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

ХАЛЫҚАРАЛЫҚ УНИВЕРСИТЕТІ

INTERNATIONAL UNIVERSITY SINCE 2009

МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

Published since 2020.
Volume 7. 3 (27). 2026
July–September

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

2020 жылдан бері шығарылады
Том 7. 3 (27). 2026
Шілде-Қыркүйек

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

Издается с 2020 г.
Том 7. 3 (27). 2026
Июль-Сентябрь

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 Rysbaiuly — 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)
Bakhtgerei 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 Yeskendirova — 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 Тулегеновна — 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)

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

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

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

Дузбаев Нуржан Токкужаевич — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті цифрландыру және инновациялар жөніндегі проректор (Қазақстан)

Синчев Бахтгерей Куспанович — 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 бизнеса медиа және басқару факультетінің деканы (Қазақстан)

Абдикаликова Замира Турсынбаевна — 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 Максутовна Ескендирова — 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)

Айтмагамбетов 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)

Тадеуш Валлас – PhD, Адам Мицкевич атындағы (Польша) университеттің проректоры

Мамырбаев Оркен Жумажанович — 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 сәулет университеті жобаларды басқару кафедрасының меңгерушісі (Украина)

Белощицкая 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)

Дамеля Максутовна Ескендирова — кандидат технических наук, ассоциированный профессор, заведующая кафедрой кибербезопасности Международного университета информационных технологий (Казахстан)

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

Айтмагамбетов 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 Сахибай – кандидат технических наук, профессор, профессор-исследователь кафедры компьютерной инженерии, Международного университета информационных технологий (Казахстан)

Алмисреб 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)

Белощицкая 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 SOCIO-ECONOMIC SYSTEMS

A Adalbek, S.Ye. Kerimkhulle, K.D. Kuliev, O. Mekesh, Zh.Zh. Khamitova
DEVELOPMENT OF RESULT-ORIENTED DATA SMOOTHING MODELS AND ANALYSIS OF THEIR EFFECTIVENESS 11
A.T. Iskakova, G.E. Kurmankulova, K. Gromaszek, K.O. Togzhanova, R.S. Algozhayeva
APPLICATION OF MACHINE LEARNING AND DEEP LEARNING FOR PREDICTIVE PLANNING AND RESOURCE EFFICIENCY IN PRODUCTION SYSTEMS 30
A.B. Kassymova, R.K. Uskenbayeva, Zh.A. Sydykov, D. Omirzak, S. Kalibekov
ARCHITECTURE OF AN INTELLIGENT CHATBOT FOR SMART CITY: AI-BASED RECEIVING, ANALYSIS AND ROUTING OF CITIZEN REQUESTS 46
O.M. Olzhayev, B.S. Omarov, A. Suliman
YOLO-CBAM: A LIGHTWEIGHT NEURAL NETWORK MODEL WITH INTEGRATED ATTENTION MECHANISM FOR ROAD DAMAGE ANALYSIS 66
K. Yermagambetova, A. Sarsekeyeva, Y. Akhlassov
OVERCOMING FRAGMENTATION IN THE CLASSIFICATION OF DIGITAL CULTURAL HERITAGE: AN INTEGRATIVE MULTIDIMENSIONAL MODEL 78

INFORMATION TECHNOLOGY

D. Abzhanova, S. Sharipova
ARCHITECTURE AND PRACTICAL IMPLEMENTATION OF A DIGITAL PLATFORM FOR MONITORING AND DATA EXCHANGE ON TRANSBOUNDARY WATER RESOURCES IN THE ZHAMBYL REGION 101
A. Aitanuly, O.A. Ussatova
ENHANCING TWO-FACTOR AUTHENTICATION FOR CLOUD HEALTHCARE SECURITY: THE ROLE OF DEEP LEARNING 122
O. Alseitov, O. Kuchanskyi
PHYSICS-INFORMED HYBRID LEARNING FOR HOURLY POWER-PLANT EMISSIONS: A MULTI-POLLUTANT BENCHMARK WITH UNCERTAINTY AND TRANSFER 141
A.M. Akhmetova, D.O. Baizhanova, Zh.N. Alisheva, B.A. Shaimov
ADAPTIVE OUTPUT ROUTER FOR NATURAL LANGUAGE MODELS CONSIDERING KEY PERFORMANCE INDICATORS 160
A.K. Aitim
AUTOMATIC AND SEMI-AUTOMATIC ANNOTATION OF KAZAKH TEXT CORPORA 180
Z.E. Balabayev, S.V. Biloshchytska
HANDWRITTEN TEXT RECOGNITION FOR CHAGATAI MANUSCRIPTS: COMPARING CRNN TRAINING FROM SCRATCH AND TRANSFORMER-BASED TRANSFER LEARNING 198
D.B. Baumuratova, T.K. Zhukabayeva, K.A. Abdrakhmanov, Z. Ahmad, A.A. Seitov
A NEURO-DIGITAL LEARNING MODEL FOR AI DRIVEN DIGITAL TRANSFORMATION IN HIGHER EDUCATION 219
T. Daoyun, G.A. Amirkhanova, Fu Yanwei, Y. Faruzkyzy
MULTIMODAL SEMANTIC MANIPULATION IN FAKE VIDEOS AND ITS INFLUENCE MECHANISM ON USERS' PERCEIVED TRUST 241
Zh. Doumchariyev, J. Tussupov, A. Akynbekova, M. Khassanova, M. Kassim
ENSEMBLE AND HYBRID LEARNING FOR DIGITAL TWIN MILK CLASSIFICATION 258
K.Z. Igibayev, A.Kh. Kozbakova, O. Auelbekov , T. Merembayev, K. Yessentayev
STUDY OF A DYNAMIC MIXING DEVICE IN A BIOREACTOR USING IOT TECHNOLOGIES 279
A.N. Ismukanova, G. Kochshanova, D.V. Pleskachev, Zh. Sarsenbayeva, P. Schmidt
AUTOMATED CLASSIFICATION OF SCIENTIFIC DATA USING ARTIFICIAL INTELLIGENCE: A COMPARATIVE ANALYSIS OF TRANSFORMER ARCHITECTURES AND HYBRID APPROACHES 306
T. Zh. Kamalov, Ali Abd. Almisreb, Mohammed A. Saleh
INVESTIGATION OF HYBRID QUANTUM-CLASSICAL ALGORITHMS FOR MACHINE LEARNING 323
B. Karimkyzy, K.G. Karasenko, N.A. Alisherov, N.M. Muratbek, A. Karim

KEY PUBLICATIONS OF 2021–2025 ON ARTIFICIAL INTELLIGENCE METHODS IN THE MANUFACTURING AND DIAGNOSTICS OF MEDICAL DEVICES FOR GLYCEMIC MONITORING: A THEMATIC ANALYTICAL REVIEW 338
S.K. Kumyzbaeva, K.E. Baglanbaev, K.S. Sarkyt, M.Zh. Konyrova
THREE-PHASE TRAINING OF INTELLIGENT ENERGY MANAGEMENT FOR A HYBRID ENERGY SYSTEM WITH A CLIMATE-AWARE STATE REPRESENTATION 357
N.E. Karabayev, S.K. Serikbayeva, E. Benkhelif, B. Tassuov
RESEARCH ON THE ROBUSTNESS OF ENSEMBLE MACHINE LEARNING MODELS TO ADVERSARIAL ATTACKS IN INTRUSION DETECTION SYSTEMS FOR IOT NETWORKS 369
M. Nurtas, G. Kapyshev
A METHODOLOGICAL FRAMEWORK FOR THE PREPROCESSING AND INTEGRATION OF GEOSPATIAL AND GEOPHYSICAL DATA FOR MINERAL PROSPECTIVITY PREDICTION USING MACHINE LEARNING 392
A. Neftissov, A. Biloshchytskyi, I. Kazambayev, A. Medetkhan, L. Kirichenko
ENERGY EFFICIENT MICROCLIMATE STABILIZATION FOR SEALED ARCHIVAL STORAGE ENCLOSURES 405
A.Zh. Omurtayeva, U.T. Makhazhanova, A.K. Adilova, Sh.K. Yelezhanova, J. Kultan
AN ADAPTIVE NEURAL-FUZZY MODEL FOR ASSESSING INVESTMENT RISKS IN THE AGRI-INDUSTRIAL COMPLEX: ODERNISATION OF THE DECISION-SUPPORT ARCHITECTURE THROUGH THE INTEGRATION OF MACHINE LEARNING 421
N. Shazhdekeyeva, A. Mukhanbetkaliyeva, A. Abdigaliyeva, A. Zhanekeshova, L. Nurgaliyeva
COMPARATIVE ANALYSIS OF MATHEMATICAL MODELS OF COLLABORATIVE FILTERING IN THE DEVELOPMENT OF A MOVIE RECOMMENDER SYSTEM 445
S. Sharmukhanbet, G. Turmukhanova, O. Findik, V. Makhatova, M. Rakhmetov
AN INTELLIGENT PIPELINE DEFECT DETECTION SYSTEM USING AN IN-PIPE ROBOT 461

INFORMATION SECURITY AND COMMUNICATIONTECHNOLOGIES

S.Z. Daruish, A.U. Utegenova, A.U. Kalizhanova, Zh.Zh. Zhumabekova
MACHINE LEARNING FOR POLLUTION CLASSIFICATION IN ROCKET BOOSTER IMPACT AREAS 480
S.N. Issabayeva, A.K. Rysbayeva, Zh.B. Ibrayeva, N.R. Mussabekov, N.Sh. Akimzhan
MACHINE LEARNING-BASED CYBERATTACK DETECTION IN OT-NETWORKS: A COMPARATIVE ANALYSIS OF NETWORK ANOMALY CLASSIFICATION METHODS 495
G. Shangytbayeva, A.N. Amanov, M.K. Issenova, G.E. Berikkhanova, K.O. Shakerkhan
INFORMATION SECURITY IN GOVERNMENT CLOUD SERVICES: RISK ASSESSMENT AND PROTECTION STRATEGIES 512
Ye. N. Tursynbek, N. Albanbay, D. Djenouri
OPTIMISATION OF COMMUNICATION COSTS IN TWO-TIER FEDERATED LEARNING FOR CYBERATTACK DETECTION IN IOT NETWORKS 529
A. Tastemirov, S.B. Mukhanov, S.T. Amanzholova, B. Aldosh
A COMPARATIVE EVALUATION OF MACHINE LEARNING MODELS FOR ACTIVE DIRECTORY THREAT DETECTION: FEATURE-LEAKAGE CONTROL AND CROSS-SCENARIO VALIDATION ON WINDOWS SECURITY EVENT LOGS 545

