

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



«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ  
«ХАЛЫҚ» ЖҚ

# Х А Б А Р Л А Р Ы

## ИЗВЕСТИЯ

РОО «НАЦИОНАЛЬНОЙ  
АКАДЕМИИ НАУК РЕСПУБЛИКИ  
КАЗАХСТАН»  
ЧФ «Халық»

## N E W S

OF THE ACADEMY OF SCIENCES  
OF THE REPUBLIC OF  
KAZAKHSTAN  
«Halyk» Private Foundation

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

**4 (348)**

**OCTOBER – DECEMBER 2023**

**PUBLISHED SINCE JANUARY 1963**  
**PUBLISHED 4 TIMES A YEAR**

ALMATY, NAS RK



## **ЧФ «ХАЛЫҚ»**

В 2016 году для развития и улучшения качества жизни казахстанцев был создан частный Благотворительный фонд «Халық». За годы своей деятельности на реализацию благотворительных проектов в областях образования и науки, социальной защиты, культуры, здравоохранения и спорта, Фонд выделил более 45 миллиардов тенге.

Особое внимание Благотворительный фонд «Халық» уделяет образовательным программам, считая это направление одним из ключевых в своей деятельности. Оказывая поддержку отечественному образованию, Фонд вносит свой посильный вклад в развитие качественного образования в Казахстане. Тем самым способствуя росту числа людей, способных менять жизнь в стране к лучшему – профессионалов в различных сферах, потенциальных лидеров и «великих умов». Одной из значимых инициатив фонда «Халық» в образовательной сфере стал проект *Ozgeris powered by Halyk Fund* – первый в стране бизнес-инкубатор для учащихся 9-11 классов, который помогает развивать необходимые в современном мире предпринимательские навыки. Так, на содействие малому бизнесу школьников было выделено более 200 грантов. Для поддержки талантливых и мотивированных детей Фонд неоднократно выделял гранты на обучение в Международной школе «Мирас» и в *Astana IT University*, а также помог казахстанским школьникам принять участие в престижном конкурсе «*USTEM Robotics*» в США. Авторские работы в рамках проекта «Тәлімгер», которому Фонд оказал поддержку, легли в основу учебной программы, учебников и учебно-методических книг по предмету «Основы предпринимательства и бизнеса», преподаваемого в 10-11 классах казахстанских школ и колледжей.

Помимо помощи школьникам, учащимся колледжей и студентам Фонд считает важным внести свой вклад в повышение квалификации педагогов, совершенствование их знаний и навыков, поскольку именно они являются проводниками знаний будущих поколений казахстанцев. При поддержке Фонда «Халық» в южной столице был организован ежегодный городской конкурс педагогов «*Almaty Digital Ustaz*».

Важной инициативой стал реализуемый проект по обучению основам финансовой грамотности преподавателей из восьми областей Казахстана, что должно оказать существенное влияние на воспитание финансовой грамотности и предпринимательского мышления у нового поколения граждан страны.

Необходимую помощь Фонд «Халык» оказывает и тем, кто особенно остро в ней нуждается. В рамках социальной защиты населения активно проводится работа по поддержке детей, оставшихся без родителей, детей и взрослых из социально уязвимых слоев населения, людей с ограниченными возможностями, а также обеспечению нуждающихся социальным жильем, строительству социально важных объектов, таких как детские сады, детские площадки и физкультурно-оздоровительные комплексы.

В копилку добрых дел Фонда «Халык» можно добавить оказание помощи детскому спорту, куда относится поддержка в развитии детского футбола и карате в нашей стране. Жизненно важную помощь Благотворительный фонд «Халык» оказал нашим соотечественникам во время недавней пандемии COVID-19. Тогда, в разгар тяжелой борьбы с коронавирусной инфекцией Фонд выделил свыше 11 миллиардов тенге на приобретение необходимого медицинского оборудования и дорогостоящих медицинских препаратов, автомобилей скорой медицинской помощи и средств защиты, адресную материальную помощь социально уязвимым слоям населения и денежные выплаты медицинским работникам.

В 2023 году наряду с другими проектами, нацеленными на повышение благосостояния казахстанских граждан Фонд решил уделить особое внимание науке, поскольку она является частью общественной культуры, а уровень ее развития определяет уровень развития государства.

Поддержка Фондом выпуска журналов Национальной Академии наук Республики Казахстан, которые входят в международные фонды Scopus и Wos и в которых публикуются статьи отечественных ученых, докторантов и магистрантов, а также научных сотрудников высших учебных заведений и научно-исследовательских институтов нашей страны является не менее значимым вкладом Фонда в развитие казахстанского общества.

**С уважением,  
Благотворительный Фонд «Халык»!**

#### **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:**

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

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

**ЖҮСІПОВ 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 Жабағаұлы**, 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), Нанокұрылымды 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/>*

## ГЛАВНЫЙ РЕДАКТОР:

**МУТАНОВ Галимкаир 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 Сагпаева (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 Тлексабул 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 Жабағевич**, 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/>*

#### **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-ЖК**, 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/>*

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

ISSN 1991-346X

Volume 4. Number 348 (2023). 117–131

<https://doi.org/10.32014/2023.2518-1726.223>

MPHTH 81.96.00

© **M. Bolatbek, K. Baisylbaeva, M. Sagynay\*, Sh. Mussiraliyeva,  
A. Zhumakhanova, 2023**

Al-Farabi Kazakh National University, Kazakhstan.

E-mail: [sagynaymoldir11@gmail.com](mailto:sagynaymoldir11@gmail.com)

### **DEVELOPMENT OF A PARSER PROGRAM FOR THE ACCUMULATION OF DESTRUCTIVE TEXTS AIMED AT YOUNG PEOPLE IN THE INTERNET SPACE**

**Bolatbek Milana** — PhD., senior lecturer of the Department of Information Systems of Al-Farabi Kazakh National University, 050040, Almaty, 71 Al-Farabi Avenue

E-mail: [bolatbek.milana@gmail.com](mailto:bolatbek.milana@gmail.com), <http://orcid.org/0000-0002-2153-180X>;

**Baisylbayeva Kymbat** — Master of science, Lecturer of the department “Information systems” of Al-Farabi Kazakh National University, 050040, Almaty, 71 Al-Farabi Avenue

E-mail: [baisylbaeva.k@gmail.com](mailto:baisylbaeva.k@gmail.com), <https://orcid.org/0000-0001-9753-0398>;

**Sagynay Moldir** — lecturer of the Department of Information Systems of Al-Farabi Kazakh National University, 050040, Almaty, 71 Al-Farabi Avenue

E-mail: [sagynaymoldir11@gmail.com](mailto:sagynaymoldir11@gmail.com), <http://orcid.org/0009-0004-1377-5742>;

**Mussiraliyeva Shynar** — Candidate of Physical and Mathematical Sciences, Head of the Department Information Systems of Al-Farabi Kazakh National University, 050040, Almaty, 71 Al-Farabi Avenue

E-mail: [mussiraliyevash@gmail.com](mailto:mussiraliyevash@gmail.com), <http://orcid.org/0000-0001-5794-3649>;

**Zhumakhanova Aygerim** — lecturer of the Department of Information Systems of Al-Farabi Kazakh National University, 050040, Almaty, 71 Al-Farabi Avenue

E-mail: [aygerim129@gmail.com](mailto:aygerim129@gmail.com), <https://orcid.org/0009-0008-0210-4037>.

**Abstract.** The article raises the urgent problem of the rapid spread of aggressive information in the global network, which threatens both society and national security. The Internet has become the main source of information, but extremist and terrorist groups skillfully use web technologies to spread their ideology and organize risky operations. This creates an urgent need to study, prevent and combat extremism on the Internet. The article also emphasizes the importance of understanding the various aspects of extremism and its dynamic nature. The authors identify four main classifiers of aggressive information on the Internet: bullying, racism, violent extremism and Nazism. The article discusses the importance of creating software for extracting information from unstructured or semi-structured data, such as text documents and databases. In the article, the authors emphasize the importance of collecting and systematizing information in various social networks (YouTube, VKontakte, Telegram) by creating a parser. This can be useful for analysis,

research, and other purposes. In a specific case, it is envisaged to create a corpus for collecting publicly available extremist data in the Kazakh language from various open sources, including social networks and video hosting.

**Keywords:** bullying, racism, violent extremism, nazism, api, html, parser

© М.А. Болатбек, К.Д. Байсылбаева, М. Сағынай\*, Ш.Ж. Мусиралиева,  
А.Н. Жумаханова, 2023

Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан, Алматы.

E-mail: sagynaymoldir11@gmail.com

## **ИНТЕРНЕТ КЕҢІСТІГІНДЕГІ ЖАСТАРҒА БАҒЫТТАЛҒАН ДЕСТРУКТИВТІ МӘТІНДЕРДІ ЖИНАҚТАУҒА ҚАЖЕТТІ ПАРСЕР БАҒДАРЛАМАСЫН ӘЗІРЛЕУ**

**Болатбек Милана** — Әл-Фараби атындағы Қазақ ұлттық университеті «Ақпараттық жүйелер» кафедрасының PhD., аға оқытушысы, 050040, Алматы, Әл-Фараби даңғылы, 71

E-mail: bolatbek.milana@gmail.com, <http://orcid.org/0000-0002-2153-180X>;

**Байсылбаева Кымбат** — Әл-Фараби атындағы Қазақ ұлттық университеті «Ақпараттық жүйелер» кафедрасының оқытушысы, 050040, Алматы, Әл-Фараби даңғылы, 71

E-mail: baisylbaeva.k@gmail.com, <https://orcid.org/0000-0001-9753-0398>;

**Сағынай Мәлдір** — Әл-Фараби атындағы Қазақ ұлттық университеті «Ақпараттық жүйелер» кафедрасының оқытушысы, 050040, Алматы, Әл-Фараби даңғылы, 71

E-mail: sagynaymoldir11@gmail.com, <http://orcid.org/0009-0004-1377-5742>;

**Мусиралиева Шынар** — физика-математика ғылымдарының кандидаты, «Ақпараттық жүйелер» кафедрасының меңгерушісі, Әл-Фараби атындағы Қазақ ұлттық университеті, 050040, Алматы, Әл-Фараби даңғылы, 71

E-mail: mussiraliyevash@gmail.com, <http://orcid.org/0000-0001-5794-3649>;

**Жумаханова Айгерім Нурланқызы** — Әл-Фараби атындағы Қазақ ұлттық университеті «Ақпараттық жүйелер» кафедрасының оқытушысы, 050040, Алматы, Әл-Фараби даңғылы, 71

E-mail: aygerim129@gmail.com, <https://orcid.org/0009-0008-0210-4037>.

