

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN  
ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ  
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН  
KAZAKHSTAN



**INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION  
TECHNOLOGIES**

Published since 2020.  
Volume 7. 1 (25). 2026  
January–March

**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ  
ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

2020 жылдан бері шығарылады  
Том 7. 1 (25). 2026  
Қаңтар-Наурыз

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

Издается с 2020 г.  
Том 7. 1 (25). 2026  
Январь-Март

Свидетельство о постановке на учет периодического печатного издания в Министерство информации и общественного развития Республики Казахстан № KZ82VPY00020475, выданное от 20.02.2020 г.

Зарегистрировано в Международном центре регистрации серийных изданий ISSN (ЮНЕСКО, Париж, Франция). ISSN 2708–2032 (print), ISSN 2708–2040 (online)

Журнал входит в Перечень научных изданий, рекомендуемых КОКНВО МНВО РК для публикации основных результатов научной деятельности.

#### EDITOR-IN-CHIEF:

**Kateryna Kolesnikova** — Doctor of Technical Sciences, professor, Vice-Rector for Research, International Information Technology University (Kazakhstan)

#### DEPUTY EDITOR-IN-CHIEF:

**Madina Ipalakova** — Candidate of Technical Sciences, associate professor, Director of the Research Department, International Information Technology University (Kazakhstan)

#### EDITORIAL BOARD:

**Abdul Razak** — PhD, professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)

**Lucio Tommaso De Paolis** — Director of the R&D Department of the AVR Laboratory, Department of Engineering for Innovation, University of Salento (Italy)

**Liz Bacon** — Professor, Deputy Vice-Chancellor, Abertay University (United Kingdom)

**Michele Pagano** — PhD, Professor, University of Pisa (Italy)

**Mukhtarbay Otelbayev** — Doctor of Physical and Mathematical Sciences, professor, academician of the National Academy of Sciences of the Republic of Kazakhstan, professor of the Department of Mathematical and Computer Modeling, International Information Technology University (Kazakhstan)

**Bolatbek Rysbauly** — Doctor of Physical and Mathematical Sciences, professor, professor of the Department of Computing and Data Science, Astana IT University (Kazakhstan)

**Yevgeniya Daineko** — PhD, research professor, Department of Information Systems, International Information Technology University (Kazakhstan)

**Nurzhan Duzbayev** — PhD, associate professor, Vice-Rector for Digitalization and Innovation, International Information Technology University (Kazakhstan)

**Bakhtgerai Sinchev** — Doctor of Technical Sciences, professor, Department of Information Systems, International Information Technology University (Kazakhstan)

**Nurgul Seilova** — Candidate of Technical Sciences, Dean of the Faculty of Computer Technologies and Cybersecurity, International Information Technology University (Kazakhstan)

**Ardak Mukhamediyeva** — Candidate of Economic Sciences, Dean of the Faculty of Business, Media and Management, International Information Technology University (Kazakhstan)

**Zamira Abdikalikova** — PhD, associate professor, Head of the Department of Mathematical and Computer Modeling, International Information Technology University (Kazakhstan)

**Yerlan Shildibekov** — PhD, associate professor, Head of the Department of Economics and Business, International Information Technology University (Kazakhstan)

**Damilya Yeskendiroya** — Candidate of Technical Sciences, associate professor, Head of the Department of Cybersecurity, International Information Technology University (Kazakhstan)

**Aigul Niyazgulova** — Candidate of Philological Sciences, Professor, Head of the Department of Media Communications and History of Kazakhstan, International Information Technology University (Kazakhstan)

**Altai Aitmagambetov** — Candidate of Technical Sciences, Professor, Department of Radio Engineering, Electronics and Telecommunications, International Information Technology University (Kazakhstan)

**Yelena Bakhtiyarova** — Candidate of Technical Sciences, associate professor, Head of the Department of Radio Engineering, Electronics and Telecommunications, International Information Technology University (Kazakhstan)

**Kanibek Sansyzbay** — PhD, research professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)

**Sakhybay Tynymbayev** — Candidate of Technical Sciences, Professor, Research Professor, Department of Computer Engineering, International Information Technology University (Kazakhstan)

**Ali Abd Almisreb** — PhD, associate professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)

**Mohamed Ahmed Hamada** — PhD, associate professor, Department of Information Systems, International Information Technology University (Kazakhstan)

**Yang Im Chu** — PhD, Professor, Gachon University (South Korea)

**Tadeusz Wallas** — PhD, Vice-Rector, Adam Mickiewicz University (Poland)

**Orken Mamyrbayev** — PhD, Deputy Director for Science, RSE Institute of Information and Computational Technologies, Committee for Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Kazakhstan)

**Sergey Bushuyev** — Doctor of Technical Sciences, professor, Director of the Ukrainian Project Management Association "UKRNET," Head of the Department of Project Management, Kyiv National University of Construction and Architecture (Ukraine)

**Svetlana Beloshitskaya** — Doctor of Technical Sciences, professor, Department of Computing and Data Science, Astana IT University (Kazakhstan)

#### MANAGING EDITOR

**Raushan Mrzabayeva** — Master of Science, editor, International Information Technology University (Kazakhstan)

---

International Journal of Information and Communication Technologies

Periodicity: 4 times a year.

Languages: Kazakh, Russian, English

DOI prefix: 10.54309

ISSN 2708-2032 (print)

ISSN 2708-2040 (online)

Thematic focus: "Information technology"; "Digital technologies in the development of socio-economic systems"; "Information security and communication technologies".

Distribution: Materials are distributed under the Creative Commons Attribution 4.0

Journal website: <https://journal.iitu.edu.kz>

Owner: International Information Technology University JSC (Almaty).

Copyright: © International Journal of Information and Communication Technologies, 2026

---

## РЕДАКЦИЯ

## БАС РЕДАКТОР:

**Колесникова Катерина Викторовна** — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университетінің ғылыми-зерттеу қызметі жөніндегі проректор (Қазақстан)

## БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

**Ипалакова Мадина Түлегеновна** — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университетінің ғылыми-зерттеу қызметі жөніндегі департамент директоры (Қазақстан)

## РЕДАКЦИЯЛЫҚ АЛҚА:

**Разак Абдул** — PhD, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының профессоры (Қазақстан)

**Лучино Томмазо де Паолис** — Саленто Университеті (Италия) инновация және технологиялық инжиниринг департаменті AVR зертханасының зерттеу және әзірлеу бөлімінің директоры

**Лиз Бэкон** — профессор, Абертей Университеті (Ұлыбритания) вице-канцлерінің орынбасары

**Микеле Пагано** — PhD, Пиза Университетінің (Италия) профессоры

**Өтелбаев Мұхтарбай Өтелбайұлы** — физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының профессоры (Қазақстан)

**Рысбайұлы Болатбек** — физика-математика ғылымдарының докторы, профессор, Есептеу және деректер ғылымдары департаментінің профессоры, Astana IT University (Қазақстан)

**Дайнеко Евгения Александровна** — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессор-зерттеушісі (Қазақстан)

**Дузаев Нуржан Тоқсуажасвич** — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті цифрландыру және инновациялар жөніндегі проректор (Қазақстан)

**Синчев Бахтгерей Кусанович** — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессоры (Қазақстан)

**Сейлова Нургуль Абадуллаевна** — техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университеті компьютерлік технологиялар және киберқауіпсіздік факультетінің деканы (Қазақстан)

**Мухамедиева Ардак Габитовна** — экономика ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті бизнес медиа және басқару факультетінің деканы (Қазақстан)

**Абикаликова Замира Тұрсынбаевна** — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының меңгерушісі (Қазақстан)

**Шильдибеков Ерлан Жаржанович** — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті экономика және бизнес кафедрасының меңгерушісі (Қазақстан)

**Дамеля Максұтовна Ескендірова** — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының меңгерушісі (Қазақстан)

**Ниязгулова Айгуль Аскарбековна** — филология ғылымдарының кандидаты, доцент, профессор, Халықаралық ақпараттық технологиялар университеті медиакоммуникация және Қазақстан тарихы кафедрасының меңгерушісі (Қазақстан)

**Айтмағамбетов Алтай Зуфарович** — техника ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының профессоры (Қазақстан)

**Бахтиярова Елена Ажибековна** — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының меңгерушісі (Қазақстан)

**Канибек Сансызбай** — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының профессор-зерттеушісі (Қазақстан)

**Тынымбаев Сахибай** — техника ғылымдарының кандидаты, профессор, Халықаралық ақпараттық технологиялар университеті компьютерлік инженерия кафедрасының профессор-зерттеушісі (Қазақстан)

**Алмисреб Али Абд** — PhD, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының қауымдастырылған профессоры (Қазақстан)

**Мохамед Ахмед Хамада** — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының қауымдастырылған профессоры (Қазақстан)

**Янг Им Чу** — PhD, Гачон университетінің профессоры (Оңтүстік Корея)

**Тадеш Валлас** — PhD, Адам Мицкевич атындағы (Польша) университеттің проректоры

**Мамырбаев Оркен Жұмажанович** — PhD, ҚР ҒЖБМ Ғылым комитеті ақпараттық және есептеу технологиялары институты ӨМК директорының ғылым жөніндегі орынбасары (Қазақстан)

**Бушуев Сергей Дмитриевич** — техника ғылымдарының докторы, профессор, Украинаның "УКРНЕТ" жобаларды басқару қауымдастығының директоры, Киев ұлттық құрылыс және сулест университеті жобаларды басқару кафедрасының меңгерушісі (Украина)

**Белюшицкая Светлана Васильевна** — техника ғылымдарының докторы, доцент, Astana IT University есептеу және деректер ғылымы кафедрасының профессоры (Қазақстан)

## ЖАУАПТЫ РЕДАКТОР:

**Мрзабаева Раушан Жалиевна** — магистр, Халықаралық ақпараттық технологиялар университетінің редакторы (Қазақстан)

Халықаралық ақпараттық және коммуникациялық технологиялар журналы

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Префикс DOI: 10.54309

Мерзімділігі: жылына 4 рет.

Басылым тілі: қазақ, орыс, ағылшын.

Тақырып бағыты: "Ақпараттық технологиялар"; "Ақпараттық қауіпсіздік және коммуникациялық технологиялар"; "Әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технология".

Журнал сайты: <https://journal.iitu.edu.kz>

Тарату: материалдар Creative Commons Attribution 4.0 лицензиясы бойынша таратылады

Меншік иесі: АҚ «Халықаралық ақпараттық технологиялар университеті» (Алматы қ.).

Авторлық құқық: © Халықаралық ақпараттық және коммуникациялық технологиялар журналы, 2026

**РЕДАКЦИЯ**

**ГЛАВНЫЙ РЕДАКТОР:**

**Колесникова Катерина Викторовна** — доктор технических наук, профессор, проректор по научно-исследовательской деятельности Международного университета информационных технологий (Казахстан)

**ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:**

**Ипалакова Мадина Тулегеновна** — кандидат технических наук, ассоциированный профессор, директор департамента по научно-исследовательской деятельности Международного университета информационных технологий (Казахстан)

**РЕДАКЦИОННАЯ КОЛЛЕГИЯ:**

**Разак Абдул** — PhD, профессор кафедры кибербезопасности Международного университета информационных технологий (Казахстан)

**Лучио Томмазо де Паолис** — директор отдела исследований и разработок лаборатории AVR департамента инноваций и технологического инжиниринга Университета Саленто (Италия)

**Лиз Бэкон** — профессор, заместитель вице-канцлера Университета Абертей (Великобритания)

**Микеле Пагано** — PhD, профессор Университета Пизы (Италия)

**Отелбаев Мухтарбай Отелбайулы** — доктор физико-математических наук, профессор, академик НАН РК, профессор кафедры математического и компьютерного моделирования Международного университета информационных технологий (Казахстан)

**Рысбайулы Болатбек** — доктор физико-математических наук, профессор, профессор Astana IT University (Казахстан)

**Дайнеко Евгения Александровна** — PhD, профессор-исследователь кафедры информационных систем Международного университета информационных технологий (Казахстан)

**Дузбаев Нуржан Токкужаевич** — PhD, ассоциированный профессор, проректор по цифровизации и инновациям Международного университета информационных технологий (Казахстан)

**Синчев Бахтгерей Куспанович** — доктор технических наук, профессор, профессор кафедры информационных систем Международного университета информационных технологий (Казахстан)

