

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

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ

әл-Фараби атындағы Қазақ ұлттық университетінің

Х А Б А Р Л А Р Ы

ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ
НАУК РЕСПУБЛИКИ КАЗАХСТАН
Қазақстан Республикасының
Ғылым Академиясының
Әл-Фараби атындағы
Қазақ ұлттық университетінің

NEWS

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

PHYSICO-MATHEMATICAL SERIES

1 (345)

JANUARY – MARCH 2023

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

БАС РЕДАКТОР:

МУТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының бас директорының м.а. (Алматы, Қазақстан), **Н=5**

РЕДАКЦИЯ АЛҚАСЫ:

КАЛИМОЛДАЕВ Мақсат Нұрәліұлы (бас редактордың орынбасары), физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институты бас директорының кеңесшісі, зертхана меңгерушісі (Алматы, Қазақстан), **Н=7**

МАМЫРБАЕВ Өркен Жұмажанұлы (ғалым хатшы), Ақпараттық жүйелер саласындағы техника ғылымдарының (PhD) докторы, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институты директорының ғылым жөніндегі орынбасары (Алматы, Қазақстан), **Н=5**

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

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

СМОЛАРЖ Анджей, Люблин политехникалық университетінің электроника факультетінің доценті (Люблин, Польша), **Н=17**

ӘМІРҒАЛИЕВ Еділхан Несіпханұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Жасанды интеллект және робототехника зертханасының меңгерушісі (Алматы, Қазақстан), **Н=12**

КИЛАН Әлімхан, техника ғылымдарының докторы, профессор (ғылым докторы (Жапония), ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының бас ғылыми қызметкері (Алматы, Қазақстан), **Н=6**

ХАЙРОВА Нина, техника ғылымдарының докторы, профессор, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының бас ғылыми қызметкері (Алматы, Қазақстан), **Н=4**

ОТМАН Мохаммед, PhD, Информатика, коммуникациялық технологиялар және желілер кафедрасының профессоры, Путра университеті (Селангор, Малайзия), **Н=23**

НЫСАНБАЕВА Сауле Еркебұланқызы, техника ғылымдарының докторы, доцент, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының аға ғылыми қызметкері (Алматы, Қазақстан), **Н=3**

БИЯШЕВ Рустам Гакашевич, техника ғылымдарының докторы, профессор, Информатика және басқару мәселелері институты директорының орынбасары, Ақпараттық қауіпсіздік зертханасының меңгерушісі (Қазақстан), **Н=3**

КАПАЛОВА Нұрсұлту Алдажарқызы, техника ғылымдарының кандидаты, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының киберқауіпсіздік зертханасының меңгерушісі (Алматы, Қазақстан), **Н=3**

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

МИХАЛЕВИЧ Александр Александрович, техника ғылымдарының докторы, профессор, Беларусь Ұлттық Ғылым академиясының академигі (Минск, Беларусь), **Н=2**

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

«ҚР ҰҒА Хабарлары. Информатика сериясы».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

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

Тақырыптық бағыты: *физика-математикалық сериясы*.

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

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 218 бөл., тел.: 272-64-39*

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

© Қазақстан Республикасының Ұлттық ғылым академиясы, 2023
Типографияның мекен-жайы: «Аруна» ЖК, Алматы қ., Мұратбаев көш., 75.

Главный редактор:

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

Редакционная коллегия:

КАЛИМОЛДАЕВ Максат Нурадилович, (заместитель главного редактора), доктор физико-математических наук, профессор, академик НАН РК, советник генерального директора «Института информационных и вычислительных технологий» КН МНВО РК, заведующий лабораторией (Алматы, Казахстан), **Н=7**

МАМЫРБАЕВ Оркен Жумажанович, (ученый секретарь), доктор философии (PhD) по специальности «Информационные системы», заместитель директора по науке РГП «Институт информационных и вычислительных технологий» Комитета науки МНВО РК (Алматы, Казахстан), **Н=5**

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

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

СМОЛАРЖ Анджей, доцент факультета электроники Люблинского политехнического университета (Люблин, Польша), **Н=17**

АМИРГАЛИЕВ Едилхан Несипханович, доктор технических наук, профессор, академик Национальной инженерной академии РК, заведующий лабораторией «Искусственного интеллекта и робототехники» (Алматы, Казахстан), **Н=12**