**Аннотация.** Мақала қоғамға да, ұлттық қауіпсіздікке де қауіп төндіретін жаһандық желіде агрессивті ақпаратты жылдам таратудың өзекті мәселесін көтереді. Интернет ақпараттың негізгі көзіне айналды, бірақ экстремистік және террористік топтар өздерінің идеологиясын тарату және қауіпті операцияларды ұйымдастыру үшін веб-технологияларды шебер қолданады. Бұл онлайн экстремизмді зерттеу, алдын алу және онымен күресу үшін шұғыл қажеттілікті тудырады. Мақала сонымен қатар экстремизмнің әртүрлі аспектілерін және оның динамикалық сипатын түсінудің маңыздылығын көрсетеді. Авторлар интернеттегі агрессивті ақпараттың төрт негізгі жіктеуішін анықтайды: буллинг, расизм, зорлық-зомбылық экстремизмі және нацизм. Мақалада мәтіндік құжаттар және дерекқорлар сияқты құрылымдалмаған немесе жартылай құрылымдалған деректерден ақпарат алу үшін бағдарламалық парсер құрудың маңыздылығы талқыланады. Мақалада авторлар парсер жасау арқылы әртүрлі әлеуметтік желілерде (YouTube, Вконтакте, телеграм) ақпаратты жинау және жүйелеудегі маңыздылығын көрсетеді. Бұл талдау,



зерттеу және басқа мақсаттар үшін пайдалы болуы мүмкін. Нақты жағдайда элеуметтік желілер мен бейнехостингтерді қоса алғанда, әртүрлі ашық көздерден қазақ тілінде жалпыға қолжетімді экстремистік деректерді жинау үшін корпус құру қарастырылады.

**Түйін сөздер:** буллинг, расизм, зорлық-зомбылық экстремизмі, нацизм, API, HTML, парсер

*Берілген зерттеу Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті қаржыландыратын “Жастар экстремизмін анықтау және заманауи ақпараттық кеңістікте жастардың қауіпсіздігін қамтамасыз етуге арналған модельдер мен әдістерді әзірлеу” жобасы аясында орындалды (грант AP19576868).*

© М.А. Болатбек, К.Д. Байсылбаева, М. Сагынай\*, Ш.Ж. Мусиралиева,  
А.Н. Жумаханова, 2023

Казахский национальный университет имени аль-Фараби,  
Алматы, Казахстан.

E-mail: sagynaymoldir11@gmail.com

#### РАЗРАБОТКА ПРОГРАММЫ ПАРСЕРА ДЛЯ СБОРА ДЕСТРУКТИВНЫХ ТЕКСТОВ, ОРИЕНТИРОВАННЫХ НА МОЛОДЕЖЬ В ИНТЕРНЕТ-ПРОСТРАНСТВЕ

**Болатбек Милана** — PhD., старший преподаватель кафедры «Информационные системы» Казахского национального университета имени Аль-Фараби, 050040, проспект Аль-Фараби 71, Алматы, Казахстан

E-mail: bolatbek.milana@gmail.com, <http://orcid.org/0000-0002-2153-180X>;

**Байсылбаева Кымбат** — магистр естественных наук, преподаватель кафедры «Информационные системы» Казахского национального университета имени Аль-Фараби, 050040, проспект Аль-Фараби 71, Алматы, Казахстан

E-mail: baisylbaeva.k@gmail.com, <https://orcid.org/0000-0001-9753-0398>;

**Сагынай Молдир** — преподаватель кафедры «Информационные системы» Казахского национального университета имени Аль-Фараби, 050040, проспект Аль-Фараби 71, Алматы, Казахстан

E-mail: sagynaymoldir11@gmail.com, <http://orcid.org/0009-0004-1377-5742>;

**Мусиралиева Шынар** — кандидат физико-математических наук, заведующая кафедрой «Информационные системы» Казахского национального университета им. Аль-Фараби, 050040, проспект Аль-Фараби 71, Алматы, Казахстан

E-mail: mussiraliyevash@gmail.com, <http://orcid.org/0000-0001-5794-3649>;

**Жумаханова Айгерим Нурлановна** — преподаватель кафедры «Информационные системы» Казахского национального университета имени Аль-Фараби, 050040, проспект Аль-Фараби 71, Алматы, Казахстан

E-mail: aygerim129@gmail.com, <https://orcid.org/0009-0008-0210-4037>.

**Аннотация.** Статья поднимает актуальную проблему быстрого распространения агрессивной информации в глобальной сети, которая угрожает как обществу, так и национальной безопасности. Интернет стал основным источником информации, но экстремистские и террористические

группы умело используют веб-технологии для распространения своей идеологии и организации рискованных операций. Это создает острую потребность в изучении, профилактике и борьбе с экстремизмом в интернете. Статья также подчеркивает важность понимания различных аспектов экстремизма и его динамического характера. Авторы выделяют четыре основных классификатора агрессивной информации в интернете: буллинг, расизм, насильственный экстремизм и нацизм. В статье обсуждается необходимость создания программного обеспечения для извлечения информации из неструктурированных или полуструктурированных данных, таких как текстовые документы и базы данных. В статье авторы подчеркивают важность сбора и систематизации информации в различных социальных сетях (YouTube, Вконтакте, телеграм) путем создания парсера. Это может быть полезно для анализа, исследований и других целей. Авторами предусматривается создание корпуса для сбора общедоступных экстремистских данных на казахском языке из различных открытых источников, включая социальные сети и видеохостинги.

**Ключевые слова:** буллинг, расизм, насильственный экстремизм, нацизм, API, HTML, парсер

### **Кіріспе**

Жаһандық желіде агрессивті ақпараттың жылдам таралуы қоғам үшін де, билік үшін де өзекті мәселе болып табылады. Қазіргі таңда Интернет негізгі ақпарат көзі ретінде пайда болып, мәтіндік деректер қорының ең кең таралған түріне айналды. Алайда экстремистік және террористік топтар ақпарат тарату, идеологияны насихаттау және радикалды миссияларды жүзеге асыру мақсатында веб-технологияларды шебер пайдалануды үйренді. Бұл ұлттық қауіпсіздікке елеулі қатер туғызады.

Экстремистік ұйымдар қылмыстық әрекеттер үшін Интернет ұсынатын шексіз мүмкіндіктерді белсенді пайдаланады. Оларға қылмыстарды дайындау және орындау, қоғамдық қауіпті операцияларды ұйымдастыру үшін құпия ақпарат алмасуды жүргізу, экстремистік кәсіпорындарды қаржыландыру сияқты орталықтандырылмаған қаржылық желілерді пайдалану жатады. Оның үстіне, олар арнайы құрылған веб-сайттар мен басқа интернет-ресурстарда өздерінің идеологиясын насихаттау үшін мұқият жоспарланған ақпараттық операцияларға қатысады. Демек, экстремизм мәселесі соңғы жылдары күшейіп, мемлекеттік маңызы бар және Қазақстанның ұлттық қауіпсіздігіне тікелей қатер төндіретін мәселеге айналды. Бұл күрделі процесс онлайн экстремизмнің салдарын анықтау, алдын алу және азайту, тиімді және уақтылы қарсы шараларды жүзеге асыру үшін шұғыл зерттеу қажеттілігін көрсетеді.

Экстремизм — қоғамдағы негізгі немесе қалыптасқан нормалар мен құндылықтардан жиі ауытқып кететін экстремалды және радикалды сенімдермен, әрекеттермен немесе идеологиялармен сипатталатын күрделі

және көп қырлы құбылыс. Ол өзінің экстремистік мақсаттарына жету үшін зорлық-зомбылық немесе басқа да зиянды құралдарды қолдануды қоса алғанда, үлкен күшке баруға дайын болуды білдіреді. Экстремистік идеологиялар саясатты, дінді, әлеуметтік мәселелерді немесе ұлтшылдықты қоса алғанда, әртүрлі салаларды қамтуы мүмкін және олар әртүрлі жолдармен көрінуі мүмкін. Экстремизмді түсіну оның зорлық-зомбылық, әлеуметтік толқулар немесе ұлттық қауіпсіздікке қатер төндіруі мүмкін ықтимал салдарын жою үшін өте маңызды [1].

Экстремизм динамикалы және үнемі өзгеріп, жаңа белгілерге ие болып отырады. Қазіргі уақытта бұл лаңкестік әрекеттердің көбеюіне, желілік және еркін байланысқан құрылыммен сипатталатын ұйымдасқан террористік желілердің құрылуына әкеліп соқты, ал Интернет бұл ұйымдардың ақпарат алмасуы мен ілгерілетуінің негізгі арнасы ретінде қызмет етеді. Олардың негізгі байланыс арналары — әлеуметтік желілер, электрондық пошта және веб-ресурстар. YouTube, VKontakte және Telegram сияқты әлеуметтік желілердің пайда болуы экстремистік, зорлық-зомбылық және радикалды мазмұнның таралуын жеделдетті. Демек, интернетте террористік және экстремистік ақпаратты қалыптастыратын және тарататын жеке тұлғалардан, топтардан және желілік қоғамдастықтардан туындайтын қауіптерді анықтау қажеттілігі туындайды. Бұл байланыс тақырыптарын қадағалауды және мінез-құлықты бақылауды және болжауды талап етеді.