**Сейлова Нургуль Абадуллаевна** — кандидат технических наук, декан факультета компьютерных технологий и кибербезопасности Международного университета информационных технологий (Казахстан)

**Мухамедиева Ардак Габитовна** — кандидат экономических наук, декан факультета бизнеса медиа и управления Международного университета информационных технологий (Казахстан)

**Абдикаликова Замира Турсынбаевна** — PhD, ассоциированный профессор, заведующая кафедрой математического и компьютерного моделирования Международного университета информационных технологий (Казахстан)

**Шильдибеков Ерлан Жаржанович** — PhD, ассоциированный профессор, заведующий кафедрой экономики и бизнеса Международного университета информационных технологий (Казахстан)

**Дамеуш Максустовна Ескендирова** — кандидат технических наук, ассоциированный профессор, заведующая кафедрой кибербезопасности Международного университета информационных технологий (Казахстан)

**Ниязгулова Айгуль Аскарбековна** — кандидат филологических наук, доцент, профессор, заведующая кафедрой медиакоммуникации и истории Казахстана Международного университета информационных технологий (Казахстан)

**Айтмаганбетов Алтай Зуфарович** — кандидат технических наук, профессор кафедры радиотехники, электроники и телекоммуникаций Международного университета информационных технологий (Казахстан)

**Бахтиярова Елена Ажибековна** — кандидат технических наук, ассоциированный профессор, заведующая кафедрой радиотехники, электроники и телекоммуникаций Международного университета информационных технологий (Казахстан)

**Канибек Сансызбай** — PhD, ассоциированный профессор, профессор-исследователь кафедры кибербезопасности, Международного университета информационных технологий (Казахстан)

**Тынымбаев Сахбай** — кандидат технических наук, профессор, профессор-исследователь кафедры компьютерной инженерии, Международного университета информационных технологий (Казахстан)

**Алимсиреб Али Абд** — PhD, ассоциированный профессор кафедры кибербезопасности Международного университета информационных технологий (Казахстан)

**Мохамед Ахмед Хамада** — PhD, ассоциированный профессор кафедры информационных систем Международного университета информационных технологий (Казахстан)

**Янг Им Чу** — PhD, профессор университета Гачон (Южная Корея)

**Тадеуш Валлас** — PhD, проректор университета имен Адама Мицкевича (Польша)

**Мамырбаев Оркен Жумажанович** — PhD, заместитель директора по науке РГП Института информационных и вычислительных технологий Комитета науки МНВО РК (Казахстан)

**Бушуев Сергей Дмитриевич** — доктор технических наук, профессор, директор Украинской ассоциации управления проектами «УКРНЕТ», заведующий кафедрой управления проектами Киевского национального университета строительства и архитектуры (Украина)

**Белошицкая Светлана Васильевна** — доктор технических наук, доцент, профессор кафедры вычислений и науки о данных Astana IT University (Казахстан)

**ОТВЕТСТВЕННЫЙ РЕДАКТОР:**

**Мрзабаева Раушан Жалиевна** — магистр, редактор Международного университета информационных технологий (Казахстан)

Международный журнал информационных и коммуникационных технологий

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Префикс DOI: 10.54309

Периодичность: 4 выпусков в год.

Язык издания: казахский, русский, английский.

Тематическая направленность: "Информационные технологии"; "Информационная безопасность и коммуникационные технологии"; "Цифровые технологии в развитии социо-экономических систем".

Сайт журнала: <https://journal.iitu.edu.kz>

Распространение: материалы распространяются по лицензии Creative Commons Attribution 4.0

Собственник: АО «Международный университет информационных технологий» (г. Алматы).

Авторские права: © Международный журнал информационных и коммуникационных технологий, 2026

## CONTENTS

### DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF SOCIOECONOMIC SYSTEMS

|                                                                                                                                                                                                      |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>A.B. Zhalgas, Y.N. Kalpakov, B.Ye. Amirgaliyev</b><br>MACHINE LEARNING-DRIVEN OPTIMIZATION OF LOGISTICS IN SMART CITIES: A CASE STUDY OF ASTANA .....                                             | 11 |
| <b>L. Kurmangazyeva, Sh. Kodanova, M. Urazgaliyeva, O. Findik, S. Iskakova</b><br>INTEGRATING FUZZY LOGIC AND ARTIFICIAL INTELLIGENCE IN OPTIMIZING BUSINESS PROCESS AUTOMATION DECISIONS .....      | 26 |
| <b>Y. Mailybayev, U. Adilbayeva, R. Amanova</b><br>ORGANIZATION OF AN ONLINE SURVEY OF PARTICIPANTS IN THE EDUCATIONAL PROCESS AND ANALYSIS OF THE RESULTS BASED ON THE MODIFIED DELPHI METHOD ..... | 48 |
| <b>V.A. Takizhanov, A.Z. Ibragimov, A. Shalakhmetov</b><br>SIMULATION-BASED ROBUSTNESS ASSESSMENT OF ASTANA'S BUS NETWORK UNDER RANDOM AND TARGETED FAILURES .....                                   | 63 |
| <b>G. Hamada, M. Hamada</b><br>LEVERAGING MACHINE LEARNING MODELS TO BOOST THE MANIPULATION OF UNSTRUCTURED DECISIONS AND UNSTRUCTURED DATA IN DIGITAL BUSINESS FIRMS .....                          | 78 |

### INFORMATION TECHNOLOGY

|                                                                                                                                                                                                                           |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>M. Zh. Aitimov, G. K. Muratova, Zh. K. Bissenbayeva, I.M. Bapiyev, M. Kassim</b><br>SEMANTIC COMPLETENESS IN KAZAKH-LANGUAGE EXTRACTIVE QA THROUGH ONTOLOGY AND RETRIEVAL MECHANISMS .....                             | 96  |
| <b>A.Aitim, A. Auyezova</b><br>ENCODING NP-COMPLETE PROBLEM INSTANCES INTO NATURAL LANGUAGE FOR LLM PROCESSING .....                                                                                                      | 109 |
| <b>O.N. Akylbekov, Y.T. Dauletbek, A.N. Moldagulova, G.S. Zakariya, D.A. Gura</b><br>MACHINE LEARNING METHODS FOR ANALYSING THREE-DIMENSIONAL SPATIAL DATA IN KAZAKHSTAN'S LAND USE PLANNING .....                        | 126 |
| <b>S.Zh. Aliaskarov, R.K. Uskenbayeva, A. Razaque, A.B. Kassymova, A.M. Anartayeva</b><br>TOWARDS EFFICIENT BIG DATA ANALYTICS IN REGIONAL SYSTEMS: PRACTICAL INSIGHTS FROM HYBRID ARCHITECTURE DEPLOYMENT .....          | 146 |
| <b>Zh. Bilyalova, N. Shazhdekeyeva, K. Uteulyeva, T.Akhmurzina, L. Nurgaliyeva</b><br>A MATHEMATICAL MODEL FOR DECISION-MAKING UNDER UNCERTAINTY HYBRID BASED ON FUZZY LOGIC AND GENERATIVE ARTIFICIAL INTELLIGENCE ..... | 165 |
| <b>A. Ismailova, G. Yessenbayeva, K. Kadyrkulov, R. Moldasheva, A. Amangeldi</b><br>DEVELOPMENT OF A HYBRID DEEP LEARNING MODEL FOR MULTICLASS CLASSIFICATION OF MICROSCOPIC IMAGES OF BACTERIA .....                     | 186 |
| <b>G. Kalman, J. Kultan, A.N. Ismukamova, N.M. Ausilova, Y.V. Makhatova</b><br>A DOMAIN-KNOWLEDGE-BASED MODEL FOR REFERENCE RESOLUTION IN LOW-RESOURCE LANGUAGES.....                                                     | 199 |
| <b>Y. Kamen, Zh. Yessendauletova, L. Fazylova, M. Rakhimzhanova, A.M. Nedzved</b><br>USING NEURAL NETWORKS FOR OBJECTIVE ASSESSMENT OF ATTENTION IN CHILDREN BASED ON EEG DATA .....                                      | 216 |
| <b>A.Ye. Kulakayeva, Ye.A. Bakhtiyarova, G.T. Jakanova, Sh. Nursultan</b><br>COMPARATIVE ANALYSIS OF VARIOUS RADIO WAVE PROPAGATION MODELS FOR MOBILE NETWORK COVERAGE PREDICTION .....                                   | 231 |
| <b>A. Kalykulova, K. Bakirov, A. Shoman, A. Nugumanova</b><br>FORMATION OF A UNIFIED DATA SET FOR ANALYZING THE RELATIONSHIP BETWEEN DOG BREEDS AND DISEASES .....                                                        | 247 |
| <b>D. Kantayeva, A. Seisenbekkyzy, S. Biloshchytska</b><br>COMPARATIVE EVALUATION MODEL BASED ON EXPERT SYSTEMS FOR ABSTRACTIVE SUMMARIZATION OF SCIENTIFIC TEXTS .....                                                   | 269 |
| <b>N. Mordvintsev, Zh. Oralbekova, A. Bakiyeva</b><br>DEEP LEARNING-BASED VEHICLE MAKE AND MODEL RECOGNITION: DATASET CHALLENGES                                                                                          |     |

|                                                                                                                                                                                                       |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| AND INFLUENCING FACTORS IN URBAN TRAFFIC .....                                                                                                                                                        | 269 |
| <b>M.B. Nurpeissova, Sh.K. Aitkazinova, A.M. Abenov, N.S. Donenbayeva</b><br>METHODOLOGY FOR TRANSFORMING SATELLITE COORDINATES INTO A TOPOCENTRIC<br>RECTANGULAR COORDINATE SYSTEM .....             | 315 |
| <b>L. Naizabayeva, G. Turken, R.A. Druba</b><br>EFFECTIVE METHOD FOR OPTIMAL CLUSTERING IN UNSUPERVISED LEARNING .....                                                                                | 328 |
| <b>S. Nurgaliyeva, Zh. Seitkaliyev, A. Sakhipov, M. Kali, R. Hrmo</b><br>YIELD FORECASTING FOR FRUITS AND VEGETABLES USING RANDOM FOREST ON LARGE-<br>SCALE AGROCLIMATIC DATA .....                   | 339 |
| <b>A. Nurgalykov, S. Sapakova, D. Zhalmukhanbetov, G. Nurgazy, A. Myrzabayev</b><br>LANGUAGE ASYMMETRY IN LLMS: THE IMPACT OF NON-ENGLISH DATA ON ACCURACY AND<br>RELIABILITY .....                   | 355 |
| <b>A. Ospanov, P. Alonso-Jordá, A. Zhumadillayeva</b><br>BLOCKCHAIN-ENABLED ERP WAREHOUSE INTEGRATION WITH IOT DIMENSIONERS AND<br>MACHINE LEARNING–OPTIMIZED DIMENSIONAL WEIGHT RECONCILIATION ..... | 371 |
| <b>A.A. Sakhipov, R.B. Seıitbek</b><br>EVENT-DRIVEN MICROSERVICES FOR INCIDENT DETECTION AND RESPONSE IN INTELLIGENT<br>TRAFFIC SYSTEM .....                                                          | 387 |
| <b>A. Tursynova, B. Omarov</b><br>EVOLUTION OF THE EFFECTIVENESS OF CONDITIONAL GAN FOR THE SYNTHESIS OF CT<br>IMAGES OF THE BRAIN IN THE DIAGNOSIS OF STROKE .....                                   | 400 |
| <b>G. Yusupova, K.S. Shadinova, D. Ussipbekova, Zh.Zh. Azhibekova, P. Schmidt</b><br>DETERMINATION OF SOIL PROFILE STRATIFICATION AT 0–200 CM DEPTH USING A MULTILEVEL<br>STACKING MODEL .....        | 414 |

