

ISSN 2518-1726 (Online),  
ISSN 1991-346X (Print)

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ

әл-Фараби атындағы Қазақ ұлттық университетінің

# Х А Б А Р Л А Р Ы

---

---

## ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ  
НАУК РЕСПУБЛИКИ КАЗАХСТАН  
Казахский национальный  
университет имени аль-Фараби

## N E W S

OF THE ACADEMY OF SCIENCES  
OF THE REPUBLIC OF  
KAZAKHSTAN  
al-Farabi Kazakh National University

**SERIES**  
**PHYSICS AND INFORMATION TECHNOLOGY**

**2 (346)**

**APRIL – JUNE 2023**

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

#### **БАС РЕДАКТОР:**

**МУТАНОВ Ғалымқайыр Мұтанұлы**, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), **Н=5**

#### **БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:**

**МАМЫРБАЕВ Өркен Жұмажанұлы**, ақпараттық жүйелер мамандығы бойынша философия докторы (Ph.D), ҚР БҒМ Ғылым комитеті «Ақпараттық және есептеуші технологиялар институты» РМК жауапты хатшысы (Алматы, Қазақстан), **Н=5**

#### **РЕДАКЦИЯ АЛҚАСЫ:**

**КАЛИМОЛДАЕВ Мақсат Нұрәділұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі (Алматы, Қазақстан), **Н=7**

**БАЙГУНЧЕКОВ Жұмаділ Жанабайұлы**, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Сатпаев университетінің Қолданбалы механика және инженерлік графика кафедрасы, (Алматы, Қазақстан), **Н=3**

**ВОЙЧИК Вальдемар**, техника ғылымдарының докторы (физика), Люблин технологиялық университетінің профессоры (Люблин, Польша), **Н=23**

**БОШКАЕВ Қуантай Авгазыұлы**, Ph.D. Теориялық және ядролық физика кафедрасының доценті, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=10**

**QUEVEDO Nemandó**, профессор, Ядролық ғылымдар институты (Мехико, Мексика), **Н=28**

**ЖҮСПОВ Марат Абжанұлы**, физика-математика ғылымдарының докторы, теориялық және ядролық физика кафедрасының профессоры, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=7**

**КОВАЛЕВ Александр Михайлович**, физика-математика ғылымдарының докторы, Украина ҰҒА академигі, Қолданбалы математика және механика институты (Донецк, Украина), **Н=5**

**РАМАЗАНОВ Тілекқабұл Сәбитұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университетінің ғылыми-инновациялық қызмет жөніндегі проректоры, (Алматы, Қазақстан), **Н=26**

**ТАКИБАЕВ Нұрғали Жабғаулы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=5**

**ТИГИНЯНУ Ион Михайлович**, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), **Н=42**

**ХАРИН Станислав Николаевич**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Қазақстан-Британ техникалық университеті (Алматы, Қазақстан), **Н=10**

**ДАВЛЕТОВ Асқар Ербуланович**, физика-математика ғылымдарының докторы, профессор, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=12**

**КАЛАНДРА Пьетро**, Ph.D (физика), Наноқұрылымды материалдарды зерттеу институтының профессоры (Рим, Италия), **Н=26**

**«ҚР ҰҒА Хабарлары. Физика және информатика сериясы».**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **№ 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *физика және ақпараттық коммуникациялық технологиялар сериясы.*

Қазіргі уақытта: *«ақпараттық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*  
*<http://www.physico-mathematical.kz/index.php/en/>*

---

© Қазақстан Республикасының Ұлттық ғылым академиясы, 2023

Типографияның мекен-жайы: «Аруна» ЖК, Алматы қ., Мұратбаев көш., 75.

## ГЛАВНЫЙ РЕДАКТОР:

**МУТАНОВ Галимжаир Мутанович**, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МОН РК (Алматы, Казахстан), **Н=5**

## ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

**МАМЫРБАЕВ Оркен Жумажанович**, доктор философии (PhD) по специальности Информационные системы, ответственный секретарь РГП «Института информационных и вычислительных технологий» Комитета науки МОН РК (Алматы, Казахстан), **Н=5**

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

**КАЛИМОЛДАЕВ Максат Нурадилович**, доктор физико-математических наук, профессор, академик НАН РК (Алматы, Казахстан), **Н=7**

**БАЙГУНЧЕКОВ Жумадил Жанабаевич**, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сатпаева (Алматы, Казахстан), **Н=3**

**ВОЙЧИК Вальдемар**, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), **Н=23**

**БОШКАЕВ Куантай Авгазыевич**, доктор Ph.D, преподаватель, доцент кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=10**

**QUEVEDO Hemando**, профессор, Национальный автономный университет Мексики (UNAM), Институт ядерных наук (Мехико, Мексика), **Н=28**

**ЖУСУПОВ Марат Абжанович**, доктор физико-математических наук, профессор кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=7**

**КОВАЛЕВ Александр Михайлович**, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), **Н=5**

**РАМАЗАНОВ Тлеккабул Сабитович**, доктор физико-математических наук, профессор, академик НАН РК, проректор по научно-инновационной деятельности, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=26**

**ТАКИБАЕВ Нурғали Жабагаевич**, доктор физико-математических наук, профессор, академик НАН РК, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=5**

**ТИГИНЯНУ Ион Михайлович**, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), **Н=42**

**ХАРИН Станислав Николаевич**, доктор физико-математических наук, профессор, академик НАН РК, Казахстанско-Британский технический университет (Алматы, Казахстан), **Н=10**

**ДАВЛЕТОВ Аскар Ербуланович**, доктор физико-математических наук, профессор, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=12**

**КАЛАНДРА Пьетро**, доктор философии (Ph.D, физика), профессор Института по изучению наноструктурированных материалов (Рим, Италия), **Н=26**

**«Известия НАН РК. Серия физика и информатики».**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ 16906-Ж** выданное 14.02.2018 г.

Тематическая направленность: *серия физика и информационные коммуникационные технологии.* В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*

*<http://www.physico-mathematical.kz/index.php/en/>*

---

© Национальная академия наук Республики Казахстан, 2023  
Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75.

#### **EDITOR IN CHIEF:**

**MUTANOV Galimkair Mutanovich**, doctor of technical Sciences, Professor, Academician of NAS RK, acting director of the Institute of Information and Computing Technologies of SC MES RK (Almaty, Kazakhstan), **H=5**

#### **DEPUTY EDITOR-IN-CHIEF**

**MAMYRBAYEV Orken Zhumazhanovich**, Ph.D. in the specialty information systems, executive secretary of the RSE “Institute of Information and Computational Technologies”, Committee of Science MES RK (Almaty, Kazakhstan) **H=5**

#### **EDITORIAL BOARD:**

**KALIMOLDAYEV Maksat Nuradilovich**, doctor in Physics and Mathematics, Professor, Academician of NAS RK (Almaty, Kazakhstan), **H=7**

**BAYGUNCHEKOV Zhumadil Zhanabayevich**, doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), **H=3**

**WOICIK Waldemar**, Doctor of Phys.-Math. Sciences, Professor, Lublin University of Technology (Lublin, Poland), **H=23**

**BOSHKAYEV Kuantai Avgazievich**, PhD, Lecturer, Associate Professor of the Department of Theoretical and Nuclear Physics, Al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=10**

**QUEVEDO Hemando**, Professor, National Autonomous University of Mexico (UNAM), Institute of Nuclear Sciences (Mexico City, Mexico), **H=28**

**ZHUSSUPOV Marat Abzhanovich**, Doctor in Physics and Mathematics, Professor of the Department of Theoretical and Nuclear Physics, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=7**

**KOVALEV Alexander Mikhailovich**, Doctor in Physics and Mathematics, Academician of NAS of Ukraine, Director of the State Institution «Institute of Applied Mathematics and Mechanics» DPR (Donetsk, Ukraine), **H=5**

**RAMAZANOV Tlekkabul Sabitovich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Vice-Rector for Scientific and Innovative Activity, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=26**

**TAKIBAYEV Nurgali Zhabagaevich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=5**

**TIGHINEANU Ion Mikhailovich**, Doctor in Physics and Mathematics, Academician, Full Member of the Academy of Sciences of Moldova, President of the AS of Moldova, Technical University of Moldova (Chisinau, Moldova), **H=42**

**KHARIN Stanislav Nikolayevich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Kazakh-British Technical University (Almaty, Kazakhstan), **H=10**

**DAVLETOV Askar Erbulanovich**, Doctor in Physics and Mathematics, Professor, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=12**

**CALANDRA Pietro**, PhD in Physics, Professor at the Institute of Nanostructured Materials (Monterotondo Station Rome, Italy), **H=26**

#### **News of the National Academy of Sciences of the Republic of Kazakhstan.**

**Series of physics and informatics.**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. 16906-Ж**, issued 14.02.2018  
Thematic scope: *series physics and information technology*.