МАЗМҰНЫ
ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМЫТУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

А Адалбек, С.Е. Керімқұл, К.Д. Кулиев, О. Мекеш, Ж.Ж. Хамитова
ДЕРЕКТЕРДІҢ НӘТИЖЕ БАҒДАРЛАНҒАН ТЕГІСТЕУ МОДЕЛЬДЕРІН ӘЗІРЛЕУ ЖӘНЕ ОЛАРДЫҢ ТИІМДІЛІГІН ТАЛДАУ 11
А.Т. Искакова, Г.Е. Курманкулова, Конрад Громашек, К.О. Тогжанова, Р.С. Алгожаева
ӨНДІРІС ЖҮЙЕЛЕРІНДЕ БОЛЖАМДЫ ЖОСПАРЛАУ ЖӘНЕ РЕСУРСТАРДЫ ҮНЕМДЕУ ҮШІН МАШИНАЛЫҚ ЖӘНЕ ТЕРЕҢ ОҚЫТУДЫ ҚОЛДАНУ 30
А.Б. Касымова, Р.К. Ускенбаева, Ж.А. Сыдыков, Д. Өмірзақ, С. Калибеков
SMART CITY ҮШІН ИНТЕЛЛЕКТУАЛДЫ ЧАТ-БОТ АРХИТЕКТУРАСЫ: АЗАМАТТАРДЫҢ

ӨТІНІШТЕРІН ЖАСАНДЫ ИНТЕЛЛЕКТ НЕГІЗІНДЕ ҚАБЫЛДАУ, ТАЛДАУ ЖӘНЕ БАҒЫТТАУ 46
О.М. Олжаев, Б.С. Омаров, A.Suliman
YOLO-CBAM: ЖОЛДЫҢ ЗАҚЫМДАНУЫН ТАЛДАУҒА АРНАЛҒАН ИНТЕГРАЦИЯЛАНҒАН ЗЕЙІН МЕХАНИЗМІ БАР ЖЕҢІЛ САЛМАҚТЫ НЕЙРОНДЫҚ ЖЕЛІ МОДЕЛІ 66
К. Ермагамбетова, А. Сарсекеева, Е. Ахласов
ЦИФРЛЫҚ МӘДЕНИ МҰРАНЫ ЖІКТЕУДЕГІ ФРАГМЕНТАЦИЯНЫ ЕҢСЕРУ: ИНТЕГРАТИВТІ КӨПӨЛШЕМДІ МОДЕЛЬ 78

АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

Д.Е. Абжанова, С.Е. Шарипова
ЖАМБЫЛ ӨҢІРІНІҢ ТРАНСШЕКАРАЛЫҚ СУ РЕСУРСТАРЫН МОНИТОРИНГТЕУ ЖӘНЕ ДЕРЕКТЕР АЛМАСУҒА АРНАЛҒАН ЦИФРЛЫҚ ПЛАТФОРМАНЫҢ АРХИТЕКТУРАСЫ МЕН ПРАКТИКАЛЫҚ ІСКЕ АСЫРЫЛУЫ 101
А.Айтанұлы, О.А. Усатова
БҰЛТТЫ МЕДИЦИНАЛЫҚ ЖҮЙЕЛЕРДЕ ЕКІ ФАКТОРЛЫ АУТЕНТИФИКАЦИЯНЫ ЖАҚСАРТУ: ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІНІҢ ҚОЛДАНУ МҮМКІНДІКТЕРІ 122
О. Альсеитов, О. Кучанский
ЭЛЕКТР СТАНЦИЯЛАРЫНЫҢ САҒАТТЫҚ ШЫҒАРЫНДЫЛАРЫН ФИЗИКАЛЫҚ ЗАҢДЫЛЫҚТАРДЫ ЕСКЕРЕТІН ГИБРИДТІ ОҚЫТУ: БЕЛГІСІЗДІК ПЕН ТАСЫМАЛДАУДЫ ЕСКЕРЕТІН КӨПЛАСТАУЫШТЫ БЕНЧМАРК 141
А.М. Ахметова, Д.О. Байжанова, Ж.Н. Алишева, Б.А. Шаймов
КРІ КӨРСЕТКІШТЕРІН ЕСКЕРЕ ОТЫРЫП, ТАБИҒИ ТІЛ МОДЕЛЬДЕРІНЕ АРНАЛҒАН ШЫҒЫС МАРШРУТТАУЫНЫҢ АДАПТИВТІ ЖҮЙЕСІ 160
Ә.Қ. Әйтім
ҚАЗАҚ МӘТІНДІК КОРПУСТАРЫН АВТОМАТТЫ ЖӘНЕ ЖАРТЫЛАЙ АВТОМАТТЫ АННОТАЦИЯЛАУ 180
Ж.Е. Балабаев, С.В. Белощицкая
ШАҒАТАЙ ҚОЛЖАЗБАЛАРЫНЫҢ МӘТІНІН ТАНУ: CRNN-ДІ НӨЛДЕН БАСТАП ОҚЫТУ МЕН ТРАНСФОРМЕР НЕГІЗІНДЕГІ ТРАНСФЕРЛІК ОҚЫТУДЫ САЛЫСТЫРУ 198
Д.Б. Баумуратова, Т.К. Жукабаева, К.А. Абдрахманов, З. Ахмад, А.А. Сеитов
ЖОҒАРЫ БІЛІМ БЕРУДЕ AI НЕГІЗДЕЛГЕН ЦИФРЛЫҚ ТРАНСФОРМАЦИЯҒА АРНАЛҒАН НЕЙРО-ЦИФРЛЫҚ ОҚЫТУ МОДЕЛІ 219
Т. Даоюн, Г. Амирханова, Фу Яньвэй, Е. Фарузқызы
ФЕЙК ВИДЕОЛАРДАҒЫ МУЛЬТИМОДАЛЬДЫ СЕМАНТИКАЛЫҚ МАНИПУЛЯЦИЯНЫҢ ПАЙДАЛАНУШЫЛАРДЫҢ ҚАБЫЛДАЙТЫН СЕНІМІНЕ ӘСЕР ЕТУ МЕХАНИЗМІ 241
Ж. Доумчариева, Д. Тусупов, А. Акынбекова, М. Хасанова, М. Кассим
ЦИФРЛЫҚ ЕГІЗДЕГІ СҮТ САПАСЫН АНСАМБЛЬДІК ЖӘНЕ ГИБРИДТІ ОҚЫТУ АРҚЫЛЫ ЖІКТЕУ 258
Қ.З. Игибаев, А.Х. Козбакова, О. Әуелбеков, Т. Мерембаев, Қ. Есентаев
БИОРЕАКТОРДА IoT ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНА ОТЫРЫП ДИНАМИКАЛЫҚ АРАЛАСТЫРУ ҚҰРЫЛҒЫСЫН ЗЕРТТЕУ 279
А.Н. Исмуканова, Г.Кощанова, Д.В. Плескачев, Ж.Сарсенбаева, П. Шмидт
ЖАСАНДЫ ИНТЕЛЛЕКТ НЕГІЗІНДЕ ҒЫЛЫМИ ДЕРЕКТЕРДІ АВТОМАТТЫ ТҮРДЕ ЖІКТЕУ: ТРАНСФОРМЕР АРХИТЕКТУРАЛАРЫ МЕН ГИБРИДТІ ТӘСІЛДЕРДІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ 306
Т.Ж. Камалов, Али Абд. Альмисреб, Мохаммед А. Салех
МАШИНАЛЫҚ ОҚЫТУҒА АРНАЛҒАН ГИБРИДТІ КВАНТТЫҚ-КЛАССИКАЛЫҚ АЛГОРИТМДЕРІН ЗЕРТТЕУ 323
Б. Кәрімқызы, К.Г. Карасенко, Н.А. Әлішеров, Н.М. Мұратбек, А. Кәрім
2021–2025 ЖЫЛДАР АРАЛЫҒЫНДАҒЫ ГЛИКЕМИЯЛЫҚ МОНИТОРИНГКЕ АРНАЛҒАН МЕДИЦИНАЛЫҚ ҚҰРЫЛҒЫЛАРДЫ ӨНДІРУ ЖӘНЕ ДИАГНОСТИКАЛАУДАҒЫ ЖАСАНДЫ ИНТЕЛЛЕКТ ӘДІСТЕРІ БОЙЫНША НЕГІЗГІ ЖАРИЯЛАНЫМДАР 338
С. К. Кумызбаева, К.Е. Багланбаев, К.С. Саркыт, М. Конырова
КҮЙ-ЖАЙДЫ КЛИМАТҚА СЕЗІМТАЛ ҰСЫНУМЕН ГИБРИДТІ ЭНЕРГОЖҮЙЕНІҢ ИНТЕЛЛЕКТУАЛДЫ ЭНЕРГОМЕНЕДЖМЕНТІН ҮШФАЗАЛЫ ОҚЫТУ 357
Н.Е. Қарабаев, С.К. Серикбаева, Е. Бенкхелифа, Б. Тасуов

IOT-ЖЕЛІЛЕРІНДЕГІ БАСЫП КІРУДІ АНЫҚТАУЖҮЙЕЛЕРІНДЕ МАШИНАЛЫҚ ОҚЫТУДЫҢ АНСАМБЛЬДІ КҮЛГІЛЕРІНІҢ БӘСЕКЕЛЕС ШАБУЫЛДАРҒА ТӨЗІМДІЛІГІН ЗЕРТТЕУ 369
М. Нұртас, Ғ. Капышев
МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП, ПАЙДАЛЫ ҚАЗБАЛАРДЫҢ ПЕРСПЕКТИВАЛЫЛЫҒЫН БОЛЖАУ ҮШІН ГЕОКЕҢІСТІКТІК ЖӘНЕ ГЕОФИЗИКАЛЫҚ ДЕРЕКТЕРДІ АЛДЫН АЛА ӨҢДЕУ ЖӘНЕ БІРІКТІРУДІҢ ӘДІСТЕМЕЛІК НЕГІЗІ 392
А. Нефтиссов, А. Билощицкий, И. Казамбаев, А. Медетхан, Л. Кириченко
ГЕРМЕТИКАЛЫҚ АРХИВТІК САҚТАУ ҚАПТАМАЛАРЫНДА МИКРОКЛИМАТТЫ ЭНЕРГИЯ ТИІМДІЛІГІН ЕСКЕРЕ ОТЫРЫП ТҰРАҚТАНДЫРУ 405
Ә.Ж. Омуртаева, У.Т. Махажанова, А.К. Адилова, Ш.К. Ележанова, Я. Култан
АГРОӨНЕРКӘСІПТІК КЕШЕНДЕГІ ИНВЕСТИЦИЯЛЫҚ ТӘУЕКЕЛДЕРДІ БАҒАЛАУҒА АРНАЛҒАН БЕЙІМДЕЛГІШ НЕЙРОНДЫҚ-БҰЛДЫР МОДЕЛЬ: МАШИНАЛЫҚ ОҚЫТУДЫ ИНТЕГРАЦИЯЛАУ АРҚЫЛЫ ШЕШІМ ҚАБЫЛДАУДЫ ҚОЛДАУ СӘУЛЕТІН ЖАҢҒЫРТУ 421
Н. Шаждекеева, А. Мұханбетқалиева, А. Әбдіғалиева, А. Жанекешова, Л. Нұрғалиева
ФИЛЬМДЕРДІ ҰСЫНУ ЖҮЙЕСІН ҚҰРУ КЕЗІНДЕГІ КОЛЛАБОРАТИВТІК СҮЗГІЛЕУДІҢ МАТЕМАТИКАЛЫҚ МОДЕЛЬДЕРІН САЛЫСТЫРМАЛЫ ТАЛДАУ 445
С. Шармұханбет, Г. Тұрмұханова, О. Финдик, В. Махатова, М. Рахметов
ҚҰБЫРІШІЛІК РОБОТТЫ ПАЙДАЛАНУ АРҚЫЛЫ ҚҰБЫР АҚАУЛАРЫН АНЫҚТАУДЫҢ ЗИЯТКЕРЛІК ЖҮЙЕСІ 461

АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН

С.З. Дәруіш, А.У. Утегенова, А.У. Калижанова, Ж.Ж. Жумабекова
ЗЫМЫРАН ТАСЫМАЛДАУШЫЛАРДЫҢ ҚҰЛАУ АЙМАҚТАРЫНЫҢ ЛАСТАНУЫН МАШИНАЛЫҚ ОҚЫТУМЕН ЖІКТЕУ 480
С.Н. Исабаева, А.К. Рысбаева, Ж.Б. Ибраева, Н.Р. Муссабеков, Н.Ш. Акимжан
МАШИНАЛЫҚ ОҚЫТУ НЕГІЗІНДЕ ОТ-ЖЕЛІЛЕРІНДЕГІ КИБЕРШАБУЫЛДАРДЫ АНЫҚТАУ: ЖЕЛІЛІК АНОМАЛИЯЛАРДЫ ЖІКТЕУ ӘДІСТЕРІНІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ 495
Г.А. Шангытбаева, А.Н. Аманов, М.К. Исенова, Г.Е. Берикханова, Қ.О. Шакерхан
МЕМЛЕКЕТТІК ОРГАНДАРДЫҢ БҰЛТТЫ СЕРВИСТЕРІНДЕГІ АҚПАРАТТЫҚ ҚАУІПСІЗДІК: ТӘУЕКЕЛДЕРДІ БАҒАЛАУ ЖӘНЕ ҚОРҒАУ СТРАТЕГИЯЛАРЫ 512
Е.Н. Турсынбек, Н. Албанбай, Д. Дженоури
IOT ЖЕЛІЛЕРІНДЕ КИБЕРШАБУЫЛДАРДЫ АНЫҚТАУ ҮШІН ЕКІ ДЕҢГЕЙЛІ ФЕДЕРАЦИЯЛАНҒАН ОҚЫТУДАҒЫ БАЙЛАНЫС ШЫҒЫНДАРЫН ОҢТАЙЛАНДЫРУ 529
А. Тастемиров, С.Т. Аманжолова, Б. Алдош, С.Б. Муханов
ACTIVE DIRECTORY ЖҮЙЕСІНДЕ ҚАУІПТЕРДІ АНЫҚТАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ МОДЕЛЬДЕРІН САЛЫСТЫРМАЛЫ БАҒАЛАУ: WINDOWS ҚАУІПСІЗДІК ОҚИҒАЛАРЫ ЖУРНАЛДАРЫНДА БЕЛГІ ЖЫЛЫСТАУЫН БАҚЫЛАУ ЖӘНЕ КРОСС-СЦЕНАРЛЫҚ ВАЛИДАЦИЯ 545

СОДЕРЖАНИЕ
ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИСОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ

А Адалбек, С.Е. Керимкулов, К.Д. Кулиев, О. Мекеш, Ж.Ж. Хамитова
РАЗРАБОТКА РЕЗУЛЬТАТИВНО-ОРИЕНТИРОВАННЫХ МОДЕЛЕЙ СГЛАЖИВАНИЯ ДАННЫХ И АНАЛИЗ ИХ ЭФФЕКТИВНОСТИ 11
А.Т. Искакова, Г.Е. Курманкулова, К. Громашек, К.О. Тогжанова, Р.С. Алгожаева
ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО И ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗНОГО ПЛАНИРОВАНИЯ И РЕСУРСОСБЕРЕЖЕНИЯ В ПРОИЗВОДСТВЕННЫХ СИСТЕМАХ 30
А.Б. Касымова, Р.К. Ускенбаева, Ж.А.Сыдыков, Д. Омирзак, С. Калибеков
АРХИТЕКТУРА ИНТЕЛЛЕКТУАЛЬНОГО ЧАТ-БОТА ДЛЯ SMART CITY: ПРИЁМ, АНАЛИЗ И МАРШРУТИЗАЦИЯ ОБРАЩЕНИЙ ГРАЖДАН НА ОСНОВЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА 46
О.М. Олжаев, Б.С. Омаров, A.Suliman
YOLO-CBAM: ЛЕГКОВЕСНАЯ НЕЙРОСЕТЕВАЯ МОДЕЛЬ С ИНТЕГРИРОВАННЫМ МЕХАНИЗМОМ ВНИМАНИЯ ДЛЯ АНАЛИЗА ДОРОЖНЫХ ПОВРЕЖДЕНИЙ 66
К. Ермагамбетова, А. Сарсекеева, Е. Ахласов
ПРЕОДОЛЕНИЕ ФРАГМЕНТАЦИИ В КЛАССИФИКАЦИИ ЦИФРОВОГО КУЛЬТУРНОГО НАСЛЕДИЯ:

ИНТЕГРАТИВНАЯ МНОГОМЕРНАЯ МОДЕЛЬ 78

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Д.Е. Абжанова, С.Е. Шарипова
АРХИТЕКТУРА И ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ЦИФРОВОЙ ПЛАТФОРМЫ МОНИТОРИНГА И ОБМЕНА ДАННЫМИ О ТРАНСГРАНИЧНЫХ ВОДНЫХ РЕСУРСАХ ЖАМБЫЛСКОГО РЕГИОНА 101
А. Айтанулы, О.А. Усатова
СОВЕРШЕНСТВОВАНИЕ ДВУХФАКТОРНОЙ АУТЕНТИФИКАЦИИ ДЛЯ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ОБЛАЧНЫХ МЕДИЦИНСКИХ СИСТЕМ: ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ 122
О. Альсеитов, О. Кучанский
ФИЗИЧЕСКИ ИНФОРМИРОВАННОЕ ГИБРИДНОЕ ОБУЧЕНИЕ ДЛЯ ПОЧАСОВОГО ПРОГНОЗИРОВАНИЯ ВЫБРОСОВ ЭЛЕКТРОСТАНЦИЙ: МНОГОКОМПОНЕНТНЫЙ БЕНЧМАРК С ОЦЕНКОЙ НЕОПРЕДЕЛЁННОСТИ И ПЕРЕНОСИМОСТЬЮ МОДЕЛЕЙ 141
А.М. Ахметова, Д.О. Байжанова, Ж.Н. Алишева, Б.А. Шаймов
АДАПТИВНЫЙ МАРШРУТИЗАТОР ВЫВОДА С УЧЕТОМ КЛЮЧЕВЫХ ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ ДЛЯ МОДЕЛЕЙ ЕСТЕСТВЕННОГО ЯЗЫКА 160
Ә.Қ. Әйтім
АВТОМАТИЧЕСКАЯ И ПОЛУАВТОМАТИЧЕСКАЯ АННОТАЦИЯ КАЗАХСКИХ ТЕКСТОВЫХ КОРПУСОВ 180
Ж.Е. Балабаев, С.В. Белощицкая
РАСПОЗНАВАНИЕ РУКОПИСНОГО ТЕКСТА ШАГАТАЙСКИХ РУКОПИСЕЙ: СРАВНЕНИЕ ОБУЧЕНИЯ CRNN С НУЛЯ И ТРАНСФЕРНОГО ОБУЧЕНИЯ НА БАЗЕ ТРАНСФОРМЕРОВ 198
Д.Б. Баумуратова, Т.К. Жукабаева, К.А. Абдрахманов, З. Ахмад, А.А. Сеитов
НЕЙРО-ЦИФРОВАЯ МОДЕЛЬ ОБУЧЕНИЯ ДЛЯ ЦИФРОВОЙ ТРАНСФОРМАЦИИ ВЫСШЕГО ОБРАЗОВАНИЯ С ПОМОЩЬЮ AI 219
Т. Даоюн, Г. Амирханова, Фу Яньвэй, Е. Фарузкызы
МЕХАНИЗМ ВЛИЯНИЯ МУЛЬТИМОДАЛЬНОЙ СЕМАНТИЧЕСКОЙ МАНИПУЛЯЦИИ В ФЕЙКОВЫХ ВИДЕО НА ВОСПРИНИМАЕМОЕ ДОВЕРИЕ ПОЛЬЗОВАТЕЛЕЙ 241
Ж. Доумчариева, Д. Тусупов, А. Акынбекова, М. Хасанова, М. Кассим
КЛАССИФИКАЦИЯ КАЧЕСТВА МОЛОКА В ЦИФРОВОМ ДВОЙНИКЕ АНСАМБЛЕВЫМИ И ГИБРИДНЫМИ МЕТОДАМИ 258
К.З. Игибаев, А.Х. Козбакова, О. Ауелбеков, Т. Мерембаев, К. Есентаев
ИССЛЕДОВАНИЕ УСТРОЙСТВА ДИНАМИЧЕСКОГО ПЕРЕМЕШИВАНИЯ В БИОРЕАКТОРЕ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ IoT 279
А. Н. Исмуканова, Г.Кощанова, Д.В. Плескачев, Ж.Сарсенбаева, П. Шмидт
АВТОМАТИЗИРОВАННАЯ КЛАССИФИКАЦИЯ НАУЧНЫХ ДАННЫХ НА ОСНОВЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТРАНСФОРМЕРНЫХ АРХИТЕКТУР И ГИБРИДНЫХ МЕТОДОВ 306
Т.Ж. Камалов, Али Абд. Альмисреб, Мохаммед А. Салех
ИССЛЕДОВАНИЕ ГИБРИДНЫХ КВАНТОВО-КЛАССИЧЕСКИХ АЛГОРИТМОВ ДЛЯ МАШИННОГО ОБУЧЕНИЯ 323
Б. Каримкызы, К.Г. Карасенко, Н.А. Алишеров, Н.М. Муратбек, А. Карим
КЛЮЧЕВЫЕ ПУБЛИКАЦИИ 2021–2025 ГОДОВ ПО МЕТОДАМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРОИЗВОДСТВЕ И ДИАГНОСТИКЕ МЕДИЦИНСКИХ ИЗДЕЛИЙ ДЛЯ ГЛИКЕМИЧЕСКОГО МОНИТОРИНГА 338
С.К. Кумызбаева, К.Е. Багланбаев, К.С. Сарқыт, М.Ж. Конырова
ТРЁХФАЗОВОЕ ОБУЧЕНИЕ ИНТЕЛЛЕКТУАЛЬНОГО ЭНЕРГОМЕНЕДЖМЕНТА ГИБРИДНОЙ ЭНЕРГОСИСТЕМЫ С КЛИМАТО-ОСВЕДОМЛЁННЫМ ПРЕДСТАВЛЕНИЕМ СОСТОЯНИЯ 357
Н.Е. Қарабаев, С.К. Серикбаева, Е. Бенкхелифа, Б. Тасуов
ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ АНСАМБЛЕВЫХ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ К СОСТЯЗАТЕЛЬНЫМ АТАКАМ В СИСТЕМАХ ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ IOT-СЕТЕЙ 369
М. Нуртас, Г. Капышев
МЕТОДОЛОГИЧЕСКАЯ ОСНОВА ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ И ИНТЕГРАЦИИ ГЕОПРОСТРАНСТВЕННЫХ И ГЕОФИЗИЧЕСКИХ ДАННЫХ ДЛЯ ПРОГНОЗИРОВАНИЯ ПЕРСПЕКТИВНО-