## INFORMATION SECURITY AND COMMUNICATION TECHNOLOGIES

|                                                                                                                                                                                                                               |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>S.A. Adilzhanova, M.Zh. Sakypbekova, L.Sh. Cherikbaeva, G.A. Tyulepberdinova, G.T. Zhubanysheva</b><br>SYSTEMATIC ANALYSIS OF RISK ASSESSMENT METHODS AND MODELS IN INFORMATION<br>SECURITY.....                           | 427 |
| <b>T. K. Zhukabayeva, D.B. Baumuratova, E. Benkhelifa , N.A. Niyetbayeva</b><br>EDGE COMPUTING-BASED TECHNIQUE FOR CONSTRUCTION OF ATTACK DETECTION MEANS<br>IN CYBER-PHYSICAL SYSTEMS OF INDUSTRIAL INTERNET-OF-THINGS ..... | 453 |
| <b>N.E. Karabayev, S.K. Serikbayeva, Y.M. Mardenov, B. Tassuov, M. Fajkus</b><br>DETECTION OF CYBER ATTACKS IN TRANSPORT NETWORKS BASED ON MACHINE LEARNING<br>METHODS .....                                                  | 475 |
| <b>B.A. Kumalakov, A.O. Dargulova</b><br>A HYBRID FRAMEWORK FOR RESUME-JOB MATCHING SYSTEM .....                                                                                                                              | 494 |
| <b>V. Makhatova, B. Dzhugembayeva, A. Gabdulova, L. Nurgaliyeva, A. Abdigaliyeva</b><br>MATHEMATICAL MODEL FOR OPTIMAL SENSOR SELECTION IN SIEM SYSTEMS USING THE<br>ANALYTIC HIERARCHY PROCESS .....                         | 509 |

## МАЗМУНЫ

### ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

|                                                                                                                                                                                                                            |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>А.Б. Жалғас, Е.Н. Калпаков, Б.Е. Амиргалиев</b><br>АҚЫЛДЫ ҚАЛАЛАРДАҒЫ ЛОГИСТИКАНЫ МАШИНАЛЫҚ ОҚЫТУҒА НЕГІЗДЕЛГЕН<br>ОҢТАЙЛАНДЫРУ: АСТАНАНЫҢ ЖАҒДАЙЫН ЗЕРТТЕУ.....                                                        | 11 |
| <b>Л.Қурманғазиева, Ш. Қоданова, М. Уразғалиева, О. Findik, С. Искакова</b><br>ЖАСАНДЫ ИНТЕЛЛЕКТ ПЕН АЙҚЫН ЕМЕС ЛОГИКАНЫ БІРІКТІРУ АРҚЫЛЫ БИЗНЕС-<br>ПРОЦЕСТЕРДІ АВТОМАТТАНДЫРУ ШЕШІМДЕРІН ОҢТАЙЛАНДЫРУ .....              | 26 |
| <b>Е. Майлыбаев, У. Адилбаева, Р. Аманова</b><br>ҰЙЫМДАСТЫРЫЛҒАН ОНЛАЙН САУАЛНАМА АРҚЫЛЫ БІЛІМ БЕРУ ПРОЦЕСІНЕ<br>ҚАТЫСУШЫЛАРДЫҢ ПІКІРЛЕРІН ЖИНАУ ЖӘНЕ НӘТИЖЕЛЕРІН МОДИФИКАЦИЯЛАНҒАН<br>ДЕЛЬФИ ӘДІСІ НЕГІЗІНДЕ ТАЛДАУ ..... | 48 |
| <b>В.А. Такижанов, А.Ж. Ибрагимов, А. Шалахметов</b>                                                                                                                                                                       |    |

|                                                                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| МОДЕЛЬДЕУ НЕГІЗІНДЕ АСТАНАНЫҢ АВТОБУС ЖЕЛІСІНІҢ ТҮРАҚТЫЛЫҒЫН БАҒАЛАУ:<br>КЕЗДЕЙСОҚ ЖӘНЕ МАҚСАТТЫ ІСТЕН ШЫҒУЛАР ЖАҒДАЙЫНДА .....                                                                             | 63 |
| <b>Г. Хамада, М. Хамада</b><br>ЦИФРЛЫҚ БИЗНЕС-ҰЙЫМДАРЫНДА ҚҰРЫЛЫМДАЛМАҒАН ШЕШІМДЕР МЕН ҚҰРЫЛЫМДАЛ-<br>МАҒАН ДЕРЕКТЕРМЕН ЖҰМЫС ІСТЕУ ТИІМДІЛІГІН АРТТЫРУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ<br>МОДЕЛЬДЕРІН ПАЙДАЛАҢУ ..... | 78 |

### АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

|                                                                                                                                                                                                                                 |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>М.Ж. Айтимов, Г.К. Муратова, Ж.К. Бисенбаева, И.М. Бапиев, М. Кассим</b><br>ОНТОЛОГИЯ ЖӘНЕ ІЗДЕУ МЕХАНИЗМДЕРІ АРҚЫЛЫ ҚАЗАҚ ТІЛІНДЕГІ ЭКСТРАКЦИЯЛЫҚ<br>ҚА-ДАҒЫ СЕМАНТИКАЛЫҚ ТОЛЫҚТЫҚ .....                                    | 96  |
| <b>Ә. Әйтiм, Ә. Әуезова</b><br>LLM ӨНДЕУІ ҮШІН NP-ТОЛЫҚ ЕСЕПТЕРДІҢ ЭКЗЕМПЛЯРЛАРЫН ТАБИҒИ ТІЛДЕ КОДТАУ .....                                                                                                                     | 109 |
| <b>О.Н. Ақылбеков, Е.Т. Даулетбек, А.Н. Молдагулова, Г.С. Закария, Д.А. Гура</b><br>ҚАЗАҚСТАННЫҢ АУМАҚТЫҚ ЖОСПАРЛАУЫНДАҒЫ ҮШ ӨЛШЕМДІ КЕҢІСТІКТІК<br>МӘЛІМЕТТЕРДІ ТАЛДАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІ .....                     | 126 |
| <b>С.Ж. Алиаскаров, Р.К. Ускенбаева, А. Разак, А.Б. Касымов, А.М. Анартаева</b><br>АЙМАҚТЫҚ ЖҮЙЕЛЕРДЕГІ ҮЛКЕН ДЕРЕКТЕРДІ ТИІМДІ ТАЛДАУҒА ҚАРАЙ: ГИБРИДТІ<br>АРХИТЕКТУРАНЫ ЕНГІЗУДІҢ ПРАКТИКАЛЫҚ ТҮСІНІКТЕР.....                 | 146 |
| <b>Ж. Билялова, Н. Шаждекеева, К. Өтеулиева, Т. Ахмурзина, Л. Нұрғалиева</b><br>АЙҚЫН ЕМЕС ЛОГИКА ЖӘНЕ ГЕНЕРАТИВТІ ЖАСАНДЫ ИНТЕЛЛЕКТ НЕГІЗІНДЕ БЕЛГІСІЗДІК<br>ЖАҒДАЙЫНДА ШЕШІМ ҚАБЫЛДАУДЫҢ ГИБРИДТІ МАТЕМАТИКАЛЫҚ МОДЕЛІ .....  | 165 |
| <b>А.А. Исмаилова, Г.Р. Есенбаева, К.К. Кадиркулов, Р.Н. Молдашева, А. Амангелді</b><br>РОСКОПИЯЛЫҚ БЕЙНЕЛЕРІН КӨПКЛАССТЫ ЖІКТЕУГЕ АРНАЛҒАН ГИБРИДТІ ТЕРЕҢ ОҚЫТУ<br>МОДЕЛІН ӨЗІРЛЕУ .....                                       | 186 |
| <b>Г. Қалман, К. Ярослав, А.Н. Исмуканова, Н.М. Аусилова, В.Е. Махатова</b><br>ПӘНДІК САЛА БІЛІМ НЕГІЗІНДЕ РЕУСРСТАРЫ АЗ ТІЛДЕРДЕГІ РЕФЕРЕНЦИЯНЫ ШЕШУДІҢ<br>МОДЕЛІ.....                                                         | 199 |
| <b>Е.Г. Кәмен, Ж.Т. Есендаулетова, Л.С. Фазылова, М.Б. Рахимжанова, А.М. Недзьведь</b><br>ЭЭГ ДЕРЕКТЕРІ БОЙЫНША БАЛАЛАРДЫҢ ЗЕЙІНІН ОБЪЕКТИВТІ БАҒАЛАУ ҮШІН НЕЙРОНДЫҚ<br>ЖЕЛІЛЕРДІ ҚОЛДАНУ .....                                 | 216 |
| <b>А.Е. Кулакаева, Е.А. Бахтиярова, Г.Т. Джаканова, Ш. Нурсултан</b><br>ҰЯЛЫ БАЙЛАНЫС ЖЕЛІЛЕРІНІҢ ҚАМТУ АЙМАҒЫН БОЛЖАУҒА АРНАЛҒАН ӨРТҮРЛІ<br>РАДИОТӨЛҚЫН ТАРАЛУ МОДЕЛЬДЕРІНІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ .....                        | 231 |
| <b>А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова</b><br>ИТ ТҰҚЫМДАРЫ МЕН АУРУЛАРЫ АРАСЫНДАҒЫ БАЙЛАНЫСТЫ ТАЛДАУҒА АРНАЛҒАН<br>БЫРЫҒАЙ ДЕРЕКТЕР ЖИЫНТЫҒЫН ҚАЛЫПТАСТЫРУ .....                                                 | 247 |
| <b>Д. Кантаева, А. Сейсенбекқызы, С. Билощницкая</b><br>ҒЫЛЫМИ МӘТІНДЕРДІ АБСТРАКТИВТІ РЕФЕРАТТАУҒА АРНАЛҒАН ЭКСПЕРТТІК ЖҮЙЕЛЕРГЕ<br>НЕГІЗДЕЛГЕН САЛЫСТЫРМАЛЫ БАҒАЛАУ МОДЕЛІ .....                                              | 269 |
| <b>Н. Мордвинцев, Ж. Оралбекова, А. Бакиева</b><br>ТЕРЕҢ ОҚЫТУҒА НЕГІЗДЕЛГЕН КӨЛІК ҚҰРАЛДАРЫНЫҢ МАРКАСЫ МЕН МОДЕЛІН ТАҢУ: ҚА-<br>ЛАЛЫҚ КӨЛІК ҚОЗҒАЛЫСЫНДАҒЫ ДЕРЕКТЕР ЖИЫНТЫҒЫНЫҢ МӘСЕЛЕЛЕРІ ЖӘНЕ ЫҚПАЛ<br>ЕТУШІ ФАКТОРЛАР ..... | 285 |
| <b>М.Б. Нұрпейісова, Ш.Қ. Айтқазина, А.М. Абенев, Н.С. Дөненбаева</b><br>СПУТНИКТІК КООРДИНАТТАРДЫ ТОПОЦЕНТРЛІК ТІК БҰРЫШТЫ КООРДИНАТТАР ЖҮЙЕСІНЕ<br>ТҮРЛЕНДІРУДІҢ ӘДІСТЕМЕСІ .....                                             | 315 |
| <b>Л. Найзабаева, Г. Түркен, Р.А. Друба</b><br>БАҚЫЛАНБАЙТЫН ОҚЫТУДАҒЫ ОҢТАЙЛЫ КЛАСТЕРЛЕУДІҢ ТИІМДІ ӘДІСІ .....                                                                                                                 | 328 |
| <b>С. Нұрғалиева, Ж. Сенткалиев, А. Сахипов, М. Кали, Р. Грмо</b><br>ҮЛКЕН КӨЛЕМДЕГІ АГРОКЛИМАТТЫҚ ДЕРЕКТЕР НЕГІЗІНДЕ RANDOM FOREST ӘДІСІН<br>ПАЙДАЛАНА ОТЫРЫП, ЖЕМІС-КӨКӨНІС ДАҚЫЛДАРЫНЫҢ ӨНІМДІЛІГІН БОЛЖАУ .....             | 339 |
| <b>А. Нұрғалықов, С. Сапакова, Д. Жалмұханбетов, Г. Нұрғазы, А. Мырзабаев</b><br>ҮЛКЕН ТІЛДІК МОДЕЛЬДЕРДЕГІ ТІЛДІК АСИММЕТРИЯ: АҒЫЛШЫН ТІЛІНЕН БАСҚА<br>ДЕРЕКТЕРДІҢ ДӘЛДІК ПЕН СЕНІМДІЛІККЕ ӘСЕРІ .....                         | 355 |
| <b>А. Оспанов, П. Алонсо-Хорда, А. Жұмаділлаева</b>                                                                                                                                                                             |     |

|                                                                                                                                                                                         |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| БЛОКЧЕЙН-ТЕХНОЛОГИЯСЫМЕН ЫҚПАЛДАС ERP ҚОЙМА ЖҮЙЕСІН ІОТ ДИМЕНСИОНЕРЛЕР ЖӘНЕ МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ ОПТИМИЗАЦИЯЛАНҒАН ӨЛШЕМДІ САЛМАҚ ЕСЕПТЕУМЕН ИНТЕГРАЦИЯЛАУ .....                      | 371 |
| <b>А.А. Сахипов, Р.Б. Сейітбек</b><br>ОҚИҒАҒА БАҒДАРЛАНҒАН МИКРОҚЫЗМЕТТЕР ЖҮЙЕСІ АРҚЫЛЫ АҚЫЛДЫ ТРАФИК ЖҮЙЕЛЕРІНДЕ ОҚИҒАЛАРДЫ АНЫҚТАУ ЖӘНЕ ШАРАЛАР ҚОЛДАНУ .....                         | 387 |
| <b>А. Турсынова, Б. Омаров</b><br>ИНСУЛЬТТИ ДИАГНОСТИКАЛАУ КЕЗІНДЕ МИ КТ-БЕЙНЕЛЕРІН СИНТЕЗДЕУГЕ АРНАЛҒАН ШАРТТЫ ҒАН ТИІМДІЛІГІНІҢ ЭВОЛЮЦИЯСЫ .....                                      | 400 |
| <b>Г.М. Юсупова, К.С. Шадинова, Д.И. Усипбекова, Ж.Ж. Ажибекова, Р. Schmidt</b><br>ТОПЫРАҚ ПРОФИЛІНІҢ 0–200 СМ ТЕРЕНДІКТЕГІ СТРАТИФИКАЦИЯСЫН КӨПДЕНГЕЙЛІ СТЕКИНГ-МОДЕЛІМЕН АНЫҚТАУ..... | 414 |