Экстремистік нанымдар мен әрекеттер әр түрлі болуы мүмкін, және экстремалды көзқарасты ұстанатын барлық адамдар немесе топтар зорлық-зомбылыққа бармайды. Экстремизмді анықтау және онымен күресу көбінесе әлеуметтік, саяси, экономикалық және психологиялық факторлардың күрделі өзара әрекетін қамтиды. Төтенше нанымдарды ұстану мен зорлық-зомбылық немесе заңсыз әрекеттерді ажырата білу маңызды, өйткені біріншісі әдетте сөз бостандығымен қорғалады, ал екіншісі маңызды құқықтық және қауіпсіздік мәселелерін тудырады. Экстремизммен күресу әдетте радикалдануға қарсы тұру, толеранттылықты насихаттау және заң үстемдігін қамтамасыз ету бойынша күш-жігерді қоса алғанда, көп қырлы көзқарасты талап етеді.

Интернет экстремизмді деструктивті мәтін ретінде қарастырған жағдайда оны бағыты бойынша әр түрлі топтарға және классификаторларға жіктеуге болады. Дегенмен біз бұл зерттеуде төрт негізгі классификаторлармен жұмыс жасаймыз. Олар:

Буллинг;

Расизм;

Зорлық-зомбылық экстремизмі;

Ұлттық экстремизм.

Буллинг – бұл бір адам немесе адамдар тобы басқа адамға зиян келтіру, қорқыту немесе бақылау үшін ауызша, физикалық немесе психологиялық тактикаларды қолданатын күштердің теңгерімсіздігін қамтитын қайталанатын және қасақана агрессивті мінез-құлық. Буллинг мектептерде,

жұмыс орындарында немесе желіде әртүрлі жағдайларда орын алуы мүмкін және жәбірленушіге эмоционалды, физикалық немесе психологиялық зиян келтіруі мүмкін.

Расизм — адамдардың нәсіліне немесе этникалық тегіне байланысты тән артықшылығына немесе кемшілігіне сенуді қамтитын сенім жүйесі немесе көзқарастар жиынтығы. Бұл нәсілдік немесе этникалық шығу тегіне байланысты адамдарды немесе топтарды кемсітуге, немқұрайлы қарауға әкеледі. Нәсілшілдік жүйесінің жеке адамдар мен қауымдастық үшін зиянды салдары болуы мүмкін.

Зорлық-зомбылық экстремизмі — саяси, діни немесе идеологиялық мақсаттарға жету үшін зорлық-зомбылық немесе заңсыз құралдарды қолдануға ықпал ететін экстремистік идеологияларды мақұлдау немесе қолдауды білдіреді. Бұл көбінесе радикалдануды және адамдарды терроризм актілерін, зорлық-зомбылықты немесе заңсыз әрекеттердің басқа түрлерін насихаттайтын экстремистік топтарға тартуды қамтиды. Зорлық-зомбылық экстремистері ұлттық қауіпсіздік пен қоғамдық қауіпсіздікке қатер төндіруі мүмкін.

Ұлттық экстремизм (нацизм) — ұлттық социализмнің қысқартылған сөзі, 1930 және 1940 жылдардағы Германиядағы нацистік партияның сенімдері мен тәжірибесімен байланысты әсіре оңшыл саяси идеология. Нацизмге шектен шыққан ұлтшылдық, нәсілшілдік, антисемитизм, авторитаризм және арийлік нәсілдік тазалыққа ұмтылу тән. Ол Холокост пен Екінші дүниежүзілік соғыста басты рөл атқарды, миллиондаған адамдардың, соның ішінде еврейлердің, сығандардың, мүгедектердің және басқалардың қуғынға ұшырауына және жаппай өлтірілуіне әкелді. Нацизм тарихтағы ең жиіркенішті идеологиялардың бірі ретінде кеңінен айыпталады [2-3].

#### *Әдебиеттерге шолу*

[4] мақалада авторлар интернет-технологиялардың дамуымен желіде адам өміріне қауіп төндіретін және мемлекет заңдарын бұзатын көптеген ақпарат пайда болатынын атап көрсетеді. Бұл веб-сайттар санының халықтың жартысынан көбінен тез өсіп келе жатқандығын және күрделі мәселе болып табылатын үлкен көлемдегі ақпаратты өңдеуге шұғыл қажеттілік туындағаны туралы айтады. Бұл мәселені ішінара шешу үшін қазіргі уақытта мәтіндік корпус белсенді түрде жасалуда. Бұл жұмыс жарамсыз мазмұнды анықтауда маңызды құрал болып табылатын мәтіндік корпуссты құруды сипаттайды.

[5] мақалада авторлар қоғамның ақпараттық-психологиялық қауіпсіздігіне онлайн-ортадағы жағымсыз сөздердің көбеюіне байланысты үнемі қауіп төніп тұрғанын атап өтті. Атап айтқанда, нәсіліне, этникалық тегіне, жынысына, гендерлік сәйкестігіне, дініне, жасына, мүгедектігіне және зорлық-зомбылықты насихаттауды қоса алғанда осындай санаттарға бөлінген. Практикалық тәжірибе, жасанды интеллекттің әртүрлі әдістерінің арқасында балағат сөздер мен сөз тіркестерін іздеу және анықтау барысында енді адамның ең аз қатысуымен жүзеге асырылуы мүмкін екенін растайды. Бұдан басқа, зерттеу балағат сөздердің шығу тегін ғана емес, сонымен қатар

тіл мен кибербуллингті қоса алғанда, арам сөздердің әртүрлі категорияларын ажыратуға мүмкіндік беретін ұғымдарды ұсынады. Авторлар әлеуметтік желілердегі балағат сөздерді автоматтандырылған сүзу үшін пайдалануға болатын деректер жиынтығына қол жеткізу үшін машиналық оқыту әдістерін қолданады. Жұмыс мәтіндік дерекқорды талдау мысалында жалған мәлімдемелерді анықтау және жіктеу үшін деректерді өндіруді пайдалануды ұсынады.

[6] мақалада әлеуметтік желілерде қазақ тіліндегі балағат пікірлерді автоматтандырылған түрде жинау мүмкіндігі зерттеледі. Қазіргі уақытта онлайн-орталарда теріс комментаторлар санының өсуі байқалуда, зерттеушілер Қазақстанда ұсынылған түрлі әлеуметтік желілер мен бұқаралық ақпарат құралдарының осындай пікірлері бар дерекқорды жинады. Авторлардың зертеуінде, машиналық оқыту әдістерін қолдана отырып, әлеуметтік медиа түсініктемелерінде қолданылатын қорлайтын лексиканың шығу тегін талдап қана қоймайды, сонымен қатар әртүрлі жағымсыз пікірлерді жіктейді және әрі қарай зерттеу үшін автоматтандырылған деректер жиынтығына қол жеткізуге мүмкіндік береді.

[7] мақалада қазіргі таңда әлемдік қоғамдастықтағы басты проблемалардың бірі-радикализм мен экстремизм екендігін айтады. Бұл мақалада қазіргі қоғамда кең таралған құбылыстар болып табылатын "радикализм" және "экстремизм" терминдеріне талдау жасалынған. Мақала осы терминдердің әрқайсысының нақты сипаттамаларын бөліп көрсетеді және олардың шығу тегі мен қоғамдағы эволюциясын талдайды. Жұмыстың негізгі мақсаты — "радикализм" терминін әртүрлі елдердің ғалымдарының, соның ішінде Батыс, Ресей және отандық зерттеушілердің зерттеулері негізінде анықтауға және оны "экстремизм" ұғымынан ажыратуға тырысады. Осы мақсатқа жету үшін "радикализм" терминін қолдану тарихы қарастырылған, сонымен қатар оның Ұлыбритания мен Германиядағы даму ерекшеліктері талданған. "Радикализм" терминін қолданатын саяси жағдай да қарастырылып, бағаланған.

### **Қолданылатын әдістер**

Парсер бағдарламалар веб-беттер, мәтіндік құжаттар, дерекқорлар және т.б. сияқты құрылымдық немесе жартылай құрылымдық деректерден ақпаратты алу үшін жасалады. Оларды деректерді жинау, мазмұнды талдау немесе белгілі бір тапсырмаларды автоматтандыру үшін пайдалануға болады.

Парсер бағдарламасын жасау әдетте келесі қадамдарды қамтиды:

Ең алдымен парсердің мақсатын нақты анықтап алу керек. Мысалы, бұл жаңалықтар сайтынан жаңалықтар тақырыптарын, электрондық коммерция дүкенінен өнім бағасын немесе әлеуметтік медиа деректерін алу немесе видео жазбалардың астындағы комментарийлерді алу және т.б. болуы мүмкін.

Парсер бағдарлама өзі жұмыс істейтін бастапқы деректердің құрылымы мен пішімін түсінуі керек. Мысалы, егер веб-беттерді талдайтын болса, онда олардың HTML құрылымын түсіну керек. Егер мәтіндік құжаттармен жұмыс жасайтын болса, онда алдымен олардың форматын анықтап алу қажет.

Парсерді жасау үшін әдетте Python сияқты әртүрлі бағдарламалау тілдерін және веб-сайттармен жұмыс істеуге арналған кітапханаларды (мысалы, BeautifulSoup немесе Scrapy) немесе мәтінді өңдеуге (мысалы, тұрақты өрнектер) пайдаланады.

Белгілі бір ережелерге сәйкес бастапқы көздерден қажетті деректерді шығаратын кодты жазу қажет. Қажетті ақпаратты дұрыс шығарғанына көз жеткізу үшін парсерді әртүрлі кірістерде тексеру керек. Қателіктер шығып жатса, онда, кодты жөндеу қажеттілігі туындайды. .

Деректер шығарылғаннан кейін оны дерекқорда сақтау, талдауды орындау немесе мақсатыңызға сәйкес басқа әрекеттерді орындауға болады.