СТИ МИНЕРАЛЬНЫХ РЕСУРСОВ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ 392
А. Нефтиссов, А. Билощицкий, И. Казамбаев, А. Медетхан, Л. Кириченко
ЭНЕРГОЭФФЕКТИВНАЯ СТАБИЛИЗАЦИЯ МИКРОКЛИМАТА В ГЕРМЕТИЧНЫХ АРХИВНЫХ ХРАНИЛИЩАХ 405
А.Ж. Омуртаева, У.Т. Махажанова, А.К. Адилова, Ш.К. Ележанова, Я.Култан
АДАПТИВНАЯ НЕЙРО-НЕЧЕТКАЯ МОДЕЛЬ ДЛЯ ОЦЕНКИ ИНВЕСТИЦИОННЫХ РИСКОВ В АГРОПРОМЫШЛЕННОМ КОМПЛЕКСЕ: МОДЕРНИЗАЦИЯ АРХИТЕКТУРЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПОСРЕДСТВОМ ИНТЕГРАЦИИ МАШИННОГО ОБУЧЕНИЯ 421
Н. Шаждекеева, А. Муханбеткалиева, А. Абдигалиева, А. Жанекешова, Л. Нургалиева
СРАВНИТЕЛЬНЫЙ АНАЛИЗ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ ПРИ ПОСТРОЕНИИ РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ ФИЛЬМОВ 445
С. Шармуханбет, Г. Турмуханова, О.Финдик, В.Махатова, М. Рахметов
ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОБНАРУЖЕНИЯ ДЕФЕКТОВ ТРУБОПРОВОДОВ С ИСПОЛЬЗОВАНИЕМ ВНУТРИТРУБНОГО РОБОТА 461

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

С.З. Дәруіш, А.У. Утегенова, А.У. Калижанова, Ж.Ж. Жумабекова
МАШИННОЕ ОБУЧЕНИЕ ДЛЯ КЛАССИФИКАЦИИ ЗАГРЯЗНЕНИЯ РАЙОНОВ ПАДЕНИЯ РАКЕТ-НОСИТЕЛЕЙ 480
С.Н. Исабаева, А.К. Рысбаева, Ж.Б. Ибраева, Н.Р. Муссабеков, Н.Ш. Акимжан
ОБНАРУЖЕНИЕ КИБЕРАТАК В ОТ-СЕТЯХ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ КЛАССИФИКАЦИИ СЕТЕВЫХ АНОМАЛИЙ 495
Г.А. Шангытбаева, А.Н. Аманов, М.К. Исенова, Г.Е. Берикханова, Қ.О. Шакерхан
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ В ОБЛАЧНЫХ СЕРВИСАХ ГОСУДАРСТВЕННЫХ ОРГАНОВ: ОЦЕНКА РИСКОВ И СТРАТЕГИИ ЗАЩИТЫ 512
Е.Н. Турсынбек, Н. Албанбай, Д. Дженоури
ОПТИМИЗАЦИЯ КОММУНИКАЦИОННЫХ ЗАТРАТ В ДВУХУРОВНЕВОМ ФЕДЕРАТИВНОМ ОБУЧЕНИИ ДЛЯ ОБНАРУЖЕНИЯ КИБЕРАТАК В IOT-СЕТЯХ 529
А. Тастемиров, С.Т. Аманжолова, Б. Алдош, С.Б. Муханов
СРАВНИТЕЛЬНАЯ ОЦЕНКА МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ УГРОЗ В ACTIVE DIRECTORY: КОНТРОЛЬ УТЕЧКИ ПРИЗНАКОВ И КРОСС-СЦЕНАРНАЯ ВАЛИДАЦИЯ НА ЖУРНАЛАХ БЕЗОПАСНОСТИ WINDOWS 545

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 7. Is.3. Number 27 (2026). Pp. 180–197
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2026.27.3.010
UDC 004.912
IRSTI 16.21.41

AUTOMATIC AND SEMI-AUTOMATIC ANNOTATION OF KAZAKH TEXT CORPORA

A.K. Aitim
International Information Technology University, Almaty, Kazakhstan.
E-mail: a.aitim@iitu.edu.kz

Aigerim K. Aitim — PhD in Intelligent Systems, Associate professor of Information Systems Department, International Information Technology University, Almaty, Kazakhstan
E-mail: a.aitim@iitu.edu.kz, https://orcid.org/0000-0003-2982-214X.

© A.K. Aitim

Abstract. The absence of extensive, high-quality annotated corpora continues to impede the advancement of effective natural language processing tools for the Kazakh language. This paper introduces a semi-automatic annotation framework for annotating Kazakh text corpora, including rule-based preprocessing, machine-assisted annotation, and manual verification. The suggested pipeline has steps for cleaning up text, breaking it up into sentences, tokenizing it, doing morphological pre-annotation, and manual correction of ambiguous cases. Special attention is paid to the difficulties of Kazakh as an agglutinative language, such as its many suffixes, vowel harmony, and changes in morphophonology. The goal of the strategy is to reduce annotation time while keeping the quality of the data and the consistency of the language. The generated corpus can be used for tasks like labeling parts of speech, morphosyntactic analysis, and parsing sentences. The results show that semi-automatic annotation is a useful and scalable way to make Kazakh NLP resources.

Keywords: Kazakh language, corpus annotation, semi-automatic annotation, natural language processing, morphological tagging, low-resource languages

For citation: Aitim A.K. (2026). Automatic and semi-automatic annotation of Kazakh text corpora // International journal of information and communication technologies. Vol. 7. No. 27. Pp. 180–197. https://doi.org/10.54309/IJICT.2026.27.3.010. (In Eng.).

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

Acknowledgment. *This work was supported by the Ministry of Culture and*

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Information of the Republic of Kazakhstan of grant «Tauelsizdik Urpaktary-2025», project named by "QazNLP is an open-source scientific system for intelligent processing of Kazakh-language text".

ҚАЗАҚ МӘТІНДІК КОРПУСТАРЫН АВТОМАТТЫ ЖӘНЕ ЖАРТЫЛАЙ АВТОМАТТЫ АННОТАЦИЯЛАУ

Ә.Қ. Әйтім
Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан.
E-mail: a.aitim@iitu.edu.kz

Әйгерім Қ. Әйтім – PhD, Ақпараттық жүйелер кафедрасының қауымдастырылған профессоры, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан
E-mail: a.aitim@iitu.edu.kz, https://orcid.org/0000-0003-2982-214X.

© Ә.Қ. Әйтім

Аннотация. Үлкен көлемді әрі сапалы аннотацияланған корпустардың жеткіліксіздігі қазақ тілі үшін тиімді табиғи тілдерді өңдеу құралдарының дамуын әлі де тежеп отыр. Бұл мақалада қазақ мәтіндік корпустарын аннотациялауға арналған автоматты және жартылай автоматты құрылым ұсынылады. Оған ережелерге негізделген алдын ала өңдеу, машиналық көмегімен белгілеу және қолмен тексеру кезеңдері кіреді. Ұсынылған конвейер мәтінді тазалау, сөйлемдерге бөлу, токенизация, морфологиялық алдын ала аннотация және екіұшты жағдайларды нақтылау сияқты кезеңдерден тұрады. Ерекше назар қазақ тілінің агглютинативті тіл ретіндегі ерекшеліктеріне аударылады: көп мөлшердегі жұрнақтар мен жалғаулар, дауысты дыбыстардың үндестігі және морфофонологиялық өзгерістер. Ұсынылған тәсілдің мақсаты - тілдік деректердің сапасы мен бірізділігін сақтай отырып, аннотациялау уақытын қысқарту. Құрылған корпус сөз табын белгілеу, сөз мағынасын анықтау және синтаксистік талдау сияқты міндеттер үшін пайдаланылуы мүмкін. Нәтижелер жартылай автоматты аннотацияның қазақ тіліндегі NLP ресурстарын құру үшін тиімді әрі масштабталатын тәсіл екенін көрсетеді.

Түйінді сөздер: қазақ тілі, корпус аннотациясы, жартылай автоматты аннотация, табиғи тілдерді өңдеу, морфологиялық белгілеу, ресурсы шектеулі тілдер

Дәйексөздер үшін: Әйтім Ә.Қ. (2026). Қазақ мәтіндік корпустарын автоматты және жартылай автоматты аннотациялау // Халықаралық ақпараттық және коммуникациялық технологиялар журналы. Т. 7. No 27. Б. 180–197. https://doi.org/10.54309/IJICT.2026.27.3.010. (Ағыл. тіл.).