Currently: *included in the list of journals recommended by the CCSES MES RK in the direction of «information and communication technologies».*

Periodicity: *4 times a year.*

Circulation: *300 copies.*

Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*

*<http://www.physico-mathematical.kz/index.php/en/>*

---

© National Academy of Sciences of the Republic of Kazakhstan, 2023

Address of printing house: ST «Aruna», 75, Muratbayev str, Almaty.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN  
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 237–251

<https://doi.org/10.32014/2023.2518-1726.196>

УДК 004.032.26

МРПТИ 28.23.37

©**A.B. Toktarova**<sup>1\*</sup>, **B.S. Omarov**<sup>1</sup>, **Zh.Zh. Azhibekova**<sup>3</sup>, **G.I. Beissenova**<sup>4</sup>,  
**R.B. Abdrakhmanov**<sup>5</sup>, 2023

<sup>1</sup>Khoja Akhmet Yassawi International Kazakh-Turkish University,  
Turkistan, Kazakhstan;

<sup>2</sup>AL-Farabi Kazakh National University, Almaty, Kazakhstan;

<sup>3</sup>Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan;

<sup>4</sup>Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan;

<sup>5</sup>International University of Tourism and Hospitality, Turkestan, Kazakhstan.

E-mail: [toktar.aigerim@list.ru](mailto:toktar.aigerim@list.ru)

## ANALYSIS OF HATE SPEECH WORDS IN ONLINE CONTENT BY USING DATA MINING

**Toktarova Aigerim** — doctoral student, department of Computer engineering, faculty of Engineering, Khoja Akhmet Yassawi International Kazakh – Turkish University, Turkistan, Kazakhstan  
E-mail: [toktar.aigerim@list.ru](mailto:toktar.aigerim@list.ru), <https://orcid.org/0000-0002-6265-9236>;

**Omarov Batyrkhan** — PhD, department of Information System, faculty of Information technologies AL-Farabi Kazakh National University, Almaty, Kazakhstan  
E-mail: [batyahan@gmail.com](mailto:batyahan@gmail.com), <https://orcid.org/0000-0002-8341-7113>;

**Azhibekova Zhanar** — candidate of pedagogical sciences, department of Information Communication Technologies, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan  
E-mail: [zhanar\\_azhibekova@mail.ru](mailto:zhanar_azhibekova@mail.ru), <https://orcid.org/0000-0002-4396-1261>;

**Beissenova Gulbakhram** — candidate of pedagogical sciences, department of Computer science, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan  
E-mail: [g\\_beissenova@mail.ru](mailto:g_beissenova@mail.ru), <https://orcid.org/0000-0003-0545-0652>;

**Abdrakhmanov Rustam** — Cand. of Techn.Sc., Ass.Prof., International University of Tourism and Hospitality Turkestan, Kazakhstan  
E-mail: [abdrakhmanov.rustam@iuth.edu.kz](mailto:abdrakhmanov.rustam@iuth.edu.kz), <https://orcid.org/0000-0003-0385-0684>.

**Abstract.** The state of ensuring the information and psychological security of society is declining due to the daily increase in destructive words on the Internet. That is, categories such as advocacy of violence based on race, ethnicity, national origin, gender, gender identity, religion, age, disability or illness are defined. It shows that the search and identification of non-speech words can be done with minimal effort from the point of view of the subject, since this can now be achieved automatically using various methods. The existing methods (expert analysis, intellectual data processing methods) are considered, the advantages and

disadvantages of these methods are noted. In addition, he not only explores the roots of profanity, but also proposes concepts to help distinguish between types of offensive language, such as profanity and cyberbullying. We use machine learning approaches to access datasets that we can use to automatically collect social media profanity. To solve this problem, Data Mining is proposed to search for offensive words, such as swear words in text databases. It can be seen that the presented method is distinguished by the ability to automatically fill in the vocabulary through the system when identifying unfamiliar words classified as non-standard. The correctness of vocabulary replenishment is measured using various indicators. We are creating a set of Internet users in Kazakhstan who use social networks and media to share their opinion. The description of the proposed research work is given from data collection to classification and identification of sets of words with destructive content.

**Key words:** hate speech, analysis, detection, online content, Data Mining

© **А.Б. Тоқтарова<sup>1\*</sup>, Б.С. Омаров<sup>2</sup>, Ж.Ж. Ажибекова<sup>3</sup>, Г.И. Бейсенова<sup>4</sup>,  
Р.Б. Абдрахманов<sup>5</sup>, 2023**

<sup>1</sup>Қожа Ахмет Ясауи атындағы Халықаралық Қазақ-Түрік Университеті,  
Түркістан, Қазақстан;

<sup>2</sup>Әл-Фараби атындағы Қазақ Ұлттық университеті Алматы, Қазақстан;

<sup>3</sup>С.Ж. Асфендияров атындағы Қазақ ұлттық медицина университеті,  
Алматы, Қазақстан;

<sup>4</sup>М. Ауезов атындағы Оңтүстік Қазақстан Университеті,  
Шымкент, Қазақстан;

<sup>5</sup>Халықаралық туризм және меймандостық университеті.  
E-mail: toktar.aigerim@list.ru

## **ОНЛАЙН КОНТЕНТТЕГІ БЕЙӘДЕП СӨЗДЕР МӘЛІМЕТТЕР ҚОРЫН DATA MINING АРҚЫЛЫ АНАЛИЗДЕУ**

**Тоқтарова Айгерім** — Докторант, кафедра Компьютерлік инженерия, Қожа Ахмет Ясауи атындағы Халықаралық Қазақ-Түрік Университеті, Түркістан, Қазақстан  
E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>;

**Омаров Батырхан** — Phd, кафедра Ақпараттық жүйелер, факультет Ақпараттық технологиялар, Әл-Фараби атындағы Қазақ Ұлттық университеті, 050040, Алматы, Қазақстан Алматы, Қазақстан,

E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

**Ажибекова Жанар** — п.ғ.к., С.Ж. Асфендияров атындағы Қазақ ұлттық медицина университеті Алматы, Қазақстан Республикасы

E-mail: zhanar\_azhibekova@mail.ru, <https://orcid.org/0000-0002-4396-1261>;

**Бейсенова Гүльбахрам** — п.ғ.к., М. Ауезов атындағы Оңтүстік Қазақстан Университеті, Шымкент, Қазақстан

E-mail: g\_beissenova@mail.ru, <https://orcid.org/0000-0003-0545-0652>;

**Абдрахманов Рустам** — техн.ғ.к., доцент м.а., Халықаралық туризм және меймандостық университеті, Түркістан, Қазақстан

E-mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>.

**Аннотация.** Қоғамның ақпараттық-психологиялық қауіпсіздігін қамтамасыз ету жағдайын интернеттегі деструктивті сөздердің күнделікті тоқтаусыз өсуі төмендетіп отыр. Яғни, нәсіліне, этникалық тегіне, ұлттық шығу тегіне, жынысына, гендерлік сәйкестігіне, дініне, жасына, мүгедектігіне немесе ауруына негізделген зорлық-зомбылықты насихаттау сияқты санаттары анықталып отыр. Ғадауат сөздерді іздеу және анықтау субъект тұрғысынан минимальді түрде күш жұмсауға болатынын көрсетіп отыр, себебі қазіргі таңда автоматты түрде түрлі әдістерді қолдану арқылы қол жеткізуге болады. қолданыстағы әдістер (эксперттік талдау, деректерді өндеудің интеллектуалды әдістері) қарастырылады, бұл әдістердің артықшылықтары мен кемшіліктері атап өтіледі. Сонымен қатар, бейәдеп тілдің түп-тамырын зерттеп қана қоймайды, кемсіту тіл мен кибербуллинг сияқты ғадауат сөздер түрлерін ажыратуға көмектесетін ұғымдарды ұсынады. Біз әлеуметтік желідегі балағат сөздерді автоматты түрде жинау үшін пайдалана алатын деректер жиынына қол жеткізу үшін машиналық оқыту тәсілдерін қолданамыз. Мұны шешу үшін мәтіндік мәліметтер қорындағы балағаттау сөздерді мысалға ала отырып, ғадауат мазмұнды сөздерді іздеу үшін Data Mining ұсынылады. Көрсетілген әдіс нормаға сай емес деп жіктелген бейтаныс сөздерді анықтауда жүйе арқылы сөздік қорды өздігінен толтыру мүмкіндігімен ерекшеленетінін байқауға болады. Әртүрлі көрсеткіштерді пайдалану арқылы сөздікті толықтырудың дұрысытығы өлшенеді. Қазақстандағы интернет пайдаланушылары деректер жинағын жасап жатырмыз, олар өз пікірлерімен бөлісу үшін әлеуметтік желілер мен медианы пайдаланады. Ұсынылып отырған зерттеу жұмысының сипаттамасы деректерді жинаудан бастап деструктивті мазмұндас сөздер топтамасын жіктеуге және оларды анықтауға дейін берілген.

**Түйін сөздер:** бейәдеп мазмұнды сөздер, анықтау, онлайн контент, анализдеу, Data Mining

© **А.Б. Токтарова<sup>1\*</sup>, Б.С. Омаров<sup>2</sup>, Ж.Ж. Ажибекова<sup>3</sup>, Г.И. Бейсенова<sup>4</sup>,  
Р.Б. Абдрахманов<sup>5</sup>, 2023**

<sup>1</sup>Международный Казахско – Турецкий университет им. Ходжа Ахмета Ясауи, Туркестан, Казахстан;

<sup>3</sup>Казахский Национальный университет им. аль-Фараби, Алматы, Казахстан;

<sup>3</sup>Казахский Национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Казахстан,

<sup>4</sup>Южно – Казахстанский университет имени Мухтара Ауезова, Шымкент, Казахстан;

<sup>5</sup>Международный университет туризма и гостеприимства.