Парсердің 3 түрі қолданылады:

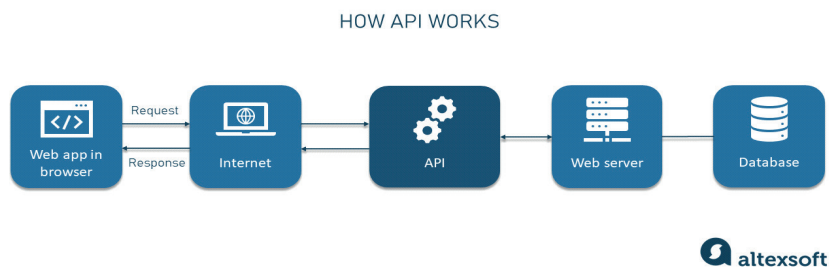
1) API арқылы

Бұл тәсіл өте ыңғайлы, жылдам және барлық керекті ақпаратты сенімді, қатесіз және артық ақпараттарымен қоса жинауға көмектеседі.

Парсер – әртүрлі сайттарда орналастырылған ақпаратты жинауға және жүйелеуге арналған программа. Деректер көзі мәтін мазмұны, веб-сайттың HTML коды, тақырыптар, мәзір элементтері, дерекқорлар және басқа элементтер болуы мүмкін. Ақпаратты жинау процесі парсинг деп аталады [8].

Бұл зерттеу деректердің айтарлықтай көлемін қажет етеді. Қазақ тілінде жалпыға қолжетімді экстремистік мәліметтерді анықтауға арналған корпусты құруда әртүрлі ашық ақпарат көздерінен, соның ішінде YouTube, ВКонтакте және Telegram желілерінен ақпарат жинауға арналған парсер әзірленді. API арқылы қолжетімді болған кезде осы көздерден ақпаратты алуға мүмкіндік беретін кірістірілген әдістер қолданылды.

API – ол қолданба әзірлеушісі мен осы қолданба өзара әрекеттесуі мүмкін орта арасындағы байланыстырушы құрал. Мұндай әдістер қол жетімді болмаған кезде, HTTP сұраулары арқылы қажетті ақпаратты алуға болады. 1-суретте API-дің жұмыс жасау реті көрсетілен [9].



1-сурет. API-дің жұмыс жасау сызбасы (altexsoft.com)

2) HTTP GET, POST сұраныстар арқылы

Сайттың html түпнұсқасын GET сұраныс арқылы алып оны BeautifulSoup4 және т.б. кітапханалармен оларды жинауға болады.

Кемшіліктері:

-кейде шамадан тыс көп сұраныстар жібергендіктен бұғаттауы мүмкін.  
-жұмыс атқару жылдамдығы орташа келеді  
-APIге қарағанда сенімділігі төмен болады және қандай да бір қателіктер жіберуі мүмкін.

-Lazy down форматындағы сайттардан жинау мүмкін емес.

### 3) Selenium немесе Bot Scraper

Бұл көбінесе get сұраныстардағы Lazy down сияқты проблемаларды шешеді. Себебі, бот парсер қосылғанда, адам браузерді қосқандай ашылады, керекті сайттарға кіреді. Солайша әрбір керек элементті браузерді ашқанда жинайды.

Кемшіліктері:

-өте баяу.

-сенімділігі get сұраныстарға қарағанда төмен және өте көп құлайды(сайтты қосуда қателіктер туындайды)

-4000–5000 шақты сайттан ақпарат жинау 1 күнді алуы мүмкін(сайтқа байланысты)

-оны қосу үшін кемінде 1 компьютер, 1 қолданушыны қосу қажет, яғни оған сырттан ештеңе кедергі келтірмей қосылып туруы керек. Ол дегеніміз параллельді түрде ақпарат алу қиын кейде тіпті мүмкін емес. Себебі, бот компьютердегі браузерді адам қосқандай ашады.

Lazy down — сайт толық қосылғаннан кейін кейбір элементтері уақыт өте келе жиналады. Сондықтан get сұраныстар кезінде олардың көрінбей қалуы әбден мүмкін.

Парсер бағдарламалары әдетте сөздіктерді тікелей іздемейді. Олар бағдарлама кодының синтаксисін талдауға және одан ақпарат алуға арналған, бірақ сөздіктермен немесе сөздіктер сияқты деректер құрылымдарымен жұмыс істеуге арналмаған. Дегенмен, талдау мақсатына байланысты сөздіктер және басқа деректер құрылымдары парсерлерде әртүрлі мақсаттарда пайдаланылуы мүмкін:

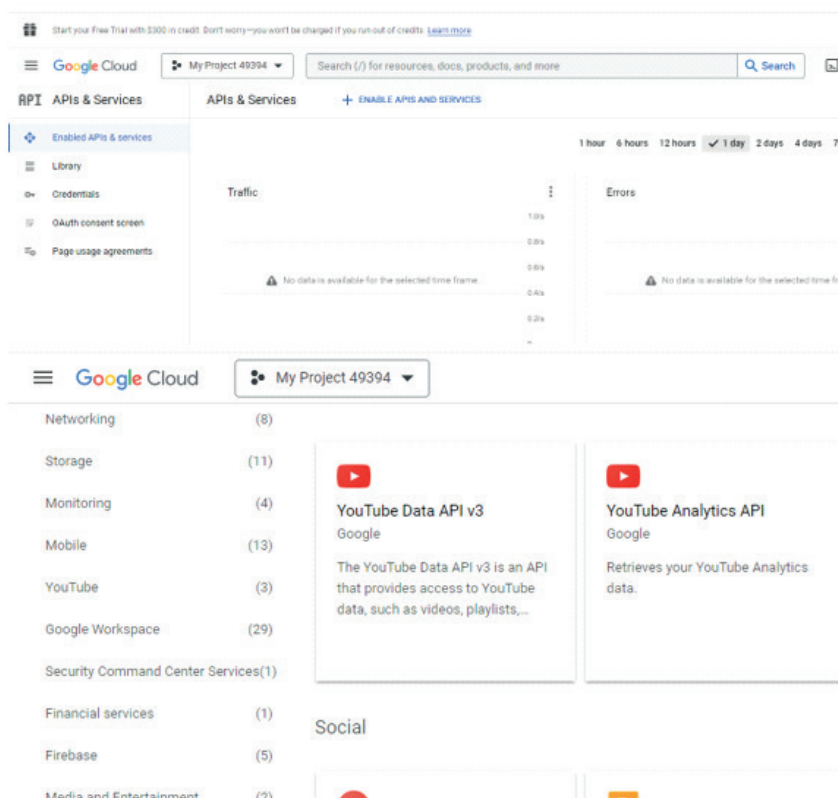
Кейбір бағдарламалау тілдері (мысалы, Python) деректерді сақтау және өңдеу үшін сөздіктер сияқты кірістірілген деректер құрылымдарын қамтамасыз етеді. Мұндай тілдердегі бағдарламаларды талдау кезінде талдаушы тілдің өзі қамтамасыз ететін сөздіктерді тануға және олармен жұмыс істеуге конфигурациялануы мүмкін.

Жалпы, парсерлерде сөздіктерді іздеу және жұмыс істеу тәсілі жобаның нақты талаптарына және қолданылатын бағдарламалау тіліне байланысты. Кейбір жағдайларда бұл бағдарлама кодын талдау кезінде сөздіктермен жұмыс істеуге қажетті функционалдылықты қамтамасыз ету үшін арнайы код жазуды немесе үшінші тарап кітапханаларын пайдалануды қажет етуі мүмкін.

Деректерді талдау негізінде деректер көзінен қажетті ақпаратты шығаратын талдаушы немесе сценарий жасау керек. Бұл HTML өңдеу үшін код жазуды (веб-скрепинг жағдайында), API интерфейстерімен жұмыс істеуді, файлдарды оқуды және т.б. қамтуы мүмкін.

Дұрыс жұмыс істейтініне көз жеткізу үшін парсерді әртүрлі деректерде тексеру қажет. Қателерді туындаған жағдайда оған түзетулер енгізу қажет болады.

API интерфейсі мен әлеуметтік желілердің өзара әрекеттесуі сұраулар кітапханасын пайдалану арқылы жеңілдетілді. Әлеуметтік желілерден деректерді талдау үшін Python бағдарламалау тілі таңдалды. API қажетті деректерді шығарып алуда маңызды рөл атқарды. Бұл жоба серверге HTTPS сұраулары арқылы әлеуметтік желі дерекқорынан қажетті ақпаратты алуға мүмкіндік беретін жүйені құруды қамтиды. 2-суретте YouTube Data API қолданылуы келтірілген.



2-сурет. YouTube Data API қолдану

YouTube парсерді қолдану барысы:

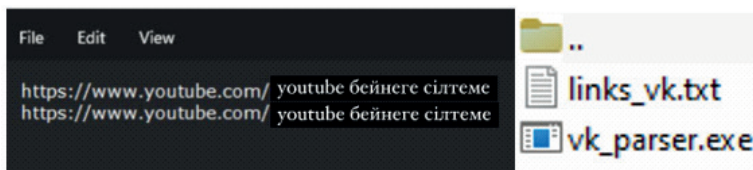
Парсерді жүктеу;

Бос абзацтармен бөлінген сілтемелерді қою (3-сурет);

Ехе файлын іске қосу (youtube\_parser.exe);

Программа өздігінен аяқталғанша күту. Барлық ақпарат dataset\_youtube ішінде CSV файлы түрінде сақталады.





3-сурет. YouTube парсерді қолдану

VK парсердің жұмыс жасау принципі YouTube парсерге ұқсас. Оның қолданылу барысы:

Парсерді жүктеу;

Бос абзацтармен бөлінген домендерді қою(4-сурет);

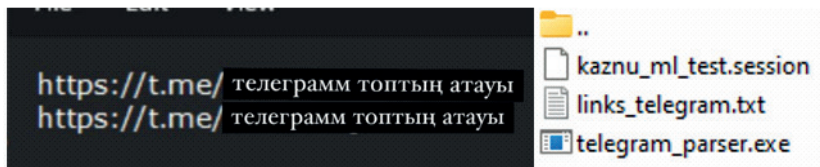
Ехе файлын іске қосу (vk\_parser.exe);

Программа өздігінен аяқталғанша күту. Барлық ақпарат dataset\_vk ішінде CSV файлы түрінде сақталады.