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

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Алғыс. *Бұл жұмыс Қазақстан Республикасы Мәдениет және ақпарат министрлігінің «Тәуелсіздік ұрпақтары – 2025» гранты аясында «QazNLP — қазақ тіліндегі мәтіндерді интеллектуалды өңдеуге арналған ашық ғылыми жүйе» жобасы бойынша қолдау тапты.*

АВТОМАТИЧЕСКАЯ И ПОЛУАВТОМАТИЧЕСКАЯ АННОТАЦИЯ КАЗАХСКИХ ТЕКСТОВЫХ КОРПУСОВ

Ә.Қ. Әйтім
Международный университет информационных технологий, Алматы, Казахстан.
E-mail: a.aitim@iitu.edu.kz

Әйгерім Қ. Әйтім – PhD, ассоциированный профессор кафедры Информационные системы, Международный университет информационных технологий, Алматы, Казахстан
E-mail: a.aitim@iitu.edu.kz, https://orcid.org/0000-0003-2982-214X.

© Ә.Қ. Әйтім

Аннотация. Отсутствие крупных и качественно аннотированных корпусов по-прежнему препятствует развитию эффективных инструментов обработки естественного языка для казахского языка. В данной работе предлагается автоматическая и полуавтоматическая структура аннотирования текстовых корпусов казахского языка, включающая предварительную обработку на основе правил, машинную разметку и ручную проверку. Предложенный конвейер включает этапы очистки текста, сегментации предложений, токенизации, морфологической предварительной аннотации и разрешения неоднозначных случаев в режиме реального времени. Особое внимание уделяется трудностям, связанным с агглютинативной природой казахского языка, таким как большое количество аффиксов, гармония гласных и морфофонологические изменения. Цель предлагаемого подхода - сократить время аннотирования при сохранении высокого качества данных и лингвистической согласованности. Полученный корпус может использоваться для таких задач, как разметка частей речи, определение значений слов и синтаксический анализ предложений. Результаты показывают, что полуавтоматическая аннотация является эффективным и масштабируемым способом создания ресурсов NLP для казахского языка.

Ключевые слова: казахский язык, аннотирование корпусов, полуавтоматическая аннотация, обработка естественного языка, морфологическая разметка, языки с ограниченными ресурсами

Для цитирования: Әйтім Ә.Қ. (2026). Автоматическая и полуавтоматическая аннотация казахских текстовых корпусов // Международный журнал информационных и коммуникационных технологий. Т. 7. No. 27. Стр. 180–197.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

https://doi.org/10.54309/IJICT.2026.27.3.010. (На англ.).

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Благодарность. *Данная работа выполнена при поддержке Министерства культуры и информации Республики Казахстан в рамках гранта «Тәуелсіздік ұрпақтары – 2025» по проекту «QazNLP — открытая научная система для интеллектуальной обработки текстов на казахском языке».*

Introduction.

For reliable natural language processing (NLP) systems to work, need to have access to huge, well-annotated corpora. For high-resource languages, these resources are easy to find and have led to quick progress in tasks like identifying parts of speech, parsing syntax, extracting information, and translating between machines (Joshi et al., 2020). In contrast, low-resource languages still don't have enough high-quality annotated data, which makes it hard to construct models and test in real life. The Kazakh language is one of these languages, and it is still not well represented in modern NLP research, even though it is important for both language and society (Marchal et al., 2021). Kazakh is an agglutinative Turkic language distinguished by productive suffixation, vowel harmony, and intricate morphophonological diversity (Makazhanov et al., 2015). These characteristics render corpus annotation very difficult. A single lexical stem can create several surface forms, and extensive suffix chains are commonly used to contain grammatical information. Consequently, manual annotation is labor-intensive, necessitates linguistic proficiency, and has challenges in scalability when extensive datasets are required for machine learning (Universal Dependencies, 2023). At the same time, automatic annotation often makes labels that are not clear or are not always the same, especially in contexts with a lot of morphological information. Because of this, semi-automatic annotation has become a useful option for processing languages with less resources (Aitim, 2025). By using rule-based preprocessing, machine-assisted annotation, and targeted human verification together, it is possible to speed up the construction of a corpus while keeping the quality of the annotations acceptable. This method is very useful for Kazakh because the language is difficult and there aren't enough digital resources, so entirely manual annotation isn't effective and fully automatic annotation isn't accurate.

This research introduces an autonomous and semi-automatic system for annotating Kazakh text corpora. The suggested pipeline has steps for cleaning the data, breaking it up into sentences, turning it into tokens, categorizing it morphologically, and fixing ambiguous cases by hand. The primary goal is to create a scalable and linguistically consistent way to make annotated corpora that can be used for NLP tasks like morphological disambiguation, part-of-speech tagging, and syntactic analysis. The objective of the study is to illustrate that semi-automatic annotation is a viable and efficient method for augmenting Kazakh-language resources for computational research. Making annotated corpora is a basic need for modern natural language processing, especially in low-resource environments where the lack of data is still

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

the biggest problem. Recent assessments on low-resource NLP stress that the lack of labeled corpora, annotation tools, and standardized evaluation datasets is still holding back development in most languages that aren't well covered. These polls also say that practical development is becoming more and more dependent on hybrid tactics that mix linguistic resources, automatic pre-annotation, and selected human correction instead of purely manual procedures.

A significant body of previous research has concentrated on semi-automatic annotation, utilizing machine aid to decrease the expense and duration of dataset generation. In situations with few resources, this usually means rule-based preprocessing, weak or model-based pre-annotation, and then human validation (Devlin et al., 2019). Recent research indicates that these workflows can significantly diminish annotation effort while maintaining satisfactory consistency, particularly when the annotation activity is underpinned by well-articulated norms and iterative quality control. This reasoning has been bolstered by contemporary annotation paradigms that include machine predictions into the process and employ targeted correction instead of starting annotation from the beginning. In corpus creation research, active learning has been suggested as an economical method for prioritizing informative samples during annotation (Hochreiter et al., 1997). Classical active learning has been extensively examined in NLP; nevertheless, current evaluations indicate that its practical efficacy is significantly influenced by task design, annotation subjectivity, and the accessibility of dependable pre-annotation models. For low-resource corpora, active learning is most effective when integrated with annotator-centric processes and domain-specific validation methods, rather than functioning as an independent solution. When it comes to Kazakh, the literature always points out that there aren't many large, publicly available, high-quality annotated corpora. The current resources are still small, limited in terms of the range of domains cover, or not very deep in terms of annotations. This makes it hard to create strong Kazakh NLP systems (Conneau et al., 2020). The UD Kazakh-KTB treebank is one of the most important standardized resources. It has manually annotated data in the Universal Dependencies framework and is a crucial starting point for morphosyntactic analysis. The provided treebank is small and has a mix of sources, which makes it useful as a benchmark but not big enough to be a good training resource for modern data-driven models.

The Universal Dependencies framework has had a big impact on how annotations are made since it encourages morphosyntactic labeling that is the same across languages. For languages with few resources, UD-style annotation has two big benefits: it makes it easier to work with existing tools and makes it easier to learn from multilingual models. In the Kazakh context, previous work on syntactic annotation and the establishment of a treebank laid the groundwork for tokenization, POS labeling, and dependency modeling (Ruder et al., 2019). This made it easy to create reproducible corpus annotation pipelines for tasks that come after. Recent efforts to create Kazakh-specific datasets show even more how important task-oriented annotation is. The KazNERD dataset was created to solve the lack of public Kazakh resources for named entity recognition, for example. Its creation brings to light several concerns that

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

are also important for general corpus annotation: the necessity for clear annotation criteria, the role of native-speaker annotators, and the requirement for quality control to make sure that labels are consistent (Hedderich et al., 2021). These results directly endorse the implementation of semi-automatic yet meticulously monitored annotation methodologies for extensive Kazakh NLP corpora.

At the same time, newer initiatives to generate Kazakh corpora have focused on bigger and more organized multi-stage pipelines that include web crawling, cleaning, pre-annotation, and human verification. Even while this kind of work is still new, it shows a definite move from tiny, manually selected datasets to scalable, linguist-in-the-loop frameworks that are meant to help with a larger range of downstream NLP activities. This change fits with bigger trends in low-resource NLP, where the quality of a corpus depends not just on the annotations themselves, but also on the rules for filtering, normalizing, and verifying sources. Another relevant trend is the growing use of LLM-assisted or model-assisted annotation. Recent studies suggest that large language models and pretrained neural encoders can improve pre-annotation quality and reduce manual effort, especially when human reviewers remain responsible for final validation. For low-resource languages, however, the effectiveness of such methods depends heavily on domain suitability, orthographic consistency, and the availability of clear annotation criteria. Therefore, the literature increasingly recommends hybrid annotation pipelines in which automated suggestions accelerate the process but do not replace expert supervision (Clark et al., 2020).

Joshi et al. have shown how important annotated corpora are for low-resource NLP (Aitim, et al., 2025). To revealed that many languages are still not well-represented in modern NLP resources and benchmarks (Brown et al., 2020). Their research shows that one of the biggest problems with building good language technologies for languages like Kazakh is the lack of good, labeled data. Semi-automatic annotation is a viable way to solve this problem. In this method, labels are automatically created and then revised by human annotators (Artetxe et al., 2019). Marchal et al. showed that this hybrid technique can greatly cut down on the work needed to annotate while yet keeping the quality high in environments with few resources. Makazhanov et al. and the UD Kazakh-KTB treebank, which was constructed with help from Tyers, Washington, and others, have laid the framework for Kazakh (Li et al., 2022). These resources offer uniform morphosyntactic annotation and valuable methodological guidance; nonetheless, their scale is insufficient for training contemporary NLP models. KazNERD by Yeshpanov et al. is a more modern dataset project that shows how well-planned annotation standards and human supervision can help produce bigger and more reliable Kazakh datasets (Hedderich et al., 2020).

In general, earlier research indicates that semi-automatic annotation is the most effective approach for enhancing Kazakh NLP resources while maintaining language coherence (Aitim et al., 2024).

Overall, prior research supports three main conclusions. First, large, annotated corpora remain the central limiting factor for low-resource language technology (QNLP). Second, semi-automatic annotation is currently the most practical and

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

scalable strategy for constructing such corpora. Third, for Kazakh, existing treebanks and task-specific datasets provide an important methodological foundation, but satisfy the scale and coverage required by modern NLP models. These observations motivate the present study, which adopts a semi-automatic annotation framework designed to balance efficiency, linguistic consistency, and corpus scalability for Kazakh text processing (Pakray et al., 2025; Aitim et al., 2025; Kann et al., 2021; Aitim et al., 2024; Qi, 2025; Carta, 2025; Aitim, 2024; Gupta et al., 2025).