E-mail: toktar.aigerim@list.ru

## **АНАЛИЗ НЕОБРАЗНЫХ СЛОВ В ОНЛАЙН-КОНТЕНТЕ С ПОМОЩЬЮ DATA MINING**

**Токтарова Айгерім** — докторант, кафедра Компьютерная инженерия, факультет Инженерия, Международный Казахско-Турецкий университет им. Ходжа Ахмета Ясауи, Туркестан, Казахстан

E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>;

**Омаров Батырхан** — Phd, кафедра Информационных систем, факультет Информационных технологии, Казахский Национальный университет им. аль-Фараби, Алматы, Казахстан

E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

**Ажибекова Жанар** — к.п.н., Казахский Национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Республика Казахстан

E – mail: zhanar\_ajhibekova@mail.ru, <https://orcid.org/0000-0002-4396-1261>;

**Бейсенова Гульбахрам** — к.п.н., Южно – Казахстанский университет имени Мухтара Ауезова, Шымкент, Казахстан

E – mail: g\_beissenova@mail.ru, <https://orcid.org/0000-0003-0545-0652>;

**Абдрахманов Рустам** — к.т.н., и.о. доцента, Международный университет туризма и гостеприимства Туркестан, Казахстан

E – mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>.

**Аннотация.** Состояние обеспечения информационно-психологической безопасности общества снижается из-за ежедневного увеличения деструктивных слов в сети Интернет. То есть определены такие категории, как пропаганда насилия по признаку расы, этнической принадлежности, национального происхождения, пола, гендерной идентичности, религии, возраста, инвалидности или болезни. Он показывает, что поиски идентификация неречевых слов может осуществляться с минимальными усилиями с точки зрения испытуемого, так как теперь это может быть достигнуто автоматически с использованием различных методов. Рассмотрены существующие методы (экспертный анализ, интеллектуальные методы обработки данных), отмечены преимущества и недостатки этих методов. Кроме того, он не только исследует корни нецензурной лексики, но также предлагает концепции, помогающие различать типы оскорбительного языка, такие как ненормативная лексика и киберзапугивание. Мы используем подходы машинного обучения для доступа



к наборам данных, которые мы можем использовать для автоматического сбора ненормативной лексики в социальных сетях. Для решения этой проблемы предлагается Data Mining для поиска оскорбительных слов, например, матерных слов в текстовых базах данных. Видно, что представленный метод отличается возможностью автоматического заполнения словарного запаса через систему при выявлении незнакомых слов, отнесенных к нестандартным. Правильность пополнения словарного запаса измеряется с помощью различных показателей. Мы создаем набор пользователей Интернета в Казахстане, которые используют социальные сети и СМИ, чтобы поделиться своим мнением. Описание предлагаемой исследовательской работы дано от сбора данных до классификации и выявления наборов слов деструктивного содержания.

**Ключевые слова:** нецензурная речь, анализ, онлайн контент, определение, Data Mining

### **Кіріспе**

Қазіргі таңдағы ғылым мен техниканың даму прогресі адам қызметтерінің біршама салаларына жаңа ақпараттық технологияларды үздіксіз енгізілуде. Сонымен қатар, интернеттің дамуы сол жердегі мәтін түрінде берілген ақпарат көлемі мен мазмұнының бақыланбайтын экспоненциалды өсуіне жеткізді (Канхото, 2020). Ақпараттық қауіпсіздікті сақтау мақсатында әлеуметтік желілеріндегі заңсыз ақпаратты, оның ішінде бөліп қарастырсақ терроризм, есірткінің заңсыз стабылымы, наразылық акцияларын немесе жаппай тәртіпсіздіктерді ұйымдастыру, мемлекеттік рәміздерді қорлау сөздері, балағат сөздерді және т.б. қоғамға қауіпті деректері бар мәліметтерді талдаудың маңызы бар (Кумар, 2020). Осыған байланысты, сондай-ақ «жабық» интернет кеңістігін әлеуетті құру жағдайында деструктивті ақпаратты іздеу және автоматты түрде анықтау. Сонымен қатар, бүгінгі күнге дейін қоғамның ақпараттық-психологиялық қауіпсіздігін сақтау үшін заң жүзіндегі іс-шаралар жеткіліксіз екендігін көруге болады (Джайн, 2020).

Сонымен, автоматтандырылған және мамандандырылған іздеу жүйелерін іске асыру және оларды санаттарына қарай жіктеу қажеттілігі туындайды. Осылайша, ақпараттық ресурстардың үнемі өсуі қауіпті мазмұнды сөздерді анықтауда автоматтандыру процесі еңбек шығындарын төмендетумен қатар, субъективті іс — әрекеттер желісін және адам факторының әсерінен келетін қауіптер ықтималдығын азайтады (Мохан, 2020). Әлеуметтік желілердің белсенді дамуы, интернетті «қылмыстық» қауіптерге ерекше назар аударуды талап етеді. Ақпаратты қорғаудың келесі жолдары бар (Мохан, 2020; Мюльхофф, 2020):

- қауіптің алдын алу — олардың туындау мүмкіндігін болжау және ақпараттық қауіпсіздікті қамтамасыз ететін алдын алу шаралары;
- қауіп-қатерлерді анықтау нақты немесе қауіптердің пайда болу

ықтималдығын жүйелі талдау мен бақылауда ұстау, сонымен қатар, олардың алдын алу бойынша дер кезінде шаралар қабылдау;

- қылмыстық әрекеттерді оқшаулау және қауіпті немесе нақты қылмыстық әрекеттерді жою шараларын қабылдау;

Шартты түрде ақпараттық — технологияларды пайдалануда сыртқы қауіптерді «дәстүрлі» және «жаңа» деп бөліп қарастыруға болады. (Муссиралиева, 2021) «Дәстүрлі» сыртқы қауіптер ретінде спам, фишинг, компьютерлік вирустар, трояндар және желілік шабуылдар жатқызуға болады. Келесі тарауда «жаңа» сыртқы қауіптер қарастырылады.

### **Әдістер мен материалдар**

«Жаңа» ақпараттық – технологиялық қауіптер. Расында, интернет-коммуникация тақырыбы біздің уақытта өзекті болып отыр. Жоғары дәрежеде интернет-технологиялардың дамуы неге әкелді деген сұраққа біржақты жауап беру мүмкін болмай отыр. Әдетте халықаралық желі интернет миллиондаған компьютерлік желілердің ғаламдық байланысуы. Басқаша айтқанда, бұл дүние жүзіндегі барлық компьютерлік желілердің желісі немесе жай ғана желілер желісі. Интернет ғаламның кез келген нүктесінде бірден ақпарат алмасуға мүмкіндік бере алады. Көбіне, қысқа хабарламалар, жылдам хабар алмасу немесе чаттар үшін электрондық пошта арқылы интернет адамдарға тұлғааралық және бұқаралық коммуникация деңгейінде басқа адамдармен байланысуға мүмкіндік береді. Сондай-ақ ол өз пайдаланушыларына World Wide Web-те қолжетімді ақпарат көлеміне қол жеткізуді ұйымдастыра алады (Неупане, 2020). Кез келген энциклопедиядан да үлкен деректер әлемнің бір бөлігінен жер шарының ең алыс бөлігіне жарық жылдамдығымен тасымалдана алады. Осылайша, қолданушылар ақпаратты таңқаларлық жылдамдықпен жүктеп немесе жүктей алады. І суретте әлеуметтік желілерді пайдаланатын белсенді тұтынушылар саны (миллиондаған) көрсетілген (Саллум, 2020).



Сур 1. Әлеуметтік желілерді пайдаланатын белсенді тұтынушылар саны (миллион)  
(Fig 1. Number of active users using social networks (millions))

Технологияға негізделген және сөз мазмұнына негізделген киберқылмыс деп киберқылмысты екі түрге бөліп қарастыруға болады. Сөз мазмұнына негізделген қылмысты жыныстық қудалау, киберқорқыту, балалар порнографиясы бойынша кез келген террористік ұйым жүзеге асырады. Ал, аккаунтты бұзып кіру, тыңшылық және зиянды кодты енгізу технологияға негізделген қылмыстың мысалдары болып табылады (Саркер, 2020).

*Ransomware (бопсалаушы) бағдарламалары*

Бұл бағдарламалар операциялық жүйенің жұмысын блоктайды және/немесе қатты дискідегі барлық файлдарды шифрлап тастайды (Сингх, 2020). Блоктау немесе шифрлау процесі аяқталғаннан кейін пайдаланушыдан құлыпты ашу / шифрды шешу кодын алу үшін белгілі бір соманы (әдетте 10–20 мың рубль аймағында) аударуды сұрайды. Файлдардың шифрын ашу мүмкіндігін растау ретінде шабуылдаушылар оларға 2–3 шифрланған файлды жіберуді ұсынады, содан кейін оларды бастапқы түрінде қайтарады. Қауіпті операциялық жүйе бұғатталған кезде оны жай ғана қайта орнатуға болады, бірақ барлық файлдарды шифрлаған жағдайда бұл көмектеспейді.