КЕЙЛАН Алимхан, доктор технических наук, профессор (Doctor of science (Japan)), главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=6**

ХАЙРОВА Нина, доктор технических наук, профессор, главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=4**

ОТМАН Мохамед, доктор философии, профессор компьютерных наук, Департамент коммуникационных технологий и сетей, Университет Путра Малайзия (Селангор, Малайзия), **Н=23**

НЫСАНБАЕВА Сауле Еркебулановна, доктор технических наук, доцент, старший научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=3**

БИЯШЕВ Рустам Гакашевич, доктор технических наук, профессор, заместитель директора Института проблем информатики и управления, заведующий лабораторией информационной безопасности (Казахстан), **Н=3**

КАПАЛОВА Нурсулу Алдажаровна, кандидат технических наук, заведующий лабораторией кибербезопасности РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=3**

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

МИХАЛЕВИЧ Александр Александрович, доктор технических наук, профессор, академик НАН Беларуси (Минск, Беларусь), **Н=2**

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

«Известия НАН РК. Серия информатики».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

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

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

Тематическая направленность: *серия физика-математическая.*

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

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 218, тел.: 272-64-39*

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

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

Chief Editor:

MUTANOV Galimkair Mutanovich, doctor of technical sciences, professor, academician of NAS RK, acting General Director of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H=5**

EDITORIAL BOARD:

KALIMOLDAYEV Maksat Nuradilovich, (Deputy Editor-in-Chief), Doctor of Physical and Mathematical Sciences, Professor, Academician of NAS RK, Advisor to the General Director of the Institute of Information and Computing Technologies of the CS MES RK, Head of the Laboratory (Almaty, Kazakhstan), **H = 7**

Mamyrbayev Orken Zhumazhanovich, (Academic Secretary, PhD in Information Systems, Deputy Director for Science of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H = 5**

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

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

SMOLARJ Andrej, Associate Professor Faculty of Electronics, Lublin polytechnic university (Lublin, Poland), **H= 17**

AMIRGALIEV Edilkhan Nesipkhanovich, Doctor of Technical Sciences, Professor, Academician of NAS RK, Head of the Laboratory of Artificial Intelligence and Robotics (Almaty, Kazakhstan), **H= 12**

KEILAN Alimkhan, Doctor of Technical Sciences, Professor (Doctor of science (Japan)), chief researcher of Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), **H= 6**

KHAIROVA Nina, Doctor of Technical Sciences, Professor, Chief Researcher of the Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), **H= 4**

OTMAN Mohamed, PhD, Professor of Computer Science Department of Communication Technology and Networks, Putra University Malaysia (Selangor, Malaysia), **H= 23**

NYSANBAYEVA Saule Yerkebulanovna, Doctor of Technical Sciences, Associate Professor, Senior Researcher of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H= 3**

BIYASHEV Rustam Gakashevich, doctor of technical sciences, professor, Deputy Director of the Institute for Informatics and Management Problems, Head of the Information Security Laboratory (Kazakhstan), **H= 3**

KAPALOVA Nursulu Aldazarhovna, Candidate of Technical Sciences, Head of the Laboratory cybersecurity, Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H=3**

KOVALYOV Alexander Mikhailovich, Doctor of Physical and Mathematical Sciences, Academician of the National Academy of Sciences of Ukraine, Institute of Applied Mathematics and Mechanics (Donetsk, Ukraine), **H=5**

MIKHALEVICH Alexander Alexandrovich, Doctor of Technical Sciences, Professor, Academician of the National Academy of Sciences of Belarus (Minsk, Belarus), **H=2**

TIGHINEANU Ion Mihailovich, Doctor of Physical and Mathematical Sciences, Academician, President of the Academy of Sciences of Moldova, Technical University of Moldova (Chisinau, Moldova), **H=42**

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

Series of informatics.

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

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

Thematic scope: *series physical-mathematical series.*

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

Periodicity: *4 times a year.*

Circulation: *300 copies.*

Editorial address: *28, Shevchenko str., of. 218, Almaty, 050010, tel. 272-64-39*

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

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

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

NEWS OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X

Volume 1, Number 345 (2023), 121-135
<https://doi.org/10.32014/2023.2518-1726.173>

ӨОЖ 004.9

© G. Kalman¹, M.G. Esmaganbet², M.M. Zhamankarin^{2*},
A.I. Gabdulina², D.V. Pleskachev^{2*}, 2023

¹L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;

²Abay Myrzakhmetov Kokshetau University, Kokshetau, Kazakhstan.