## **АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН**

|                                                                                                                                                                                                                                        |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>С.А. Адилжанова, М.Ж. Сақыпбекова, Л.Ш. Черикбаева, Г.А. Түлепбердинова, Г.Т. Жубанышева</b><br>АҚПАРАТТЫҚ ҚАУІПСІЗДІКТЕ ТӘУЕКЕЛДЕРДІ БАҒАЛАУ ӘДІСТЕРІ МЕН МОДЕЛЬДЕРІН ЖҮЙЕЛІ ТАЛДАУ.....                                           | 427 |
| <b>Т.К. Жукабаева, Д. Б. Баумуратова, Е. Бенкхелифа, Н.А. Ниетбаева</b><br>ШЕКАРАЛЫҚ ЕСЕПТЕУЛЕРДІ ҚОЛДАНА ОТЫРЫП, ЗАТТАРДЫҢ ӨНЕРКӘСІПТІК ИНТЕРНЕТІНІҢ КИБЕРФИЗИКАЛЫҚ ЖҮЙЕЛЕРІНДЕГІ ШАБУЫЛДАРДЫ АНЫҚТАУ ҚҰРАЛДАРЫН ҚҰРУ ӘДІСТЕМЕСІ..... | 453 |
| <b>Н.Е. Қарабаев, С.К. Серикбаева, Е.М. Марденов, Б. Тасуов, М. Файқус</b><br>МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІНЕ НЕГІЗДЕЛГЕН КӨЛІК ЖЕЛІЛЕРІНДЕГІ КИБЕРШАБУЫЛДАРДЫ АНЫҚТАУ .....                                                                | 475 |
| <b>Б.А. Кумалаков, А.О. Даргулова</b><br>ТҮЙІНДЕМЕЛЕР МЕН ВАКАНСИЯЛАРДЫ АВТОМАТТАНДЫРЫЛҒАН СӘЙКЕСТЕНДІРУГЕ НЕГІЗДЕЛГЕН ГИБРИДТІ ҮМІТКЕРЛЕРДІ ІРІКТЕУ ЖҮЙЕСІ .....                                                                      | 494 |
| <b>В. Махатова, Б. Джугембаева, А. Габдулова, Л. Нургалиева, А. Абдигалиева</b><br>ИЕРАРХИЯЛАРДЫ ТАЛДАУ ӘДІСІ НЕГІЗІНДЕ SIEM ЖҮЙЕЛЕРІНДЕ ОҢТАЙЛЫ СЕНСОРДЫ ТАҢДАУДЫҢ МАТЕМАТИКАЛЫҚ МОДЕЛІ .....                                         | 509 |

## **СОДЕРЖАНИЕ ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИ СОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ**

|                                                                                                                                                                                                                         |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>А.Б. Жалғас, Е.Н. Калпаков, Б.Е. Амиргалиев</b><br>ОПТИМИЗАЦИЯ ЛОГИСТИКИ В УМНЫХ ГОРОДАХ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ: НА ПРИМЕРЕ АСТАНЫ .....                                                                       | 11 |
| <b>Л. Курмангазиева, Ш. Коданова, М. Уразғалиева, О. Финдик, С. Исакова</b><br>ИНТЕГРАЦИЯ НЕЧЕТКОЙ ЛОГИКИ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ ОПТИМИЗАЦИИ РЕШЕНИЙ ПО АВТОМАТИЗАЦИИ БИЗНЕС-ПРОЦЕССОВ .....                   | 26 |
| <b>Е. Майлыбаев, У. Адилбаева, Р. Аманова</b><br>СБОР МНЕНИЙ УЧАСТНИКОВ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПОСРЕДСТВОМ ОРГАНИЗОВАННОГО ОНЛАЙН-АНКЕТИРОВАНИЯ И АНАЛИЗ РЕЗУЛЬТАТОВ НА ОСНОВЕ МОДИФИЦИРОВАННОГО МЕТОДА ДЕЛЬФИ ..... | 48 |
| <b>В.А. Такижанов, А.Ж. Ибрагимов, А. Шалахметов</b><br>ОЦЕНКА УСТОЙЧИВОСТИ АВТОБУСНОЙ СЕТИ АСТАНЫ НА ОСНОВЕ МОДЕЛИРОВАНИЯ ПРИ СЛУЧАЙНЫХ И ЦЕЛЕНАПРАВЛЕННЫХ ОТКАЗАХ .....                                               | 63 |
| <b>Г. Хамада, М. Хамада</b><br>ИСПОЛЬЗОВАНИЕ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ РАБОТЫ С НЕСТРУКТУРИРОВАННЫМИ РЕШЕНИЯМИ И НЕСТРУКТУРИРОВАННЫМИ ДАННЫМИ В ЦИФРОВЫХ БИЗНЕС-ОРГАНИЗАЦИЯХ .....         | 78 |

## **ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

|                                                                             |
|-----------------------------------------------------------------------------|
| <b>М.Ж. Айтимов, Г.К. Муратова, Ж.К. Бисенбаева, И.М. Бапиев, М. Кассим</b> |
|-----------------------------------------------------------------------------|

|                                                                                                                                                                                                                                        |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| СЕМАНТИЧЕСКАЯ ПОЛНОТА В КАЗАХСКОЯЗЫЧНОМ EXTRACTIVE QA ЧЕРЕЗ ОНТОЛОГИЮ И RETRIEVAL-МЕХАНИЗМЫ .....                                                                                                                                      | 96  |
| <b>А. Айтим, А. Ауезова</b><br>КОДИРОВАНИЕ ЭКЗЕМПЛЯРОВ NP-ПОЛНЫХ ЗАДАЧ В ЕСТЕСТВЕННОМ ЯЗЫКЕ ДЛЯ ОБРАБОТКИ LLM .....                                                                                                                    | 109 |
| <b>О.Н. Акылбеков, Е.Т. Даулетбек, А.Н. Молдагулова, Г.С. Закария, Д.А. Гура</b><br>МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ТРЁХМЕРНЫХ ПРОСТРАНСТВЕННЫХ ДАННЫХ В ТЕРРИТОРИАЛЬНОМ ПЛАНИРОВАНИИ КАЗАХСТАНА .....                           | 126 |
| <b>С.Ж. Алиаскаров, Р.К. Ускенбаева, А. Разак, А.Б. Касымова, А.М. Анартаева</b><br>НА ПУТИ К ЭФФЕКТИВНОЙ АНАЛИТИКЕ БОЛЬШИХ ДАННЫХ В РЕГИОНАЛЬНЫХ СИСТЕМАХ: ПРАКТИЧЕСКИЕ ВЫВОДЫ ИЗ ВНЕДРЕНИЯ ГИБРИДНОЙ АРХИТЕКТУРЫ .....               | 146 |
| <b>Ж. Билялова, Н. Шаждекеева, К. Утеулиева, Т. Ахмурзина, Л. Нурғалиева</b><br>ГИБРИДНАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ НА ОСНОВЕ НЕЧЕТКОЙ ЛОГИКИ И ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ..... | 165 |
| <b>А.А. Исмаилова, Г.Р. Есенбаева, К.К. Кадиркулов, Р.Н. Молдашева, А. Амангелды</b><br>РАЗРАБОТКА ГИБРИДНОЙ МОДЕЛИ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ МНОГОКЛАССОВОЙ КЛАССИФИКАЦИИ МИКРОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ БАКТЕРИЙ .....                    | 186 |
| <b>Г. Калман, К. Ярослав, А.Н. Исмуханова, Н.М. Аусилова, В.Е. Махатова</b><br>МОДЕЛЬ НА ОСНОВЕ ЗНАНИЙ ПРЕДМЕТНОЙ ОБЛАСТИ ДЛЯ РАЗРЕШЕНИЯ КОРЕФЕРЕНЦИИ В МАЛОРЕСУРСНЫХ ЯЗЫКАХ .....                                                     | 199 |
| <b>Е.Г. Камен, Ж.Т. Есендаулетова, Л.С. Фазылова, М.Б. Рахимжанова, А.М. Недзведь</b><br>ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБЪЕКТИВНОЙ ОЦЕНКИ ВНИМАНИЯ У ДЕТЕЙ ПО ДАННЫМ ЭЭГ .....                                                     | 216 |
| <b>А.Е. Кулакаева, Е.А. Бахтиярова, Г.Т. Джаканова, Ш. Нурсултан</b><br>СРАВНИТЕЛЬНЫЙ АНАЛИЗ РАЗЛИЧНЫХ МОДЕЛЕЙ РАСПРОСТРАНЕНИЯ РАДИОВОЛН ДЛЯ ПРОГНОЗИРОВАНИЯ ПОКРЫТИЯ СЕТЕЙ МОБИЛЬНОЙ СВЯЗИ .....                                      | 231 |
| <b>А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова</b><br>ФОРМИРОВАНИЕ ЕДИНОГО НАБОРА ДАННЫХ ДЛЯ АНАЛИЗА ВЗАИМОСВЯЗИ МЕЖДУ ПОРОДАМИ СОБАК И ЗАБОЛЕВАНИЯМИ .....                                                                     | 247 |
| <b>Д. Кантаева, А. Сейсенбеккызы, С. Билощицкая</b><br>СРАВНИТЕЛЬНАЯ МОДЕЛЬ ОЦЕНКИ НА ОСНОВЕ ЭКСПЕРТНЫХ СИСТЕМ ДЛЯ АБСТРАКТИВНОГО РЕФЕРИРОВАНИЯ НАУЧНЫХ ТЕКСТОВ .....                                                                  | 269 |
| <b>Н. Мордвинцев, Ж. Оралбекова, А. Бакиева</b><br>РАСПОЗНАВАНИЕ МАРКИ И МОДЕЛИ ТРАНСПОРТНЫХ СРЕДСТВ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ: ПРОБЛЕМЫ НАБОРОВ ДАННЫХ И ФАКТОРЫ, ВЛИЯЮЩИЕ НА РАБОТУ СИСТЕМЫ В УСЛОВИЯХ ГОРОДСКОГО ДВИЖЕНИЯ .....  | 285 |
| <b>М.Б. Нурпеисова, Ш.К. Айтказинова, А.М. Абенов, Н.С. Доненбаева</b><br>МЕТОДИКА ПРЕОБРАЗОВАНИЯ СПУТНИКОВЫХ КООРДИНАТ В ТОПОЦЕНТРИЧЕСКУЮ ПРЯМОУГОЛЬНУЮ СИСТЕМУ КООРДИНАТ .....                                                       | 315 |
| <b>Л. Найзабаева, Г. Туркен, Р.А. Друба</b><br>ЭФФЕКТИВНЫЙ МЕТОД ОПТИМАЛЬНОЙ КЛАСТЕРИЗАЦИИ В ОБУЧЕНИИ БЕЗ УЧИТЕЛЯ .....                                                                                                                | 328 |
| <b>С. Нурғалиева, Ж. Сенткалиев, А. Сахипов, М. Кали, Р. Грмо</b><br>ПРОГНОЗИРОВАНИЕ УРОЖАЙНОСТИ ФРУКТОВ И ОВОЩЕЙ С ИСПОЛЬЗОВАНИЕМ RANDOM FOREST НА ОСНОВЕ КРУПНОМАСШТАБНЫХ АГРОКЛИМАТИЧЕСКИХ ДАННЫХ .....                             | 339 |
| <b>А. Нурғалыков, С. Сапакова, Д. Жалмуханбетов, Г. Нургазы, А. Мырзабаев</b><br>ЯЗЫКОВАЯ АСИММЕТРИЯ В БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЯХ: ВЛИЯНИЕ ДАННЫХ НА НЕАНГЛИЙСКИХ ЯЗЫКАХ НА ТОЧНОСТЬ И НАДЕЖНОСТЬ .....                                  | 355 |
| <b>А. Оспанов, П. Алонсо-Хорда, А. Жумадилаева</b><br>ИНТЕГРАЦИЯ СКЛАДСКИХ МОДУЛЕЙ ERP-СИСТЕМ С ИСПОЛЬЗОВАНИЕМ БЛОКЧЕЙНА, ИОТ-ДИМЕНСИОНЕРОВ И ОПТИМИЗИРОВАННОГО МАШИНЫМ ОБУЧЕНИЕМ РАСЧЁТА ГАБАРИТНОГО ВЕСА .....                       | 371 |
| <b>А.А. Сахипов, Р.Б. Сейитбек</b><br>СОБЫТИЯ-ОРИЕНТИРОВАННЫЕ МИКРОСЕРВИСЫ ДЛЯ ОБНАРУЖЕНИЯ И РЕАГИРОВАНИЯ НА ИНЦИДЕНТЫ В ИНТЕЛЛЕКТУАЛЬНЫХ ТРАНСПОРТНЫХ СИСТЕМАХ .....                                                                  | 387 |
| <b>А. Тұрсынова, Б. Омаров</b><br>ЭВОЛЮЦИЯ ЭФФЕКТИВНОСТИ УСЛОВНОГО GAN ДЛЯ СИНТЕЗА КТ-ИЗОБРАЖЕНИЙ ГОЛОВНО-                                                                                                                             |     |