4-сурет. VK парсерді қолдану

Telegram парсердің де жұмыс жасау принципі алдыңғы парсерлерге ұқсас(5-сурет).



5-сурет. Telegram парсерді қолдану

### Нәтижелер

Мәтіндерді жинақтау мақсатында қауіпті деген топтар, парақшалар қарастырылды. Олардың саны шамамен 98. Қауіпті топтарды қарау мақсатында Телеграм, Вконтакте, Youtube сынды әлеуметтік желілер алынды. Осы аталған топтардың сілтемесін алып парсер бағдарламаға берді, соның нәтижесінде аталған топтардан кілттік сөздер бойынша қауіпті, расизм, зорлық-зомбылық, нацизм және буллинг сипатындағы барлық жазбалар алынды.

Осы көрсетілген қауіпті сілтемелер негізінде төрт бағыттағы мәтіндер алынды. Төмендегі суреттерде әрқайсы көрсетілген.

violent,мен талибтерді қолдаймын. махсуд кіші өз тәуелсіз мемлекеттерінді құрындар деп жүгірді.				
violent,Талибан мұны Ауғанстандағы тәжіктерге дұрыс жасайды.				
violent,алға талибтер бұл тәжіктердің емес сіздердің жерлеріңіз.				
violent,жарайсындар талибтер әйелдер өз орнын білу керек еркектермен тең болмайсындар				
violent,мен БҰҰ мен батыстық гуманитарийларға қарағанда талибтерге көбірек сенемін.				
violent,қырғыздар мен орыстар көтеріліске шығады.				
violent,Талибан алға Урааа Кабул біздікі.				
violent,батыстық вируспен уланған Масуд. Талибан жеңсін.				
violent,америкалықтардың бастарын шабу керек.				
violent,толибтерді қолдаймын.				
violent,Талибанды күтемін мына елді жойғым келеді.				
violent,талибтерді құрметтеймін барлық жерді жаулап алса екен.				
violent,талибан барлығын бейбітшілікке әкеледі кәпірлерді өлтіріңдер.				
violent,өлтiре берсiн бiр бiрiн кiмге керек олар.				
violent,ТАЛИБАНИ бауырлар жарайсындар сатқын билікті құрту керек.				
violent,тыйым салынған Талибанға рұқсат беретін ұақыт келді.				
violent,кім Талибанға қарсы болса ол исламға қарсы деген сөз.				
violent,талибтер жарайсындар барлық елдер үшін БҰҰ болдындар.				
violent,Талибан мықты жарайсындар				

6-сурет. Зорлық-зомбылық экстремизмі бағытындағы мәтіндер

Аталған топтағы мәтіндер келесідей сипаттарға ие болып келеді: олар зорлық-зомбылық және қорлайтын сөздерді қолданады, ондай мәтіндер адамдардың бойында үрей, ашу сынды реакцияларды туындатуы мүмкін, адамдарға немесе жануарларға, қоршаған ортаға қатысты зорлық-зомбылық, азаптау сияқты әрекеттерді қолдануды қолдайтын мәтіндер кіреді. Зорлық-зомбылық экстремизмі бағытындағы мәтіндерді бақылаусыз тарату сол мәтіндерді оқитын адамдарға зиянын тигізіп, эмоционалды күйзеліске ұшыратуы мүмкін.

bullying,Оның ақыл-есі кем болғаны соншалық қарапайым заттарды да түсінбейді				
bullying,Ақылсыздықтан құрып бара жатқан бұл кісіге менің жаным ашиды				
bullying,Мен мүгедектерді жұмысқа алмаймын олар баю және тиімсіз				
bullying,Саған мына арнайы мүгедектер орындығына отыру керек. Сен қарапайым орындықта отыра алмайсың.				
bullying,Сен тым қараңғы жерге бара алмайсың себебі соқырсың				
bullying,Біз сізге баспананы ұсынбаймыз өйткені біздің компания мүмкіндігі шектеулі адамдарды алмайды				
bullying,Мен мүгедектермен жұмыс істегім келмейді бұл өте қиын				
bullying,Сіз бұл спортпен айналыса алмайсыз себебі сіздің мүгедектігіңіз сізді шектейді.				
bullying,Сіз өзіңіздің мүгедектігіңізге байланысты сақтандыруды ала алмайсыз өйткені сіз біздің компания үшін өте үлкен				
bullying,Оның өмірдегі сәтсіздіктерінің себебі – оның соқырлығы.				
bullying,Қандай жиіркенішті жүздер				
bullying,Бұл бөлмені пайдалана алмайсыз себебі ол мүгедектер үшін арнайы жасақталмаған.				
bullying,Біз сіздің қажеттіліктеріңізге сай бейімделген жұмыс орнын ұсына алмаймыз				
bullying,Сізге ауырсыну жайлы шағым айтудың қажеті жоқ себеі сіз мүгедексіз				
bullying,Мүгедектер – қоғамға жүк көмекке лайық емес				
bullying,Мүмкіндігі шектеулі адамдар өз мансабында ең танымал бола алмайды				
bullying,Саған өзінің мақсаттарың бен армандарыңа емес мүгедектігіңе назар аударуың керек.				
bullying,Мүгедектердің толыққанды қарым-қатынасы мен отбасы болуы мүмкін емес.				
bullying,Мүгедек бола тұра қалай саяхаттады жоспарлап отырғаныңды білмеймін. Мүмкін үйде отырғаның дұрыс шығар				

7-сурет. Буллинг бағытындағы мәтіндер

Буллинг бағытындағы мәтіндер негізінен қандай бір адамдар немесе адамдар тобын қорлауға, тілдеуге, балағаттауға, кемсітуге бағытталған мәтіндерді қамтиды. Негізгі құрамы қорлайтын сөздерден, келеке-мазақ ететін пікірлерден, жалған ақпарат тарататын мәтіндерден тұрады. Буллинг бағытындағы мәтіндерді таратумен күресу онлайн орталардағы қауіпсіздік пен әлеуметтік-психологиялық жай-күйді тұрақты ұстап тұру үшін өте маңызды.

racism,Негроидты шөрке осылай болады ма?									
racism,Оған қоса ол негр									
racism,Шөрке									
racism,негр									
racism,негрлер әлемде еңбек етуде									
racism,қайдағы жоқ шөрке біздің хабарламаларды өшіреді									
racism,Ол сен шөркесің									
racism,Сен сонда шөркесің бе?									
racism,Есі дұрыс Дашалар бар ма? Шөрке Дашалар бар сау бол									
racism,Барша негрлерге сәлем									
racism,Күлкілі негр									
racism,Хохолдар									
racism,Негрден 2 метр алшақ жүріңіз									
racism,Сондай кішкентай болса да негр									
racism,Енді су перісі қара ма? Қандай сұмдық									
racism,Дисней ақымақтары неге әртис қара нәсілді? Ертегіде ақ қой									
racism,Диснейде дальтониктер пайда болға ба әлде соқырлар әртістерді жинап жатыр ма?									
racism,Бұл негр неге өзіне көп алып тұр?									
racism,Негр. Қызық қай жерде судың астында осылай күйдіреді?									
racism,Мен қара адамдарды көргенде әрқашан қара комедияны қосамын									

## 8-сурет. Расизм бағытындағы мәтіндер

Расизм бағытындағы мәтіндер қандай да бір расалық топтарға қатысты дискриминациялық көзқарастарды қамтиды, яғни қандай бір нәсілді алалау, тіл тигізу, төмендету. Расизм мәтіндер мынадай сипатқа ие: белгілі бәр нәсілге қарсы өшпенділік, дұшпандылық сезімін ояту, кемсітетін сөздер қолдану, ақ нәсілділер мен қара нәсілділерді салыстырып, ақ нәсілділердің үстемдігі туралы пікір тарату, қандай да бір расалық топтарға қатысты жалған ақпарат тарату. Расизм бағытындағы мәтіндер қоғам үшін өте қауіпті және саналы адамдар тарапынан қабылданбайтын мәтіндер, ол әділдік, әділеттілік деген сияқты адамгершілікке жататын ұғымдарға қарама-қайшы болып келеді.

nazism,Кез келген басқа ақ емес нәсілділерді жою керек Гитлер жасасын! Брейвик жасасын! Брентон Таррант сәлем!									
nazism,Сіз неміс овчаркасын алып оны Голден ретривермен шағыстырсаңыз ешкім қаламайтын және оны есіруге немесе сатуға ты									
nazism,Палестинкалықтар ақымақ, Олар құрметтейтін жалғыз нәрсе ол зорлық зомбылық пен қаталдық.									
nazism,Жеріміз Гаитиге айналмасын... жабайы аңдардың аузын жабу керек.									
nazism,Америка ақ нәсілдің өлтірушісі. ...Мен сіздің қызыл ақ және көк шүберекпен дәретханамды бұзбас едім									
nazism,Корона вирусын тоқтату барлық заңсыз жүргендерді депортациялау шекараларды жабу иммиграцияны қазір тоқтату									
nazism,еврейлер өтірік айтады ақтар өлуде									
nazism,Америка ұлт ретінде тек террорды білдіреді									
nazism,Екінші нөмір – сионистік Израиль мемлекеті және оның дүние жүзіндегі бұқаралық ақпарат құралдары									
nazism,Ал үшінші орында Сауд Арабиясы тұр. Сауд Арабиясы Интерфитна бастамашысы болып табылады Ирактағы шити орталық									
nazism,Америкалық кеме суға батып бара жатыр. Ал оны шатастыратын дәл осы сайқымазақтар. Бізге ештеңе істеудің қажеті жоқ.									
nazism,Жолымыздан таймасаңдар бізді жайымызға қалдырмасаңдар сендерді өртеп жібереміз.									
nazism,еврейлер адамзаттың қас жауы									
nazism,ядролық қаруды 100% дінге сенбейтін елде ғана қолдануға рұқсат етілген									
nazism,Евреймен қалай әрекет ету керек? Сіз еврейді өлтіресіз. Үнділер жағдайында олардың кәсіпорындарын бомбалау арқылы									

## 9-сурет. Нацизм бағытындағы мәтіндер

Нацизм бағытындағы мәтіндер қандай да бір ұлт және оның өкілдерін жоғары қоюмен сипатталады. Қандай да бір ұлттың өзге ұлттан артықшылығын, басымдылығын, олардың өзіндік таңбалары мен ұрандарының мықтылығын,

көшбасшыларын мадақтау арқылы деңгейлерінің жоғарылығын көрсеткісі келетін бағытта жазылған мәтіндер нацизм бағытындағы мәтіндердің негізгі сипаттарының бірі болып табылады. Бұл бағыттағы мәтіндерде өзінен басқа ұлттардың құндылықтары төмен кейде тіпті жарамсыз деген пікір айтылады. Нацизм бағытындағы мәтіндер көп жағдайда заңсыз әрі ұлтаралық алауыздықты тудырады, көп қасірет пен әскери қақтығыстарға әкеліп соғады.