E-mail: zhamankarin@yandex.kz

COREFERENCE SOLUTION USING THE CLUSTERING METHOD

Kalman Gulzhamal — graduate student at the Department of Information Systems. L.N. Gumilyov Eurasian National University, Astana, Kazakhstan.

E-mail: guljamal14@gmail.com, DOI: <https://orcid.org/0000-0001-8863-9447>;

Esmaganbet Musatai Galymuly — Abay Myrzakhmetov Kokshetau University candidate of physics and mathematics, professor. Kokshetau, Kazakhstan.

E-mail: esmakhan_kz@list.ru, <https://orcid.org/0009-0002-3293-1734>;

Zhamankarin Maxut — Doctor of profile. Abay Myrzakhmetov Kokshetau University. 020000. Kokshetau, Kazakhstan.

E-mail: zhamankarin@yandex.kz, ORCID 0000-0002-5254-9797;

Gabdulina Anara Ibrahimovna — Master of Humanities. Lecturer at the “Foreign Languages and Translation Studies” Department. Abay Myrzakhmetov Kokshetau University. 020000. Kokshetau, Kazakhstan.

E-mail: kali-anara@mail.ru, <https://orcid.org/0009-0002-1236-260X>;

Pleskachev Denis Vladimirovich — Teacher, Master of Technical Sciences Department of Informational System and Informatics. Abay Myrzakhmetov Kokshetau University. 020000. Kokshetau, Kazakhstan.

E-mail: denispleskachev@mail.ru, <https://orcid.org/0000-0001-6277-021X>.

Abstract. In linguistics, reference is a comparison of a textual expression with some non—linguistic object and an event in the real or abstract world. Determining referential relations in linguistics has been studied for a long time for other languages, but there are still very few studies for the Kazakh language. Considering these issues, we set the goal of solving reference relations in the Kazakh language. The purpose of the studied system here is to solve coreference relations in the Kazakh language, that is, to cluster personal names related to persons (Type of person). In other words, the task is to combine all the parts of the name in the text (that is, the title, first name, last name, patronymic of each person mentioned in the text). To achieve the goal, we used tomita-parser, keyword dictionary, grammar

for extracting full name, grammar for extracting vocabulary names, clustering, pairwise model, feature vector, pairwise weight vector. Algorithm Performance was evaluated using traditional evaluation metrics, the obtained results compared to other methods, the Tomita-parser algorithm was 0.87%, and the clustering algorithm was 0.81%, the results were given in the form of a table.

Keywords: coreference, clustering, Tomita-parser, reference, anaphora

© Г. Қалман¹, М.Ф. Есмағанбет², М.М. Жаманқарин^{2*}, А.И. Габдулина^{2*}
Д.В. Плескачев^{2*}, 2023

¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;

²А.Мырзахметов атындағы Көкшетау университеті, Көкшетау, Қазақстан.

E-mail: zhamankarin@yandex.kz

КЛАСТЕРЛЕУ ӘДІСІН ҚОЛДАНЫП КОРЕФЕРЕНЦИЯН ШЕШУ

Қалман Гүлжамал — 8D06103. «Ақпараттық жүйе» мамандығының докторанты. Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті. 010000. Астана, Қазақстан.

E-mail: guljamal14@gmail.com. DOI:https://orcid.org/0000-0001-8863-9447;

Есмағанбет Мұсатай Ғалымұлы — Физика математика ғылымдарының кандидаты, профессор. Абай Мырзахметов атындағы Көкшетау университеті. 020000. Көкшетау, Қазақстан.

E-mail: esmakhan_kz@list.ru, https://orcid.org/0009-0002-3293-1734;

Жаманқарин Максұт Мұхамбетназарұлы — бейіні бойынша доктор. Абай Мырзахметов атындағы Көкшетау университеті. 020000. Көкшетау, Қазақстан.

E-mail: zhamankarin@yandex.kz. ORCID 0000-0002-5254-9797;

Габдулина Анара Ибрагимовна — Гуманитарлық ғылымдар магистрі, шетел тілдері және аударма ісі кафедрасының оқытушысы. А. Мырзахметов атындағы Көкшетау университеті.