Ransomware (бопсалаушы) бағдарламалары қамтамасыз ету эпидемиясы кезінде (2012–2013) антивирустық компаниялар блоктан шығару қызметтерін ұсынды. Қызметтердің жұмысы ransomware кодындағы осалдықтарға және бағдарламаның әрекеті және ол берген сұраулар туралы ақпарат негізінде құлыпты ашу кодын таңдау мүмкіндігіне негізделген. Дегенмен, содан бері бұл зиянды бағдарламалар жетілдірілді және антикодты талдау құралдарымен жабдықталған, сондықтан антивирустық қызметтерді пайдаланудың жетістігі айтарлықтай төмендеді («ескі» зиянды кодты жұқтырған кезде ғана жұмыс істейді) (Юань, 2021).

*«Қаламаған» контент.* Мұндай мазмұнда пайдаланушыға жағымсыз және оның сұрауына сәйкес келмейтін кез келген ақпаратты қамтылуы, өйткені көптеген сайттар тез басылатын терезелерді көрсетеді (Мюльхофф, 2020). Бұл қажетсіз жарнама, порнография, онлайн казинода ойнау немесе кейбір ақылы онлайн қызметтерді пайдалану туралы ұсыныс және т.б. болуы мүмкін. Сондай-ақ қажетсіз мазмұнда іздеу жүйесінің сұрауы бойынша табылған сайттар болуы мүмкін. Мысалы, егер бала «Қызыл телпек» сұрауына кірсе, онда Яндекс шығарған алғашқы сілтемелер «ужас, қиял, триллер» санатына жататын 2011 жылғы фильммен байланысты болады.

*«Қаламаған» контент ретінде flood, offtop, flame және holivar (Хуанг, 2019).* flood — мағынасыз, форумдар мен чаттарда қажетсіз, мағынасыз ақпаратты орналастыру. Offtopic — форумдар мен чаттарда талқыланатын тақырыпқа қатысы жоқ ақпаратты орналастыру, мұндай хабарламалар ағылшын тілінен «offtopic» деп аталады. off — from, away және тақырып — тақырып, сұрақ, яғни сөзбе-сөз аудармасы — тақырыптан тыс сұрақ. Ағылшын тілінен алынған flame — жалын, құмарлық — интернет форумдары мен чаттардағы хабарламалар, бұл сөздік соғыс, көбінесе бұл соғыс түрі даудың бастапқы себебімен байланысты болмауы мүмкін (Саллум, 2020). Holivar ағылшын

тілінен қасиетті соғыс — бұл Интернет форумдарындағы және чаттардағы хабар алмасудың қасиетті соғысы, бұл қатысушылар бір-біріне ұқсас бірнеше баламалардың (компьютерлік бағдарламалар, технологиялар, актерлер, музыкалық топтар және т.б.) біреуінің артықшылығын дәлелдеуге тырысатын мағынасыз пікірталас. (Неупане, 2020) .

*Интернеттегі алаяқтық.* Бұндай қауіптер виртуалды кеңістікке «сәтті» түрде өткен дәстүрлі алаяқтықтың барлық түрлерін қамтиды. Әдетте, қолданушылар келесі жағдайларға тап болады:

1. «Тізбек хаттары» – адресаттың қомақты жүлденің иегері болғаны туралы немесе белгілі бір қайырымды адамның немесе кенеттен табылған алыс туысы оған белгілі бір соманы бергісі келетіні туралы ақпарат көрсетілген электрондық пошта немесе әлеуметтік желідегі хабарламалар оның. «Жүлдені» алу үшін «қайырымдылық» белгісі ретінде белгілі бір соманы шотқа аударуын сұрайды.

2. «Балаға көмектесу» - ауыр науқас балаға немесе басқа да қайырымдылық мақсаттарға қаражат жинау. Көбінесе шабуылдаушылар қажеттіліктері үшін қаражат жинау ұйымдастырылған нақты адамның деректерін пайдаланады, бірақ олар қайырымдылықты аудару үшін өз реквезиттерін жазады.

3. Кейбір заңсыз немесе жартылай заңды ресурстарға, мысалы, пираттық ақпараттарға қол жеткізу үшін төлем жасау. Әртүрлі төлем әдістері ұсынылуы мүмкін, бірақ ол аяқталғаннан кейін пайдаланушыға ақпаратқа қол жеткізе алмайды, ал SMS арқылы төлем жасаған жағдайда, пайдаланушы белгісіз қызметке жазылуы мүмкін, ол үшін оның ұялы телефонының шотынан жүйелі түрде қаражат жұмсалады.

4. Кіру жарнасын төлеуді талап ететін әртүрлі жобаларға қатысуды ұсынады.

5. Алдын ала төлем шартымен тауарды сату, егер іс жүзінде тауар сатып алушыға келмесе немесе арзанырақ зат жіберуі де мүмкін.

*Троллинг және кибербуллинг.* Интернеттегі троллинг — бұл басқа пайдаланушыларды тікелей қорлау, арандатушылық немесе қорлайтын хабарламаларды жариялау, қарым-қатынас нормалары мен ережелерін өрескел бұзу сияқты кез келген агрессивті әрекет. Кибербуллинг — өзін қорғау қиынға соғатын жәбірленушіге қарсы бірнеше рет қайталанатын және электрондық байланыс нысандарын пайдалана отырып, бір топ адамдар немесе бір адам жасайтын агрессивті қасақана әрекет. Шетелдік психологтар білім беру қабырғаларындағы білім алушыларға зорлық-зомбылық көрсету жағдайын анықтады — буллинг (мектептегі бұзақылық). Шетелдегі қорқыту мәселесін С.М. Агого, К. Ли, Э. Роланд және т.б авторлар зерттеулер жүргізген. Интернет арқылы қорқытуды жеңудің негізгі шараларының тізімі әзірленді, психологиялық тетіктері анықталып, жеке мінездемелері берілді. Ақпараттық технологиялардың дамуымен қазіргі жасөспірімнің өмірінде елеулі өзгерістер орын алуда: виртуалды шындық пайда болды, онда жана коммуникация және тұлғааралық қарым-қатынастар олар үшін әдеттен тыс

деңгейге көтеріледі. Кибербуллингті алғаш рет Билл Белси анықтаған. Оның пікірінше, кибербуллинг — ақпараттық-коммуникациялық технологияларды, мысалы, электрондық поштаны, ұялы телефонды, жеке интернет-сайттарды адамның немесе топтың басқа адамдарды қорлауға бағытталған қасақана, қайталанатын және дұшпандық әрекеттері үшін пайдалану.

Кибербуллинг анықтамасына сүйене отырып, оның дәстүрлі боулингтен негізгі айырмашылығын бөліп көрсетуге болады: адамға қарсы бағытталған барлық әрекеттер виртуалды кеңістікте орын алады. Бірақ бұл жалғыз айырмашылық емес. Виртуалды кеңістікте баламалы «Мен-бейнесін» құру мүмкін болады, нәтижесінде нақты «Мен-бейне» деформациялануы мүмкін; жасөспірім өз әрекеттеріне жауапты болу ықтималдылығы төмен. Іс жүзінде кибербуллинг келесідей болуы мүмкін:

- жәбірленушінің электрондық поштасына жүйелі түрде қорлайтын хабарламаларды жіберу немесе жәбірленушінің әлеуметтік желідегі парақшасында жариялау;

- Әлеуметтік желіде киберқорқытулар тобы жәбірленушіні қорлайтын хабарламалар жариялайтын «жек көрушілік қабырғасын» құру;

- әлеуметтік желіде немесе ұялы телефонда жеке фотосуреттерді, хат-хабарлардың скриншоттарын жариялау, жәбірленушіге жағымсыз көзқараста көрсету, оның ар-намысы мен қадір-қасиетіне нұқсан келтіру, осыған байланысты «мамандандырылған» сайттардың пайда болуы. «экс сүйікті адамыңыздан кек алу» ұраны, бір кездері жақын адамның интимдік фотосуреттерін орналастыру ұсынылады;

- жәбірленушінің қатысуымен кибербуллинг жасаушы адам ашық немесе жасырын түсірілген және жәбірленушіні жағымсыз жағдайда, мысалы, мас күйінде, киім ауыстыратын бөлмеде, душта киім ауыстырып жатқанда бейнелейтін фотосуреттер мен бейнелерді әлеуметтік желілерде орналастыру; дәретханада және т.б.; Жәбірленушіні зорлау, ұрып-соғу, қорлау немесе қорлауды бейнелейтін бейнелер мен фотосуреттер туралы бөлек айту керек (бұл мәселе әсіресе жасөспірімдерге қатысты)

### **Нәтижелер**

Ғадауат сөзді іздеу және анықтау мысалын қолдана отырып, деструктивті ақпаратты автоматты түрде анықтау мәселесін шешу үшін мәтіндегі нақты сөздерді іздеуде қолданылып жүрген іздеу әдістері талданды. Бұл әдістердің қолданылу аясы кең екендігі анықталды:

- нәтижесі «қара тізімдерді» құру болып табылатын ақпаратты сараптамалық өңдеуге толығымен негізделген әдістер;

- жұмыста бізді қызықтыратын автоматтандырылған әдістер, соның ішінде тақырыптық іздеу (сөздікте) және деректерді интеллектуалды өңдеу әдістері.