|                                                                                                                             |     |
|-----------------------------------------------------------------------------------------------------------------------------|-----|
| ГО МОЗГА ПРИ ДИАГНОСТИКЕ ИНСУЛЬТА .....                                                                                     | 400 |
| <b>Г.М. Юсупова, К.С. Шадинова, Д.И. Усипбекова, Ж.Ж. Ажибекова, П. Шмидт</b>                                               |     |
| ОПРЕДЕЛЕНИЕ СТРАТИФИКАЦИИ ПОЧВЕННОГО ПРОФИЛЯ НА ГЛУБИНЕ 0–200 СМ С<br>ИСПОЛЬЗОВАНИЕМ МОДЕЛИ МНОГОУРОВНЕВОГО НАЛОЖЕНИЯ ..... | 414 |

## **ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ**

|                                                                                                                                                                                                                                   |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>С.А. Адилжанова, М.Ж. Сакыпбекова, Л.Ш. Черикбаева, Г.А. Тюлепбердинова, Г.Т. Жубанышева</b><br>СИСТЕМАТИЧЕСКИЙ АНАЛИЗ МЕТОДОВ И МОДЕЛЕЙ ОЦЕНКИ РИСКОВ ИНФОРМАЦИОННОЙ<br>БЕЗОПАСНОСТИ .....                                    | 427 |
| <b>Т.К. Жукабаева, Д.Б. Баумуратова, Е. Бенкхелифа, Н.А. Ниембаева</b><br>МЕТОДИКА ПОСТРОЕНИЯ СРЕДСТВ ОБНАРУЖЕНИЯ АТАК В КИБЕРФИЗИЧЕСКИХ СИСТЕМАХ<br>ПРОМЫШЛЕННОГО ИНТЕРНЕТА ВЕЩЕЙ С ИСПОЛЬЗОВАНИЕМ ГРАНИЧНЫХ<br>ВЫЧИСЛЕНИЙ ..... | 453 |
| <b>Н.Е. Карабаев, С.К. Серикбаева, Е.М. Марденов, Б. Тасуов, М. Файкус</b><br>ОБНАРУЖЕНИЕ КИБЕРАТАК В ТРАНСПОРТНЫХ СЕТЯХ НА ОСНОВЕ МЕТОДОВ МАШИННОГО<br>ОБУЧЕНИЯ .....                                                            | 475 |
| <b>Б.А. Кумалаков, А.О. Даргулова</b><br>ГИБРИДНЫЙ ПОДХОД К АВТОМАТИЗИРОВАННОМУ ПОДБОРУ КАНДИДАТОВ НА ОСНОВЕ<br>СОПОСТАВЛЕНИЯ РЕЗЮМЕ И ВАКАНСИЙ .....                                                                             | 494 |
| <b>В. Махатова, Б. Джугембаева, А. Габдулова, Л. Нургалиева, А. Абдигалиева</b><br>МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ВЫБОРА ОПТИМАЛЬНОГО СЕНСОРА В SIEM-СИСТЕМАХ<br>СРЕДСТВАМИ МЕТОДА АНАЛИЗА ИЕРАРХИЙ .....                                  | 509 |

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Vol. 7. Is.1. Number 25 (2026). Pp. 269–284

Journal homepage: <https://journal.iitu.edu.kz><https://doi.org/10.54309/IJICT.2026.25.1.016>

## COMPARATIVE EVALUATION MODEL BASED ON EXPERT SYSTEMS FOR ABSTRACTIVE SUMMARIZATION OF SCIENTIFIC TEXTS

***D. Kantayeva\*, A. Seisenbekkyzy, S. Biloshchytska***

Astana IT University, Astana, Kazakhstan.

E-mail: [d.kantayeva@astanait.edu.kz](mailto:d.kantayeva@astanait.edu.kz)

**Dina Kantayeva** — Master of Technical Science, PhD student, Astana IT University, Astana, Kazakhstan

E-mail: [d.kantayeva@astanait.edu.kz](mailto:d.kantayeva@astanait.edu.kz), <https://orcid.org/0000-0002-4097-5078>;

**Arailym Seisenbekkyzy** — Master of Technical Science, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0009-0005-4907-8639>;

**Svitlana Biloshchytska** — Doctor of Technical Sciences, Associate Professor, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0000-0002-0856-5474>.

© D. Kantayeva, A. Seisenbekkyzy, S. Biloshchytska

**Abstract.** The growing number of scientific publications requires effective methods of summarizing scientific texts. Traditional extractive approaches do not allow preserving the logical structure, while abstract generalization methods are difficult to process highly structured documents. In this study, a comparative assessment model is proposed to improve the quality of abstract generalization of scientific texts. This study presents a comparative evaluation of BART-base and BART-large models. A corpus of 17,038 documents from arXiv and PubMed was used for fine-tuning. Evaluation with ROUGE and BLEU metrics shows that BART-large outperforms BART-base, achieving ROUGE-1 = 0.2224, ROUGE-2 = 0.0695, and BLEU = 0.29. In contrast, BART-base reached ROUGE-1 = 0.1932 and BLEU = 0.25. BART-large demonstrated better phrase-level coherence.

**Keywords:** Abstractive summarization, BART, ROUGE, NLP, text generation, model evaluation

**For citation:** Kantayeva D., Seisenbekkyzy A., Biloshchytska S. (2026). Comparative evaluation model based on expert systems for abstractive summarization of scientific texts // International journal of information and communication technologies. Vol. 7. No. 26. Pp. 269–284. <https://doi.org/10.54309/IJICT.2026.25.1.016>. (In Eng.).

**Conflict of interest:** The authors declare that there is no conflict of interest.



# ҒЫЛЫМИ МӘТІНДЕРДІ АБСТРАКТИВТІ РЕФЕРАТТАУҒА АРНАЛҒАН ЭКСПЕРТТІК ЖҮЙЕЛЕРГЕ НЕГІЗДЕЛГЕН САЛЫСТЫРМАЛЫ БАҒА- ЛАУ МОДЕЛІ

*Д. Кантаева, А. Сейсенбекқызы, С. Билощицкая\**

Astana IT University, Астана, Қазақстан.

E-mail: d.kantayeva@astanait.edu.kz

**Кантаева Дина** — техника ғылымдарының магистрі, докторант, Astana IT University, Астана, Қазақстан

E-mail: d.kantayeva@astanait.edu.kz, <https://orcid.org/0000-0002-4097-5078>;

**Сейсенбекқызы Арайлым** — техника ғылымдарының магистрі, Astana IT University, Астана, Қазақстан

<https://orcid.org/0009-0005-4907-8639>;

**Билощицкая Светлана** — техника ғылымдарының докторы, қауымдастырылған профессор, Astana IT University, Астана, Қазақстан

<https://orcid.org/0000-0002-0856-5474>.

© Д. Кантаева, А. Сейсенбекқызы, С. Билощицкая

**Аннотация.** Ғылыми жарияланымдар санының артуы ғылыми мәтіндерді рефераттаудың тиімді әдістерін қолдануды қажет етеді. Дәстүрлі экстрактивті тәсілдер мәтіннің логикалық құрылымын сақтауға мүмкіндік бермейді, ал абстрактивті жалпылау әдістері құрылымы күрделі құжаттарды өңдеу барысында қиындықтарға тап болады. Осы зерттеуде ғылыми мәтіндерді абстрактивті жалпылау сапасын арттыруға бағытталған салыстырмалы бағалау моделі ұсынылады. Зерттеуде BART-base және BART-large модельдеріне салыстырмалы бағалау жүргізілді. Модельдерді қосымша оқыту үшін arXiv және PubMed дерекқорларынан алынған 17 038 құжаттан тұратын корпус пайдаланылды. ROUGE және BLEU метрикалары арқылы жүргізілген бағалау нәтижелері BART-large моделінің BART-base моделінен жоғары нәтиже көрсеткенін анықтады: ROUGE-1 = 0,2224, ROUGE-2 = 0,0695 және BLEU = 0,29. Ал BART-base моделі ROUGE-1 = 0,1932 және BLEU = 0,25 нәтижелеріне қол жеткізді. Сонымен қатар, BART-large моделі сөз тіркестері деңгейінде жоғарырақ мәтіндік үйлесімділік көрсетті.

**Түйінді сөздер:** абстрактивті рефераттау, BART, ROUGE, NLP, мәтін генерациясы, модельдерді бағалау

**Дәйексөздер үшін:** Кантаева Д., Сейсенбекқызы А., Билощицкая С. (2026). Ғылыми мәтіндерді абстрактивті рефераттауға арналған эксперттік жүйелерге негізделген салыстырмалы бағалау моделі // Халықаралық ақпараттық және коммуникациялық технологиялар журналы. Т. 7. No. 26. Б. 269–284. <https://doi.org/10.54309/IJICT.2026.25.1.016>. (Ағыл. тіл.).

**Мүдделер қақтығысы:** Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдейді.

# СРАВНИТЕЛЬНАЯ МОДЕЛЬ ОЦЕНКИ НА ОСНОВЕ ЭКСПЕРТНЫХ СИСТЕМ ДЛЯ АБСТРАКТИВНОГО РЕФЕРИРОВАНИЯ НАУЧНЫХ ТЕКСТОВ

*Д. Кантаева, А. Сейсенбеккызы, С. Билощицкая\**

Astana IT University, Астана, Казахстан.

E-mail: d.kantayeva@astanait.edu.kz

**Кантаева Дина** — магистр технических наук, докторант, Astana IT University, Астана, Казахстан

E-mail: d.kantayeva@astanait.edu.kz, <https://orcid.org/0000-0002-4097-5078>;

**Сейсенбеккызы Арайлым** — магистр технических наук, Astana IT University, Астана, Казахстан

<https://orcid.org/0009-0005-4907-8639>;

**Билощицкая Светлана** — доктор технических наук, доцент, Astana IT University, Астана, Казахстан

<https://orcid.org/0000-0002-0856-5474>.

