

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. 2 (26). 2026  
April–June

**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ  
ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

2020 жылдан бері шығарылады  
Том 7. 2 (26). 2026  
Сәуір-Маусым

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

Издается с 2020 г.  
Том 7. 2 (26). 2026  
Апрель-Июнь

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)

**Bakhtgerci 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)  
**Луччо 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)  
**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 Токсулжаевич** — 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 Абдуллаевна** — 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)  
**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)  
**Дамелия Максустовна Ескендрова** — 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)  
**Талеуш 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)  
**Бушув Сергей Дмитриевич** — техника ғылымдарының докторы, профессор, Украинаның "УКРНЕТ" жобаларды басқару қауымдастығының директоры, Киев ұлттық құрылыс және сулет 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)

**Дамелия Максуговна Ескендрова** — 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)

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

**Алимурабаев 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)

**Белошницкая 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 социально-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

**D. Abzhanova, A. Biloshchytski**

A MODEL AND METHOD FOR MANAGING DATA ON EMISSIONS FROM STATIONARY SOURCES OF POLLUTION IN AN INTELLIGENT ENVIRONMENTAL MONITORING SYSTEM .....9

**A. Slanbekova, M. Rakhimzhanova, A. Zhanibekova, A. Alimagambetova, M. Xudoyberganov**

EARLY DETECTION OF HYDROLOGICAL HAZARDS BASED ON SPATIOTEMPORAL ANALYSIS .....25

## INFORMATION TECHNOLOGY

**F.N. Abdraimova, A.A. Kereibayeva, D.S. Dyussenova, D.A. Aliyeva, T.Zh. Toktarova**

AI TECHNOLOGIES IN LANGUAGE EDUCATION: PRACTICAL ASPECTS AND CHALLENGES OF STUDENT USAGE .....36

**G. Azieva, M. Yessenova, A. Abzhapparova, G. Abdikerimova, P. Schmidt**

HYBRID STACKING FRAMEWORK FOR CROP CLASSIFICATION USING UAV DATA .....50

**A.K. Aitim**

JOINT MORPHOLOGICAL DISAMBIGUATION AND POS TAGGING FOR AGGLUTINATIVE LANGUAGES .....62

**S.A. Yesniyazova, S.T. Kaimov**

PREDICTIVE MAINTENANCE OF HEAVY-DUTY TRUCKS USING EXPLAINABLE MACHINE LEARNING .....78

**T. Imanbekova, Zh. Ibrayeva, G. Jakanova, G. Askanbay**

DATA COMPRESSION ALGORITHM BASED ON WAVELET TRANSFORMER; ANALYSIS AND IMPLEMENTATION IN MATLAB .....92

**B.Z. Kenzhegulov, Zh.T. Bilyalova, K.N. Uteuliyeva, L. Nurgaliyeva, Sh.S. Nurzhanova**

A MATHEMATICAL AND ALGORITHMIC APPROACH TO THE DEVELOPMENT OF AN INTELLIGENT TEXT-TO-SQL SYSTEM BASED ON LARGE LANGUAGE MODELS .....110

**N.Sh. Maxutova, J.A. Tussupov, A.A. Shekerbek, Zh.E. Kenzhebayeva, Q.O. Rakhimov**

MACHINE LEARNING FOR COMPREHENSIVE EVALUATION OF CARDIOVASCULAR DISEASE RISK AND BIOCHEMICAL ALTERATIONS: FOCUS ON ASPARTATE AMINOTRANSFERASE .....131

**O.S. Salykova, V.A. Madin, B.R. Salykov, D.N. Komarov, N.V. Manuilov**

INTEGRATION OF MEMS ACCELEROMETER SENSOR MODULES IN INDUSTRIAL MONITORING SYSTEMS .....146

**R. Taberkhan, M.A. Sambetbayeva, G. Kalman**

KAZCAUSAL: THE FIRST CORPUS-BASED ANNOTATION OF CAUSAL RELATIONSHIPS IN THE KAZAKH LANGUAGE .....160

**S.Tynymbayev, S.E. Mamanova, R. Berdybayev, Zh.E. Temirbekova, T. Chinibayeva**

DIVIDING DEVICES WITH PRELIMINARY PREPARATION OF MULTIPLES OF THE DIVISOR .....172

**K.N. Uteuliyeva, B.Z. Kenzhegulov, T.A. Karazhigitova, H.İ. Bülbül, Z.Zh. Zhanuzakova**

MATHEMATICAL AND ALGORITHMIC APPROACHES TO THE DEVELOPMENT OF A COLLABORATIVE FILTERING-BASED RECOMMENDER SYSTEM .....188

**S. Sharmukhanbet, G. Turmukhanova, O. Findik, V. Makhatova, L. Kurmangazyeva**

HIGH-PRECISION ROBOTIC ASSEMBLY UNDER VARIABLE ILLUMINATION: A ROBUST MECHATRONIC ARCHITECTURE FOR VISUAL SERVOING .....209

## INFORMATION SECURITY AND COMMUNICATIONTECHNOLOGIES

**A. Amirbay, Z. Amanbaikyzy, K. Maxutova, A. Mukhanova, M. Kassim**

MACHINE LEARNING ALGORITHM FOR EARLY DETECTION OF AUTISM SPECTRUM DISORDERS IN CHILDREN BASED ON MULTIMODAL ANALYSIS OF EYE MOVEMENTS AND FACIAL EXPRESSIONS .....227

**K. Baisylbayeva, Sh. Mussiraliyeva, Zh. Yeltay**

DETECTION OF EXTREMIST IDEOLOGY IN THE KAZAKH LANGUAGE: ANNOTATION CHALLENGES AND DEEP LEARNING APPROACHES .....242

**M.A. Bolatbek, A.M.Usmanova, K.B. Bagitova, G.B. Baispay**

DEVELOPMENT AND RESEARCH OF A METHOD FOR ANALYZING NETWORK TRAFFIC TO IDENTIFY A CYBER THREAT .....	261
<b>D.I. Prokopovych-Tkachenko, N.K. Zhumagalieva, D.N. Shchytyov, N.F. Mormul, D.A. Cherkaskyi</b> FUZZY MODEL FOR EVALUATING INFORMATION SECURITY PARAMETERS OF INFORMATION SYSTEMS UNDER INCOMPLETE AND QUALITATIVE DATA: CONSTRUCTION METHODOLOGY, RULE BASE TUNING, AND DEMONSTRATION CASE FOR ORGANIZATIONS .....	279
<b>E.A. Pustovoy, O.A. Pustovaya, A.N. Raushanova, I.S. Zaurbekov</b> EVALUATION OF THE EFFECTIVENESS OF SYNTHESIS OF STOCHASTIC MODELS WITH CONTROLLED PROPERTIES .....	305
<b>Y. Serzhan, T. Umarov, A. Abilbayeva</b> FRAUD DETECTION IN CREDIT CARD TRANSACTIONS USING MACHINE LEARNING: A COMPARATIVE ANALYSIS .....	321

## МАЗМҰНЫ

### ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

<b>Д.Е. Абжанов, А.А. Белоощицкий</b> ЭКОЛОГИЯЛЫҚ МОНИТОРИНГТІҢ ЗИЯТКЕРЛІК ЖҮЙЕСІНДЕГІ СТАЦИОНАРЛЫҚ ЛАСТАНУ КӨЗ-ДЕРІНІҢ ШЫҒАРЫНДЫЛАРЫ ТУРАЛЫ ДЕРЕКТЕРДІ БАСҚАРУДЫҢ МОДЕЛІ МЕН ӘДІСІ .....	9
<b>А.Е. Сланбекова, М.Б. Рахимжанова, А.И. Жанибекова, А.З. Алимагамбетова, М. Худойбергенов</b> КЕҢІСТІКТІК-УАҚЫТТЫҚ (SPATIOTEMPORAL) ТАЛДАУ НЕГІЗІНДЕ ГИДРОЛОГИЯЛЫҚ ҚАУІП-ҚАТЕРДІ ЕРТЕ АНЫҚТАУ .....	25

### АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

<b>Ф.Н. Абдранмова, А.А. Керейбаева, Д.С. Дюсенова, Д.А. Алиева, Т.Ж. Токтарова</b> ТІЛ БІЛІМІНДЕ ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫ: СТУДЕНТТЕР ҚОЛДАНУЫНЫҢ ПРАКТИКАЛЫҚ АСПЕКТІЛЕРІ МЕН МӘСЕЛЕЛЕРІ .....	36
<b>Г.Т. Азиева, М.Б. Есенова, А.К. Абжаппарова, Г.Б. Абдикеримова, Р. Schmidt</b> UAV ДЕРЕКТЕРІ НЕГІЗІНДЕ АУЫЛ ШАРУАШЫЛЫҒЫ DAҚЫЛДАРЫН ЖІКТЕУГЕ АРНАЛҒАН ГИБРИДТІ СТЕКИНГ МОДЕЛІ .....	50
<b>Ә.Қ. Әйтiм</b> АГГЛЮТИНАТИВТІ ТІЛДЕРГЕ АРНАЛҒАН МОРФОЛОГИЯЛЫҚ ДИЗАМБИГУАЦИЯ МЕН POS-ТАҢ-БАЛАУДЫ БІРЛЕСІП МОДЕЛЬДЕУ .....	62
<b>С.А. Есниязова, С.Т. Каимов</b> ТҮСІНДІРІЛЕТІН МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП АУЫР ЖҮК КӨЛІКТЕРІНЕ БОЛЖАМДЫ ТЕХНИКАЛЫҚ ҚЫЗМЕТ КӨРСЕТУ .....	78
<b>Т.Д. Иманбекова, Ж.Б. Ибраева, Г.Т. Джаканова, Г.Т. Асқанбай</b> МӨЛІМЕТТЕРДІ ВЕЙВЛЕТ-ТҮРЛЕНДІРГІШТІҢ НЕГІЗІНДЕ ҚЫСУ АЛГОРИТМІ; MATLAB ОРТАСЫНДА ТАЛДАУ ЖӘНЕ ІСКЕ АСЫРУ .....	92
<b>Б.З. Кенжегулов, Ж.Т. Билялова, К.Н. Утеулиева, Л. Нурғалиева, Ш.С. Нуржанова</b> ҮЛКЕН ТІЛДІК МОДЕЛЬДЕР НЕГІЗІНДЕ ИНТЕЛЛЕКТУАЛДЫ TEXT-TO-SQL ЖҮЙЕСІН ӨЗІРЛЕУДІҢ МАТЕМАТИКАЛЫҚ-АЛГОРИТМДІК ТӘСІЛІ .....	110
<b>Н.Ш. Максұтова, Ж.А. Тусупов, А.Ә. Шекербек, Ж.Е. Кенжебаева, К.О. Рахимов</b> ЖҮРЕК-ҚАН ТАМЫРЛАРЫ АУРУЛАРЫНЫҢ ҚАУІП-ҚАТЕРІН ЖӘНЕ БИОХИМИЯЛЫҚ ӨЗГЕРІСТЕРДІ КЕШЕНДІ БАҒАЛАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ: АСПАРТАМИНОТРАНСФЕРАЗАҒА ЕРЕКШЕ НАЗАР .....	131
<b>О.С. Салықова, В.А. Мадин, Б.Р. Салықов, Д.Н. Комаров, Н.В. Мануилов</b> ӨНЕРКӘСІПТІК МОНИТОРИНГ ЖҮЙЕЛЕРІНДЕГІ MEMS-АКСЕЛЕРОМЕТРЛЕРДІҢ СЕНСОРЛЫҚ МОДУЛЬДЕРІН ИНТЕГРАЦИЯЛАУ .....	146
<b>Р. Таберхан, М.А. Самбетбаева, Г. Қалман</b> KAZCAUSAL: ҚАЗАҚ ТІЛІНДЕГІ СЕБЕП-САЛДАРЛЫҚ ҚАТЫНАСТАРДЫҢ АЛҒАШҚЫ КОРПУСТЫҚ АННОТАЦИЯСЫ .....	160
<b>С. Тынымбаев, С.Е. Маманова, Р. Бердібаев, Ж.Е. Темірбекова, Т. Чинибаева</b> БӨЛГІШТІҢ ЕСЕЛІ МӘНДЕРІН АЛДЫН АЛА ДАЙЫНДАУМЕН ЖҮЗЕГЕ АСЫРЫЛАТЫН БӨЛУ ҚҰРЫЛҒЫЛАРЫ .....	172



<b>К.Н. Утеулиева, Б.З. Кенжегулов, Т.А. Каражигитова, Х. Булбул, З.Ж. Жанузакова</b> КОЛЛАБОРАТИВТІК СҮЗГІЛЕУ НЕГІЗІНДЕГІ ҰСЫНЫМДЫҚ ЖҮЙЕНІ ӨЗІРЛЕУДІҢ МАТЕМАТИКАЛЫҚ-АЛГОРИТМДІК ТӘСІЛДЕРІ .....	188
<b>С. Шармуханбет, Г. Тұрмуханова, О. Финдик, В. Махатова, Л. Курмангазиева</b> АЙНЫМАЛЫ ЖАРЫҚ ЖАҒДАЙЫНДАҒЫ ЖОҒАРЫ ДӘЛДІКТІ РОБОТТЫҚ ҚҰРАСТЫРУ: ВИЗУАЛДЫ СЕРВОТЕЖЕУДІҢ ТӨЗІМДІ МЕХАТРОНИКАЛЫҚ АРХИТЕКТУРАСЫ .....	209

### АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН

<b>А. Амирбай, З. Аманбайқызы, К. МаксUTOBA, А. Муханова, М. Kassim</b> КӨЗ ҚОЗҒАЛЫСТАРЫ МЕН БЕТ МИМИКА БЕЛГІЛЕРІН МУЛЬТИМОДАЛЬДЫ ТАЛДАУҒА НЕГІЗ- ДЕЛГЕН БАЛАЛАРДАҒЫ АУТИЗМ СПЕКТРІНІҢ БҰЗЫЛЫСТАРЫН ЕРТЕ АНЫҚТАУҒА АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМІ .....	227
<b>К.Д. Байсылбаева, Ш.Ж. Мусиралиева, Ж. Елтай</b> ҚАЗАҚ ТІЛІНДЕГІ ЭКСТРЕМИСТІК ИДЕОЛОГИЯНЫ АНЫҚТАУ: АННОТАЦИЯЛАУ МӘСЕЛЕЛЕРІ ЖӘНЕ ТЕРЕҢ ОҚЫТУ ТӘСІЛДЕРІ .....	242
<b>М.А. Болатбек, А.М. Усманова, Қ.Б. Багитова, Г.Б. Байспай</b> КИБЕР ҚАУІПТІ АНЫҚТАУ ҮШІН ЖЕЛІЛІК ТРАФИКТІ ТАЛДАУ ӘДІСІН ӨЗІРЛЕУ ЖӘНЕ ЗЕРТТЕУ .....	261
<b>Д.И. Прокопович-Ткаченко, Н.К. Жумагалиева, Д.Н. Щитов, Н.Ф. Мормуль, Д.А. Черкасский</b> ТОЛЫҚ ЕМЕС ЖӘНЕ САПАЛЫҚ ДЕРЕКТЕР ЖАҒДАЙЫНДА АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІҢ АҚПА- РАТТЫҚ ҚАУІПСІЗДІК ПАРАМЕТРЛЕРІН БАҒАЛАУДЫҢ БҰЛЫҢҒЫР МОДЕЛІ: ҚҰРУ ӘДІСТЕМЕСІ, ЕРЕЖЕЛЕР БАЗАСЫН БАПТАУ ЖӘНЕ ҰЙЫМДАРҒА АРНАЛҒАН ДЕМОНСТРАЦИЯЛЫҚ КЕЙС .....	279
<b>Е.А. Пустовой, О.А. Пустовая, А.Н. Раушанова, И.С. Заурбеков</b> БАСҚАРЫЛАТЫН ҚАСИЕТТЕРІ БАР СТОХАСТИКАЛЫҚ МОДЕЛЬДЕРДІ СИНТЕЗДЕУДІҢ ТИМДІЛІГІН БАҒАЛАУ .....	305
<b>Е. Сержан, Т. Умаров, А. Әбілбаева</b> МАШИНАЛЫҚ ОҚУ ӘДІСІ АРҚЫЛЫ КРЕДИТ КАРТА ОПЕРАЦИЯЛАРЫНДАҒЫ АЛАЯҚТЫҚТЫ АНЫҚТАУ: САЛЫСТЫРМАЛЫ ТАЛДАУ .....	321

