epoch	**75%**
bert-base-multilingual-uncased – 1 epoch	47%
bert-base-multilingual-uncased – 3 epochs	52%
GPT-4o-mini	64%
BgGPT-7B	47%

We can see that fine-tuning Llama-3.2-3B on our dataset improves its performance significantly.

Additional tests were conducted using another MCQA datasets — EXAMS. EXAMS [34] is benchmark dataset for cross-lingual and multilingual question answering for high school examinations. It contains 1100 train examples in Bulgarian language. We used these train examples to test the different models including fine-tuned Llama-3.2-3B model on BGRACE. The results are shown in Table 6.

Table 6: EXAMS train evaluation results

Model Name	**Accuracy**
Llama-3.1-8B	45%
Llama-3.2-3Bs	20%
Llama-3.2-3B (fine-tuned) – 1 epoch	72%
GPT-4o-mini	**74%**
BgGPT-7B	60%

We can see, that after fine-tuning Llama-3.2-3B on our dataset it improves its performance significantly and can achieve state-of-the-art results similar to the much bigger GPT-4o-mini model.

5. CONCLUSION

In this work we translate SQuAD and RACE Question Answering datasets from English to Bulgarian language. We further train different transformer models and evaluate their performance. We also conduct evaluation on different large language models using zero-shot and system prompting. Our tests show, that our datasets can be effectively used to improve the performance of the transformer models on QA tasks in Bulgarian language. Large models like GPT-4o can perform quite well on Bulgarian language QA tasks, but much smaller models like Llama-3.2-3B need additional fine-tuning and our datasets can help in achieving state-of-the-art results for these models.

ACKNOWLEDGMENT

This paper is partially supported by project MUPD23-FMI-009 of the Scientific Fund of the Paisii Hilendarski University of Plovdiv, Bulgaria.

REFERENCES

[1] Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S. and Avila, R., Gpt-4 technical report, (2023), arXiv preprint arXiv: 2303.08774.

[2] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F. and Rodriguez, A., Llama: Open and efficient foundation language models, (2023), arXiv preprint arXiv: 2302.13971.

[3] Xue, L., mt5: A massively multilingual pre-trained text-to-text transformer, (2020), arXiv preprint arXiv: 2010.11934.

[4] Chung, H.W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S. and Webson, A., Scaling instruction-finetuned language models, *Journal of Machine Learning Research*, **25**(70) (2024), 1-53.

[5] Su, D., Xu, Y., Winata, G.I., Xu, P., Kim, H., Liu, Z. and Fung, P., Generalizing question answering system with pre-trained language model fine-tuning, In: *Proceedings of the 2nd workshop on machine reading for question answering*, (2019), 203-211.

[6] Phogat, K.S., Puranam, S.A., Dasaratha, S., Harsha, C. and Ramakrishna, S., Fine-tuning Smaller Language Models for Question Answering over Financial Documents, (2024), arXiv preprint arXiv: 2408.12337.

[7] Ye, J., Yang, Y., Zhang, Q., Gui, T., Huang, X., Wang, P., Shi, Z. and Fan, J., Empirical Insights on Fine-Tuning Large Language Models for Question-Answering, (2024), arXiv preprint arXiv: 2409.15825.

[8] Rajpurkar, P., Squad: 100,000+ questions for machine comprehension of text, (2016), arXiv preprint arXiv: 1606.05250.

[9] Lai, G., Xie, Q., Liu, H., Yang, Y. and Hovy, E., Race: Large-scale reading comprehension dataset from examinations, (2017), arXiv preprint arXiv: 1704.04683.

[10] Cambazoglu, B.B., Sanderson, M., Scholer, F. and Croft, B., A review of public datasets in question answering research, In: *ACM SIGIR Forum*, **54**(2) (2011), 1-23, New York, NY, USA: ACM.

[11] Wang, Z., Modern question answering datasets and benchmarks: A survey, (2022), arXiv preprint arXiv: 2206.15030.

[12] Pandya, H.A. and Bhatt, B.S., Question answering survey: Directions, challenges, datasets, evaluation matrices, (2021), arXiv preprint arXiv: 2112.03572.

[13] Qamar, F., Latif, S. and Shah, A., Techniques, datasets, evaluation metrics and future directions of a question answering system, *Knowledge and Information Systems*, **66**(4) (2024), 2235-2268.

[14] Chandra, A., Fahrizain, A. and Laufried, S.W., A survey on non-english question answering dataset, (2021), arXiv preprint arXiv: 2112.13634.

[15] Hládek, D., Staš, J., Juhár, J. and Koctúr, T., Slovak dataset for multilingual question answering, *IEEE Access*, **11** (2023), 32869-32881.

[16] Mozannar, H., Hajal, K.E., Maamary, E. and Hajj, H., Neural Arabic question answering, (2019), arXiv preprint arXiv: 1906.05394.

[17] Tahsin Mayeesha, T., Md Sarwar, A. and Rahman, R.M., Deep learning based question answering system in Bengali, *Journal of Information and Telecommunication*, **5**(2) (2021), 145-178.

[18] Macková, K. and Straka, M., Reading comprehension in Czech via machine translation and cross-lingual transfer, In: *Text, Speech, and Dialogue: 23rd International Conference*, TSD 2020, Brno, Czech Republic, September 8–11, 2020, Proceedings 23 (pp. 171-179), Springer International Publishing.