E-mail: kali-anara@mail.ru, https://orcid.org/0009-0002-1236-260X;

Плескачев Денис Владимирович — Техника ғылымдарының магистрі. Ақпараттық жүйелер және Информатика кафедра Оқытушысы. Абай Мырзахметов атындағы Көкшетау университеті.

E-mail: denispleskachev@mail.ru, https://orcid.org/0000-0001-6277-021X.

Аннотация. Тіл білімінде референция — мәтіндік өрнекті қандай да бір тілдік емес объектімен нақты немесе дерексіз дүниедегі оқиғамен салыстыру. Дәстүрлі тіл білімі референция сөз тіркестерінің екі негізгі класын қарастырады: лексикалық толық формалар (атаулы сөз тіркестері және т.б.) және қысқартылған формалар (мысалы, есімдіктер, сілтеу есімдігі, жіктеу есімдігі, өздік есімдіктер). Референцияны шешудің міндеті белгілі бір тілден тыс нысан туралы нақты мәтіндік сілтемені осы мәтіндегі басқа сілтемелермен анықтау. Тіл біліміндегі референциялық қатынасты шешу басқа тілдер үшін көптен зерттеліп келеді, ал қазақ тілі үшін әлі күнге дейін зерттеулер өте аз. Біз осы мәселелерді ескере отырып, қазақ тіліндегі сілтемелік қатынастарды шешу мақсатын қойдық. Бұндағы зерттелетін жүйенің мақсаты қазақ тіліндегі кореференциялық қатынастарды шешу яғни тұлғаларға қатысты жалқы есімдерді кластерлеу (Тұлға түрі). Басқаша айтқанда, тапсырма мәтінде бар есімнің барлық бөліктерін (яғни, мәтінде көрсетілген әрбір адамның лауазымы,

аты, тегі, әкесінің аты) біріктіру болып табылады. Мақсатқа жету үшін tomita-parser, кілттік сөздер сөздігі, толық атауды алуға арналған грамматика, сөздік атауларды шығаруға арналған грамматика, кластерлеу, жұптық модель, мүмкіндік векторы (Вектор признаков), жұптық салмақ векторын қолдандық. Алгоритмінің Өнімділігін дәстүрлі бағалау метрикасының көмегімен бағаланды, алынған нәтижелер басқа әдістермен салыстырғанда томиита — парсер алгоритмі 0.87%, ал кластерлеу алгоритмі 0.81% құрады, нәтижелер кесте түрінде белгіленді.

Түйін сөздер: кореференция, кластерлеу, томиита-парсер, референция, анафора

© Г. Қалман¹, М.Ғ. Есмағанбет², М.М. Жаманкарин^{2*}, А.Г. Габдулина^{2*},
Д.В. Плескачев^{2*}, 2023

¹Евразийского национального университета им. Л.Н. Гумилева,
Астана, Қазақстан;

²Кокшетауский университет имени Абая Мырзахметова,
Кокшетау, Қазақстан.

E-mail: zhamankarin@yandex.kz

РЕШЕНИЕ КОРЕФЕРЕНЦИИ С ПОМОЩЬЮ МЕТОДА КЛАСТЕРИЗАЦИИ

Қалман Гулжамал — Докторант, кафедра Информационные системы, факультет Информационных технологий. ЕНУ им. Л.Н. Гумилева. 010000. Астана, Қазақстан.

E-mail: guljamal14@gmail.com. DOI: <https://orcid.org/0000-0001-8863-9447>;

Есмағанбет Мұсатай Ғалымұлы — кандидат физико-математических наук, профессор. Кокшетауский университет имени Абая Мырзахметова. 020000. Кокшетау, Қазақстан.

E-mail: esmakhan_kz@list.ru, <https://orcid.org/0009-0002-3293-1734>;

Жаманкарин Мақсұт Мухамбетназарұлы — доктор по профилю. Кокшетауский университет имени Абая Мырзахметова. 020000. Кокшетау, Қазақстан.

E-mail: zhamankarin@yandex.kz. ORCID 0000-0002-5254-9797;

Габдулина Анара Ибрагимовна — Магистр гуманитарных наук, преподаватель кафедры иностранных языков и переводческого дела. Кокшетауский университет им. Абая Мырзахметова. 020000. Кокшетау, Қазақстан.

