Brief information about the project

Name of the project	AP19676342, Multiclassification of ideological trends of cyber extremism in the Kazakh language using artificial intelligence methods
Relevance	Extremists can use the Internet to spread their ideas, recruit new supporters, coordinate actions, and even plan and commit crimes. The Internet provides them with ample opportunities for communication and propaganda, which makes it important to work on controlling content and countering extremist groups in the online environment. Defining the ideology of extremist speech is of great importance in the modern world, especially in the context of online communications. The relevance of this definition is due to a number of factors. Governments and public organizations strive to counter extremism and terrorism, and for this it is important to clearly identify the ideologies underlying them. The spread of extremist ideas on the Internet can lead to radicalization and negative consequences. Identifying and combating such speech helps to create a safe online space. Cooperation between countries and international organizations in the fight against extremism requires a clear definition of ideologies in order to ensure uniform standards and approaches. Thus, the relevance of defining the ideologies of extremist speech is undeniable in the modern information society.
Purpose	The aim of the project is to study the spread of destructive cyber propaganda using artificial intelligence methods in social networks and messengers, create multiclass models in ideological directions of extremist content in text, audio and video publications and create the most active cyber propaganda of religious content, models and methods for identifying destructive communities.
Objectives	1. Development of a data collection module. 2. Creation of a corpus for multiclassing the ideological direction of extremist content. 3. Creation of a multiclassification model that defines the ideological orientation of extremist content in text publications of social networks and messengers (propaganda of destructive religious movements, radicalization and involvement in extremist and terrorist organizations). 4. Creation of multiclass models to determine the ideological orientation of extremist content in audio and video publications of social networks and messengers 5. Creation of a multiclassication model for social networks and messengers that defines extremist (religious extremism, political extremism, xenophobia) content in text, audio, and video publications. 6. Creation of a hybrid model for identifying the most active cyber propaganda in social networks and messengers

	7. Development of a model, algorithms and methods for identifying communities in social networks based on a given set of parameters. 8. Creation of a chatbot with a dialogue in the Kazakh language for counseling on religious issues. 9. Development of software that implements the developed methods and models.
Expected and achieved results	Achieved results: A data collection module has been developed. The data collection module was developed using API technologies for searching data on social networks Telegram, VKontakte, Twitter, Youtube. More than 400 groups with signs of destructive beliefs were analyzed in selected social networks. The data collection module implements the functionality of replenishing the database in accordance with the list of keywords and based on the selected time interval, requesting the group ID. A multiclassical corpus of extremist content and ideological orientation has been created. To determine the ideological extremist orientation that collects text data from social networks and news sites, the rules and categories of extremist and neutral texts were defined. Lists of keywords for each class, rules, and lists of groups for data collection have been compiled. The case is sorted manually according to the rules. As a result of the annotation, the following designations were assigned to the classes: Propaganda (0), Radicalization (1), Recruitment (2), Neutral (3). The assembled corpus was divided into training and testing kits in the ratio of 80% and 20% in order to apply machine learning methods. There are more than 2000 texts in each class. Preprocessing algorithms such as tokenization, punctuation purification, purification from the most common words, and purification from stop words were applied to the corpus data. Statistics and visualization of the case have been compiled. Word2vec, bag of words and n-gram algorithms and models were used, and the corpus was prepared for machine learning. A model of multiclassification of the ideological direction of extremist content in text publications of social networks and messengers (propaganda of destructive religious movements, radicalization and involvement in extremist and terrorist organizations) has been created. word2vec and tf-idf algorithms were applied to the corpus data. Since the texts in web resources are mostly in an unstructured state and are filled in by various users, there are many spelling errors. Therefore, first of all, a Spell Checker-based method was proposed to correct spelling errors in the Kazakh language. The mentioned method is a very useful feature of any search engine. The simplest solution is to sort all dictionary positions as the editorial distance increases and display only the first few positions. The Levenshtein distance can be used as an editorial distance. Levenshtein distance-shows the minimum number of attempts to enter/delete/change characters required to convert the source string to the target string. To determine the ideological direction in the text, a comparative analysis of classical machine algorithms such as Logistic Regression, KNN, SVM, Naive Bayes, Decision Tree, Random Forest, Gradient Boosting was