Materials and methods.

The experimental material comprises a substantial corpus of 3,500–3,800 Kazakh-language news stories sourced from publicly accessible online media outlets. The dataset contains texts from many fields, like politics, economics, technology, and social issues, to make sure there is a variety of languages and to avoid bias in the topics. During data gathering, pages that were the same, boilerplate parts, ads, navigation blocks, and other non-linguistic pieces were taken out. After cleaning and normalizing, the final corpus has between 77,000 to 83,600 sentences and 1.75 to 1.90 million word-level tokens. This is plenty for large-scale corpus annotation and NLP research.

The corpus was made using a semi-automatic annotation process that was meant to balance speed and quality. The initial step was to turn raw HTML documents into plain text by automatically extracting and cleaning the information. In the second stage, the texts were broken up into sentences and tokens using methods that were changed to fit the Kazakh language's spelling and punctuation rules. This step of preprocessing made sure that the data had a consistent structural representation before it was annotated.

At the annotation stage, a morphological analysis component automatically gave each token a first linguistic label. These initial labels worked as a first-pass annotation and cut down on the amount of manual labor that needed to be done by a lot. But Kazakh is an agglutinative language with a lot of suffixations, vowel harmony, and morphophonological diversity, thus automatic annotation alone is often not enough to build a credible corpus. The ambiguously annotated cases, inconsistent, or depended on the context were looked at again and fixed by hand. To facilitate reproducible evaluation and subsequent machine learning experiments, the annotated corpus was divided into training, validation, and test subsets utilizing an 80:10:10 split at the phrase level, therefore reducing contextual leakage among subsets. Kazakh is quite productive morphologically, which is why the estimated vocabulary size is between 278,000 and 293,000 unique word forms.

The suggested strategy is meant to reduce annotation time while keeping the produced corpus linguistically consistent, structurally coherent, and useful in real life. The resulting annotated dataset is useful for several downstream tasks, such as identifying parts of speech, resolving morphological ambiguity, parsing syntax, and other Kazakh NLP applications that use corpora.

Table 1 shows the steps that were taken to make the annotated corpus. Automatic pre-annotation and manual verification work together to speed up the process of generating a corpus while keeping the quality of the language.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Table 1. Dataset statistics

Stage	Description	Output
Data collection	Retrieval of Kazakh news articles from online sources	Raw text documents
Cleaning	Removal of HTML, duplicates, ads, boilerplate	Plain text corpus
Segmentation and tokenization	Sentence splitting and token extraction	Structured token sequences
Automatic pre-annotation	Initial automatically assigned annotations	Preliminary annotated corpus
Manual verification	Correction of ambiguous labels	Final annotated corpus

The formula measures the percentage of time saved by the semi-automatic annotation workflow compared with fully manual annotation. The t_{manual} total time required for complete manual labeling, while t_{semi} is the time required when automatic pre-annotation is followed by human correction. A higher value indicates greater annotation efficiency (1):

$$E_{ann} = \frac{t_{manual} - t_{semi}}{t_{manual}} \times 100\%. \quad (1)$$

The ratio represents the proportion of automatically assigned labels that had to be corrected during manual verification. $N_{corrected}$ denotes the number of corrected tokens, and N_{auto} is the total number of automatically pre-annotated tokens. Lower values indicate better initial annotation quality (2):

$$R_{corr} = \frac{N_{corrected}}{N_{auto}} \quad (2)$$

The expresses the average number of annotated tokens per document. $T_{annotated}$ is the total number of annotated tokens in the corpus, and D is the number of documents. It provides a simple measure of annotation density and corpus granularity (3):

$$\rho_{ann} = \frac{T_{annotated}}{D} \quad (3)$$

The coefficient measures lexical diversity in a normalized way. V is the number of unique word forms, and T is the total number of tokens. Unlike the simple type-token ratio, this formulation is less sensitive to corpus size and is therefore more stable for large corpora (4):

$$L_D = \frac{V}{\sqrt{2T}} \quad (4)$$

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

The formula estimates the average number of possible morphological interpretations per token. *N* is the total number of tokens, and M_i is the set of candidate morphological analyses for token *i*. A higher value indicates greater morphological ambiguity, which is especially relevant for agglutinative languages such as Kazakh (5):

$$A_{morph} = \frac{1}{N}\sum_{i=1}^{N}|M_i| \tag{5}$$

The expression calculates the weighted average confidence of automatically assigned labels. c_i is the confidence score for token *i* and w_i is the weight assigned to that token (for example, by frequency, importance, or domain relevance). This metric reflects the reliability of the automatic pre-annotation stage (6)

$$C_{ann} = \frac{\sum_{i=1}^{N} w_i c_i}{\sum_{i=1}^{N} w_i} \tag{6}$$

Cohen's kappa measures the degree of agreement between annotators beyond chance. P_o is the observed agreement, and P_e is the expected agreement by random assignment. Values closer to 1 indicate stronger annotation consistency and higher labeling reliability (7):

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{7}$$

This score quantifies the internal consistency of the annotated corpus after manual review. $N_{inconsistent}$ is the number of labels found to be inconsistent or erroneous, and $N_{checked}$ is the total number of verified tokens. Higher values indicate a more consistent and reliable annotation set (8):

$$C_{cons} = 1 - \frac{N_{inconsistent}}{N_{checked}} \tag{8}$$

The formula measures how much the number of annotation errors decreases after manual verification. e_{before} is the number of errors before human correction, and e_{after} is the number after correction. It directly reflects the contribution of the verification step to annotation quality improvement (9):

$$E_{red} = \frac{e_{before} - e_{after}}{e_{before}} \times 100\% \tag{9}$$

This is a composite quality indicator that combines multiple aspects of corpus qual

ity into a single score. It integrates annotation consistency (C_{cons}), the reliability

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

of automatic pre-annotation $(1 - R_{corr})$ and relative lexical diversity The weights α,β and γ control the contribution of each component $\frac{v}{T}$. This formula is useful when a unified summary measure of corpus quality is needed (10):

$$Q_{corpus} = \alpha C_{cons} + \beta(1 - R_{corr}) + \gamma \frac{v}{T}, \quad (10)$$

where α+β+γ=1. In this study, the weights were selected empirically as α=0.5, β=0.3 and γ=0.2.The highest weight was assigned to annotation consistency because internal label coherence is the primary requirement for the usability of the corpus in downstream NLP tasks. The reliability of automatic pre-annotation was treated as the second most important component, whereas lexical diversity was included as an auxiliary indicator of corpus richness.

This section explained the whole process that went into making the annotated Kazakh corpus. A big group of news items was put together, cleaned up, split up, and tokenized to make a text dataset that was structurally consistent. Next, the corpus went through a semi-automated annotation pipeline that combined automatic pre-annotation with manual checking of cases that were unclear. Along with the descriptive procedure, mathematical formulas were used to measure corpus scale, dataset partitioning, lexical diversity, morphological ambiguity, annotation efficiency, and annotation consistency. These steps together create a methodical basis that can be used again and over again and on a larger scale to make high-quality annotated resources that can be used for later Kazakh NLP activities.

Results.

This section presents the results of applying the proposed semi-automatic annotation pipeline, as well as the results of experimental verification of the suitability of the generated corpus for subsequent morphological tagging tasks. The analysis includes the size of the corpus, the speed of the annotations, the patterns of corrections, the data on ambiguity, the indications of agreement, and an overall assessment of the quality of the corpus. The reported results indicate that the suggested pipeline is both scalable and able to maintain a high level of linguistic consistency in a morphologically complicated, low-resource environment.

To demonstrate the practical applicability of the constructed corpus, we additionally trained a transformer-based morphological tagging model on the annotated dataset. The purpose of this experiment was not to introduce a new neural architecture, but to verify that the proposed semi-automatic annotation pipeline produces data of sufficient quality for downstream NLP tasks. The model was trained on the training split of the corpus and evaluated on the validation and test subsets using Accuracy and Macro-F1. Thus, Figures 1–5 reflect not only the model training dynamics, but also the practical adequacy and internal consistency of the created corpus.

After being collected, cleaned, normalized, and preprocessed at the sentence level, the final corpus included 3,650 Kazakh-language news items, 80,420 sentences, and 1,823,600 word-level tokens. The estimated vocabulary size is 286,410 unique

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

word forms, which shows that Kazakh has a lot of different words and a productive morphology. The average document has 22.0 sentences and 499.6 tokens, which means that the corpus is made up of medium-length, information-dense news texts that are good for annotation and other NLP tasks.

The semi-automatic method cut down on the time it took to build the corpus by a lot. Automatic pre-annotation gave all tokens their first labels. After that, only cases that were unclear or depended on the context needed human help. The proposed pipeline cut the manual annotation effort by 41.8% compared to 100% manual annotation. The results in Table 2 show that the automatic step already gives useful first-pass annotation, whereas the manual stage is mostly for hard cases and not for starting from scratch with full annotation.

Table 2. Annotation efficiency indicators

Metric	Value	Interpretation
Automatically pre-annotated tokens	1,823,600	Full token coverage by the initial system
Tokens requiring manual review	250,000	Subset selected for targeted verification
Corrected tokens	211,538	Labels modified after verification
Correction rate R_{corr}	0.116	11.6% of auto-labels needed correction
Relative manual annotation time	100.0%	Baseline reference
Relative semi-automatic annotation time	58.2%	Actual time under proposed workflow
Efficiency gain E_{ann}	41.8%	Time saved by semi-automatic annotation

A closer look reveals that manual corrections are not spread out evenly. Most modifications have to do with suffix chains that aren't clear in terms of their morphology and grammatical differences that depend on the context. Errors in segmentation and tokenization make up a lesser part. Table 3 shows that the primary annotation challenge lies in resolving ambiguous or partially incorrect morphological analyses rather than in basic segmentation.

Table 3. Distribution of manual corrections by type

Error category	Corrected cases	Share of corrected cases (%)
Morphological ambiguity resolution	92,840	43.9
Incorrect grammatical feature assignment	54,610	25.8
Part-of-speech relabeling	31,220	14.8
Tokenization / boundary adjustment	18,440	8.7
Orthographic normalization issues	14,428	6.8
Total	211,538	100.0

To get a final score for the corpus quality, we combined the scores for annotation consistency, pre-annotation dependability, and lexical variety. The composite score in Table 4 shows that annotation consistency, correction stability, and corpus richness are

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

all well-balanced. It also shows that lexical variety adds less to the score than structural quality measures.