### СОДЕРЖАНИЕ

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

<b>Д.Е. Абжанова, А.А. Белошицкий</b> МОДЕЛЬ И МЕТОД УПРАВЛЕНИЯ ДАННЫМИ О ВЫБРОСАХ СТАЦИОНАРНЫХ ИСТОЧНИКОВ ЗАГРЯЗНЕНИЯ В ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА .....	9
<b>А.Е. Сланбекова, М.Б. Рахимжанова, А.И. Жанибекова, А.З. Алимагамбетова, М. Худойбергенов</b> РАННЕЕ ВЫЯВЛЕНИЕ ГИДРОЛОГИЧЕСКИХ ОПАСНОСТЕЙ НА ОСНОВЕ ПРОСТРАНСТВЕННО- ВРЕМЕННОГО (SPATIOTEMPORAL) АНАЛИЗА .....	25

#### ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

<b>Ф.Н. Абдраимова, А.А. Керейбаева, Д.С. Дюсенова, Д.А. Алиева, Т.Ж. Токтарова</b> ТЕХНОЛОГИИ ИИ В ЯЗЫКОВОМ ОБРАЗОВАНИИ: ПРАКТИЧЕСКИЕ АСПЕКТЫ И ПРОБЛЕМЫ ПРИМЕНЕНИЯ СТУДЕНТАМИ .....	36
<b>Г.Т. Азиева, М.Б. Есенова, А.К. Абжаппарова, Г.Б. Абдикеримова, P. Schmidt</b> ГИБРИДНАЯ МОДЕЛЬ СТЕКИНГА ДЛЯ КЛАССИФИКАЦИИ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР ПО ДАННЫМ UAV .....	50
<b>Ә.Қ. Әйтiм</b> СОВМЕСТНАЯ МОРФОЛОГИЧЕСКАЯ ДИЗАМБИГУАЦИЯ И POS-РАЗМЕТКА ДЛЯ АГГЛЮТИНАТИВНЫХ ЯЗЫКОВ .....	62
<b>С.А. Есниязова, С.Т. Каимов</b> ПРЕДИКТИВНОЕ ТЕХНИЧЕСКОЕ ОБСЛУЖИВАНИЕ ТЯЖЕЛЫХ ГРУЗОВИКОВ С ИСПОЛЬЗОВАНИ- ЕМ ОБЪЯСНИМОГО МАШИННОГО ОБУЧЕНИЯ .....	78
<b>Т.Д. Иманбекова, Ж.Б. Ибраева, Г.Т. Джаканова, Г.Т. Асқанбай</b>	

АЛГОРИТМ СЖАТИЯ ДАННЫХ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАТЕЛЯ: АНАЛИЗ И РЕАЛИЗАЦИЯ В МАТЛАВ .....	92
<b>Б.З. Кенжегулов, Ж.Т. Билялова, К.Н. Утеулиева, Л. Нургалиева, Ш.С. Нуржанова</b>	
МАТЕМАТИКО-АЛГОРИТМИЧЕСКИЙ ПОДХОД К РАЗРАБОТКЕ ИНТЕЛЛЕКТУАЛЬНОЙ TEXT-TO-SQL СИСТЕМЫ НА ОСНОВЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ .....	110
<b>Н.Ш. МаксUTOва, Д.А. Тусупов, А.А. Шекербек, Ж.Е. Кенжебаева, К.О. Рахмтов</b>	
МАШИННОЕ ОБУЧЕНИЕ ДЛЯ КОМПЛЕКСНОЙ ОЦЕНКИ РИСКА СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ И БИОХИМИЧЕСКИХ ИЗМЕНЕНИЙ: АКЦЕНТ НА АСПАРТАМИНОТРАНСФЕРАЗЕ ...	131
<b>О.С. Салыкова, В.А. Мадин, Б.Р. Салыков, Д.Н. Комаров, Н.В. Мануйлов</b>	
ИНТЕГРАЦИЯ СЕНСОРНЫХ МОДУЛЕЙ MEMS-АКСЕЛЕРОМЕТРОВ В СИСТЕМАХ ПРОМЫШЛЕННОГО МОНИТОРИНГА .....	146
<b>Р. Таберхан, М.А. Самбетбаева, Г. Калман</b>	
КАЗСАUSAL: ПЕРВАЯ КОРПУСНАЯ АННОТАЦИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ НА КАЗАХСКОМ ЯЗЫКЕ .....	160
<b>С. Тынымбаев, С.Е. Маманова, Р. Бердибаев, Ж.Е. Темирбекова, Т. Чинибаева</b>	
УСТРОЙСТВА ДЕЛЕНИЯ ЧИСЕЛ С ПРЕДВАРИТЕЛЬНОЙ ПОДГОТОВКОЙ КРАТНЫХ ДЕЛИТЕЛЮ .....	172
<b>К.Н. Утеулиева, Б.З. Кенжегулов, Т.А. Каражигитова, Х.Бюльбюль, З.Ж. Жанузакова</b>	
МАТЕМАТИКО-АЛГОРИТМИЧЕСКИЕ ПОДХОДЫ К РАЗРАБОТКЕ РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ НА ОСНОВЕ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ .....	188
<b>С. Шармуханбет, Г. Турмуханова, О.Финдик, В.Махатова, Л. Курмангазиева</b>	
ВЫСОКОТОЧНАЯ РОБОТИЗИРОВАННАЯ СБОРКА ПРИ ПЕРЕМЕННОЙ ОСВЕЩЁННОСТИ: РОБАСТНАЯ МЕХАТРОННАЯ АРХИТЕКТУРА ВИЗУАЛЬНОГО СЕРВОУПРАВЛЕНИЯ .....	209

## ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

<b>А. Амирбай, З. Аманбайкызы, К. МаксUTOва, А. Муханова, М. Kassim</b>	
АЛГОРИТМ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАННЕГО ВЫЯВЛЕНИЯ РАССТРОЙСТВ АУТИСТИЧЕСКОГО СПЕКТРА У ДЕТЕЙ НА ОСНОВЕ МУЛЬТМОДАЛЬНОГО АНАЛИЗА ДАННЫХ ДВИЖЕНИЯ ГЛАЗ И МИМИЧЕСКИХ СИГНАЛОВ .....	227
<b>К.Д. Байсылбаева, Ш.Ж. Мусиралиева, Ж.Елтай</b>	
ОБНАРУЖЕНИЕ ЭКСТРЕМИСТСКОЙ ИДЕОЛОГИИ НА КАЗАХСКОМ ЯЗЫКЕ: ПРОБЛЕМЫ АННОТИРОВАНИЯ И МЕТОДЫ ГЛУБОКОГО ОБУЧЕНИЯ .....	242
<b>М.А. Болатбек, А.М. Усманова, К.Б. Багитова, Г.Б. Байспай</b>	
РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДА АНАЛИЗА СЕТЕВОГО ТРАФИКА ДЛЯ ВЫЯВЛЕНИЯ КИБЕРУГРОЗЫ .....	261
<b>Д.И. Прокопович-Ткаченко, Н.К. Жумагалиева, Д.Н. Щитов, Н.Ф. Мормуль, Д.А. Черкасский</b>	
НЕЧЕТКАЯ МОДЕЛЬ ОЦЕНИВАНИЯ ПАРАМЕТРОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ В УСЛОВИЯХ НЕПОЛНЫХ И КАЧЕСТВЕННЫХ ДАННЫХ: МЕТОДИКА ПОСТРОЕНИЯ, НАСТРОЙКА БАЗЫ ПРАВИЛ И ДЕМОСТРАЦИОННЫЙ КЕЙС ДЛЯ ОРГАНИЗАЦИЙ .....	279
<b>Е.А. Пустовой, О.А. Пустовая, А.Н. Раушанова, И.С. Заурбеков</b>	
ОЦЕНКА ЭФФЕКТИВНОСТИ СИНТЕЗА СТОХАСТИЧЕСКИХ МОДЕЛЕЙ С УПРАВЛЯЕМЫМИ СВОЙСТВАМИ .....	305
<b>Е. Сержан, Т. Умаров, А. Абильбаева</b>	
ВЫЯВЛЕНИЕ МОШЕННИЧЕСТВА С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ ПРИ ОПЕРАЦИЯХ С КРЕДИТНЫМИ КАРТАМИ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ .....	321



INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Vol. 7. Is.2. Number 26 (2026). Pp. 160–171

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

<https://doi.org/10.54309/IJICT.2026.26.2.011>

IRSTI / FTAXP / MPHTI 004.656

## KAZCAUSAL: THE FIRST CORPUS-BASED ANNOTATION OF CAUSAL RELATIONSHIPS IN THE KAZAKH LANGUAGE

*R. Taberkhan<sup>1\*</sup>, M.A. Sambetbayeva<sup>1</sup>, G. Kalman<sup>2</sup>*

<sup>1</sup>L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;

<sup>2</sup>K.K. Ualikhanov Kokshetau University, Kokshetau, Kazakhstan.

E-mail: [roman.tn82@gmail.com](mailto:roman.tn82@gmail.com)

**Taberkhan Roman** — PhD student, L.N. Gumilyov Eurasian National University, Department of Information Systems, Astana, Kazakhstan

E-mail: [roman.tn82@gmail.com](mailto:roman.tn82@gmail.com), <https://orcid.org/0000-0002-3375-4947>;

**Madina Sambetbayeva** — PhD, Associate Professor, Department of Information Systems, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

<https://orcid.org/0000-0001-9358-1614>;

**Gulzhamal Kalman** — PhD, Associate Professor, Department of Information and Communication Technologies, Sh. Sh. Ualikhanov Kokshetau University, Kokshetau, Kazakhstan

<https://orcid.org/0000-0001-8863-9447>.

© R. Taberkhan, M.A. Sambetbayeva, G. Kalman

**Abstract.** In this article, the KAZCausal corpus attempts for the first time to annotate causal relationships in modern Kazakh. The study examines the structural and methodological foundations of the KAZCausal corpus. In addition, the first attempt was made to systematically record causal and resultative relationships in texts based on scientific, journalistic, and Internet information materials, and the results of this study make a significant contribution to the field of natural language processing (NLP), integrating with the field of linguistics. In the course of the study, the process of corpus creation, the annotation scheme, and linguistic criteria will be discussed as prerequisites for creating a corpus applied to syntactic and semantic cause-and-effect configurations in the modern Kazakh language. This article will examine a number of tools, models, and corpora that have been successfully used in research on other languages for contextualization in current NLP experiments. These include the Humor (High-speed Unification Morphology) annotator for morphological processing, the POS tagging model, the Multi-tasl learning and Subdomain adaptation model, Multilingual BERT, the multilingual XLM-R model, as well as the Kazakh Language Corpus-KLC with wide coverage, Penn Discourse Treebank (PDTB),



CausBank, and BECause, the main differences and innovative aspects of the KAZCausal corpus aimed at annotating causal relationships were examined, as well as features such as annotation principles, scope of application, and structural nature. In the future, the KAZCausal corpus may become the basis for application in other Turkic-language corpora, effective use of the Kazakh language in computer processing, development of semantic search technologies in information systems, and improvement of NLP models.

**Keywords:** NLP, text annotator, corpus linguistics, KazCausal, multilingual model, annotation

**For citation:** R. Taberkhan, M.A. Sambetbayeva, G. Kalman (2026). KazCausal: the first corpus-based annotation of causal relationships in the kazakh language // International journal of information and communication technologies. Vol. 7. No. 26. Pp. 160–171. <https://doi.org/10.54309/IJICT.2026.26.2.011>. (In Russ.).

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

**Acknowledgment:** *This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (AP26195165).*

## KAZCAUSAL: ҚАЗАҚ ТІЛІНДЕГІ СЕБЕП-САЛДАРЛЫҚ ҚАТЫНАСТАРДЫҢ АЛҒАШҚЫ КОРПУСТЫҚ АННОТАЦИЯСЫ

*Р. Таберхан<sup>1\*</sup>, М.А. Самбетбаева<sup>1</sup>, Г. Қалман<sup>2</sup>*

<sup>1</sup>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;

<sup>2</sup>Ш.Ш. Уәлиханов атындағы Көкшетау университеті, Көкшетау, Қазақстан.

E-mail: roman.tn82@gmail.com

**Таберхан Роман** — PhD докторант, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Ақпараттық жүйелер кафедрасы, Астана, Қазақстан

E-mail: roman.tn82@gmail.com, <https://orcid.org/0000-0002-3375-4947>;

**Мадина Самбетбаева** — PhD докторы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Ақпараттық жүйелер кафедрасының қауымдастырылған профессоры, Астана, Қазақстан

<https://orcid.org/0000-0001-9358-1614>;

**Гүлжамал Қалман** — PhD докторы, Ш.Ш. Уәлиханов атындағы Көкшетау университеті, Ақпараттық-коммуникациялық технологиялар кафедрасының қауымдастырылған профессоры, Көкшетау, Қазақстан

<https://orcid.org/0000-0001-8863-9447>.

© Р. Таберхан, М.А. Самбетбаева, Г. Қалман

**Аннотация.** Аталған мақалада KAZCausal корпусы қазіргі қазақ тіліндегі себеп-салдарлық байланысты аннотациялауға алғаш рет талпыныс жасап отыр. Зерттеуде KAZCausal корпусын құрастыру, құрылымдық дизайны және әдіснамалық негіздері қарастырылады. Сонымен қатар, ғылыми, публицистикалық, интернеттік-ақпараттық

материалдар негізіндегі мәтіндердегі себеп-салдарлық пен нәтижелік қатынастарды жүйелі түрде тіркеудің алғашқы әрекеті жасалды, әрі аталған зерттеу нәтижесі лингвистика саласымен интеграциялана отырып, табиғи тілдерді өңдеу (Natural language processing – NLP) саласына да айтарлықтай үлес қосады. Зерттеу барысында қазіргі қазақ тіліндегі синтаксистік және семантикалық себеп-салдарлық конфигурацияларға қолданылатын корпусы құрудың алғышарттары ретінде корпус құру үдерісіне, аннотация сұлбасына және лингвистикалық критерийлерге тоқталады. Бұл мақалада қазіргі NLP тәжірибелері аясында контексттеу үшін зерттеуде басқа да тілдерде сәтті қолданыс тапқан бірқатар құралдар, моделдер мен корпусар қарастырылады. Оларға морфологиялық өңдеуге арналған Humor (High-speed Unification Morphology) аннотаторы, POS tagging моделі, Multi-task learning and Subdomain adaptation моделі, Multilingual BERT, XLM-R көптілді моделі, сонымен қатар кең ауқымды қамтыған қазақ тілді Kazakh Language Corpus – KLC корпусы, Penn Discourse Treebank (PDTB), CausBank және BECause сияқты аннотация қағидаттары, қолдану саласы және құрылымдық сипаты сияқты ерекшеліктеріне шолу жасалып, себеп-салдарлық байланыстарды аннотациялауға бағытталған KAZCausal корпусының негізгі айырмашылықтары мен инновациялық аспектілері қарастырылды. KAZCausal корпусы болашақта өзге де түркі тілдес корпусарға қолдануға, қазақ тілін компьютерлік өңдеу бағытында тиімді пайдалануға, ақпараттық жүйелерде семантикалық іздеу технологияларын дамытуда және NLP моделдерін жетілдіруге негіз етуге болады.

