Brief information about the project

Name of the project	AP15473412 "Creation of a method for identifying nodes for disseminating criminal information in social networks" (0122PK00937)
Relevance	Social networks have become an integral part of our daily lives. They are used not only by individuals and legitimate businesses, but also by organized groups such as activists and cybercriminals to spread their ideas. This represents a new source of intelligence collection as it provides an inside look at the behavior of these previously closed, secretive groups. One potential opportunity for this online data source is to use public messaging on social media to identify key users in such groups. The idea of the project is to develop an algorithm for identifying nodes distributing malicious content on social networks. This is especially important for law enforcement agencies who want to track or interrogate influential people in suspicious groups.
Purpose	The goal of the project is to create a method for identifying nodes for the dissemination of criminal information in social networks on the territory of the Republic of Kazakhstan. Research on social network analysis metrics that can help identify key players in organized groups, mainly activists.
Objectives	The objectives of the project are: 1. Development of a machine learning model for detecting criminal information in social networks. 2. Building of a user graph in social networks. 3. Development of an algorithm for identifying nodes for disseminating criminal information in social networks. At the initial stage of the project, a dataset is created using a software module from public messaging based on keywords. Together with text data. The next step is to build a graph based on the collected data. User graphs will be built based on posts in groups and graphs based on comments under posts. The graphs were constructed using the matplotlib library. The last stage will be the development of an algorithm for identifying nodes distributing criminal information on social networks. Discovering influential nodes is a central research topic in social network analysis. Therefore, social network analysis metrics such as number of edges, number of vertices, graph density, average path length, Betweenness Centrality, Closeness Centrality, Degree Centrality and Chi Squared will be defined and analyzed. These properties are planned to be used in developing an algorithm for identifying the most influential nodes.
Expected and achieved results	Expected results: development of a machine learning model for detecting criminal information in online content, constructing a graph of users in social networks and developing an algorithm for identifying nodes for disseminating criminal information in social networks. Based on the results of the work, at least 2 (two) articles are planned in journals from the first three quartiles by impact factor in the Web

	of Science database or having a CiteScore percentile in the Scopus database of at least 50. The social significance of the project is associated with the creation of an algorithm for identifying nodes for the dissemination of criminal information in social networks for the domestic market. Since there is now a trend for monitoring social networks, there is a possibility of commercializing the results of the project. The target consumers of the results obtained are applied results in the form of methods and algorithms that can be used by authorized authorities to ensure information security and to combat criminal information.
Research team members with their identifiers (Scopus Author ID, Researcher ID, ORCID, if available) and links to relevant profiles	1. Baispay Gulshat Bolatkyzy, Master of Technical Sciences, Hirsch Index –2, Author ID in Scopus: 57221648127, ResearcherID ABB-9718-2021, ORCID ID: 0000-0003-4292-2938 2. Mussiraliyeva Shynar Zhenisbekovna, Candidate of Physical and Mathematical Sciences, h-index– 7, Scopus Author ID: 57202216979, ResearcherID ABA-9832-2021, ORCID ID: 0000-0001-5794-3649.
List of publications with links to them	1. S. Mussiraliyeva, G. Baispay, R. Ospanov, Z. Medetbek and K. Shalabayev, "Graphical Visualization of the Connections of Involved Users and Identifying Influential Spreaders in a Social Network," 2022 9th International Conference on Electrical and Electronics Engineering (ICEEE), 2022, pp. 311-315, doi: 10.1109/ICEEE55327.2022.9772556, https://ieeexplore.ieee.org/document/9772556 2. Ш.Ж. Мусиралиева, Г.Б. Байспай, А.К. Бекетова, Әлеуметтік желілердегі криминалдық ақпараттарды тарату тораптарын сәйкестендіру әдісін зерттеу және құру, Қазақстан Республикасы, Ұлттық инженерлік академиясының Хабаршысы, № 4 (86), 2022, 82-90 б., ISSN 2709–4693 3. Байспай Г.Б., Мусиралиева Ш.Ж., Әлеуметтік желілердегі криминалдық ақпаратты тарату тораптарын сәйкестендіру әдісін құру, «Фараби Әлемі» студенттер мен жас ғалымдардың халықаралық ғылыми конференциясы, «Қазақ университеті», Алматы, 2022, ISBN 978-601-04-5965-6 4. Байспай Г.Б., Мусиралиева Ш.Ж., Создание метода идентификации узлов распространения криминальной информации в социальных сетях, Международная конференция студентов и молодых ученых «Фараби Әлемі», «Қазақ университеті», Алматы, 2023 г., 120 стр., ISBN 978-601-04-6247-2 5. Мусиралиева Ш.Ж., Байспай Г.Б., Болатбек М.А., Сағынай М., Терейковский И.А., Веб-ресурстардағы экстремисттік мәліметтерді анықтауға арналған машиналық әдістерді оқыту және сынау үшін қазақ тіліндегі мәтіндер корпусын құру, Университет еңбектері, Республикалық журнал, №3 (92), 2023, 453-459 б, ISSN 1609-1825, DOI 10.52209/1609-1825_2023_3_453, http://tu.kstu.kz/archive/journal/26 6. Мусиралиева Ш.Ж., Азаматова Д.Т., Байспай Г.Б., Фэйк-аккаунттарды анықтау мәселесінде машиналық оқытуды қолдану, Университет еңбектері, Республикалық журнал, №4

	(93), 2023, 360-369б, ISSN 1609-1825, DOI 10.52209/1609-1825_2023_4_360, http://tu.kstu.kz/archive/journal/26 7. Байспай Г.Б., Мусиралиева Ш.Ж., Терейковский И.А., Интернеттегі әлеуметтік желілердегі қылмыстық ақпараттың ықпалды таратушыларын анықтау, VIII Международная научно-практическая конференция «Информатика и прикладная математика», ИИВТ МНВО РК, Алматы, 2023 ж., 377-386 стр., ISBN 978-601-332-165-3 Zhenisbekovna, M. S., Aslanbekkyzy, B. M., & Bolatkyzy, B. G. (2024). Investigating long short-term memory approach for extremist messages detection in Kazakh language. Expert Systems, e13595. https://doi.org/10.1111/exsy.13595
Patents	**Патент – 1:** Байспай Г.Б., Мусиралиева Ш.Ж., Болатбек М.А., Медетбек Ж.Б., Омаров Б., Оспанов Р.Қ., Патент на полезную модель «Способ и система анализа веб-контента для определения экстремистской направленности», номер патента № 7997, Удостоверение автора №110136 **Авторское свидетельсво -1:** Байспай Гүлшат Болатқызы, Мусиралиева Шынар Женисбековна, Абайұлы Ерұлан, Авторлық құқық объектісі: ЭЕМ-ге арналған бағдарлама, программа для ЭВМ под названием «Алгоритм и ПО графической визуализации связей вовлеченных пользователей в социальных сетях»