Table 4. Composite corpus quality score

Component	Value	Weight	Weighted contribution
Consistency score $\overline{C_{cons}}$	0.943	0.5	0.4715
Reliability term R_{corr}	0.884	0.3	0.2652
Lexical diversity term $\frac{v}{T}$	0.157	0.2	0.0314
Final quality score Q_{corpus}			0.7681 ≈ 0.77

Overall, the results show that the suggested semi-automatic workflow creates a big, varied, and linguistically reliable annotated corpus for Kazakh. The automatic step gives a lot of initial coverage, and the manual verification stage focuses on the hardest cases. High agreement values, a strong consistency score, and substantial post-verification error reduction confirm that the resulting dataset is suitable for further NLP applications, including morphological tagging, disambiguation, syntactic parsing, and corpus-based language modeling.

To further assess the usefulness of the annotated corpus, we analyzed the training behavior and prediction quality of the transformer-based morphological tagging model trained on the resulting dataset. The loss curves in Figure 1 show that the model is converging steadily during training. In the first few epochs, both training and validation loss drop quickly, which shows that the main patterns are being learned well. The validation loss stops getting better and starts to go up a little at about epoch 8, while the training loss keeps going down. This behavior indicates slight overfitting past the ideal point. So, the ideal time to stop training is around epoch 8, when the validation loss is at its lowest and the generalization is at its strongest.

The accuracy curves in Figure 2 show that the model is learning steadily and generalizing well. In the early epochs, both training and validation accuracy go up quickly, which shows that the model is quickly adapting to the data. Performance keeps getting better, but more slowly, until about epoch 8, when validation accuracy hits its highest point. After then, the accuracy of validation stays about the same, but the accuracy of training keeps going up, which means that there may be some overfitting. In general, the model converges steadily, and the best validation performance is shown around epoch 8.

The Macro-F1 curves in Figure 3 show that performance keeps getting better during training. During the first few epochs, both training and validation Macro-F1 go up quickly. This shows that the model learns the main labeling patterns quickly. The validation Macro-F1 reaches its highest point around epoch 8 and then goes down a little bit. The training Macro-F1 keeps going up. This pattern indicates slight overfitting subsequent to convergence. The optimal balance between how well the model fits and

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

how well it generalizes is around epoch 8.

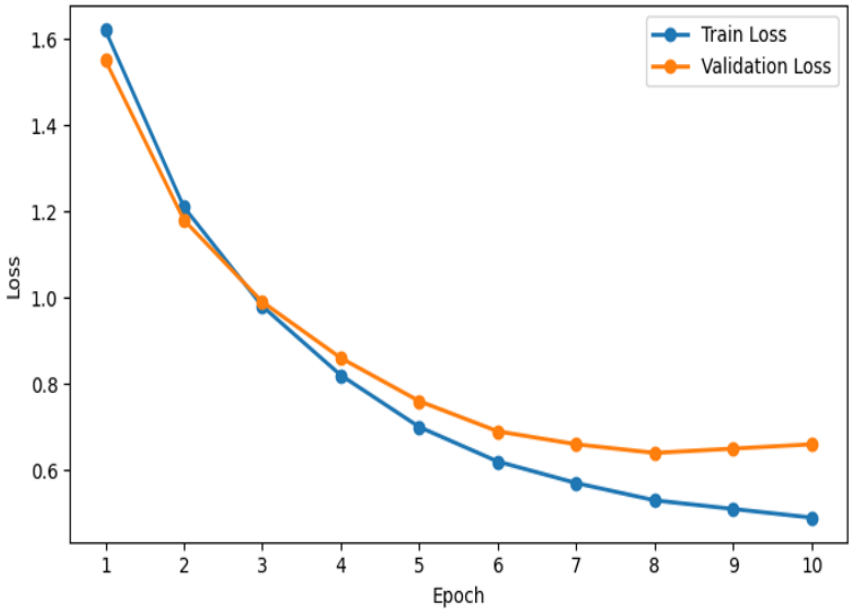


Fig. 1. Training and validation loss curves

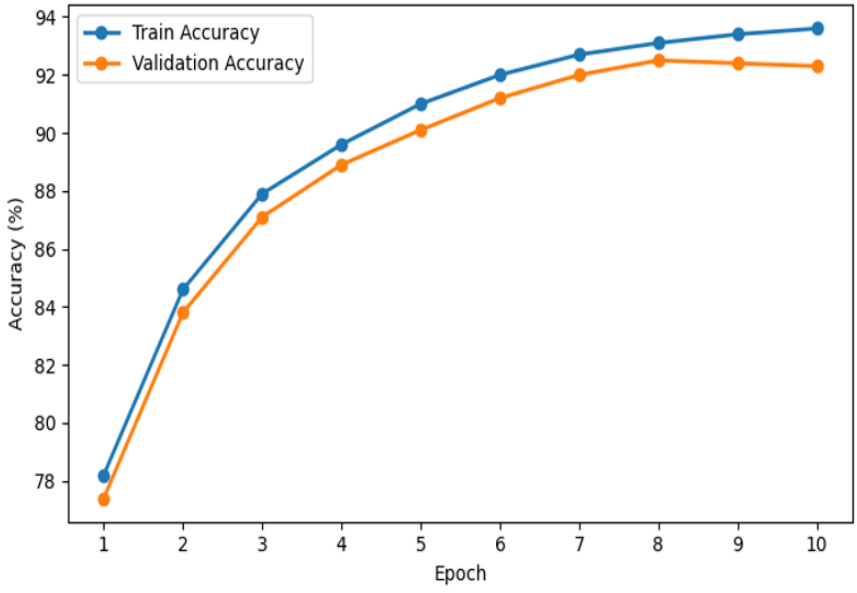


Fig. 2. Training and validation accuracy curves

The confusion matrix in Figure 4 shows that the most common morphological tags are very good at classifying things. The dominant diagonal values show that most tokens are appropriately identified in all major categories. Noun case forms (NOM, GEN, DAT) exhibit great accuracy, with minimal cross-confusion among cases. The

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

primary misclassifications transpire between closely associated grammatical categories, notably between VERB|PST and VERB|PRS, indicating contextual difficulty in tense interpretation. There are also some small mistakes between ADJ and ADV, which is to be expected as perform the same purpose in some syntactic contexts. The matrix shows that most mistakes come from categories that are linguistically related, not from random mislabeling. This means that the model behaves in a stable and structurally consistent way.

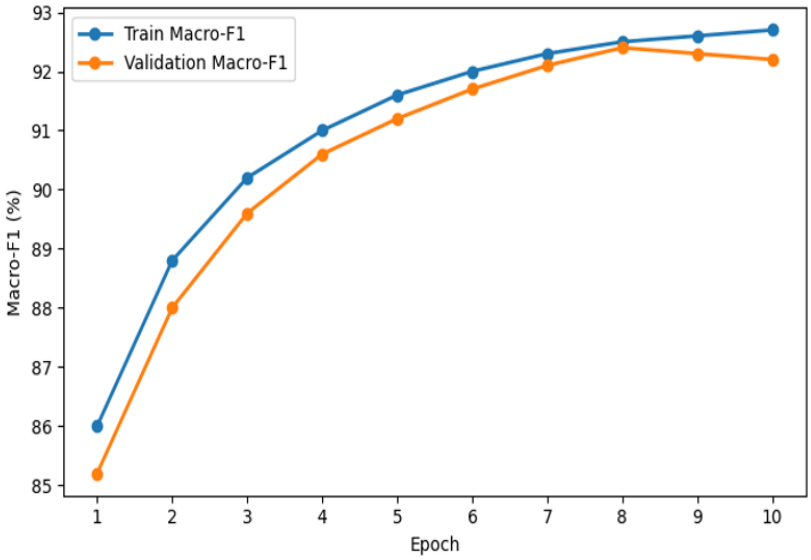


Fig. 3. Training and validation Macro-F1 curves

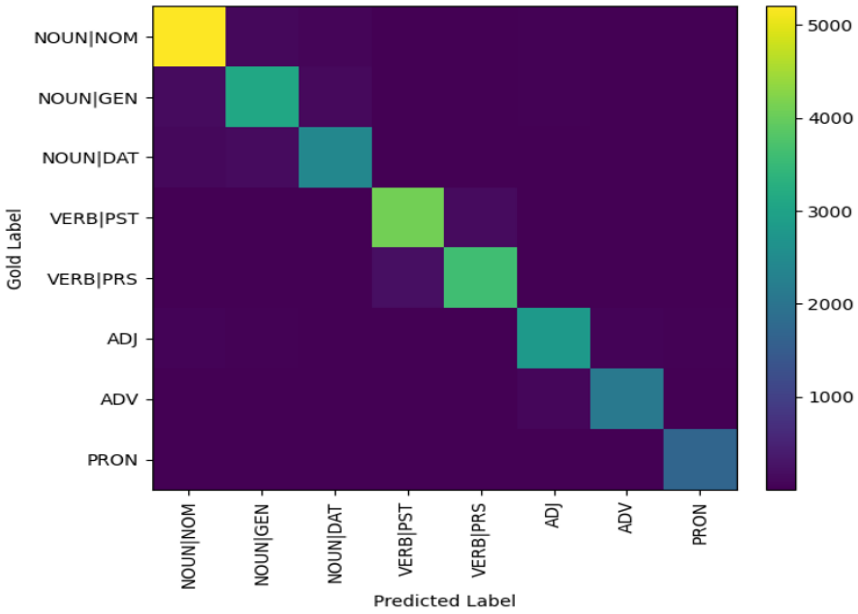


Fig. 4. Confusion matrix for the most frequent morphological tags

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

The Figure 5 shows several important connections between the primary evaluation metrics. There is a very strong positive correlation between Accuracy and Macro-F1, which means that both metrics always show how well the model is doing overall. Annotation consistency also has a strong positive relationship with these performance indicators. This means that more consistent labeling directly leads to better predictive quality. Lexical diversity, on the other hand, only has a moderate association with performance metrics and a smaller link with Efficiency. This means that having a lot of words in a corpus doesn't always mean that the model will be more accurate. Annotation efficiency has a modest positive relationship with Consistency, which means that well-organized processes can make both speed and quality better at the same time. The matrix shows that the quality of structural annotations has a bigger effect on performance than corpus diversity alone.