	carried out. A new model has been created based on Spellcheker+Stemming+TF-idf+LSTM+BERT. Hyperparameters of the Bert model: input = 128 words or tokens, linear classification = 4, bert_model_name = 'Bert-base-multilingual-uncased', num_classes = 4, max_length = 128, batch_size = 64, num_epochs = 10, learning_rate = 2e-5. In addition, 2 deep learning algorithms (BERT and LSTM, Bert+linear) were combined to solve the problem of multiclassing ideological texts. For LSTM, the model accepts a sequence of text as input along with the corresponding lengths of each sequence. It embeds text (nested size = 20), processes it through the LSTM layer (size = 64), transmits the last hidden state through fully added layers with ReLU activations, and finally uses sigmoid activation to obtain a single output value. Hyperparameters of the combined model: input = 128 words or tokens, BERT = 768, LSTM = 256, DropOut = 0.2, linear Classification = 4, bert_model_name = 'Bert-base-multilingual-uncased', num_classes = 4, max_length = 128, batch_size = 64, num_epochs = 20, learning_rate = 2E-5. Expected results: Machine and deep learning models for multiclassication of extremist orientation of text, audio and video materials in social networks and messengers (religious extremism, political extremism, xenophobia, etc.), a hybrid model that is a combination of various deep neural networks selected depending on the type of tags, designed to identify the most active cyber propaganda in social networks and messengers.
Research team members with their identifiers (Scopus Author ID, Researcher ID, ORCID, if available) and links to relevant profiles	1. Mussiraliyeva Shynar, ORCID: https://orcid.org/0000-0001-5794-3649 , Scopus profile: https://www.scopus.com/authid/detail.uri?authorId=57202216979 , Web of Science profile: https://www.webofscience.com/wos/author/record/2394890 2. Bolatbek Milana, ORCID: https://orcid.org/0000-0002-2153-180X , Scopus profile: https://www.scopus.com/authid/detail.uri?authorId=57202834055 , Web of Science profile: https://www.webofscience.com/wos/author/record/GZL-7318-2022 3. Baisylbayeva Kymbat, ORCID: https://orcid.org/0000-0001-9753-0398 , Web of Science profile: https://www.webofscience.com/wos/author/record/N-9664-2017 4. Yeltay Zhastay, Researcher ID: https://www.webofscience.com/wos/author/record/JNR-6763-2023 , ORCID: https://orcid.org/my-orcid?orcid=0000-0002-9275-7582 Scopus author ID: https://www.scopus.com/authid/detail.uri?authorId=57237959800 5. Narbayeva Saltanat, ORCID: https://orcid.org/0000-0001-5230-3781 , Scopus profile: https://www.scopus.com/authid/detail.uri?authorId=57216484412 ,

	Web of Science profile: https://www.webofscience.com/wos/author/record/KBA-1599-2024 6. Medetbek Zhanar ORCID: http://orcid.org/0000-0001-7536-5889 7. Bekkozhayeva A. 8. Aetova B.
List of publications with links to them	1. 1. Sh.Zh. Musiralieva, R.Қ. Ospanov, M.A. Bolatbek, Zh.B. Medetbek Profilirovanie polzovatelei sotsialnykh setei na osnove demograficheskikh dannykh. (KOKSON) Zhurnal Vestnik AUES, Tom 3 # 62(2023), str. 133-144. https://vestnik.aues.kz/index.php/none/issue/view/95/116 2. 2. Musiralieva Sh.Zh., Baispai G.B., Bolatbek M.A., Sagynai M., Tereikovskii I.A. Veb-resurstardағy ekstremisttik məlimetterdi anyqtauga arnalgan mashinalyq adisterdi oqytu zhane synau ushin qazaq tilindegi mətinder korpusyn quru. Zhurnal Trudy universiteta (KarGTU), #3, 2023, str 453-458. Rekomendovano KOKSON, tu.kstu.kz/archive 3. 3. Sh.Zh. Musiralieva, M.A. Bolatbek, M. Sarynai, Zh.Y. Eltai, K.B. Bagitova. Ekstremistik malimetter tusinigi zhane ekstremizmge qarsy kures zhobalaryna zhuielik sholu. (KOKSON) . Zhurnal NEWS OF THE NATIONAL ACADEMY OF SCIENCESOF THE REPUBLIC OF KAZAKHSTAN PHYSICO-MATHEMATICAL SERIES, ISSN 1991-346X Volume 3. # 347 (2023). 112–130. https://journals.nauka-nanrk.kz/physics-mathematics/article/view/5792 4. 4. Sh.Zh.Musiralieva, Zh.B.Medetbek , M.A.Bolatbek , Zhumakhanova A.N., Zh.Y.Eltai. ƏLEUMETTIK MEDIADAGY EKSTREMISTIK MAZMUNDY ANYQTAU ZhANE ZERTTEU. V materialakh konferentsii. VIII Mezhdunarodnaia nauchno-prakticheskaia konferentsiia "Informatika i prikladnaia matematika" posviashchennoi pamiati 85 d.t.n, professora Biiasheva R.G. , 26 – 27 oktiabria 2023 g., 265-272str., Almaty, Kazakhstan 5. Shynar Mussiraliyeva, Milana Bolatbek, Moldir Sagynay, Aygerim Zhumakhanova, Zhastay Yeltay, and Zhanar Medetbek. 2024. Identifying Cyber-Threatening Texts in the Kazakh Segment of Web Resources. In Proceedings of the 2023 7th International Conference on Advances in Artificial Intelligence (ICAAI '23). Association for Computing Machinery, New York, NY, USA, 68–72. https://doi.org/10.1145/3633598.3633610
Patents	-