### **Қорытынды**

Әлеуметтік желі қазіргі таңда адам өмірінің ажырамас бір бөлігі болып табылады. Әлеуметтік желінің адам өміріне әкелетін пайдасы да, зияны да аз емес. Әлеуметтік желі көмегімен зиянкестер, киберқылмыскерлер қоғамға, адам өміріне орасан қауіп келтіре алады. Берілген жұмыс бойынша Интернет кеңістігіндегі экстремистік мәліметтерді анықтауға арналған корпус құрылды. Ол келесідей бағыттағы мәтіндерді қамтиды: буллинг, нәсілшілдік, зорлық-зомбылық экстремизмі, ұлтшылдық. Алдымен Интернет кеңістігіндегі қауіпті деген топтар анықталды. Қазақ тілінде жалпыға қолжетімді экстремистік мәліметтерді анықтауға арналған корпусы құруда әртүрлі ашық ақпарат көздерінен, соның ішінде YouTube, ВКонтакте және Telegram желілерінен ақпарат жинауға арналған парсер әзірленді. Парсердің көмегімен қажетті бағыттарды қамтитын мәтіндер алынды.

### **ӘДЕБИЕТТЕР**

Ахмад С., Асгар М.З., Алотайби Ф.М., Аван И. (2019). Көңіл-күйді талдау әдістерін қолдана отырып, әлеуметтік желілерге негізделген экстремистік тиістілікті анықтау және жіктеу // адамға бағытталған есептеу және ақпараттық ғылымдар. — 2019. — Том. 9. — № 24. — Рр. 1–23. — Q1

Девяткин Д., Смирнов И., Соловьев Ф., Суворова М., Чеповский А. (2019). әлеуметтік желіде экстремистік мәтіндерді табу. Информатика және ақпараттық жүйелер бойынша көпсалалы конференция. — MCCSIS 2019. — Порту. — Рр. 344–350.

Леспекова А.А., Муканова А.С., Елибаева Г.К. (2022). Тыйым салынған контентті анықтау үшін мәтіндік корпус құру // «TURKLANG 2022», «Түркі тілдерін компьютерлік өңдеу» атты Х халықаралық конференция ЕҢБЕКТЕРІ. — 2022 ж. — 22–230 б.

Майур Г., Свати А., Шрадха П., Кетан К. (2021). Интернеттегі экстремизмді анықтау: мәліметтер жиынтығына, жіктеу әдістеріне, тексеру әдістері мен құралдарына баса назар аударып, әдебиетке жүйелі шолу. / Kіru / IEEE. — 2021. — Q1

Садиев Е.Ө. (2018). Діни экстремизм және терроризм ұғымдарының тарихи-саяси ретроспективасы // Хабаршы. Философия сериясы. Мәдениеттану сериясы. Саясаттану сериясы. — №2 (64). — 2018. — 151–160 б.

Тоқтарова А.Б., Ажибекова Ж.Ж., Сұлтан Д.Р., Керимбеков М.А. (2023). Онлайн контенттегі қазақ тілді бейәдеп пікірлерді машиналық оқытуда жинақтау // Абай атындағы ҚазҰПУ-нің ХАБАРШЫСЫ, «Физика-математика ғылымдары» сериясы. — №1(81). — 2023. — 265–272 б.

Тоқтарова А.Б., Омаров Б.С., Ажибекова Ж.Ж., Бейсенова Г.И., Абдрахманов Р.Б. (2023). Онлайн контенттегі бейәдеп сөздер мәліметтер қорын data mining арқылы анализдеу // әл-Фараби атындағы Қазақ ұлттық университетінің Хабарлары. — 2023. — №2 (346). — 237–251 б.

<https://habr.com/ru/articles/464261> 04.11.2023

## REFERENCES

- Ahmad S., Asghar M.Z., Alotaibi F.M., Awan I. (2019). Detection and classification of social media-based extremist affiliations using sentiment analysis techniques // Human- centric Computing and Information Sciences. — 2019. — Vol. 9. — № 24. — Pp. 1–23. — Q1
- Deviatkin D., Smirnov I., Solovyev F., Suvorova M., Chepovskiy A. (2019). Extremist Text Detection In Social Web. Multi Conference on Computer Science and Information Systems. — MCCSIS 2019. — Porto. — Pp. 344–350.
- Lespekova A.A., Mukanova A.S., Elibaeva G.K. (2022). creation of a text corpus to identify prohibited content // Proceedings of the X International Conference "TURKLANG 2022", "computer processing of Turkic languages". — 2022. — Pp. 223–230.
- Mayur G., Swati A., Shraddha P., Ketan K. (2021). Online Extremism Detection: A Systematic Literature Review With Emphasis on Datasets, Classification Techniques, Validation Methods, and Tools. // IEEE Access. — 2021. — Q1
- Toktarova A.B., Azhibekova ZH.ZH., Sultan D.R., Kerimbekov M.A. (2023). collection of non-Kazakh-language comments in online content in machine learning // Bulletin of Abai Kaznpu, Series "physical and Mathematical Sciences". — No. 1(81). — 2023. — Pp. 265–272.
- Toktarova A.B., Omarov B.S., Azhibekova ZH.ZH., Beisenova G.I., Abdrakhmanov R.B. (2023). analysis of the database of obscene words in online content through data mining // reports of the Al-Farabi Kazakh National University. — 2023. — №2 (346). — Pp. 237–251.
- Sadibekov E.U. (2018). historical and political retrospective of the concepts of religious extremism and terrorism // Bulletin. Philosophy series. Cultural Studies series. Political science series. — №2 (64). — 2018. — Pp. 151–160.  
<https://habr.com/ru/articles/464261> 04.11.2023

## МАЗМҰНЫ

<b>Г.Б. Абдикеримова, Р.М. Аманов, Г.Т. Азиева, А.М. Заманбекова, Қ. Жеңсқанқызы</b> <i>МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП, АККОРДТЫ ТАҢУ ТАПСЫРМАСЫНДАҒЫ ДЫБЫСТЫ ӨҢДЕУ ӘДІСТЕРІН САЛЫСТЫРМАЛЫ ТАЛДАУ</i> .....	7
<b>Л.А. Абдыкеримова, Г.Е. Мырзабекова, Г.С. Омарова, Л. Ақзуллақызы, Г.Ш. Мусагулова</b> ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ АРҚЫЛЫ ЖҮРЕК ПАТОЛОГИЯСЫН АНЫҚТАУ.....	21
<b>А.Е. Әбжанова, Е.Ә. Әбжанов, А.А. Мырзамуратова, А.Г. Батырханов, А.Б. Бексейтова</b> ҚАШЫҚТАН ЗОНДТАУ АРҚЫЛЫ АЛЫНҒАН ТОПЫРАҚ ЫЛҒАЛДЫЛЫҒЫ.....	35
<b>У.Ж. Айтимова, М.Ж. Айтимов, Э.Н. Тулегенова, А.У. Есиркепова, Ж.Т. Абилдаева</b> СУРЕТТЕН ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ ӨРТ ОШАҒЫН АНЫҚТАУ.....	50
<b>К.М. Алдабергенова, М.Ж. Жасұзақова, М.Ж. Айтимов, Н.Т. Мұстафаева, К.К. Дауренбеков</b> АУЫЛ ШАРУАШЫЛЫҒЫН ЦИФРЛАНДЫРУ: ДАМУ МҮМКІНДІКТЕРІ МЕН ПЕРСПЕКТИВАЛАРЫ.....	64
<b>А.С. Баегизова, Г.И. Мухамедрахимова, Ж.Б. Ламашева, А.З. Абдрахманова, Т.Т. Оспанова</b> ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІМЕН КЕСКІННІҢ САПАСЫН ЖАҚСARTУ.....	78
<b>Г.Т. Бекманова, А.С. Омарбекова, М.А. Кантуреева, Н.О. Байгабылов, М.М. Құдабеков</b> ӘЛЕУМЕТТАНУЛЫҚ САУАЛНАМАЛЫҚ ЗЕРТТЕУЛЕРДЕГІ АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР.....	91
<b>М.Ә. Берсүгір, Г.У. Маматова, А.А. Нурпейсова, М.Б. Онгарбаева, Ж.Т. Алтынбекова</b> ТЕКСТУРАЛЫҚ ТИПТЕГІ СУРЕТТЕРДІ ЖАҚСARTУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ.....	104
<b>М.А. Болатбек, К.Д. Байсылбаева, М. Сағынай, Ш.Ж. Мусиралиева, А.Н. Жумаханова</b> ИНТЕРНЕТ КЕҢІСТІГІНДЕГІ ЖАСТАРҒА БАҒЫТТАЛҒАН ДЕСТРУКТИВТІ МӘТІНДЕРДІ ЖИНАҚТАУҒА ҚАЖЕТТІ ПАРСЕР БАҒДАРЛАМАСЫН ӨЗІРЛЕУ.....	117
<b>М.Қ. Болсынбек, Г.Б. Абдикеримова, Г.С. Омарова, А.Б. Остаева, А.Г. Батырханов</b> ТОПЫРАҚ ДАЙЫНДАУДЫ БОЛЖАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУДЫ ПАЙДАЛАНУ.....	132
<b>Ш.К. Ележанова, А.Г. Батырханов, А.Е. Чукуров, Б.С. Хайржанова, Д.А. Тагиев</b> АҚПАРАТТЫҚ БЕЛГІСІЗДІК ТИПОЛОГИЯСЫ ЖӘНЕ АҚПАРАТТЫ ІЗДЕУ ТҮРЛЕРІ.....	151
<b>М.М. Есмагамбетова, Т.Т. Оспанова, Л.К. Бобров, Т.Л. Тен, Т.У. Есмагамбетов</b> ҒАРЫШТЫҚ ТӨТЕНШЕ ЖАҒДАЙЛАРДЫ БАҚЫЛАУ КЕСКІНДЕРІН ӨҢДЕУДЕ ТҮСТЕРДІ ӨЛШЕУ БАҒДАРЛАМАЛЫҚ ҚҰРАЛЫН ТАҢДАУ.....	161