© Д. Кантаева, А. Сейсенбеккызы, С. Билощицкая

**Аннотация.** Рост количества научных публикаций требует применения эффективных методов реферирования научных текстов. Традиционные экстрактивные подходы не позволяют сохранять логическую структуру текста, в то время как методы абстрактного обобщения испытывают трудности при обработке документов со сложной структурой. В настоящем исследовании предложена сравнительная модель оценки, направленная на повышение качества абстрактного обобщения научных текстов. В исследовании представлена сравнительная оценка моделей BART-base и BART-large. Для дообучения моделей использовался корпус, содержащий 17 038 документов из баз данных arXiv и PubMed. Оценка с использованием метрик ROUGE и BLEU показала, что модель BART-large превосходит BART-base, достигнув значений ROUGE-1 = 0,2224, ROUGE-2 = 0,0695 и BLEU = 0,29. В свою очередь, модель BART-base продемонстрировала значения ROUGE-1 = 0,1932 и BLEU = 0,25. Модель BART-large также продемонстрировала более высокую связность на уровне фраз.

**Ключевые слова:** абстрактное реферирование, BART, ROUGE, NLP, генерация текста, оценка моделей

**Для цитирования:** Кантаева Д., Сейсенбеккызы А., Билощицкая С. (2026). Сравнительная модель оценки на основе экспертных систем для абстрактного реферирования научных текстов // Международный журнал информационных и коммуникационных технологий. Т. 7. No. 25. Стр. 269–284. <https://doi.org/10.54309/IJICT.2026.25.1.016>. (На англ.).

**Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

## Introduction.

The rapid growth of digitalization and the introduction of the Internet of Things and intelligent systems in areas such as fitness (Hladkyi et al., 2021), healthcare (Ananthula et al., 2024) and smart offices (Neftissov et al., 2023) have significantly increased the volume of text data. This highlights the need for automated approaches to information processing and synthesis. Existing large-scale datasets for abstracting usually cover small texts. The average length of articles in the CNN/Daily Mail dataset is about 600 words (Hermann et al., 2015). Some studies confirm the effectiveness of hybrid models combining statistical and semantic methods. (Larsen et al., 2010; Babenko et al., 2024; Biloshchytska et al., 2025). In this study, we develop a model for high-quality generalization of long texts.

The relevance of developing a method of abstract abstracting of scientific texts in English is due to the need to optimize scientific information. While research is mainly focused on extractive approaches (Nallapati et al., 2016; Vanderwende et al., 2007), modern neural network models such as seq2seq have shown promising results (Nallapati et al., 2016; See, et al., 2017; Li et al., 2017). However, when processing long and structured texts, they face difficulties in maintaining context (Shao et al., 2017).

The research hypothesis. The use of an integrated approach combining artificial intelligence and deep learning technologies will make it possible to develop an effective method for abstracting scientific texts in English. It is expected that this method will significantly improve the process of extracting key information from scientific texts.

Goals and objectives of the study. The purpose of the research is to develop and test a method of abstract abstracting of scientific texts in English, which provides automatic creation of concise and informative summaries reflecting the key content of the source documents.

## Materials and methods.

The focus is on methods using neural networks and statistical approaches. There are extractive and abstract text-based summarization approaches. The output of the first approach is a summary generated from key sentences of the input text, the second generates new sentences that are not present in the input text (Jones et al., 2007).

Extractive statistical methods, such as frequency analysis of the Moon (Luhn, 1958), were effective for short texts. Later improvements with vector models (Yu et al., 2020) increased their accuracy, but they remain limited for long scientific texts.

Graph methods, such as LexRank (Erkan et al., 2004), have improved the quality of abstracting by considering the relationships between sentences, but are limited when working with scientific texts (Gunes, 2004). Later, Plaza (2008) proposed concept graphs with ontologies, which increased efficiency for specialized fields (Plaza et al., 2008). However, such methods require manual tuning that limits scalability. The summarized text generated by extractive methods is not carefully filtered and often contain unnecessary information. It may not catch overall value of the text if key information covers several sentences.

Abstract methods form new sentences, which makes summaries more informative

and coherent. The first step was the Pointer-Generator model (See et al., 2017), followed by T5 (Raffel et al., 2020) and BART (Lewis et al., 2020), which showed high generation quality, but sometimes created redundant texts. The use of attention mechanisms and hybrid architectures (Song et al., 2019; Bahdanau et al., 2014). improved the processing of semantics, and multilingual models (ruT5, mBART) confirmed their effectiveness for different languages. In general, abstract approaches are superior to extractive ones. This method handles more details in the text, but they require a lot of data for training.

The quality and relevance of the data directly determine the effectiveness of the model (Lewis et al., 2020); therefore, special attention is paid to scientific texts in English. For the study, several datasets were considered, analyzed according to the criteria of structure, relevance and suitability for abstract abstracting.

Table 1. Comparative analysis of datasets

| Dataset                              | Source            | Content                                    | Advantages                             | Limitations                                          | Suitability |
|--------------------------------------|-------------------|--------------------------------------------|----------------------------------------|------------------------------------------------------|-------------|
| Topic Modeling for Research Articles | Kaggle            | Titles, Abstracts, Keywords                | Structured metadata, thematic analysis | no full text                                         | Low         |
| Latest Research Articles             | Kaggle            | Titles, Full texts                         | Current publications                   | No scientific texts, no standard scientific sections | Average     |
| Wikipedia Promotional Articles       | Wikipedia, Kaggle | Promotional articles, Metadata, Full texts | The possibility of stylistic analysis  | Non-scientific texts                                 | Low         |
| COVID-19 Vaccine Articles            | Kaggle            | Titles, Full texts                         | Full content                           | Limited to a narrow topic                            | Average     |
| arXiv Dataset                        | arXiv, Kaggle     | Titles, Abstracts, Links to PDFs, Metadata | Extensive storage, scalability         | Text extraction is required                          | High        |
| The arXive Dataset                   | arXiv, Mendeley   | Full texts Abstracts, Sections, Metadata   | Full content, clear structure          | -                                                    | Very high   |

As shown in Table 1, most datasets (Kaggle, Wikipedia) are limited by the lack of full texts or narrow topics, whereas «The arXiv Dataset» (Mendeley Data, 2021), created on the basis of arXiv and PubMed, with full texts, annotations and section structure, proved to be the most suitable for abstract abstracting due to its scale and high-quality preparation data.

Table 2. Description of the cleaned datasets

|       | arXive      | PubMed          |             |                 |
|-------|-------------|-----------------|-------------|-----------------|
|       | Text length | Abstract length | Text length | Abstract length |
| Count | 17 038      |                 |             |                 |
| Mean  | 263.44      | 11.75           | 88.89       | 6.85            |
| Std   | 102.57      | 2.13            | 60.56       | 2.88            |
| Min   | 100.00      | 10.00           | 3.00        | 1.00            |
| 25%   | 179.00      | 10.00           | 47.00       | 5.00            |
| 50%   | 252.00      | 11.00           | 77.00       | 7.00            |
| 75%   | 338.00      | 13.00           | 113.00      | 9.00            |
| Max   | 500.00      | 20.00           | 1887.00     | 23.00           |

Combining and cleaning data from these sources resulted in the creation of a

final set of 17,038 records (see Table 2).