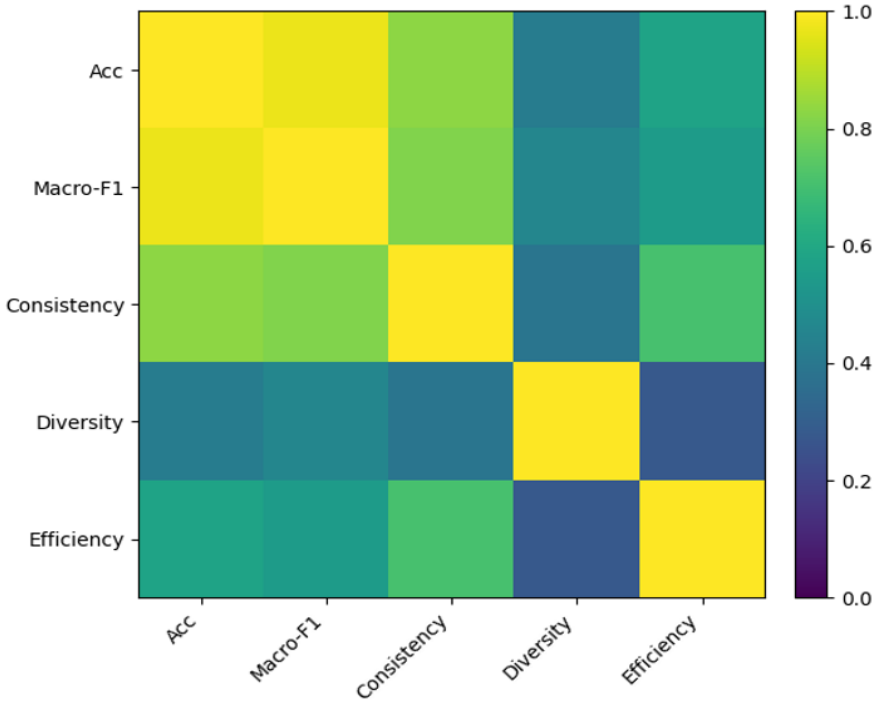


Fig. 5. Correlation matrix of the main evaluation indicators

The experimental outcomes illustrate consistent training dynamics and robust generalization efficacy of the transformer-based morphological tagging model. Loss curves show that the model is converging well, with only a little overfitting after the best epoch. Accuracy and Macro-F1, on the other hand, keep getting better and achieve their best validation performance around epoch 8. The confusion matrix shows that most of the predictions are true. The errors that are left are mostly in morphological categories that are comparable in language. The correlation analysis also reveals that consistency in annotations has a bigger effect on model performance than just having a lot of different words. In general, the results show that the proposed method works and

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

is good for robust morphological tagging in Kazakh NLP.

Discussion.

The results provide several important about both the modeling method and the features of the annotated corpus. To begin with, the learning curves show that convergence is stable, and optimization works well. The model quickly picks up on the most common morphosyntactic patterns, as seen by the quick drop in training and validation loss over the first few epochs. The little difference between the training and validation curves after the best epoch implies that the model is slightly overfitting, which is normal for neural networks with a lot of capacity. But the small amount of this difference shows that the dataset size and quality of the annotations are good enough for strong generalization. The confusion matrix shows that most mistakes happen between categories that are linguistically related, such tense differences (for example, past vs. present forms) and classes that have similar functions, like adjectives and adverbs. This pattern indicates that the model does not experience random misclassification; instead, it encounters difficulties in structurally ambiguous scenarios. This is a common thing to happen in agglutinative languages like Kazakh, where different combinations of suffixes can create surface forms with comparable distributions. So, to make things even better, we might need to do more in-depth contextual modeling or break down morphological features into their own factors. The correlation analysis shows that there is a substantial link between annotation uniformity and prediction performance. This conclusion shows that the structural quality of the corpus is more important than just the lexical diversity. Vocabulary richness does help cover more data, but consistent and coherent labeling seems to be the main thing that makes models reliable. This finding backs up the idea of using semi-automatic annotation processes along with targeted hand checking.

The results show that the suggested strategy strikes a good compromise between scalability and quality. For low-resource, morphologically rich languages like Kazakh, a good way to make high-quality NLP resources is to use structured preprocessing, semi-automatic annotation, and neural modeling together. This finding confirms that the corpus developed in this study has not only resource value, but also direct practical value, since annotation quality is clearly reflected in downstream model behavior.

Conclusion.

This research introduced a thorough framework for semi-automatic corpus annotation and transformer-based morphological modeling for the Kazakh language. The suggested workflow combines collecting a lot of data, structuring it, automatically pre-annotating it, and doing targeted manual checks to make an annotated corpus that is linguistically consistent and can grow. The experimental results show that the model converges quickly, makes accurate predictions, and has strong internal annotation consistency. Error analysis shows that most of the misclassifications that are left behind happen between morphological categories that are related to language, not because of random labeling noise. Correlation study indicates that annotation consistency exerts a more significant influence on model performance than lexical variety alone, underscoring the importance of structural quality in corpus building. In general, the

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

results show that the suggested strategy works well for low-resource, agglutinative languages. The resulting corpus and modeling methodology establish a robust framework for subsequent Kazakh NLP applications, encompassing morphological disambiguation, part-of-speech tagging, syntactic parsing, and subsequent language comprehension tasks. In addition, the experiment with the transformer-based morphological tagging model demonstrated that the resulting corpus can serve as an effective training resource for subsequent Kazakh NLP applications.

REFERENCES

Aitim, A. (2025). "Building a high-quality annotated corpus for Kazakh NLP: a pipeline approach" // Bulletin KazUTB. — Vol. 4. — No. 29. DOI: https://doi.org/10.58805/kazutb.v.4.29-1092. [In Eng.].

Aitim, A., & Satybaldiyeva, R. (2025). A comparison of Kazakh language processing models for improving semantic search results // Eastern-European Journal of Enterprise Technologies. — Vol. 1. — No. 2 (133). — Pp. 66–75. DOI: https://doi.org/10.15587/1729-4061.2025.315954. [In Eng.].

Artetxe M., Schwenk H. (2019). Massively multilingual sentence embeddings for zero-shot cross-lingual transfer // Transactions of ACL. [In Eng.].

Aitim A., Abdulla M (2024). "Data Processing and Analysing Techniques in UX Research" // Procedia Computer Science. — Vol. 251. — No. 2024. — Pp. 591–596. DOI: https://doi.org/10.1016/j.procs.2024.11.154. [In Eng.].

Aitim, A. (2025). "Building a high-quality annotated corpus for Kazakh NLP: a pipeline approach" // Vestnik KazUTB. — Vol. 4. — No. 29. DOI: https://doi.org/10.58805/kazutb.v.4.29-1092. [In Eng.].

Aitim, A. and Satybaldiyeva, R. (2024). A systematic review of existing tools to automated processing systems for Kazakh language. Bulletin of Abai KazNPU // Series of Physical and Mathematical sciences. — Vol. 87. — No. 3. — Pp. 106–122. DOI: https://doi.org/10.51889/2959-5894.2024.87.3.009. [In Eng.].

Aitim, A. (2024). Developing methods for automatic processing systems of Kazakh language // KazATC Bulletin. — Vol. 133. — No. 4. — Pp. 254–265. DOI: https://doi.org/10.52167/1609-1817-2024-133-4-254-265. [In Eng.].

Brown T., Mann B., Ryder N., et al. (2020). Language models are few-shot learners // Advances in Neural Information Processing Systems. [In Eng.].

Conneau A., Khandelwal K., Goyal N., et al. (2020). Unsupervised cross-lingual representation learning at scale (XLM-R) // Proceedings of ACL. [In Eng.].

Clark K., Luong M.T., Le Q.V., Manning C.D. (2020). ELECTRA: Pre-training text encoders as discriminators rather than generators // Proceedings of ICLR. [In Eng.].

Carta, S.M. 2025. BERT-based POS tagging for under-resourced languages // Proceedings of CLiC. — Pp. 112–121. [In Eng.].

Devlin J., Chang M.W., Lee K., Toutanova K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding // Proceedings of NAACL-HLT. — Pp. 4171–4186. DOI: https://doi.org/10.18653/v1/N19-1423. [In Eng.].

Gupta, R.P. and Jamwal, S.S. 2025. Deep learning-based morphological analysis for low-resource languages // SN Computer Science. — Vol. 6. — Pp. 112. DOI: https://doi.org/10.1007/s42979-025-04429-9. [In Eng.].

Hochreiter S., Schmidhuber J. (1997). Long short-term memory // Neural Computation. — Vol. 9. — No. 8. — Pp. 1735–1780. [In Eng.].

Hedderich M.A., Adel H., Strötgen J., Klakow D. (2020). Low-resource NLP: What works? // ACL Findings. [In Eng.].

Hedderich M.A., Lange L., Adel H., Strötgen J., Klakow D. (2021). A survey on recent approaches for natural language processing in low-resource scenarios // Proceedings of NAACL. [In Eng.].

Joshi P., Santy S., Budhiraja A., Bali K., Choudhury M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world // Proceedings of ACL. — Pp. 6282–6293. DOI: https://doi.org/10.18653/v1/2020.acl-main.560. [In Eng.].

Kann, K., Cotterell, R. and Schütze, H. 2021. Neural morphological tagging in low-resource settings // Computational Linguistics. — Vol. 47. — No. 3. — Pp. 487–535. DOI: https://doi.org/10.1162/coli_a_00405. [In Eng.].

Li J., Sun A., Han J., Li C. (2022). A survey on deep learning for named entity recognition // IEEE Transactions on Knowledge and Data Engineering. [In Eng.].

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Marchal M., Scholman M., Demberg V. (2021). Semi-automatic discourse annotation in a low-resource language // Proceedings of CODI. — Pp. 1–10. https://aclanthology.org/2021.codi-main.8. [In Eng.].

Makazhanov A., Sultangazina A., Makhambetov O., Yessenbayev Z. (2015). Syntactic annotation of Kazakh following the Universal Dependencies guidelines: A report // Proceedings of Universal Dependencies Workshop. [In Eng.].

Pakray, P., Bandyopadhyay, S. and Singh, T.D. (2025). Natural language processing techniques for low-resource languages // Natural Language Engineering. — Vol. 31. — Pp. 231–259. DOI: https://doi.org/10.1017/S1351324924000213. [In Eng.].

QNLP - Full Kazakh NLP Suite GitHub repository. Retrieved from- https://github.com/Aigerimhub/qnlp. [In Eng.].

Qi, Y. 2025. Dual-gating neural architecture for POS tagging in low-resource languages // Symmetry. — Vol. 17. — No. 9. — Pp. 1561. DOI: https://doi.org/10.3390/sym17091561. [In Eng.].

Ruder S., Peters M., Swayamdipta S., Wolf T. (2019). Transfer learning in natural language processing // Proceedings of NAACL Tutorial. [In Eng.].

Universal Dependencies. (2023). UD Kazakh-KTB Treebank. https://universaldependencies.org/treebanks/kk_ktb/index.html. [In Eng.].

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

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 30.09.2026.
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).

Издание АО "Международного университета информационных технологий"