<b>Т.К. Жукабаева, А. Адамова, Б.А. Ху Вен-Цен, Е.М. Марденов, Л.З. Жолшиева</b> СЫМСЫЗ СЕНСОР ЖЕЛІСІНДЕГІ SYBIL ЖӘНЕ WORMHOLE ШАБУЫЛДАРЫН АНЫҚТАУ.....	171
<b>А.А. Исмаилова, Ж.Т. Бельдеубаева, А.А. Нурпейсова, Г.О. Исакова, Ж.З. Жантасова</b> ӨСІМДІК АУРУЛАРЫН ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ АНЫҚТАУ.....	184
<b>А.Х. Касымова, М.Б. Есенова, М.У. Худойберганов, А.Б. Остаева, М.Г. Байбулова</b> ДАҚЫЛДАРДЫҢ АУРУЛАРЫН ЖІКТЕУ ҮШІН ТЕРЕҢ ОҚЫТУ АЛГОРИТМДЕРІН ҚОЛДАНУ.....	198
<b>А.Ұ. Мұхиядин, М.У. Мукашева, У.Т. Махажанова, А.А. Муханова, Ж.Б. Ламашева</b> ПРОГРАММАЛЫҚ ҚҰРАЛДАР КӨМЕГІМЕН ЭКСТРЕМАЛДЫ ҚАШЫҚТЫҚТАН ОҚЫТУДЫҢ ОҚУШЫЛАРҒА ӘСЕРІН ЗЕРТТЕУ.....	209
<b>Б.Б. Оразбаев, Л.Т. Салыбек, К.Н. Оразбаева, Ш.К. Коданова, С.Ш. Исакова</b> МҰНАЙДЫ АЛҒАШҚЫ ӨНДЕУДЕ ЭЛЕКТРОТҰЗСЫЗДАНДЫРЫРУ ЖӘНЕ СУСЫЗДАНДЫРУ ПРОЦЕССТЕРІН ОПТИМИЗАЦИЯЛАУ ҮШІН МОДЕЛЬДЕР ҚҰРУ ТӘСІЛІ.....	224
<b>С.К. Серикбаева, М.Қ. Болсынбек, А.Д. Абдувалова, А.Т. Абдыхалық, Д.Е. Ануарбек</b> ТОПЫРАҚ САПАСЫН БОЛЖАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНУ: АЛГОРИТМДЕР МЕН ӘДІСТЕР.....	237
<b>А.Ж. Танирбергенов, Ж.К. Тасжурекова, С.К. Серикбаева, А.А. Шораев, А.Д. Абдувалова</b> ТОЛЫҚ МӘТІНДІ ҚҰЖАТТАРДЫ ІЗДЕУДІҢ МОДЕЛІ МЕН АҚПАРАТТЫҚ ЖҮЙЕСІН ҚҰРУ ӘДІСТЕРІ.....	253
<b>А.Ә. Таурбекова, Ө.Ж. Мамырбаев, К.Ж. Тұрғанбай</b> СЕЙСМИКАЛЫҚ ҚЫЗМЕТТІ БАҒАЛАУ ҮШІН ГИДРОДИНАМИКАЛЫҚ ТҰРАҚСЫЗДЫҚ ПРОЦЕСІН ЗЕРТТЕУ.....	268
<b>Н. Т. Тұржанов, Ш. К. Ележанова, С. Н. Идрисов, Ж. К. Дюсембина</b> АҚПАРАТТЫҚ ҮДЕРІСТЕРДІҢ РЕИНЖИНИРИНГІНІҢ ИННОВАЦИЯЛЫҚ КУРСЫН ӘЗІРЛЕУ.....	290
<b>В. Шевцов, А. Исмаилова, Ж. Белдеубаева, А. Сатыбалдиева, А. Нурпейсова</b> МЛВА ГЕНОТИПТЕУДІҢ ӘДІСІ ЖӘНЕ ОНЫ ЖҮЗЕГЕ АСЫРУ АЛГОРИТМДЕРІ РЕТІНДЕГІ ГЕНОМДЫҚ ДЕРЕКТЕРДІ ПАЙДАЛАНУ.....	300
<b>А.Ә. Шекербек, А.А. Некесова, Ж.Ж. Молдашева, А.И. Онгарбаева, А.О. Тохаева</b> ФРАКТАЛДЫҚ ӘДІСПЕН ӨКПЕНІҢ ПАТОЛОГИЯЛЫҚ ЖАҒДАЙЫН ТАЛДАУ.....	313

## СОДЕРЖАНИЕ

<b>Г.Б. Абдикеримова, Р.М. Аманов, Г.Т. Азиева, А.М. Заманбекова, К. Женсканкызы</b> СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ОБРАБОТКИ ЗВУКА В ЗАДАЧЕ РАСПОЗНАВАНИЯ АККОРДОВ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ.....	7
<b>Л.А. Абдыкеримова, Г.Е. Мурзабекова, Г.С. Омарова, Л. Акзуллакызы, Г.Ш. Мусагулова</b> ОБНАРУЖЕНИЕ СЕРДЕЧНОЙ ПАТОЛОГИИ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ.....	21
<b>А.Е. Абжанова, Е.А. Абжанов, А.А. Мырзамуратова, А.Г. Батырханов, А.Б. Бексейтова</b> ВЛАЖНОСТЬ ПОЧВЫ, ПОЛУЧЕННАЯ ДИСТАНЦИОННЫМ ЗОНДИРОВАНИЕМ.....	35
<b>У.Ж. Айтимова, М.Ж. Айтимов, Э.Н. Тулегенова, А.У. Есиркепова, Ж.Т. Абилдаева</b> ОБНАРУЖЕНИЕ ОЧАГОВ ПОЖАРА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ ПО ИЗОБРАЖЕНИЮ.....	50
<b>К.М. Алдабергенова, М.Ж. Жасузакова, М.Ж. Айтимов, Н.Т. Мустафаева, К.К. Дауренбеков</b> ЦИФРОВИЗАЦИЯ СЕЛЬСКОГО ХОЗЯЙСТВА: ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ РАЗВИТИЯ.....	64
<b>А.С. Баегизова, Г.И. Мухамедрахимова, Ж.Б. Ламашева, А.З. Абдрахманова, Т.Т. Оспанова</b> УЛУЧШЕНИЕ КАЧЕСТВА ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ.....	78
<b>Г.Т. Бекманова, А.С. Омарбекова, М.А. Кантуреева, Н.О. Байгабылов, М.М. Кудабеков</b> ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В СОЦИОЛОГИЧЕСКИХ ОПРОСНЫХ ИССЛЕДОВАНИЯХ.....	91
<b>М.А. Берсугир, Г.У. Маматова, А.А. Нурпейсова, М.Б. Онгарбаева, Ж.Т. Алтынбекова</b> ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ ТЕКСТУРНОГО ТИПА.....	104
<b>М.А. Болатбек, К.Д. Байсылбаева, М. Сагынай, Ш.Ж. Мусиралиева, А.Н. Жумаханова</b> РАЗРАБОТКА ПРОГРАММЫ ПАРСЕРА ДЛЯ СБОРА ДЕСТРУКТИВНЫХ ТЕКСТОВ, ОРИЕНТИРОВАННЫХ НА МОЛОДЕЖЬ В ИНТЕРНЕТ-ПРОСТРАНСТВЕ.....	117
<b>М.К. Болсынбек, Г.Б. Абдикеримова, Г.С. Омарова, А.Б. Остаева, А.Г. Батырханов</b> ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ПОДГОТОВКИ ПОЧВЫ.....	132
<b>Ш.К. Ележанова, А.Г. Батырханов, А.Е. Чукуров, Б.С. Хайржанова, Д.А. Тагиев</b> ТИПОЛОГИЯ ИНФОРМАЦИОННОЙ НЕОПРЕДЕЛЕННОСТИ И ТИПЫ ПОИСКА ИНФОРМАЦИИ.....	151