[19] Carrino, C.P., Costa-Jussà, M.R. and Fonollosa, J.A., Automatic spanish translation of the squad dataset for multilingual question answering, (2019), arXiv preprint arXiv: 1912.05200.

[20] Cattan, O., Servan, C. and Rosset, S., On the Usability of Transformers-based models for a French Question-Answering task, (2022), arXiv preprint arXiv: 2207.09150.

[21] Gupta, S. and Khade, N., Bert based multilingual machine comprehension in english and hindi, (2020), arXiv preprint arXiv: 2006.01432.

[22] Croce, D., Brandi, G. and Basili, R., Deep Bidirectional Transformers for Italian Question Answering, In: *CLiC-it*, (2019).

[23] Lee, K., Yoon, K., Park, S. and Hwang, S.W., Semi-supervised training data generation for multilingual question answering, In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, (2018).

[24] Abadani, N., Mozafari, J., Fatemi, A., Nematbakhsh, M.A. and Kazemi, A., ParSQuAD: machine translated squad dataset for Persian question answering, In: *2021 7th International Conference on Web Research (ICWR)*, (2021), 163-168.

[25] Jurkiewicz, T., PolAQ-Polish Answers and Questions, (2020), [Online]. Available: https://github.com/tjur/polaq_master_thesis

[26] Brodzik, A., SQuAD-PL-Polish Google Translated SQuAD v2.0 Dataset, (2022), [Online]. Available: https://github.com/brodzik/SQuAD-PL

[27] Carvalho, N. R., Question-Answering Model Fine-Tuned for Portuguese, (2019), [Online]. Available: https://medium.com/nunorc/question-answering-model-fine-tuned-for-portuguese-801bb05b6119

[28] Janiake, C., Portuguese Translation of SQuAD 2.0 Dataset, (2020), [Online]. Available: https://github.com/cjaniake/squad_v2.0_pt

[29] Okazawa, S., Swedish Version of SQuAD 2.0, (2021), [Online]. Available: https://github.com/susumu2357/SQuAD_v2_sv

[30] Tiutiunnyk, S., Context-based question-answering system for the Ukrainian language, In: *Proc. 1st Masters Symp. Adv. Data Mining, Mach. Learn., Comput. Vis. (MS-AMLV)*, Lviv, Ukraine, (2020), 81-88.

[31] Devlin, J., Chang, M., Lee, K., Toutanova, K., Bert: Pre-training of deep bidirectional transformers for language understanding, (2018), arXiv preprint arXiv: 1810.04805.

[32] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V., RoBERTa: A robustly optimized BERT pretraining approach, (2019), arXiv preprint arXiv: 1907.11692.

[33] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P.J., Exploring the limits of transfer learning with a unified text-to-text transformer, Journal of machine learning research, **21**(140) (2020), 1-67.

[34] Hardalov, M., Mihaylov, T., Zlatkova, D., Dinkov, Y., Koychev, I. and Nakov, P., EXAMS: A multi-subject high school examinations dataset for cross-lingual and multilingual question answering, (2020), arXiv preprint arXiv: 2011.03080.

[35] Hamdani, S.W.A., Abbas, H., Janjua, A.R., Shahid, W.B., Amjad, M.F., Malik, J., Murtaza, M.H., Atiquzzaman, M., and Khan, A.W., Cybersecurity standards in the context of the operating system: Practical aspects, analysis, and comparisons, *ACM Computing Surveys (CSUR)*, **54**(3) (2021), 1-36.

[36] Saravanan, A. and Bama, S.S., A review on cyber security and the fifth generation cyberattacks. *Oriental journal of computer science and technology*, **12**(2) (2019), 50-56.

[37] https://doc.owncloud.com/ocis/next/availability_scaling/availability_scaling.html, last visit November 2023.

[38] Ogiriki, I., Beck, C. and Heydari, V., *Technical Analysis of Thanos Ransomware*, (2022).

[39] Hristev, R., Veselinova, M., Kolev, K., Ransomware Target: Linux. Recover Linux Data Arrays after Ransomware Attack, *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)*, **19** (2022), 78-86.

[40] Kok, S., Abdullah, A., Jhanjhi, N. and Supramaniam, M., Ransomware, threat and detection techniques: A review. *Int. J. Comput. Sci. Netw. Secur* **19**(2) (2019), 136.

[41] Tailor, J.P. and Patel, A.D., A comprehensive survey: ransomware attacks prevention, monitoring and damage control. *Int. J. Res. Sci. Innov*, **4**(15) (2017), 116-121.

[42] Hristev, R. and Veselinova, M., Using private cloud for information arrays recovery from ransomware attacks. *AIP Conference Proceedings 2505, 060006*, (2022).

[43] Golev, A., Hristev, R., Veselinova, M., Kolev, K., Crypto-ransomware attacks on Linux serveces: A data recovery method, *Intern. J. Diff. Eq. Appl.*, **21** (2022), 19-29.

[44] Hristev, R., Veselinova, M. and Kolev, K., RANSOMWARE ATTACKS ON WINDOWS SERVERS: INFECTION AND RECOVERY. *International Journal of Differential Equations and Applications*, **22**(1) (2023), 57-66.