The data set includes 17,038 records and 35 columns containing identifiers, full texts, annotations, and metrics, as shown in Figure 1. For quality control, a selective validation of the astro-ph0202198 article was performed, confirming a match with the original on the arXiv.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17038 entries, 0 to 17037
Data columns (total 35 columns):
 #   Column                                Non-Null Count  Dtype
---  --
 0   article_id                            17038 non-null   object
 1   article_text                           17038 non-null   object
 2   abstract_text                           17038 non-null   object
 3   section_names                           17038 non-null   object
 4   sections                                17038 non-null   object
 5   len_text                               17038 non-null   int64
 6   len_abstract                           17038 non-null   int64
 7   sent_ext                                17038 non-null   float64
 8   word_ext                                17038 non-null   float64
 9   word_inf                                17038 non-null   float64
10  best_rouge1_wdh                          17038 non-null   float64
11  best_rouge2_wdh                          17038 non-null   float64
12  best_summary_wdh                          17038 non-null   object
13  best_summary_vns                          17038 non-null   object
14  best_rouge1_vns                          17038 non-null   float64
15  best_rouge2_vns                          17038 non-null   float64
16  voted_summary                           17038 non-null   object
17  voted_method                             17038 non-null   object
18  voted_r1                                 17038 non-null   float64
19  voted_r2                                 17038 non-null   float64
20  best_summary_vns_initgreed               17038 non-null   object
21  best_rouge1_vns_initgreed               17038 non-null   float64
22  best_rouge2_vns_initgreed               17038 non-null   float64
23  len_sum_vns                             17038 non-null   int64
24  len_sum_wdh                             17038 non-null   int64
25  len_sum_vot                             17038 non-null   int64
26  len_sum_vns_init_greed                  17038 non-null   int64
27  best_summary_gen                         17038 non-null   object
28  best_rouge1_gen                         17038 non-null   float64
29  best_rouge2_gen                         17038 non-null   float64
30  len_sum_gen                             17038 non-null   int64
31  best_summary_gen_greedinit              17038 non-null   object
32  best_rouge1_gen_greedinit               17038 non-null   float64
33  best_rouge2_gen_greedinit               17038 non-null   float64
34  len_sum_gen_greedinit                   17038 non-null   int64
dtypes: float64(15), int64(8), object(12)
memory usage: 4.5+ MB
```

Fig. 1. Dataset structure.

After a data cleaning the texts were reduced to lowercase, excluded HTML and URLs, normalized and tokenized; signs of length and number of sentences were created.

Data analysis revealed that articles contain an average of 200–250 words, annotations – 10–12 sentences, accounting for 15–25 % of the text (see Figures 2-4).

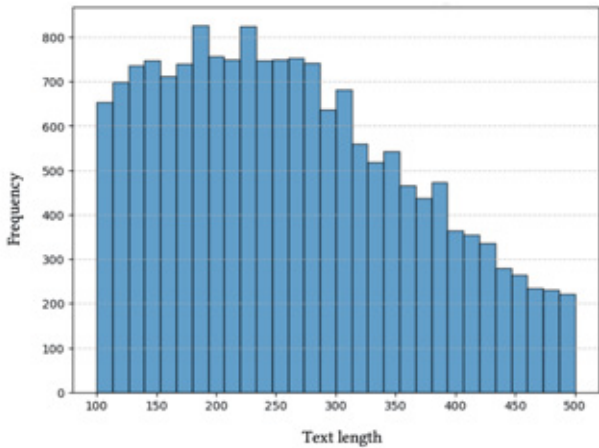


Fig. 2. Text length distribution.

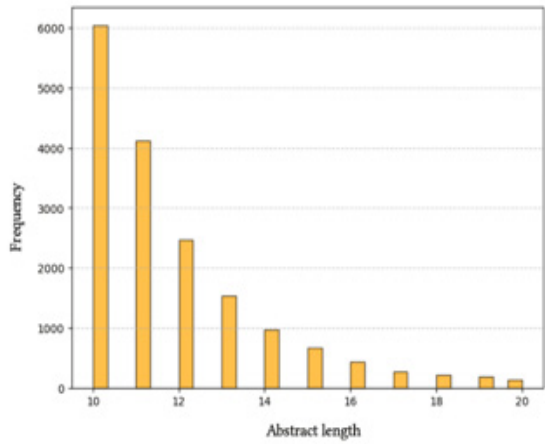


Fig. 3. Abstract length distribution.

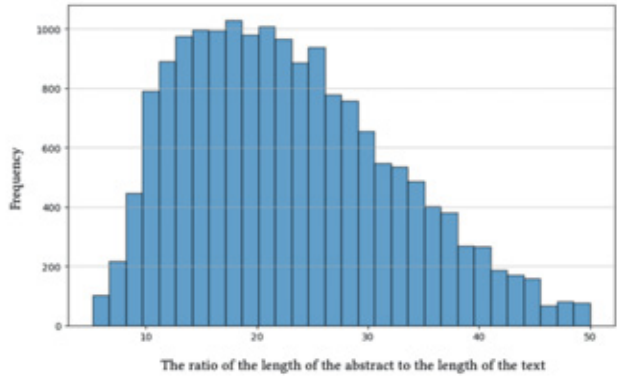


Fig. 4. The ratio of the length of the abstract to the length of the text.

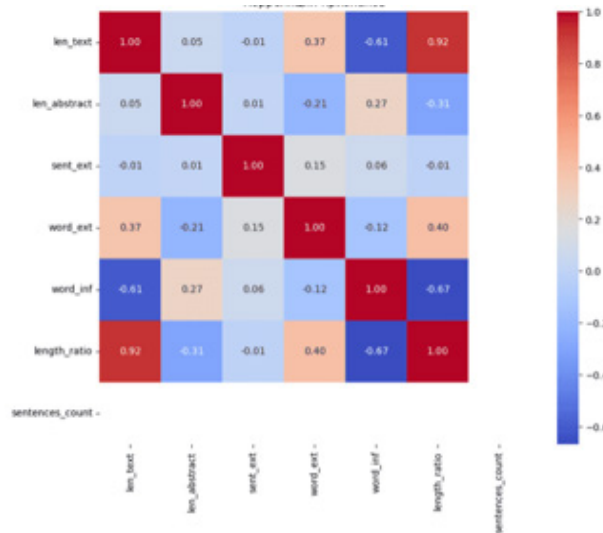


Fig. 5. Correlation of features.

As shown in Figure 5 long texts contain fewer information-rich words.

**Data preparation.** To implement the task of automatic abstracting of scientific texts, the main columns were selected: the full texts of articles and their annotations. The data was then divided into training, validation, and test sets in the proportions of 80 %, 10 %, and 10 %. The total number of examples is shown in Table 3.

Table 3. Data separation

| Dataset    | Number of examples |
|------------|--------------------|
| Train      | 17 038             |
| Validation | 1 704              |
| Test       | 1 704              |

Abstractive summarization is one of the key tasks of natural language processing and is aimed at generating annotations that convey the main ideas of the source text using new formulations. The approach is based on maximizing the conditional probability  $P(Y|X)$ , where  $X$  is the source text and  $Y$  is the abstract, and is implemented using modern deep learning architectures. The most widely used models are Seq2Seq with the attention mechanism, Transformer, BART, T5, and GPT.

A model of Sequence-to-Sequence (Seq2Seq) is used to convert the input sequence  $X = \{x_1, x_2, \dots, x_n\}$  in the output sequence  $Y = \{y_1, y_2, \dots, y_m\}$  even when the difference in their lengths. The architecture includes an encoder and a decoder interacting through a hidden representation.

The encoder generates hidden states:

$$h_t = f(h_t^{-1}, x_t) \quad (1)$$

where  $f$  is implemented via GRU or LSTM. The final state of  $h_n$  is passed to the decoder, which generates output tokens step by step, updating the state using the formula:

$$s_t = g(s_t^{-1}, y_t^{-1}, h), \quad (2)$$

where  $g$  is a nonlinear function like the function  $f$ . The probability of predicting the entire output sequence is defined as:

$$P(Y \vee X) = \prod_{t=1}^m P(y_t \vee y_1, \dots, y_{t-1}, h), \quad (3)$$

To improve the quality of predictions, an attention mechanism is used, which allows the decoder to focus on the most significant parts of the input sequence when generating each token. The significance weight of the input token  $\alpha_{ti}$  is calculated as:

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{j=1}^n \exp(e_{t,i})}, \quad (4)$$

where  $e_{ti}$  = score ( $s_t, h_i$ ) is a function estimating the importance of the  $i$ -th input token for generating the  $t$ -th output token (Zhang et al., 2024).

The main advantage of BART is the combination of a bidirectional encoder and an autoregressive decoder, which provides context understanding and generation of new formulations. The model is trained on the task of restoring distorted text  $X$  from its modified version  $X_{\sim}$ , which uses word deletion, sentence rearrangement, and noise addition.

At the encoding stage, a hidden representation is formed.:

$$H = \text{Encoder}(X), \quad (5)$$

where  $H = \{H_1, h_2, \dots, h_n\}$  represents the hidden states of the encoder. At the decoding stage, the model sequentially generates the output sequence  $Y = \{y_1, y_2, \dots, y_m\}$ . The probability of the output sequence is calculated as follows:

$$P(Y \vee X) = \prod_{t=1}^m P(y_t \vee y_1, \dots, y_{t-1}, h), \quad (6)$$

where  $P(y_t | y_1, \dots, y_{t-1}, H)$  is the probability of the next token based on previously generated tokens and the hidden representation of the encoder (Tsvetkov et al., 2008).

Generative Pre-trained Transformer (GPT) is pre-trained on large bodies of texts, which ensures its effectiveness in language processing tasks, including abstract abstracting. The basis is an autoregressive mechanism, where each  $y_t$  token is predicted based on the previous  $\{y_1, y_2, \dots, y_{t-1}\}$ :

$$P(Y) = \prod_{t=1}^m P(y_t \vee y_1, \dots, y_{t-1}), \quad (7)$$

The model is based on a transformer decoder and an attentional mechanism.:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (8)$$

where  $Q$ ,  $K$ , and  $V$  are matrices of queries, keys, and values, and  $d_k$  is the dimension of the keys. This formula provides the calculation of weighted values, which helps the model to better understand the context of the text (Zhong et al., 2024).

The BART model was chosen for the study, as it combines bidirectional encoding

and auto-correction, processes context efficiently, is resistant to data skips, and shows stable results with limited training samples.

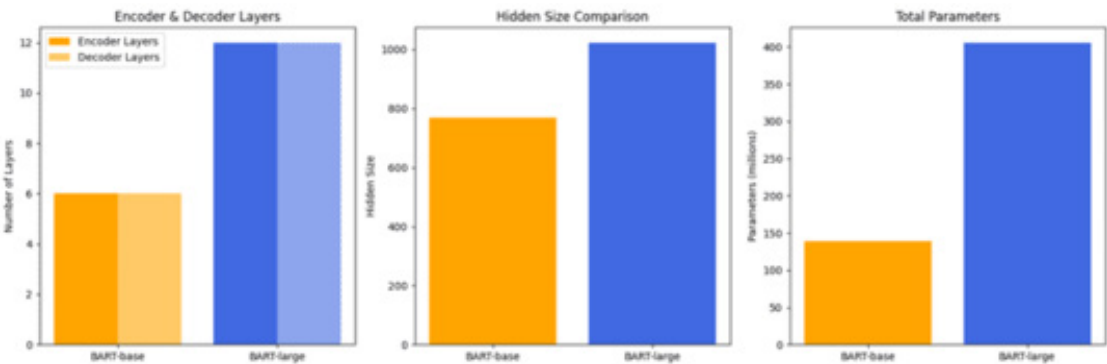


Fig. 6. Comparison of BART-base and BART-large architectures.

Depending on the computing capabilities and the requirements for the quality of text generation, BART is presented in two versions (see Figure 6): BART-base and BART-large. Each of them has its own advantages and limitations, which determines their application in various scenarios of summarization.

Table 4. Comparison of BART models

|                | BART-base                                                      | BART-large                                        |
|----------------|----------------------------------------------------------------|---------------------------------------------------|
| Specifications | 12 layers, 768 hidden, 16 heads, 139M parameters               | 24 layers, 1024 hidden, 16 heads, 406M parameters |
| Features       | Compact version optimized for limited resources                | Deep architecture, improved context understanding |
| Advantages     | Balance between quality and cost                               | Higher accuracy, better connection detection      |
| Application    | Text generation, abstract abstracting, Machine translation, QA | Long and complex texts, scientific articles       |

Due to the fixed size of the input context of the model, the maximum length of sequences was set: 512 tokens for article introductions and 150 tokens for annotations. This limitation is since BART supports up to 1024 tokens, but 512 are enough for introductions of scientific articles, and the average length of annotations does not exceed 150 words, as shown in Table 5.

Table 5. Limiting the length of sequences

| Length      | Value      | Justification                                                                                |
|-------------|------------|----------------------------------------------------------------------------------------------|
| Input data  | 512 tokens | BART supports up to 1024 tokens, but 512 is enough for introductions of scientific articles. |
| Output data | 150 tokens | The average length of scientific abstracts does not exceed 150 words.                        |

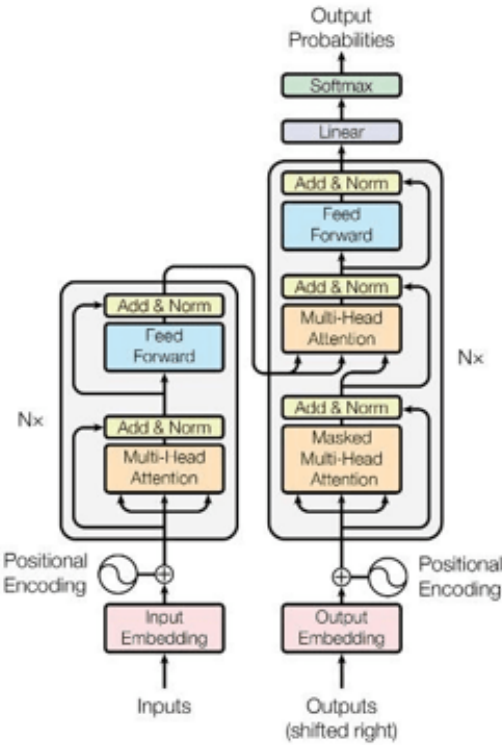


Fig. 7. Encoder-Decoder architecture (Source: (Brown et al., 2020)).

BART-base and BART-large allow you to paraphrase the source text while maintaining its structural integrity. BART-base provides a balance between quality and computational costs, which makes it convenient for practical implementation, while BART-large demonstrates maximum accuracy through deep context processing.

To assess the quality of automatic abstracting, metrics are used that compare generated annotations with reference ones. The main one is ROUGE, a standard in summation that measures match at the level of  $n$ -grams and word sequences. Additionally, the BLEU metric was used.

BLEU is one of the main metrics for evaluating the quality of automatically generated text, originally used in machine translation (Brown et al., 2020). It compares the generated text with the reference text based on the coincidence of word sequences ( $n$ -grams).

To avoid a situation where the model creates texts that are too short and have a high overlap, a brevity penalty (BP) is introduced:

$$BP = \begin{cases} 1, & \text{if } c > r \\ e^{(1-\frac{r}{c})}, & \text{if } c \leq r \end{cases}$$

where  $c$  is the length of the generated text (hypothesis), and  $r$  is the length of the reference text. If the generated text is shorter than the reference text, the penalty reduces the final BLEU score. The final formula for calculating BLEU includes a geometric mean of n-gram accuracy and a brevity penalty:

$$BLEU = BP * \exp \left( \sum_{n=1}^N \omega_n \log P_n \right)$$

where  $P_n$  is the accuracy of n-gram matches,  $\omega_n$  are the weights for different n-grams (usually assumed to be equal), and  $N$  is the maximum length of n-grams.

BLEU is focused on precision, whereas ROUGE is primarily concerned with completeness. This means that BLEU better evaluates the correspondence of the generated text to the reference in terms of the accuracy of individual phrases, while ROUGE measures how fully the source text was reflected in the final annotation.

*Recall-oriented understudy for gisting evaluation (ROUGE)*

ROUGE metrics, which are the standard in automatic abstracting tasks, were used to evaluate the quality of annotations (Lewis et al., 2019).

ROUGE-1 measures word matches (unigrams):

$$ROUGE - 1 = \frac{\sum_{\omega \in W} \min(count_{gen}(\omega), count_{ref}(\omega))}{\sum_{\omega \in W} count_{ref}(\omega)}$$

where  $count_{gen}(\omega)$  — is the number of occurrences of the word  $\omega$  in the generated text,  $count_{ref}(\omega)$  — is the number of occurrences of a word in the reference annotation, and  $W$  is the set of all words in the reference text.

The ROUGE-2 metric evaluates the coincidence of pairs of consecutive words (bigrams), which makes it possible to identify the preservation of local context and sentence coherence:

$$ROUGE - 2 = \frac{\sum_{b \in B} \min(count_{gen}(b), count_{ref}(b))}{\sum_{b \in B} count_{ref}(b)}$$

where  $B$  is the set of all bigrams in the reference text.

The ROUGE-L metric is based on the search for the Longest Common Sequence (LCS), which allows you to evaluate the grammatical and structural consistency between the generated and reference text.

This metric considers the word order and is calculated as follows:

$$ROUGE - L = \frac{LCS(gen,ref)}{length(ref)}$$