<b>М.М. Есмагамбетова, Т.Т. Оспанова, Л.К. Бобров, Т.Л. Тен, Т.У. Есмагамбетов</b> ВЫБОР ПРОГРАММНЫХ СРЕДСТВ ЦВЕТОМЕТРИИ В ОБРАБОТКЕ ИЗОБРАЖЕНИЙ КОСМИЧЕСКОГО МОНИТОРИНГА ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЙ.....	161
<b>Т.К. Жукабаева, А. Адамова, В.А. Ху Вен-Цен, Е.М. Марденов, Л.З. Жолшиева</b> ОБНАРУЖЕНИЕ SYBIL И WORMHOLE АТАК В БЕСПРОВОДНОЙ СЕНСОРНОЙ СЕТИ.....	171
<b>А.А. Исмаилова, Ж.Т. Бельдеубаева, А.А. Нурпейсова, Г.О. Исакова, Ж.З. Жантасова</b> ОБНАРУЖЕНИЕ БОЛЕЗНЕЙ РАСТЕНИЙ МЕТОДОМ ГЛУБОКОГО ОБУЧЕНИЯ.....	184
<b>А.Х. Касымова, М.Б. Есенова, М.У. Худойбергенов, А.Б. Остаева, М.Г. Байбулова</b> ПРИМЕНЕНИЕ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ БОЛЕЗНЕЙ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР.....	198
<b>А.У. Мухиядин, М.У. Мукашева, У.Т. Махажанова, А.А. Муханова, Ж.Б. Ламашева</b> ИЗУЧЕНИЕ ВЛИЯНИЯ ЭКСТРЕМАЛЬНОГО ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ НА УЧАЩИХСЯ С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНЫХ СРЕДСТВ.....	209
<b>Б.Б. Оразбаев, Л.Т. Салыбек, К.Н. Оразбаева, Ш.К. Коданова, С.Ш. Исакова</b> МЕТОД РАЗРАБОТКИ МОДЕЛЕЙ ДЛЯ ОПТИМИЗАЦИИ ПРОЦЕССОВ ЭЛЕКТРООБЕССОЛИВАНИЯ И ОБЕЗВОЖИВАНИЯ ПРИ ПЕРВИЧНОЙ ПЕРЕРАБОТКЕ НЕФТИ.....	224
<b>С.К. Серикбаева, М.К. Болсынбек, А.Д. Абдувалова, А.Т. Абдыхалык, Д.Е. Ануарбек</b> ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ КАЧЕСТВА ПОЧВЫ: АЛГОРИТМЫ И МЕТОДИКИ.....	237
<b>А.Ж. Танирбергенов, Ж.К. Тасжурекова, С.К. Серикбаева, А.А. Шораев, А.Д. Абдувалова</b> МЕТОДЫ ПОСТРОЕНИЯ МОДЕЛИ И ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОИСКА ПОЛНОТЕКСТОВЫХ ДОКУМЕНТОВ.....	253
<b>А.Ә. Taurbekova, O.Zh. Mamyrbayev, K.Zh. Doshtaev, T.K. Eginbaykyzy</b> HYDRODYNAMIC INSTABILITY MECHANISM PROCESS FOR ASSESSMENT SEISMIC ACTIVITY.....	268
<b>Н.Т. Туржанов, Ш.К. Ележанова, С.Н. Идрисов, Ж.К. Дюсембина</b> РАЗРАБОТКА ИННОВАЦИОННОГО КУРСА ПО РЕИНЖИНИРИНГУ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ.....	290
<b>В. Шевцов, А. Исмаилова, Ж. Бельдеубаева, А. Сатыбалдиева, А. Нурпейсова</b> MLVA КАК МЕТОД ГЕНОТИПИРОВАНИЯ И АЛГОРИТМЫ ЕГО РЕАЛИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ ПОЛНОГЕНОМНЫХ ДАННЫХ.....	300
<b>А.А. Шекербек, А.А. Некесова, Ж.Ж. Молдашева, А.И. Онгарбаева, А.О. Тохаева</b> АНАЛИЗ ПАТОЛОГИЧЕСКИХ СОСТОЯНИЙ ЛЕГКИХ С ИСПОЛЬЗОВАНИЕМ ФРАКТАЛЬНОГО МЕТОДА.....	313

## CONTENTS

<b>G.B. Abdikerimova, R.M. Amanov, G.T. Azieva, A.M. Zamanbekova, K. Zhengskankyzy</b> COMPARATIVE ANALYSIS OF SOUND PROCESSING METHODS IN THE CHORD RECOGNITION PROBLEM USING MACHINE LEARNING.....	7
<b>L. Abdykerimova, G. Murzabekova, G. Omarova, L. Akzullakyyzy, G. Mussagulova</b> DETECTION OF CARDIAC PATHOLOGY USING DEEP LEARNING METHODS.....	21
<b>A.E. Abzhanova, E.A. Abzhanov, A.A. Myrzamuratova, A.G. Batyrkhanov, A.B. Bekseitova</b> SOIL MOISTURE OBTAINED BY REMOTE SENSING.....	35
<b>U. Zh Aitimova, M.Zh. Aitimov, E.N. Tulegenova, A.U. Yessirkepova, Zh.T. Abildaeva</b> FIRE FOCUS DETECTION USING DEEP LEARNING METHODS FROM IMAGE.....	50
<b>K.M. Aldabergenova, M.ZH. Zhasuzakova, M.Zh. Aitimov, N.T. Mustafaeva, K.K. Daurenbekov</b> DIGITALIZATION OF AGRICULTURE: OPPORTUNITIES AND PROSPECTS FOR DEVELOPMENT.....	64
<b>A.S. Baegizova, G.I. Mukhamedrakhimova, Zh.B. Lamasheva, A.Z. Abdrakhmanova, T.T. Ospanova</b> IMPROVE IMAGE QUALITY WITH DEEP LEARNING TECHNIQUES.....	78
<b>G. Bekmanova, A. Omarbekova, M. Kantureyeva, N. Baigabylov, M. Kudabekov</b> INFORMATION TECHNOLOGIES IN SOCIOLOGICAL SURVEY RESEARCH.....	91
<b>M.A. Bersugir, G.U. Mamatova, A.A. Nurpeisova, M.B. Ongarbayeva, Zh.T. Altynbekova</b> USING MACHINE LEARNING METHODS TO IMPROVE TEXTURE-TYPE IMAGES.....	104
<b>M. Bolatbek, K. Baisylbaeva, M. Sagynay, Sh. Mussiraliyeva, A. Zhumakhanova</b> DEVELOPMENT OF A PARSER PROGRAM FOR THE ACCUMULATION OF DESTRUCTIVE TEXTS AIMED AT YOUNG PEOPLE IN THE INTERNET SPACE.....	117
<b>M. Bolsynbek, G. Abdikerimova, G. Omarova, A. Ostayeva, A. Batyrkhanov</b> APPLICATION OF MACHINE LEARNING TO PREDICT SOIL PREPARATION....	132
<b>Sh.K. Yelezhanova, A.G. Batyrkhanov, A.Y. Chukurov, B.S. Khairzhanova, J.A. Taghiyev</b> TYPOLOGY OF INFORMATION UNCERTAINTY AND TYPES OF INFORMATION RETRIEVAL.....	151
<b>M. Yesmagambetova, T. Ospanova, L. Bobrov, T. Ten, T. Yesmagambetov</b> SELECTION OF COLORIMETRY SOFTWARE TOOLS IN IMAGE PROCESSING OF SPACE MONITORING OF EMERGENCY SITUATIONS.....	161
<b>T. Zhukabayeva, A. Adamova, B. Khu Ven-Tsen, Y. Mardenov, L. Zholshiyeva</b> DETECTION OF SYBIL AND WORMHOLE ATTACKS IN A WIRELESS SENSOR NETWORK.....	171
<b>A.A. Ismailova, Zh.T. Beldeubayeva, A.A. Nurpeisova, G.O. Issakova, Zh.Z. Zhantassova</b>	

DETECTION OF PLANT DISEASES USING DEEP LEARNING METHODS.....	184
<b>A.K. Kassymova, M.B. Yessenova, M.U. Khudoyberganov, A.B. Ostayeva, M.G. Baibulova</b>	
APPLICATION OF DEEP LEARNING ALGORITHMS FOR CLASSIFICATION OF DISEASES OF AGRICULTURAL CROPS.....	198
<b>A. Mukhiyadin, M. Mukasheva, U. Makhazhanova, A. Mukhanova, Zh. Lamasheva</b>	
STUDYING THE EFFECTS OF EXTREME DISTANCE EDUCATION ON STUDENTS USING SOFTWARE TOOLS.....	209
<b>B. Orazbayev, L. Salybek, K. Orazbayeva, Sn. Kodanova, S. Iskakova</b>	
METHOD FOR DEVELOPING MODELS FOR OPTIMIZING PROCESSES OF ELECTRICAL DESALTING AND DEHYDRATION DURING PRIMARY OIL PROCESSING.....	224
<b>S.Serikbayeva, M.Bolsynbek, A. Abduvalova, A. Abdykhalyk, D. Anuarbek</b>	
APPLICATION OF MACHINE LEARNING TO PREDICT SOIL QUALITY: ALGORITHMS AND TECHNIQUES.....	237
<b>A. Tanirbergenov, Zh. Tashhurekova, S. Serikbayeva, A. Shorayev, A. Abduvalova</b>	
METHODS OF CONSTRUCTING A MODEL AND AN INFORMATION SYSTEM FOR SEARCHING FULL-TEXT DOCUMENTS.....	253
<b>A.Ə. Taurbekova, O.Zh. Mamyrbayev, K.Zh. Doshtaev, T.K. Eginbaykyzy</b>	
HYDRODYNAMIC INSTABILITY MECHANISM PROCESS FOR ASSESSMENT SEISMIC ACTIVITY.....	268
<b>N.T. Turzhanov, Sh.K. Yelezhanova, S.N. Idrissov, Zh.K. Dyusseminina</b>	
DEVELOPMENT OF AN INNOVATIVE COURSE REENGINEERING OF INFORMATION PROCESSES.....	290
<b>V. Shevtsov, A. Ismailova, Zh. Beldeubayeva, A. Satybaldiyeva, A. Nurpeisova</b>	
MLVA AS A METHOD OF GENOTYPING AND ALGORITHMS FOR ITS IMPLEMENTATION USING GENOME-WIDE DATA.....	300
<b>A.A. Shekerbek, A.A. Nekesova, Zh.Zh. Moldasheva, A.I. Ongarbayeva, A. Tokhaeva</b>	
ANALYSIS OF PATHOLOGICAL CONDITIONS OF THE LUNG USING THE FRACTAL METHOD.....	313

**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)**

Подписано в печать 28.12.2023.

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

21,0 п.л. Тираж 300. Заказ 4.