Сөздіктен сөздің дәл сәйкестігін іздеу және оны мәтінде анықтау жүйесінен тұратын сөздік іздеу деструктивті ақпаратты анықтаудың кең таралған түрлерінің бірі болып табылады. Дегенмен, мысал ретінде балағат сөздерді пайдалана отырып, деструктивті мазмұнды анықтау мәселесін шешу үшін

«таза» нысанда тақырыптық іздеуді пайдалану келесі себептерге байланысты орынсыз болып табылады:

- сөзжасамның жаңа формалары мен әдістерінің пайда болуына байланысты сөздік қорын үнемі толықтырып отыру қажеттілігі (қазіргі уақытта әртүрлі тіркестер мен сөз формаларын құрайтын 250-ден астам негізгі балағат сөздер бар);

- зерттелетін мәтінде жүйеге белгісіз оның жаңа формасының қолданылуына байланысты балағат сөздерді өткізіп жіберудің жоғары ықтималдығы.

Мәліметтерді өңдеудің интеллектуалды әдістеріне негізделген жүйелер келесі артықшылықтарға ие:

- ақпаратты адамның ойлауына жақын формада символдық (семантикалық) өңдеу;

- пайдаланушылармен қарқынды диалог жүргізуге мүмкіндік беретін дамыған коммуникациялық дағдылар, оның барысында жүйеде бар және алған білім нақтыланады;

- жүйеге сұраныстарды қалыптастыру және адамдардың қарым-қатынасына жақын табиғи тілде жауаптар (мәселелерді шешу) алу;

- өздігінен білім алу мүмкіндігі, яғни жүйемен пайдаланушының мәселелерін талдау мен шешудің жинақталған тәжірибесі негізінде жаңа білімді автоматты түрде толықтыру және алу;

- жүйенің қызмет етуінің пәндік (проблемалық) аймағындағы объективті өзгерістерге жүйенің бейімделу мүмкіндігі.

*DATA MINING және деректер қорын жинау.* Data mining ( қазақша аудармасы мәліметтерді өңдеу, деректер жинау) — әртүрлі салаларда шешім қабылдау үшін қажетті деректердегі бұрын белгісіз, өңделмеген деректер қорын кілттік сөздер арқылы пайдалы және қол жетімді білімге айландыра отырып, деректер қорын електен өткізу, деректерді қайта өңдеу болып табылады.

*Деректер жинағы.* Зерттеуіміздің біз машиналық оқыту әдістерінде пайдалануға арналған деректер жинағын құру мақсатында деректерді жинадық. Деректер әлеуметтік желі сайттарынан жиналып, екі топқа бөлінді: деструктивті идеяны қамтитын деректер және деструктивті идеясы жоқ деректер. 2-суретте деректер көзін анықтауға, талдауға және деректер жинағын әзірлеуге бөлінген зерттеуіміздің деректерді жинау процесі көрсетілген.



Сур 2. Деректер көзін жинау  
(Fig 2. collect database)

Бірінші кадамда деректер жинағын жасау үшін жинайтын деректер көзін анықтаймыз. Жоғарыда айтқанымыздай, біздің жағдайда деректер көзі әлеуметтік желі сайттары болды. Деректер көзін анықтағаннан кейін талдаушы көмегімен деректерді жүктейміз. Қажетті деректерді алғаннан кейін біз оларды екі санатқа бөлеміз. 3-суретте мәліметтерді жинаудың блок-схемасы көрсетілген. Деректерді алу үшін біз сәйкес API интерфейстері арқылы әлеуметтік желі сайттарына қосыламыз. API арқылы .json пішіміндегі деректерді жүктеп алу. Жүктеп салынғаннан кейін біз деректерді қолмен екі санатқа жіктейміз. Нәтижесінде деректерді жіктегеннен кейін біз өзімізге ыңғайлы пішінде деректер жинағын құруды бастай аламыз. Вконтакте әлеуметтік желісінен деректерді жинау үшін біз VK API қолдандық, ол зерттеу мақсатында барлық деректердің 1 % алуға мүмкіндік береді.



Сур 3. Деректерді жинау блок-схемасы  
(Fig 3. Block diagram of collecting database)

Біз Instagram деректерін жинаудың басқа әдісін қолдандық, себебі Instagram API арқылы деректерді жүктеп салуды шектейді. Келесі бөлімдерде біз Instagram-дан деректерді қалай алатынымызды түсіндіреміз. І кестеде деректерді алуға сұранысты орындау үшін қолданатын ереже базасы көрсетілген. Онда біз қолданылатын әдістерді түсіндіріп, оларға сипаттама береміз. Онда Instagram қолданушыларының жазбалары мен пікірлерін жүктеп алу үшін users.get, account.getinfo, get, copy\_history\_depth, wall.get Comment сияқты әртүрлі әдістерді қолданамыз. (Vk.com, 2023)

Кесте 1. Әдістер түрлері

Әдістер	Анықтамасы
users.get	Пайдаланушы туралы ақпарат алу
account.getInfo	Ағымдағы аккаунт туралы ақпаратты қайтару

<code>groups.getMembers</code>	Қауымдастық мүшелерінің тізімін қайтарады
<code>Get</code>	Пайдаланушыдан немесе акаунттан жазбалар тізімін қайтарады
<code>wall.getById</code>	Идентификаторлары бойынша пайдаланушылардың әлеуметтік желілерінен жазбаларды қайтарады
<code>wall.getComment</code>	Пайдаланушылар әлеуметтік желісіне пікірлер тізімін қайтарады

Сурет 4–те Instagram деректерін алуға арналған бастапқы кодты көрсетеді. (Instagram, 2023)

```

class {
  constructor() {
    this.url = 'https://www.instagram.com/';
    this.userAgent = 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.164 Safari/537.36';
    this.headers = {
      'User-Agent': this.userAgent,
      'Accept': 'text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8',
      'Accept-Language': 'en-US,en;q=0.9',
      'Accept-Encoding': 'gzip, deflate, br',
      'Connection': 'keep-alive',
      'Upgrade-Insecure-Requests': '1'
    };
    this.cookie = '';
  }

  async get(url) {
    const response = await fetch(url, {
      headers: this.headers,
      cookie: this.cookie
    });
    return response.text();
  }

  async post(url, data) {
    const response = await fetch(url, {
      method: 'POST',
      headers: this.headers,
      body: JSON.stringify(data),
      cookie: this.cookie
    });
    return response.json();
  }

  async login(username, password) {
    const loginData = {
      username: username,
      password: password
    };
    const response = await this.post('https://www.instagram.com/accounts/login/ajax/', loginData);
    this.cookie = response.cookie;
  }

  async getProfile(username) {
    const url = `https://www.instagram.com/${username}/`;
    const html = await this.get(url);
    const profileData = {
      // Extract profile information from the HTML
    };
    return profileData;
  }

  async getPosts(username) {
    const url = `https://www.instagram.com/${username}/media/`;
    const html = await this.get(url);
    const postsData = {
      // Extract post information from the HTML
    };
    return postsData;
  }
}

```

Сур 4. Тақырыптарды жинау (Fig 4. Collection of topics)

Біз онлайн платформа арқылы мәліметтер қорын алғандықтан, осы бастапқы материалды пайдаланушылар үшін қол жеткізу идеясы бірдей болуы керек. Сұраныстар кітапханасы веб-сайтпен байланыс орнатуға және онымен жұмыс істеуді бастауға мүмкіндік береді.

**Талқылаулар.**

Қолдану тілі танымал емес нәсілдік немесе этникалық топтар үшін Интернетте жасалған мәтіндерді автоматты түрде жинау және бағалау қиын. Біз әртүрлі кілттік сөздерді пайдаланып деректер қоры жинақталды.

Сур 4. Қазақ тілінде жинақталған ғадауат тілді деректер қоры (Fig 4. Kazakh language database)



Онлайн парсинг және ерікті аннотатор арқылы Instagram, Facebook- тегі қазақ тілді қауымдастықтар мен аккаунттар және Youtube платформаларынан үш немесе одан да көп сөздерден тұратын бейәдеп мазмұнды 10224 сөйлемдер таңдалып алынды.

Әзірленген корпуста ғадауат тілді мәтіндер бар құжат 3 бағаннан тұрады: хабарламаның немесе пікірдің реттік нөмірі, хабарлама немесе пікірдің өзі және хабарламаның немесе пікірдің ғадауат тілді екенін білдіретін «1» атрибутымен және «0» бейтарап мәтін екендігінен хабар беретін кодтармен белгіленген.

### Қорытынды

Бүгінгі таңда онлайн контент желілерін қолдану деңгейі күннен күнге өсуінің қарым – қатынас жасауда маңызы зор болғанымен, оның екінші беті де бар екенін ескеруіміз керек, яғни, қандай да бір бейәдеп сөздерді қолдану арқылы психологиялық зақым келтіру немесе қорқыту – үркітудің түрлері және оларды анықтаудың белгілері осы ғылыми зерттеу жұмысында келтірілді. Сонымен қатар, деструктивті мәліметтер қорын іздеуде автоматтандырылған жүйе қажет екендігі көрсетілген. Осы мақсатта қолданылған әдістер жиынтығын көрсетіп отырымыз. Іздеу жолдарын және олардың бір – бірінен ерекшелігін байқауға болады.