**Түйінді сөздер:** NLP, мәтінді аннотаторы, корпус лингвистикасы, KazCausal, көптілді модель, аннотациялау

**Дәйексөздер үшін:** Р. Таберхан, М.А. Самбетбаева, Г. Қалман (2026). KazCausal: қазақ тіліндегі себеп-салдарлық қатынастардың алғашқы корпусық аннотациясы // Халықаралық ақпараттық және коммуникациялық технологиялар журналы. Т. 7. No. 26. Б. 160–171. <https://doi.org/10.54309/IJICT.2026.26.2.011>. (Қаз. тіл.).

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

**Алғыс:** Бұл зерттеу Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті тарапынан қаржыландырылды (AP26195165).

## KAZCAUSAL: ПЕРВАЯ КОРПУСНАЯ АННОТАЦИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ НА КАЗАХСКОМ ЯЗЫКЕ

*Р. Таберхан<sup>1\*</sup>, М.А. Самбетбаева<sup>1</sup>, Г. Калман<sup>2</sup>*

<sup>1</sup>Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан;

<sup>2</sup>Университет имени Ш.Ш. Уалиханова, Кокшетау, Казахстан.

E-mail: roman.tn82@gmail.com



**Таберхан Роман** — PhD докторант, Евразийский национальный университет имени Л.Н. Гумилева, кафедра информационных систем, Астана, Казахстан

E-mail: roman.tn82@gmail.com, <https://orcid.org/0000-0002-3375-4947>;

**Мадина Самбетбаева** — Доктор PhD, Ассоциированный профессор кафедры информационных систем, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан

<https://orcid.org/0000-0001-9358-1614>;

**Гүлжамал Қалман** — Доктор PhD, Ассоциированный профессор кафедры информационных и коммуникационных технологий, Кокшетауский университет имени Ш. Ш. Уалиханова, Кокшетау, Казахстан

<https://orcid.org/0000-0001-8863-9447>.

© Р. Таберхан, М.А. Самбетбаева, Г. Қалман

**Аннотация.** В данной статье корпус KAZCausal впервые пытается аннотировать причинно-следственную связь на современном казахском языке. В исследовании рассматриваются конструктивные и методологические основы сборки корпуса KAZCausal. Кроме того, была предпринята первая попытка систематической регистрации причинно-следственных и результирующих отношений в текстах на основе научных, публицистических, интернет-информационных материалов, и результаты данного исследования вносят значительный вклад в область обработки естественных языков (Natural language processing – NLP), интегрируясь с областью лингвистики. В ходе исследования в качестве предпосылок для создания корпуса, применяемых к синтаксическим и семантическим причинно-следственным конфигурациям в современном казахском языке, будут затронуты процесс создания корпуса, схема аннотации и лингвистические критерии. В этой статье будет рассмотрен ряд инструментов, моделей и корпусов, которые успешно использовались в исследованиях на других языках для контекстуализации в рамках текущих экспериментов NLP. К ним относятся аннотатор Humor (High-speed Unification Morphology) для морфологической обработки, модель POS tagging, модель Multi-task learning and Subdomain adaptation, Multilingual BERT, многоязычная модель XLM-R, а также корпус казахоязычного Kazakh Language Corpus-KLC с широким охватом, Penn Discourse Treebank (PDTB), CausBank и BECause были рассмотрены основные различия и инновационные аспекты корпуса KAZCausal, направленные на аннотирование причинно-следственных связей, а также такие особенности, как принципы аннотации, область применения и структурный характер. Корпус KAZCausal в будущем может стать основой для применения в других тюркоязычных корпусах, эффективного использования казахского языка в направлении компьютерной обработки, развития технологий семантического поиска в информационных системах и совершенствования моделей NLP.

**Ключевые слова:** NLP, аннотатор текста, корпусная лингвистика, KazCausal, многоязычная модель, аннотация

**Для цитирования:** Р. Таберхан, М.А. Самбетбаева, Г. Қалман (2026). KazCausal: первая корпусная аннотация причинно-следственных связей на казахском



языке // Международный журнал информационных и коммуникационных технологий. Т. 7. No. 26. Стр. 160–171. <https://doi.org/10.54309/IJICT.2026.26.2.011>. (На каз.).

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

**Благодарность:** Данное исследование было профинансировано Комитетом науки Министерства науки и высшего образования Республики Казахстан (AP26195165).

### Кіріспе.

Табиғи тілдерді өңдеу (Natural language processing-NLP)-технология саласының қарқынды дамуы арқылы адам тілін модельдеу, олардың өзара қатынасу мүмкіндіктерін барынша кеңейткені белгілі. Мұндай жетістіктер, әсіресе, модельдері күрделі болып келетін деректері толық жеткілікті кең таралған тілдер үшін өте пайдалы болды. Мәселен, арнайы мәтіндік корпустарды жинау, оларды топтарға жинақтау, морфологиялық және синтаксистік талдаулар жасауға талпыныс Венгрияда Humor (High-speed Unification Morphology) аннотаторынан басталса (Proszekey et al., 1999), Орталық және Шығыс Еуропа тілдеріне зерттеу жасалынған морфологиялық-синтаксистік кодтау сызбасы MSD Copernicus 106 жобасы (Erjavec et al., 1998) жүзеге асырылды. Аталған жоба аясында Дж. Оруэллдің «1984» атты романындағы параллель аннотацияланған корпус пайда болды. Дегенмен, аталған корпустың бірқатар кемшіліктерінің ішінде ең бастысы көлемінің шағындығы еді, сондықтан ол іргелі зерттеулер жасауға мүмкіндік бермеді. Осы олқылықтың орнын толтыру мақсатымен, машиналық оқыту қолданбалары үшін сөз табын аннотациялауға арналған POS tagging (Alexin et al., 2003) көлемі кең корпусы жүзеге асырылды.

Корпустар арасындағы түрлі эмоцияларды тану ерекшеліктеріне негізделген көпміндетті оқыту мен ішкі домендік бейімдеу (Multi-task learning and Subdomain adaptation) моделі эмоцияны білдіретін конструкцияларды тану үшін бөгет жасайтын факторларды азайту мақсатын көздейді (Fu et al., 2023). Бұл моделдің басқа да кросс-корпустық сөйлеу эмоциясын тануға бағытталған зерттеулерден айырмашылығы орташа реколл (recall) көрсеткішінің жоғары екендігі болып табылады.

Алайда бұл айтылғандардың барлығы табиғи тілдерді өңдеу (NLP) барысында негізінен ресурсы мол, қолданылу аясы кең, кең таралған тілдерге негізделген зерттеулер еді. Мұндай моделдерді аз қолданыстағы, ресурсы аз тілдердің мәселесін шешу үшін қолданылатын инновациялық стратегиялар пайда бола бастады. Осындай бейімдеу тәсілін-трансферлік оқыту деп атап, Multilingual BERT, XLM-R сияқты моделі жүзеге асырылды (Pakray et al., 2024).