where  $LCS(gen,ref)$  — is the length of the largest common subsequence between the generated text and the reference annotation.

ROUGE metrics are limited by lexical overlap and do not consider semantics, which may underestimate the quality of annotations. In the future, it is worth using BERTScore and METEOR metrics that consider semantic connections.

**Results and discussion.**

To assess the quality of the annotations, the pre-trained and pre-trained versions of BART-base and BART-large were compared on the corpus of scientific articles using the ROUGE metric.

Table 6. Results of pre-trained models on the test set

|   | Model      | ROUGE-1 | ROUGE-2 | ROUGE-L | ROUGE-Lsum |
|---|------------|---------|---------|---------|------------|
| 0 | BART-base  | 0.1803  | 0.0417  | 0.0960  | 0.0961     |
| 1 | BART-large | 0.1881  | 0.0507  | 0.1088  | 0.1088     |

At the first stage of testing, the pre-trained versions of BART-base and BART-large were evaluated without additional training based on scientific articles. The results presented in Table 6 demonstrate the initial performance of the models.

Despite the more complex architecture of BART-large, its performance is only slightly higher than that of BART-base. This indicates that without additional training in a specialized corpus, models are not able to effectively consider the features of a scientific text. The main reason for this limited effectiveness is that pre-trained transformational models, including BART, are most often trained on the corpus of news articles, social networks, and other sources focused on generalized language rather than an academic style of presentation.

To improve the quality of summarization of scientific texts, both models, BART-base and BART-large, were further trained on a specialized corpus of articles. The main stages included marking up and preprocessing data, choosing a learning strategy, determining optimal hyperparameters, monitoring the learning process, and subsequent quality assessment.

The key parameters are the batch size, the learning rate, the optimization strategy, the number of epochs, and the schedule for changing the learning rate.

- 1. Batch size = 8 – the value is selected based on the balance between learning efficiency and available memory.
- 2. Learning rate = 3e-5 – the learning rate was set based on empirical observations and recommendations for setting up BART.
- 3. AdamW optimizer – was chosen as the main algorithm for updating weights because it works efficiently with large models by enabling weight decay.



4. Weight decay = 0.01 – this parameter prevents overfitting by limiting the amount of weight updates.

5. Number of epochs = 3 and 5 – the model was tested after 3 and 5 epochs, since preliminary experiments showed that with prolonged training, the model ceases to improve significantly.

In addition, a scheduler was used, which allowed the model to gradually reduce the learning rate, stabilizing the process of updating weights.

The results of BART-base retraining are presented in Table 7, which shows the ROUGE metrics after epochs 3 and 5.

Table 7. Results of the retrained BART-base model

|   | Epochs | ROUGE-1 | ROUGE-2 | ROUGE-L | ROUGE-Lsum |
|---|--------|---------|---------|---------|------------|
| 0 | 3      | 0.1932  | 0.0598  | 0.1234  | 0.1235     |
| 1 | 5      | 0.1932  | 0.0598  | 0.1234  | 0.1235     |

As can be seen from the table, the results after 5 epochs did not change compared to 3 epochs, which indicates that the model has reached its learning limit on this corpus.

Further training of the BART-large model showed a more significant increase in all metrics, which is reflected in Table 8.

Table 8. Results of the retrained BART-large model

|   | Epochs | ROUGE-1 | ROUGE-2 | ROUGE-L | ROUGE-Lsum |
|---|--------|---------|---------|---------|------------|
| 0 | 3      | 0.2224  | 0.0695  | 0.1355  | 0.1356     |

Compared to BART-base, the BART-large model has shown significant improvement, which confirms the effectiveness of a deeper architecture. The increase in ROUGE-2 is especially noticeable, which indicates a better preservation of the coherence of the text and bigram dependencies.

At the first stage, pre-trained BART-base and BART-large models were tested, which showed low ROUGE values due to lack of adaptation to scientific discourse, while BART-large gave slightly better results. After further training on the corpus of scientific articles, the indicators improved significantly: the increase in ROUGE-1 was +1.29% for BART-base and +3.43% for BART-large. Additional training (from 3 to 5 epochs) for BART-base did not increase the metrics, which indicates the efficiency limit.

Table 9. Test results of pre-trained and fine-tuned models

| Model                            | Epochs | ROUGE-1 | ROUGE-2 | ROUGE-L | ROUGE-Lsum |
|----------------------------------|--------|---------|---------|---------|------------|
| BART-base (without fine-tuning)  | –      | 0.1803  | 0.0417  | 0.0960  | 0.0961     |
| BART-large (without fine-tuning) | –      | 0.1881  | 0.0507  | 0.1088  | 0.1088     |
| BART-base (fine-tuned)           | 3      | 0.1932  | 0.0598  | 0.1234  | 0.1235     |
| BART-base (fine-tuned)           | 5      | 0.1932  | 0.0598  | 0.1234  | 0.1235     |
| BART-large (fine-tuned)          | 3      | 0.2224  | 0.0695  | 0.1355  | 0.1356     |

BART-large showed the largest increase, increasing ROUGE-2 by 1.88%, which

confirms its better ability to take into account the context and form coherent annotations. A comparison with BART-base shows that the depth of the architecture and the number of parameters directly affect the quality of the abstraction.

### **Discussion.**

The findings confirm the benefits of fine-tuning pretrained summarization models. BART-large demonstrates higher performance due to its deeper architecture and larger parameter set. BART-base is suitable for environments with limited computational resources. However, for complex scientific texts with structured sections deeper models like BART-large are preferable.

A limitation of this study is the exclusive use of lexical metrics, which are known to underestimate semantic metrics. Although ROUGE and BLEU capture surface overlap, they fail to reflect meaning preservation or factual consistency. Thus, incorporating semantic metrics such as BERTScore or BLEURT in future evaluations is essential.

Additionally, this study does not include human evaluation, which is critical for assessing fluency, coherence, and readability. The effect of summarization on literature review speeding or research indexing also remains unexplored.

### **Conclusion.**

Most abstracting models, including BART, are pre-trained on general-purpose corpora, which limits their effectiveness for scientific texts with a strict structure and specialized vocabulary. The study evaluated BART-base and BART-large, which showed low ROUGE values without adaptation and significant improvement after further training on the corpus of scientific articles. BART-large, surpassed BART-base in all metrics, especially for ROUGE-2, which reflects better text connectivity. The results confirm the need to adapt transformational models to the scientific language and further develop approaches based on semantic metrics, specialized models and postprocessing methods. The results demonstrate that fine-tuning improves the quality of summaries, particularly for BART-large, which achieved a ROUGE-1 score of 0.2224 and BLEU score of 0.29. Future work will focus on expanding comparisons to include other models (e.g., T5, PEGASUS) and conducting human evaluation.

### **REFERENCES**

- Ananthula B., Budde N., Baranovskyi O., Babenko T., Bigdan A., Lisnevskyi R. (2024). In-depth Study of Intelligent Healthcare Systems Using 5G Security, CEUR Workshop Proceedings. — Vol. 3680. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85192549346>.
- Biloshchytska S., Tleubayeva A., Kuchanskyi O., Biloshchytskyi A., Andrashko Y., Toxanov S., Mukhatayev A., Sharipova S. (2025). Text Similarity Detection in Agglutinative Languages: A Case Study of Kazakh Using Hybrid N-gram and Semantic Models. Applied Sciences. — Vol. 15. — No. 12. P. 6707. DOI: 10.3390/app15126707.
- Bahdanau D., Cho K., Bengio Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate, arXiv preprint arXiv:1409.0473.
- Babenko T., Kolesnikova K., Lisnevskyi R., Makilenov S., Landovsky Y. (2024). Definition of Cryptojacking Indicators, CEUR Workshop Proceedings. — Vol. 3680. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85192508442>
- Brown T., Mann B., Ryder N., Subbiah M., Kaplan J.D., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A.,

- Sutskever I., Amodei D. (2020). Language Models are Few-Shot Learners // arXiv.
- Batura T.V., Bakieva A.M. (2019). Methods and Systems of Automatic Text Summarization. — Novosibirsk: IPC NSU.
- Gritsenko T.Yu. (2020). Analysis of problems of automatic text summarization algorithms, *Dnevnik Nauki*. No. 5. Pp. 25–25.
- Gunes R.R. Dragomir. (1958). LexRank: Graph-based Lexical Centrality as Salience in Text Summarization // *Journal of Artificial Intelligence Research*. — Vol. 22. — Pp. 457–479.
- Hermann K.M., Kocisky T., Grefenstette E., Espeholt L., Kay W., Suleyman M., Blunsom P. (2015). Teaching machines to read and comprehend, *Advances in Neural Information Processing Systems*, Pp. 1693–1701.
- Hladkyi Y., Gladka M., Kostikov M., Lisnevskiy R. (2021). An IoT Solution: A Fitness Trainer, *CEUR Workshop Proceedings*. — Vol. 3179. — Pp. 215–226. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136276676>.
- Jones K.S. (2007). Automatic summarising: The state of the art, *Information Processing & Management*. — Vol. 43. — No. 6. Pp. 1449–1481.
- Lin C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. *Text Summarization Branches Out: Proc. ACL Workshop*. Pp. 74–81.
- Luhn H.P. (1958). The Automatic Creation of Literature Abstracts // *IBM Journal of Research and Development*. — Vol. 2. — No. 2. Pp. 159–165.
- Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed A., Levy O., Stoyanov V., Zettlemoyer L. (2019). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension // arXiv. — Vol. 36 — Pp. 8037–8050.
- Larsen B., Van I. A. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index, *Scientometrica*. — Vol. 84. — No. 3. Pp. 575–603.
- Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed A., Levy O., Stoyanov V., Zettlemoyer L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, *Proc. ACL*.
- Li P., Lam W., Bing L., Guo W., Li H. (2017). Cascaded attention based unsupervised information distillation for compressive summarization, *Proc. EMNLP*. Pp. 2071–2080.
- Neftissov A., Sarinova A., Kazambaev I., Kirichenko L., Lisnevskiy R., Rzyayeva L. (2023). Development of the Smart Office Concept, *IEEE Int. Conf. on Smart Information Systems and Technologies (SIST)*. Pp. 382–386. DOI: 10.1109/SIST58284.2023.10223512.
- Nallapati R., Zhai F., Zhou B. (2016). SummaRuNNer: A recurrent neural network-based sequence model for extractive summarization of documents. *Proc. 31st AAAI Conf. on Artificial Intelligence*. Pp. 3075–3081.
- Papineni K., Roukos S., Ward T., Zhu W. (2002). BLEU: A method for automatic evaluation of machine translation, *Proc. 40th Annual Meeting of the ACL*, Pp. 311–318.
- Plaza L., Diaz A., Gervas P. (2008). Concept-Graph Based Biomedical Automatic Summarization Using Ontologies, *Proc. Int. Conf. on Computational Linguistics*. Pp. 53–56.
- Raffel C., Shazeer N.M., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., Liu P.J. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, *J. Machine Learning Research*. — Vol. 21. — No. 140. Pp. 1–67.
- See A., Liu P.J., Manning C.D. (2017). Get to the point: Summarization with pointer-generator networks. *Proc. Association for Computational Linguistics (ACL)*. URL: <https://arxiv.org/abs/1704.04368>.
- Shao Y., Gouws S., Britz D., Goldie A., Strophe B., Kurzweil R. (2017). Generating high-quality and informative conversation responses with sequence-to-sequence models, *Proc. EMNLP*. Pp. 2210–2219.
- Song S., Huang H., Ruan T. (2019). Abstractive text summarization using LSTM-CNN based deep learning. *Multimedia Tools and Applications*. — Vol. 78. — Pp. 857–875.
- Tsvetkov V.Ya. & Domnitskaya E.V. (2008). On the methodology of data collection // *Fundamental Research*. No. 1. Pp. 52–53.
- Vanderwende L., Suzuki H., Brockett C., Nenkova A. (2007). Beyond SumBasic: Task-focused summarization with sentence simplification and lexical expansion, *Information Processing & Management*. — Vol. 43. — No. 6. Pp. 1606–1618.
- Zhang M., Li C., Wan M., Zhang X., Zhao Q. (2024). ROUGE-SEM: Better evaluation of summarization using ROUGE combined with semantics. *Expert Systems with Applications*. — Vol. 237.
- Zhong Q., Ding L., Liu J. (2024). Encoding-Enhanced Sequence-to-Sequence Pretraining for Language Understanding and Generation // *IEEE Transactions on Knowledge and Data Engineering*. — Vol. 36 — Pp. 8037–8050.

**INTERNATIONAL JOURNAL OF INFORMATION AND  
COMMUNICATION TECHNOLOGIES**

**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ  
ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

**Собственник:**

АО «Международный университет информационных  
технологий» (Казахстан, Алматы)

**Главный редактор:**

Колесникова Катерина Викторовна

**Ответственный редактор:**

Мрзабаева Раушан Жалиевна

**Компьютерная верстка:**

Калабай Замзагуль Ертугановна

Сайт журнала: <https://journal.iitu.edu.kz>

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Подписано в печать 30.03.2026.

050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).