Жинақталған мәліметтер қорын алдыңғы уақытта табиғи тілдерді өңдеу және машиналық оқыту әдістерін пайдалана отырып онлайн контентте анықтауға қолдануға болатындығын, ғадауат сөздер тізбегінің қай түріне жататындығын анықтауға болатындығын байқап отырымыз. Сонымен қатар, онлайн контенттегі ғадауат сөздерді анықтаудың моделін құруға болатындығын байқап, болашақтағы ғылыми зерттеуіміздің басты мәселесі ретінде алуға болатыны талқыланды.

### ӘДЕБИЕТТЕР

Канхото А.И., Клир Ф., 2020 — *Канхото А.И., Клир Ф.*, Искусственный интеллект и машинное обучение как бизнес-инструменты: основа для диагностики потенциала разрушения ценности. Деловые горизонты — 63(2), стр. 183–193

<https://vk.com/dev/methods> - Методы API ВКонтакте

Instagram.com — социальная сеть Instagram

Джайн А., Мандовара Дж., 2016 — *Джайн А., Мандовара Дж.*, Классификация текстов путем объединения текстовых классификаторов для повышения эффективности классификации. Международный журнал компьютерных приложений - 6 (2), стр. 2250–1797 гг.

Кумар А., Сингх Дж.П., Дживеди Ю.К., Рана Н.П., 2020 — *Кумар А., Сингх Дж.П., Дживеди Ю.К., Рана Н.П.*, Глубокая мультимодальная нейронная сеть для информативной классификации контента Твиттера во время чрезвычайных ситуаций. Анналы исследования операций, стр. 1–32.

Мохан А., Сингх А.К., Кумар Б., Дживеди Р., 2021 — *Мохан А., Сингх А.К., Кумар Б., Дживеди Р.*, Обзор методов дистанционного зондирования для обнаружения оползней с использованием машин и глубокого обучения. Сделки по новым телекоммуникационным технологиям - 32 (7), стр. 998.

Мюльхофф Р., 2020 — *Мюльхофф Р.*, Искусственный интеллект с помощью человека: или как выполнять большие вычисления в человеческом мозгу? К медиасоциологии машинного обучения. новые медиа и общество, 22 (10), стр. 1868–1884.

Муссиралиева С., Омаров Б.Ю.П., Болатбек М., 2021 — *Муссиралиева С., Омаров Б.Ю.П., Болатбек М.*, Применение методов машинного обучения для выявления религиозного экстремизма в пользовательском онлайн-контенте. Компьютеры, материалы и континуумы, - 70 (1), стр. 915–934.

Неупане Д., Сеок Дж., 2020 — *Неупане Д., Сеок Дж.*, Обнаружение и диагностика неисправностей подшипников с использованием набора данных Западного резервного университета с подходами глубокого обучения: обзор. Доступ IEEE, стр. 9178.

Саллум С.А., Альшуриде М., Эльнагар А., Шаалан К., 2020 — *Саллум С.А., Альшуриде М., Эльнагар А., Шаалан К.*, Методы машинного обучения и глубокого обучения для кибербезопасности: обзор. На Международной конференции по искусственному интеллекту и компьютерному зрению Springer, Cham. Стр. 50–57.

Саркер И.Х., Кайес А.С.М., Бадша С., Аллахтани Х., Уоттерс П., Нг А., 2020 г. — *Саркер И.Х., Кайес А.С.М., Бадша С., Аллахтани Х., Уоттерс П., Нг А.*, Наука о данных кибербезопасности: обзор с точки зрения машинного обучения. Журнал больших данных, стр. 1–29.

Сингх Дж. П., Кумар А., Рана Н.П., Дживеди Ю.К., 2020 — *Сингх Дж.П., Кумар А., Рана Н.П., Дживеди Ю.К.*, Сеть LSTM на основе внимания для оценки достоверности слухов о твитах. Границы информационных систем, стр. 1–16.

VK.com - Социальная сеть Вконтакте

Ю.М., Хуанг К., Цинь Х., Шееле К., Ян К., 2019 — *Ю.М., Хуанг К., Цинь Х., Шиле К., Ян К.*, Глубокое обучение для социальных сетей в реальном времени классификация текстов для понимания ситуации — с использованием ураганов «Сэнди», «Харви» и «Ирма» в качестве тематических исследований. Международный журнал Digital Earth, 12(11), стр. 1230–1247 гг.

Юань С.У.С., 2021 — *Юань С.У.С.*, Глубокое обучение для обнаружения внутренних угроз: обзор, проблемы и возможности. Компьютеры и безопасность, 104, стр.221.

## REFERENCES

Canhoto A.I., Clear F., 2020 — *Canhoto A.I., Clear F.* Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. Business Horizons — 63(2). Pp. 183–193.

<https://vk.com/dev/methods> - VK API methods

Instagram.com - Instagram Social Network

Jain A., Mandowara J., 2016 — *Jain A., Mandowara J.*, Text classification by combining text classifiers to improve the efficiency of classification. International Journal of Computer Application — 6(2). Pp. 2250–1797.

Kumar A., Singh J.P., Dwivedi Y.K., Rana N.P., 2020 — *Kumar A., Singh J.P., Dwivedi Y.K., Rana N.P.*, A deep multi-modal neural network for informative Twitter content classification during emergencies. Annals of Operations Research. Pp. 1–32.

Mohan A., Singh A.K., Kumar B., Dwivedi R., 2021 — *Mohan A., Singh A.K., Kumar B., Dwivedi R.*, Review on remote sensing methods for landslide detection using machine and deep learning. Transactions on Emerging Telecommunications Technologies — 32(7). P.998.

Mühlhoff R., 2020 — *Mühlhoff R.*, Human-aided artificial intelligence: Or, how to run large computations in human brains? Toward a media sociology of machine learning. new media & society, 22(10). Pp.1868–1884.

Mussiraliyeva S., Omarov B., Yoo P., Bolatbek M., 2021 — *Mussiraliyeva S., Omarov B., Yoo P., Bolatbek M.*, Applying machine learning techniques for religious extremism detection on online user contents. Computers, Materials and Continua, — 70(1). Pp.915–934.

Neupane D., Seok J., 2020 — *Neupane D., Seok J.*, Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review. IEEE Access. P. 9178.

Salloum S.A., Alshurideh M., Elnagar A., Shaalan K., 2020 — *Salloum S.A., Alshurideh M., Elnagar A., Shaalan K.*, Machine learning and deep learning techniques for cybersecurity: a review. In The International Conference on Artificial Intelligence and Computer Vision Springer, Cham. Pp. 50–57.

Sarker I.H., Kayes A.S.M., Badsha S., Alqahtani H., Watters P., Ng A., 2020 — *Sarker I. H., Kayes A.S.M., Badsha S., Alqahtani H., Watters P., Ng A.*, Cybersecurity data science: an overview from machine learning perspective. *Journal of Big data*. Pp. 1–29.

Singh J.P., Kumar A., Rana N.P., Dwivedi Y.K., 2020 — *Singh J.P., Kumar A., Rana N.P., Dwivedi Y.K.*, Attention-based LSTM network for rumor veracity estimation of tweets. *Information Systems Frontiers*. Pp. 1–16.

VK.com - Vkontakte Social Network

Yu M., Huang Q., Qin H., Scheele C., Yang C., 2019 — *Yu M., Huang Q., Qin H., Scheele C., Yang C.*, Deep learning for real-time social media text classification for situation awareness—using Hurricanes Sandy, Harvey, and Irma as case studies. *International Journal of Digital Earth*, 12(11). Pp. 1230–1247.

Yuan S., Wu X., 2021 — *Yuan S., Wu X.*, Deep learning for insider threat detection: Review, challenges and opportunities. *Computers & Security*, 104. P.-221.