Ал қазіргі қазақ тілінің корпусын жасауға талпынған зерттеуге сәйкес (Kazakh Language Corpus-KLC), аталған зерттеу жергілікті ізденушілер қауымдастығы жүзеге асырған қазақ тілінің корпусын құрастыруға бағытталған алғашқы зерттеулердің бірі болып табылады (Makhambetov et al., 2013). Оның басты ерекшелігі–135 миллионнан астам сөзді қамтыған ауқымы өте көлемді екендігі, сонымен қатар жанрлық-стильдік жағынан бес саланы қамтитындығын атауға болады. Корпустың басты негізінен басқа аннотацияланған жеке морфологиялық, синтаксистік және құрылым-

дық белгілеріне сай бөлімдер және сөйлеу тілінің аннотацияланған деректеріне арналған қосымша корпусы бар. Дегенмен, қазақ тіліндегі белгілі бір конструкцияларды аннотациялайтын модель негізіндегі корпустың болмауы зерттеу саласындағы үлкен олқылықтардың бірі болып табылады.

### **Әдістер мен материалдар.**

Қазақ тіліндегі себеп-салдарлық конструкцияларды аннотациялауға бағытталған алғашқы талпыныстардың бірі болып табылатын KAZCausal корпусын құрастыру мақсатында қазақ тіліндегі ғылыми, публицистикалық және ресми-ақпараттық және интернет мәтіндері секілді стильдік-жанрлық сипаттағы мәтіндер қолданылды. Күн сайын деректер базасы барынша толығында және қазіргі уақытта деректердің бастапқы көлемі 850 бірлікті құрайды. Корпусқа алынған материалдар pdf, docx, html форматтарынан алдын-ала өңдеу тәсілі арқылы өзара біріздендіріліп, корпус форматына келтіріледі.

Зерттеу барысында басты тәсіл-аннотациялау бірліктері, яғни сөйлем ішіндегі себеп-салдарлық қатынастарды аннотациялау жүзеге асырылады. Оған себеп компонентті (CAUSE), салдар компонентті (EFFECT), себеп-салдарлық маркерлер (CONNECTIVES) және себеп-салдарлық типтер ретінде логикалық, уақыттық, прагматикалық, шартты-себепті, мақсаттық және т.б. қарастырылды.

Келесі мәтіндерді өңдеу кезеңі тәсілі кезінде алдымен токенизация алгоритмі, ережелік модель, морфологиялық сараптама мен қателерді автоматты түзету сараптамасы аннотациялау тарапынан қолмен тексеріліп, нақтылану жүзеге асырылады.

Ал аннотациялау кезеңінде арнайы аннотациялау нұсқаулығы (annotation guidelines) даярланды. Аталған нұсқаулыққа сәйкес, қазақ тіліндегі себеп-салдарлық қатынастарды анықтау үшін алғашқы аннотация, келісім көрсеткішін есептеу, соңғы нұсқаны қалыптастыру үдерістері жасалады. Корпус Label Studio платформасында жүзеге асырылса, кіріс-шығыс форматтары JSON, XML негізінде және Python бағдарламалау тілі қолданылды.

Корпусқа жинақталған деректерді талдау үшін себеп-салдарлық маркерлердің жиілігін есептеу, себеп пен салдардың синтаксистік үлгілерін анықтау, себеп-салдарлық құрылымдар типологиясын жіктеу мен статистикалық талдау әдістері назарға алынды.

### **Нәтижелер және оларды талқылау.**

Себеп-салдарлық байланыстар табиғи тілдің мағыналық құрылымдарында маңызды орын алады және табиғи тілдерді өңдеу (NLP) саласында мәліметтерді талдаудың ең негізгі компоненттерінің бірі болып табылады. Олар мәтіндегі оқиғалар арасындағы логикалық байланыстарды көрсетіп, аргументациясының, түсіндірудің, дәлелдеу мен пайымдаудың негізін құрайды. Тілдік коммуникацияда себептің немесе салдардың нақты берілуі айтушының көзқарасын, бағалауын, ниетін, эмоциясын жеткізеді және дискурстың прагматикалық бағытын қалыптастырады. Сондықтан себеп-салдарлық байланыстарды жүйелі түрде зерттеу тек жалпы тіл білімінің ғана емес, сонымен қатар қолданбалы мәтіндік және когнитивтік лингвистиканы, қазіргі заманғы алгоритмдік талдаудың өзекті мәселелерінің бірі болып келеді.



Қазіргі кезде ағылшын, неміс, қытай және араб тілдеріндегі себеп-салдарлық байланыстарды таңбалау бойынша көптеген корпустар жасалған. Әсіресе, *Penn Discourse Treebank* (Prasad et al., 2008), *CausBank* (Li et al., 2020), *VerbNet* (Schuler, 2005), *FrameNet* (Vieu et al., 2016), *FrameNet II* (Ruppenhofer et al., 2016) және *BECause* (Dunietz et al., 2017) сияқты ресурстарда себер-салдарлық қатынастар лексикалық, синтаксистік және дискурстық деңгейлерде белгіленіп, үлкен көлемді талдаулар жасалған. Өзге тілдердегі каузалды конструкциялар ғылыми нысанның түрлі аспектілерінде қарастырылып келді, мәселен, себеп-салдарлық байланыстарды анықтауға арналған әлсіз бақылау негізінде клиникалық мәтіндерді ұсыну әдістерін анықтау (Weng et al., 2020), себеп-салдар байланыстарын анықтауда трансфер әдісін қолдану, құрылымы мен деректеріне қатысты эмпирикалық зерттеулер (Anuyah et al., 2025) жасалды. Осы жобалар табиғи тілдерді өңдеу (NLP) саласында ақпараттарды іздеу, машиналық аударма, мәтіндерді түсіндіру, себептілікті автоматты түрде анықтау сияқты көптеген міндеттердің тиімді орындалуына өте жақсы мүмкіндік береді. Дегенмен, морфологиясы күрделі агглютинативті құрылымдағы тілдер үшін осындай ресурстардың жеткіліксіздігі көрінеді. Осы кездегі қолданыстағы халықаралық схемалар агглютинативті тілдерге тән сөйлемдер құрылысы, етістік сөзжасамына толы жүйесі, есімше және көсемше формалары арқылы жасалатын полипредикативті конструкциялардың ерекшеліктерін толық қамти алмайды. Қазақ тілінде себеп-салдарлық мән –*ғандықтан*, –*ған соң*, –*ып*, –*май*, *сондықтан*, *өйткені*, *себебі*, *неге десең* деген сияқты көптеген морфологиялық және аналитикалық құралдар арқылы беріледі. Бірақ осы құралдардың жүйелі сипаттамасын, олардың семантикалық реңктерін және текстегі қызметін көрсететін арнайы корпустық ресурс қазіргі күнге дейін жасалмаған. Дегенмен, аталған салаға қатысты алғашқы зерттеулерге авторлар тарапынан талпыныс жасалды (Taberkhan et al., 2025). Осы кемшіліктің орынын толтыру үшін ұсынылып отырған KazCausal жобасы қазақ тіліндегі себеп-салдарлық байланыстарды жүйелі түрде аннотациялаудың алғашқы тәжірибелерін ұсынып отырмыз. Жоба барысында себеп пен салдарды көрсететін морфологиялық, аналитико-синтетикалық және аналитикалық құрылымдары анықталып, олардың семантикалық ерекшеліктері анықталды. Корпус көлеміне енгізілген сөйлемдер ғылыми, публицистикалық, оқу-әдістемелік және саяси мәтіндер моделдерінен алынған, бұлар қазақ тіліндегі себеп-салдарлық құрылымдардың жан-жақты қамтылуына мүмкіндік береді.

#### *Корпустық аннотация әдісі*

*KazCausal*-қазақ тіліне арнайы бейімделіп жасалған, себеп-салдарлық мағынаны білдіретін құрылымдарды олардың морфологиялық ерекшеліктері мен синтаксистік ұйымдасуын ескере отырып белгілеуге арналған алғашқы аннотациялық схема болып табылады.

Ұсынылған суретте *Label Studio* платформасы арқылы жүзеге асқан каузалды қатынастарды таңбалау үдерісінің нақты көрінісі берілген. Бұл интерфейс *KazCausal* корпусының бастапқы белгілеу кезеңдерінде қолданылған құралдардың бірі ретінде сипатталады.



Сур. 1. KazCausal корпусындағы себеп-салдарлық қатынастарды Label Studio платформасында аннотациялау интерфейсі

Аннотациялау барысында мәтіндегі себеп-салдарлық құрылымдар семантикалық және синтаксистік деңгейде сегменттеліп, арнайы әзірленген схема негізінде типологиялық классификациядан өтті. Көрсетілген мысалда «*Оқушылар сабаққа кеш келгендіктен, мұғалім сабақты кейінге шегерді*» сөйлеміндегі себептік қатынас анықталып, «*келгендіктен*» фрагменті *SYNTHETIC\_CAUSE* санатына жатқызылған. Ал «*сабақты кейінге шегерді*» бөлімі салдар компоненті ретінде *EFFECT* белгісін алған. Осы екі элементтің арасындағы байланысты *leads\_to* реляциясы көрсетеді.

Платформа мәтіндік аннотациямен қатар, оң жақ панелде белгіленген объектілер мен байланыстар туралы метадеректерлерді автоматты түрде сақтап отырады. Бұл деректердің құрылымданған түрде жинақталуына және кейінгі кезеңдерде талдау, модель дайындау немесе автоматтандырылған себеп-салдарлық ақпаратты өндіру жүйелерін әзірлеуде қолданылуына мүмкіндік береді.

*KazCausal* жобасы шеңберінде аннотациялау Label Studio негізінде жүргізіліп, көпдеңгейлі таңбалау мүмкіндіктерін тиімді пайдалануға жол ашты. Аннотацияның негізгі мақсаты-қазақ тіліндегі каузалдық қатынастардың морфологиялық, синтаксистік және семантикалық ерекшеліктерін формалды түрде белгілеу және жүйелеу болды. Осы тәсіл арқылы әрбір каузалды фрагмент анық шекаралармен белгіленіп, олардың арасындағы байланыстар жеке реляциялық жазбалар түрінде тіркелді.

Белгілеу барысында каузалды құрылымдар үш негізгі үлгіге бөлінді: синтетикалық, аналитико-синтетикалық және аналитикалық типтер. Бұл жіктеу қазақ тілінің агглютинативті жүйесіне тән грамматикалық көрсеткіштер негізінде жүзеге асты, мысалы: *-ғандықтан*, *-ған соң*, *-ғаны үшін*, *-сол үшін* және т.б. Әрбір тип өз ішінде

субкатегорияларға бөлініп, олардың қызметтік ерекшеліктері нақты таңбаланды.

Аннотациялау барысында себеп (CAUSE), салдар (EFFECT), әрекет субъектісі (SUBJECT\_CAUSE) және салдар субъектісі (SUBJECT\_EFFECT) сияқты компоненттер жеке сегменттер ретінде белгіленді. Сонымен қатар *leads\_to*, *has\_marker*, *has\_consequence\_marker*, *subject\_of* сияқты реляциялар арқылы олардың семантикалық байланыстары тіркелді. Бұл тәсіл каузалдық қатынастардың құрылымдық сипатымен қатар, мағыналық тәуелділігін де көрсетуге мүмкіндік берді.

Аннотациялау кезеңі *KazCausal* корпусының қалыптасуындағы негізгі әдістемелік қадамдардың бірі болып табылады. Формалды таңбалау нәтижесінде қазақ тіліндегі себеп-салдарлық құрылымдардың көріністері жүйеленіп, олардың негізінде төмен ресурсты тілдер үшін автоматтандырылған себеп-салдарлық ақпаратты өңдеу модельдерін әзірлеуге практикалық база жасалды.



Сур. 2. *KazCausal* корпусын аннотациялау үдерісінің жұмыс схемасы.

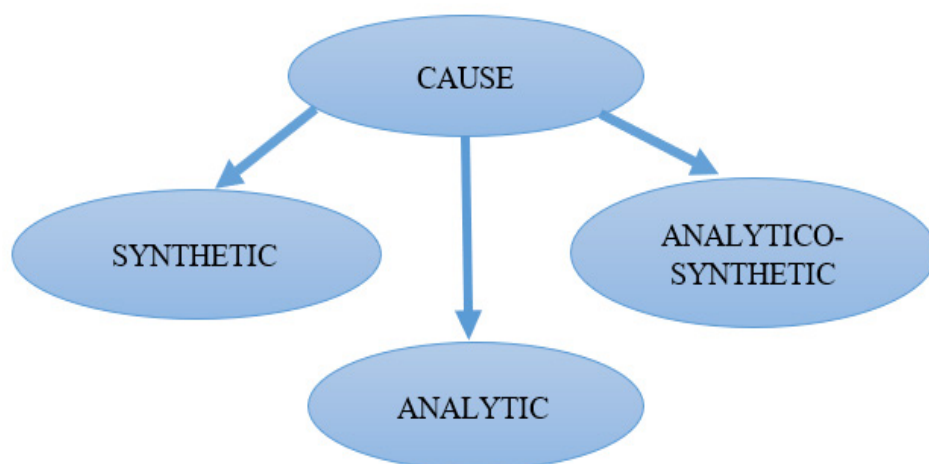
Жоғарыда келтірілген суретте (Сурет 2) корпусқа енгізілетін мәтінді таңдаудан бастап, сөйлемдерді сегменттеуге, каузалды маркерлерді анықтауға, типологиялық жіктеуге, реляциялық байланыстарды белгілеуге және аннотацияланған деректерді корпусқа енгізуге дейінгі кезеңдер көрсетілген. Бұл *workflow* аннотациялау әдіснамасының жүйелі құрылымын және деректерді формалды өңдеу

ЛОГИКАСЫН АЙҚЫНДАЙДЫ.

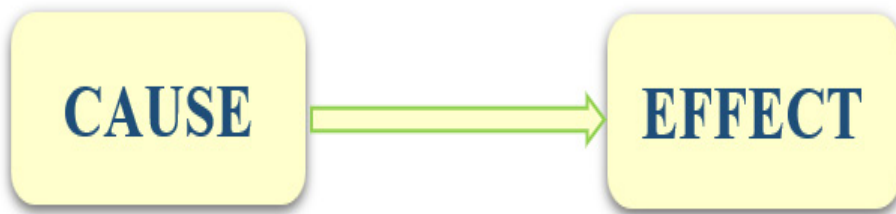


Сур. 3. KazCausal корпусындағы каузалды құрылымдардың типтік үлесі.

Сурет 3-ке сәйкес, KazCausal корпусындағы каузалды құрылымдардың типтік үлесі төмендегідей көрсеткіштерге, атап айтқанда синтетикалық (45%), аналитико-синтетикалық (32%) және аналитикалық (23%) себеп-салдарлық құрылымдардың үлесін көрсетеді. Бұл үлестірім қазақ тіліндегі каузалды мәннің басым түрде морфологиялық тәсілдер арқылы берілуін айқындайды және тілдің агглютинативті табиғатын растайды.



Сур. 4. KazCausal корпусындағы каузалды құрылымдардың таксономиялық жіктелуі.