**МАЗМҰНЫ**

<b>А. Адамова, Т. Жукабаева, Е. Марденов</b> ЗАТТАР ИНТЕРНЕТІ: ЖЕҢІЛДІК АЛГОРИТМДЕРДІҢ ДАМУЫ ЖӘНЕ БОЛАШАҒЫ.....	5
<b>Г. Алпысбай, А. Бедельбаев, О. Усагова, А. Жұмабекова, Эдзард Хофиг</b> ЗИЯНДЫ БАҒДАРЛАМАЛЫҚ ЖАБДЫҚТАРДЫ ТАЛДАУДА МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМІН ҚОЛДАНУ.....	21
<b>А.У. Алтаева, А.Ш. Каипова, А.У. Мухамеджанова, Г.К. Оспанова</b> МЕДИЦИНАДА ЧАТ-БОТТАРДЫ ҚОЛДАНУ ПЕРСПЕКТИВАЛАРЫ.....	32
<b>Г.А. Анарбекова, Н.Н. Оспанова, Д.Ж. Анарбеков</b> НОРМАЛАНҒАН КІРІС ВЕКТОРЛАРЫ: ДЕРЕКТЕРДІ ДАЙЫНДАУДЫҢ БАСТАПҚЫ КЕЗЕҢІ.....	40
<b>А.Е. Әбжанова, А.И. Такуадина, С.К. Сагнаева, С.К. Серикбаева, Г.Т. Азиева</b> ТОПЫРАҚТЫ ТЕХНИКАЛЫҚ МЕЛИОРАЦИЯЛАУ ӘДІСТЕРІНДЕ АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІ ПАЙДАЛАНУ.....	55
<b>К.Н. Әлібекова, Ж.М. Алимжанова, С.С. Байзакова</b> СЫМСЫЗ СЕНСОРЛЫҚ ЖЕЛІЛЕР ҮШІН БЛОКТЫҚ ШИФРЛАРДЫҢ ӨНІМДІЛІГІН БАҒАЛАУ.....	70
<b>К.Б. Багитова, Ш.Ж. Мүсірәлиева, М.А. Болатбек, Р.Қ. Оспанов</b> ИНТЕРНЕТТЕ ЭКСТРЕМИСТІК МАЗМҰНДЫ АНЫҚТАУҒА АРНАЛҒАН EXWEB БАҒДАРЛАМАЛЫҚ ЖАБДЫҚТАМАСЫН ӨЗІРЛЕУ.....	81
<b>А.Ш. Баракова, О.А. Усагова, А.С. Орынбаева</b> ВЕБ САЙТТАРДАҒЫ САНДЫҚ РЕСУРСТАРДЫ СТЕГАНОГРАФИЯ ӘДІСІМЕН ҚОРҒАУДЫҢ МОДЕЛІ.....	96
<b>А.С. Омарбекова, А.Е. Назырова, Н. Тасболатұлы, Б.Ш. Разахова</b> ИНТЕЛЛЕКТУАЛДЫ ELEARNING ЖҮЙЕСІНІҢ ОНТОЛОГИЯЛЫҚ МОДЕЛІ ЖӘНЕ ОҚЫТУ НӘТИЖЕЛЕРІ.....	108
<b>М.Қ. Болсынбек, Г.Б. Абдикеримова, С.К. Серикбаева, А.Ж. Танирбергенов, Ж.К. Тасжурекова</b> ТОПЫРАҚ ЖӘНЕ ТОПЫРАҚ ЭРОЗИСЫН БОЛЖАУЖЫҢ АҚПАРАТТЫҚ ЖҮЙЕЛЕРІ МЕН ӘДІСТЕРІН ЗЕРТТЕУ.....	128
<b>Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Б.А. Ху Вен-Цен</b> LSTM ЖӘНЕ GRU ҮЛГІЛЕРІ НЕГІЗІНДЕ ҚАЗАҚ ДАКТИЛЬДЕРІН ТАҢУДЫҢ ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕСІН ҚҰРУ.....	141
<b>М.Д. Кабибуллин, Б.Б. Оразбаев, К.Н. Оразбаева, С.Ш. Исакова, Ж.Ш. Аманбаева</b> КҮРДЕЛІ ХИМИЯЛЫҚ-ТЕХНОЛОГИЯЛЫҚ ЖҮЙЕЛЕР АГРЕГАТТАРЫНЫҢ МОДЕЛЬДЕРІН БАСТАПҚЫ АҚПАРАТТЫҢ ЖЕТІСПЕУШІЛІГІ МЕН АЙҚЫНСЫЗДЫҒЫ ЖАҒДАЙЫНДА ҚҰРУ.....	154

<b>М.Ж. Қалдарова, А.С. Аканова, М.Г. Гриф, У.Ж. Айтимова, А.С. Муканова</b> ТОПЫРАҚ ЖАҒДАЙЫН БАҒАЛАУ ҮШІН ҚОЛДАНЫЛАТЫН ҒАРЫШТЫҚ СУРЕТТЕРДІ ӨНДЕУ АЛГОРИТМДЕРІ МЕН ӘДІСТЕРІ.....	172
<b>К. Келесбаев, Ш. Раманкулов, М. Нуризинова, А. Паттаев, Н. Мұсахан</b> STEM ЖОБАЛЫҚ ОҚЫТУДЫҢ БОЛАШАҚ ФИЗИКА МАМАНДАРЫН ДАЯРЛАУДАҒЫ ЕРЕКШЕЛІКТЕРІ.....	193
<b>А.Е. Кулакаева, Е.А. Дайнеко, А.З. Айтмагамбетов, А.Т. Жетписбаева, Б.А. Кожаметова</b> ШАҒЫН ҒАРЫШ АППАРАТЫ ОРБИТАСЫНЫҢ СИПАТТАМАЛАРЫНЫҢ СПУТНИКТИК РАДИО МОНИТОРИНГ ЖҮЙЕСІНІҢ ПАРАМЕТРЛЕРІНЕ ӘСЕРІ ТУРАЛЫ.....	208
<b>А.Е. Назырова, Г.Т. Бекманова, А.С. Муканова, Н. Амангелді, М.Ж. Қалдарова</b> БІЛІМ БЕРУ БАҒДАРЛАМАЛАРЫ ҮШІН АВТОМАТТАНДЫРЫЛҒАН ЖҮЙЕНІ ӨЗІРЛЕУ.....	221
<b>А.Б. Тоқгарова, Б.С. Омаров, Ж.Ж. Ажибекова, Г.И. Бейсенова, Р.Б. Абдрахманов</b> ОНЛАЙН КОНТЕНТТЕГІ БЕЙӘДЕП СӨЗДЕР МӘЛІМЕТТЕР ҚОРЫН DATA MINING АРҚЫЛЫ АНАЛИЗДЕУ.....	237
<b>Ә.Б. Тынымбаев, К.С. Байшоланова, К.Е. Кубаев</b> АҚПАРАТТЫ ҚОРҒАУ ЖҮЙЕЛЕРІНДЕГІ NAVIVE BAYESIAN ЖІКІТІУШСІН ҚОЛДАНУ.....	252
<b>Г.Қ. Шаметова, А.Ә. Шәріпбай, Б.Ф. Сайлау</b> ҚОЛЖЕТІМДІЛІКТІ БАСҚАРУ ЖҮЙЕЛЕРІНДЕГІ ҚҰПИЯНЫ БӨЛҮДІҢ КРИПТОГРАФИЯЛЫҚ СҰЛБАЛАРЫН ТАЛДАУ.....	261
<b>Г.Б. Абдикеримова, А.Ә. Шекербек, М.Г. Байбулова, С.К. Абдикаримова, Ш.Ш. Жолдасова</b> КЕУДЕ ПАТОЛОГИЯСЫН АВТОКОРРЕЛЯЦИЯЛЫҚ ФУНКЦИЯ АРҚЫЛЫ АНЫҚТАУ.....	274

## СОДЕРЖАНИЕ

<b>А. Адамова, Т. Жукабаева, Е. Марденов</b> ИНТЕРНЕТ ВЕЩЕЙ: СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ЛЕГКОВЕСНЫХ АЛГОРИТМОВ.....	5
<b>Г. Алпысбай, А. Бедельбаев, О. Усагова, А. Жумабекова, Эдзарт Хофиг</b> ПРИМЕНЕНИЕ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ВРЕДОНОСНОГО ПО.....	21
<b>А.У. Алтаева, А.Ш. Каипова, А.У. Мухамеджанова, Г.К. Оспанова</b> ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ЧАТ-БОТОВ В МЕДИЦИНЕ.....	32
<b>Г.А. Анарбекова, Н.Н. Оспанова*, Д.Ж. Анарбеков</b> НОРМАЛИЗОВАННЫЕ ВХОДНЫЕ ВЕКТОРЫ: ПЕРВИЧНЫЙ ЭТАП ПОДГОТОВКИ ДАННЫХ.....	40
<b>А.Е. Абжанова, А.И. Такуадина, С.К. Сагнаева, С.К. Серикбаева, Г.Т. Азиева</b> ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ В МЕТОДАХ ТЕХНИЧЕСКИХ МЕЛИОРАЦИЙ ГРУНТОВ.....	55
<b>К.Н. Алибекова, Ж.М. Алимжанова, С.С. Байзакова</b> ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ БЛОЧНЫХ ШИФРОВ ДЛЯ БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЕЙ.....	70
<b>К.Б. Багитова, Ш.Ж. Мусиралиева, М.А. Болатбек, Р.К. Оспанов</b> РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ EXWEB ДЛЯ ВЫЯВЛЕНИЯ ЭКСТРЕМИСТСКОГО КОНТЕНТА В СЕТИ ИНТЕРНЕТ.....	81
<b>А.Ш. Баракова, О.А. Усагова, А.С. Орынбаева</b> РАЗРАБОТКА МОДЕЛИ ЗАЩИТЫ ЦИФРОВЫХ WEB РЕСУРСОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ СТЕГАНОГРАФИИ.....	96
<b>А.С. Омарбекова, А.Е. Назырова, Н. Тасболатұлы, Б.Ш. Разахова</b> ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ЭЛЕКТРОННОГО ОБУЧЕНИЯ И РЕЗУЛЬТАТЫ ОБУЧЕНИЯ.....	108
<b>М.Қ. Болсынбек, Г.Б. Абдикеримова, С.К. Серикбаева, А.Ж. Танирбергенов, Ж.К. Тасжурекова</b> ИССЛЕДОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ И МЕТОДОВ ПРОГНОЗИРОВАНИЯ ПОЧВЕННОЙ И ПОЧВЕННОЙ ЭРОЗИИ.....	128
<b>Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Б.А. Ху Вен-Цен</b> РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ КАЗАХСКИХ ДАКТИЛЬНЫХ ЖЕСТОВ НА ОСНОВЕ МОДЕЛЕЙ LSTM И GRU.....	141
<b>М.Д. Кабибуллин, Б.Б. Оразбаев, К.Н. Оразбаева, С.Ш. Искакова, Ж.Ш. Аманбаева</b> РАЗРАБОТКА МОДЕЛЕЙ АГРЕГАТОВ СЛОЖНЫХ ХИМИКО-ТЕХНОЛОГИЧЕСКИХ СИСТЕМ В УСЛОВИЯХ ДЕФИЦИТА И НЕЧЕТКОСТИ ИСХОДНОЙ ИНФОРМАЦИИ.....	154