Сур. 5. KazCausal корпусындағы себеп-салдарлық реляциялық байланыс моделі.

Схема каузалды қатынастардың негізгі категориясы (Сурет 4)-CAUSE-және оның үш құрылымдық типке (SYNTHETIC, ANALYTICO-SYNTHETIC, ANALYTIC) бөлінуін көрсетеді. Әрбір тип морфологиялық және синтаксистік көрсеткіштерге негізделген субкатегориялар арқылы сипатталады, бұл аннотациялық схеманың иерархиялық құрылымын ашады.

Себептік компонент (CAUSE) пен салдарлық компонент (EFFECT) арасындағы семантикалық тәуелділік *leads\_to* реляциясы арқылы көрсетілген (Сурет 5). Бұл модель каузалды құрылымдарды формалды түрде байланыстырудың негізін құрайды және кейінгі автоматтандырылған себеп-салдарлық ақпарат өндіру жүйелеріне қолдануға мүмкіндік береді.

### Қорытынды.

Себеп-салдарлық конструкцияларды аннотациялау, оларды автоматты түрде табу міндеті толық шешімін таппаған ғылыми зерттеу нысаны болып табылады. Себебі, аталған ғылыми бағыт ағылшын тіліндегі материалдар негізінде кеңінен зерттелгенімен, өзге де көптеген тілдер үшін, соның ішінде агглютинативті тілдер қатарына жататын, біріздендірілген аннотацияланған деректер житынтығы жоқ қазақ тілі үшін мүлдем қарастырылмаған сала деуге болады.

Мәтіндердегі себеп-салдарлық конструкцияларды анықтау-себептік қатынастарды автоматты түрде табуға бағытталған маңызды ақпараттық анықтағыш міндетін атқарады. Өйткені, қолданыстағы біртүрлі деректер жиынтығы аннотациялау нұсқаулығын қолдана отырып, олардың толық біртұтас корпуста біріздендірілуіне мүмкіндік береді. Сондықтан кең ауқымды, стандартталған жаңа деректер жиынтығынан құралған корпус өз кезегінде көптеген пәнаралық (ақпараттық жүйелер, лингвистика) міндеттердің шешілуіне мүмкіндік береді.

Себеп-салдарлық байланыстарды таңбалау және оларды формалды түрде көрсету әлемдік лингвистикада, дискурс талдауында және табиғи тілдерді өңдеу (NLP) саласында іргелі зерттеуді қажет ететін ғылыми бағыттардың бірі болып табылады. Себеп-салдарлық қатынастар мәтіндердегі оқиғалар арасындағы логикалық тәуелділіктерді көрсетіп, түсіндіру, дәлелдеу, мотивация, санкция, және эмоциялық реакция сияқты күрделі мағыналық байланыстарды анықтауға мүмкіндік береді. Алайда, бұл құбылыстарды жүйелі түрде сипаттау әр тілдің морфологиялық, синтаксистік және семантикалық ерекшеліктеріне қарай әртүрлі тәсілдерді талап етеді.

KazCausal корпусы қазақ тіліндегі себеп-салдарлық байланыстарды белгілеу-

дің формалды моделін жасауға, әрі олардың типологиясын нақтылауға және морфологиялық-семантикалық табиғатын сипаттауға бағытталған. Жоба нәтижелері қазақ тілін компьютерлік өңдеу бағытында өте пайдалы болмақ: мәтіндерді автоматты талдау, себеп-салдарлық қатынастарды автоматты түрде анықтау, ақпараттық жүйелерде семантикалық іздеу технологияларын дамыту және қазақ тіліне бейімделген NLP моделдерін жетілдіруге негіз бола алады.

## REFERENCES

- Alexin, Z. (2003). Manually annotated Hungarian corpus // Department of Informatics, University of Szeged. Available at: <http://www.inf.u-szeged.hu/~alexin/corpora>. [in Eng.].
- Anuyah, S., Vanschaik, J., Jain, P., Lehman, S., & Chakraborty, S. (2025). An empirical study of causal relation extraction transfer: Design and data // arXiv. Available at: <https://arxiv.org/abs/2503.06076>. [in Eng.].
- Dunietz J., Levin L.S., Carbonell J.G. (2017). The BECauSE Corpus 2.0: Annotating causality and overlapping relations // Proc. 11th Linguistic Annotation Workshop (LAW). – Association for Computational Linguistics. Pp. 95–104. DOI: <https://doi.org/10.18653/v1/W17-0812>. [in Eng.].
- Erjavec T., Monachini M. (Eds.) (1997). Specification and notation for lexicon encoding (MULTTEXT-EAST Project) // Copernicus Project 106. Available at: <http://nl.ijs.si/ME>. [in Eng.].
- Fu H., Zhuang Z., Wang Y., Huang C., Duan W. (2023). Cross-corpus speech emotion recognition based on multi-task learning and subdomain adaptation // Entropy. — Vol. 25. — No. 1. Pp. 124. DOI: <https://doi.org/10.3390/e25010124> – [in Eng.].
- Li Z., Ding X., Liu T., Hu J.E., Van Durme B. (2020). Guided generation of cause and effect // Proc. IJCAI-20. DOI: <https://doi.org/10.24963/ijcai.2020/447>. [in Eng.].
- Makhambetov O., Makazhanov A., Yessenbayev Z., Matkarimov B., Sabyrgaliyev I., Sharafudinov A. (2013). Assembling the Kazakh Language Corpus // Nazarbayev University Research and Innovation System. Available at: <https://nur.nu.edu.kz>. [in Eng.].
- Pakray P., Gelbukh A., Bandyopadhyay S. (2025). Natural language processing applications for low-resource languages // Natural Language Processing. — Vol. 31. — Pp. 183–197. DOI: <https://doi.org/10.1017/nlp.2024.33>. [in Eng.].
- Prasad R., Dinesh N., Lee A., Miltsakaki E., Robaldo L., Joshi A., Webber B. (2008). The Penn Discourse TreeBank 2.0 // Proc. LREC'08. – European Language Resources Association. Available at: <https://aclanthology.org/L08-1093/>. [in Eng.].
- Prószycki G., Kis B. (1999). A unification – based approach to morpho – syntactic parsing of agglutinative languages // Proc. 37th Annual Meeting of ACL. Pp. 261–268. DOI: <https://doi.org/10.3115/1034678.1034720>. [in Eng.].
- Ruppenhofer J., Ellsworth M., Petruck M.R.L., Johnson C.R., Baker C.F., Scheffczyk J. (2016). FrameNet II: Extended theory and practice // International Computer Science Institute. Available at: <https://framenet.icsi.berkeley.edu>. [in Eng.].
- Schuler K.K. (2005). VerbNet: A broad coverage, comprehensive verb lexicon // University of Pennsylvania, PhD dissertation. Available at: <https://verbs.colorado.edu/verbnets/> – [in Eng.].
- Taberkhan R., Sambetbayeva M. (2025). KazCausal: Manual annotation of causal constructions in the Kazakh language // Proc. 10th Int. Conf. on Computer Science and Engineering (UBMK). Pp. 463–467. – DOI: <https://doi.org/10.1109/UBMK67458.2025.11206832>. [in Eng.].
- Vieu L., Muller P., Candito M., Djemaa M. (2016). A General Framework for the Annotation of Causality Based on FrameNet // Proc. LREC'16. – European Language Resources Association (ELRA). Pp. 3807–3813. Available at: <https://aclanthology.org/L16-1603/>. [in Eng.].
- Weng W.H., Lin Y.W., Szolovits P., Elhadad N. (2020). Representation learning of clinical texts with weak supervision for causal relation extraction // Journal of Biomedical Informatics. — Vol. 108. — Article 103491. DOI: <https://doi.org/10.1016/j.jbi.2020.103491>. [in Eng.].



**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 30.06.2026.

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