<b>М.Ж. Калдарова, А.С. Аканова, М.Г. Гриф, У.Ж. Айтимова, А.С. Муканова</b> АЛГОРИТМЫ И МЕТОДЫ ОБРАБОТКИ КОСМИЧЕСКИХ СНИМКОВ ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ПОЧВ.....	172
<b>К. Келесбаев, Ш. Раманкулов, М. Нуризинова, А. Паттаев, Н. Мұсахан</b> ОСОБЕННОСТИ ПРОЕКТНОГО ОБУЧЕНИЯ STEM В ПОДГОТОВКЕ БУДУЩИХ СПЕЦИАЛИСТОВ ПО ФИЗИКЕ.....	193
<b>А.Е. Кулакаева, Е.А. Дайнеко, А.З. Айтмагамбетов, А.Т. Жетписбаева, Б.А. Кожаметова</b> О ВЛИЯНИИ ХАРАКТЕРИСТИК ОРБИТЫ МАЛОГО КОСМИЧЕСКОГО АППАРАТА НА ПАРАМЕТРЫ СИСТЕМЫ СПУТНИКОВОГО РАДИОМОНИТОРИНГА.....	208
<b>А.Е. Назырова, Г.Т. Бекманова, А.С. Муканова, Н. Амангелді, М.Ж. Калдарова,</b> РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ДЛЯ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ.....	221
<b>А.Б. Токгарова, Б.С. Омаров, Ж.Ж. Ажибекова, Г.И. Бейсенова, Р.Б. Абдрахманов</b> АНАЛИЗ НЕОБРАЗНЫХ СЛОВ В ОНЛАЙН-КОНТЕНТЕ С ПОМОЩЬЮ DATA MINING.....	237
<b>Ә.Б. Тынымбаев, К.С. Байшоланова, К.Е. Кубаев</b> ПРИМЕНЕНИЕ НАИВНОГО БАЙЕСОВСКОГО КЛАССИФИКАТОРА В СИСТЕМАХ ЗАЩИТЫ ИНФОРМАЦИИ.....	252
<b>Г.Қ. Шаметова, А.Ә. Шәріпбай, Б.Ғ. Сайлау</b> АНАЛИЗ КРИПТОГРАФИЧЕСКИХ СХЕМ РАСПРЕДЕЛЕНИЯ СЕКРЕТОВ В СИСТЕМАХ УПРАВЛЕНИЯ ДОСТУПОМ.....	261
<b>Г.Б. Абдикеримова, А.А. Шекербек, М.Г. Байбулова, С.К. Абдикаримова,</b> <b>Ш.Ш. Жолдасова</b> ОПРЕДЕЛЕНИЕ ГРУДНОЙ ПАТОЛОГИИ С ПОМОЩЬЮ ФУНКЦИИ АВТОКОРРЕЛЯЦИИ.....	274

**CONTENTS**

<b>A. Adamova, T. Zhukabayeva, Y. Mardenov</b> INTERNET OF THINGS: STATUS AND PROSPECTS FOR THE DEVELOPMENT OF LIGHTWEIGHT ALGORITHMS.....	5
<b>G. Alpysbay, A. Bedelbayev, O. Ussatova, A. Zhumabekova, Edzard Höfig</b> APPLICATION OF MACHINE LEARNING ALGORITHM IN THE ANALYSIS OF MALICIOUS SOFTWARE.....	21
<b>A.U. Altaeva, A.S. Kaipova, A.U. Mukhamejanova, G.K. Ospanova</b> PROSPECTS OF USING CHATBOTS IN MEDICINE.....	32
<b>G.A. Anarbekova, N.N. Ospanova, D.Zh. Anarbekov</b> NORMALIZED INPUT VECTORS: THE PRIMARY STAGE OF DATA PREPARATION.....	40
<b>A.E. Abzhanova, A.I. Takuadina, S.K. Sagnaeva, S.K. Serikbayeva, G.T. Azieva</b> THE USE OF INFORMATION SYSTEMS IN THE METHODS OF TECHNICAL SOIL RECLAMATION.....	55
<b>K. Alibekova, Zh. Alimzhanova, S.S. Baizakova</b> RATING VALUATION OF BLOCK CIPHERS FOR WIRELESS SENSOR NETWORKS.....	70
<b>K.B. Bagitova, Sh.Zh. Mussiraliyeva, M.A. Bolatbek, R.K. Ospanov</b> DEVELOPMENT OF EXWEB SOFTWARE FOR DETECTING EXTREMIST CONTENT ON THE INTERNET.....	81
<b>A.Sh. Barakova, O.A. Usatova, A.S. Orynbaeva</b> DIGITAL RESOURCES ON WEBSITES MODEL OF PROTECTION BY STEGANOGRAPHY.....	96
<b>A.S. Omarbekova, A.E. Nazyrova, N. Tasbolatuly, B.Sh. Razakhova</b> ONTOLOGICAL MODEL OF AN INTELLIGENT E-LEARNING SYSTEM AND LEARNING OUTCOMES.....	108
<b>M. Bolsynbek, G. Abdikerimova, S. Serikbayeva, A. Tanirbergenov, Zh. Taszhurekova</b> RESEARCH OF INFORMATION SYSTEMS AND METHODS OF FORECASTING SOIL AND SOIL EROSION.....	128
<b>L. Zholshiyeva, T. Zhukabayeva, Sh. Turaev, M. Berdieva, B. Khu Ven-Tsen</b> DEVELOPMENT OF AN INTELLECTUAL SYSTEM FOR RECOGNIZING KAZAKH DACTYL GESTURES BASED ON LSTM AND GRU MODELS.....	141
<b>M. Kabibullin, B. Orazbayev, K. Orazbayeva, S. Iskakova, Zh. Amanbayeva</b> DEVELOPMENT OF MODELS OF UNITS OF COMPLEX CHEMICAL-TECHNOLOGICAL SYSTEMS UNDER CONDITIONS OF DEFICIENCY AND FUZZY OF INITIAL INFORMATION.....	154
<b>M.Zh. Kaldarova, A.S. Akanova, M.G. Grif, U.Zh. Aitimova, A.S. Mukanova</b> ALGORITHM AND METHOD OF PROCESSING SPACE PHOTOS FOR ASSESSMENT OF SOIL.....	172



<b>K. Kelesbaev, Sh. Ramankulov, M. Nurizinova, A. Pattaev, N. Mussakhan</b> FEATURES OF STEAM PROJECT TRAINING IN THE PREPARATION OF FUTURE SPECIALISTS IN PHYSICS.....	193
<b>A.E. Kulakayeva, Y.A. Daineko, A.Z. Aitmagambetov, A.T. Zhetpisbaeva, B.A. Kozhakhmetova</b> ABOUT THE INFLUENCE OF THE ORBIT CHARACTERISTICS OF A SMALL SPACECRAFT ON THE PARAMETERS OF THE SATELLITE RADIO MONITORING SYSTEM.....	208
<b>A.E. Nazyrova, G.T. Bekmanova, A.S. Mukanova, N. Amangeldi, M.Zh. Kaldarova</b> DEVELOPMENT OF AN AUTOMATED SYSTEM FOR EDUCATIONAL PROGRAMS.....	221
<b>A.B. Toktarova, B.S. Omarov, Zh.Zh. Azhibekova, G.I. Beissenova, R.B. Abdrakhmanov</b> ANALYSIS OF HATE SPEECH WORDS IN ONLINE CONTENT BY USING DATA MINING.....	237
<b>A.B. Tynymbayev, K.S. Baisholanova, K.Ye. Kubaev</b> APPLICATION OF NAVIVE BAYESIAN CLASSIFIER IN INFORMATION PROTECTION SYSTEMS.....	252
<b>G.K. Shametova, A.A. Sharipbay, B.G. Sailau</b> ANALYSIS OF CRYPTOGRAPHIC SECRET DISTRIBUTION SCHEMES IN ACCESS CONTROL SYSTEMS.....	261
<b>G.B. Abdikerimova, A.A. Shekerbek, M.G. Baibulova, S.K. Abdikarimova, Sh.Sh. Zholdassova</b> CHEST PATHOLOGY DETERMINATION THROUGH AUTOCORRELATION FUNCTION.....	274

**Publication Ethics and Publication Malpractice  
the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct ([http://publicationethics.org/files/u2/New\\_Code.pdf](http://publicationethics.org/files/u2/New_Code.pdf)). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

**[www.nauka-nanrk.kz](http://www.nauka-nanrk.kz)**

**<http://physics-mathematics.kz/index.php/en/archive>**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Заместитель директор отдела издания научных журналов НАН РК *Р. Жалиқызы*

Редакторы: *М.С. Ахметова, Д.С. Аленов*

Верстка на компьютере *Г.Д. Жадыранова*

Подписано в печать 12.06.2023.

Формат 60x881/8. Бумага офсетная. Печать – ризограф.

19,0 п.л. Тираж 300. Заказ 2.