E-mail: kali-anara@mail.ru, <https://orcid.org/0009-0002-1236-260X>;

Плескачев Денис Владимирович — Преподаватель, Магистр технических наук. Кафедра Информационных систем и Информатики. Кокшетауский университет имени Абая Мырзахметова. 020000. Кокшетау, Қазақстан.

E-mail: denispleskachev@mail.ru, <https://orcid.org/0000-0001-6277-021X>.

Аннотация. В лингвистике референция — это сопоставление текстового выражения с каким-либо неязыковым объектом и событием в реальном или абстрактном мире. Традиционная лингвистика рассматривает два основных класса опорных словосочетаний: лексические полные формы (именные словосочетания и т.д.) и сокращенные формы (например, местоимения). Задача разрешения ссылок состоит в том, чтобы идентифицировать конкретную

текстовую ссылку на конкретную неязыковую сущность с другими ссылками в этом тексте. Определение референциальных отношений в языкознании давно изучается для других языков, но по казахскому языку исследований еще очень мало. Учитывая эти вопросы, мы поставили цель решить референтные отношения в казахском языке. Целью изучаемой системы здесь является решение кореферентных отношений в казахском языке, то есть кластеризация личных имен, связанных с лицами (Тип лица). Другими словами, задача состоит в том, чтобы соединить в тексте все части имени (то есть звание, имя, фамилию, отчество каждого лица, упомянутого в тексте). Для достижения цели мы использовали томита-парсер, словарь ключевых слов, грамматику для извлечения полного имени, грамматику для извлечения названий словарей, кластеризацию, парную модель, вектор признаков, вектор парных весов. Производительность алгоритма оценивалась с использованием традиционных оценочных метрик, полученные результаты составили 0,87% для алгоритма Tomita-parser и 0,81% для алгоритма кластеризации по сравнению с другими методами, результаты были представлены в виде таблицы.

Ключевые слова: кореференция, кластеризация, томита-парсер, анафора

Кіріспе

Сөйлеудің тұтастығы мен өміршеңдігі сол субъектілерді қайта еске түсірумен тікелей байланысты. Сондықтан, сөйлеу барысында біз экстралингвистикалық шындықтың белгілі бір объектілерін еске түсіруді жиі кездестіреміз, олар референт деп аталады, ал қайталау процесі референция деп аталады.

Дискурста референциялық қатынастарды орнату мәтінді автоматты талдаудың ең өзекті, бірақ модельдеуге қиын мәселелерінің бірі болып табылады. Айтылымның мағынасы және берілген айтылым мағынасының қолжетімділік дәрежесі сөйлеушінің қандай сілтемелік құралдарды пайдаланатынына байланысты өзгеруі мүмкін. Ол жалқы есім, есімдік немесе баяндауыш зат есім болуы мүмкін. Егер бір сілтенілген өрнек бұрын қолданылған екіншісіне сілтеме жасаса, онда олардың арасында анафоралық қатынас (анафора) орнатылады, өрнектердің соңғысы анафор, ал оның алдындағысы — антецедент деп аталады.

Соңғы жылдары референциялық қатынастарды шешу өте қызықты және нәтижелі зерттеулердің тақырыбы болды. Бұл зерттеулерде тұлғаларға қатысты жалқы есімдерді кластерлеуді зерттейміз. Жалқы есімдер тегі, аты, әкесінің аты және лақап аттан тұруы мүмкін. Бұл классификацияны толық қазақша жалқы есімнің стандарт құрылымына сүйене отырып пайдаландық.

Атаулы нысандарды шығару үшін біз табиғи тілдегі мәтіндерден құрылымдық деректерді шығаруға арналған еркін қолжетімді құрал Tomita Parser (Bodrova, 2016) қолданамыз. Талдаушы GLR алгоритміне негізделген (Lavie et al., 1993: 123) және контексті еркін грамматикалардың формализмын қолданады.

Табиғи тілдегі референцилық қатынасты шешу мәселесі 60 жылдардан бастап зерттеліп келеді, мұнда ағылшын, неміс, араб, қытай, орыс және т.б. көптеген тілдерде зерттелді.

Соңғы жылдардағы зерттеу жұмыстарына талдау жасайтын болсақ. (Juntao et al., 2022: 4873) жұмысында көптілді корпусты пайдалана отырып “анафор антецедент” жұбын табудың алгоритмін құрды. Корпусты пайдалану барысында әр түрлі тілдер үшін білім базасын құрып, жұптық моделдер көмегімен үміткерлерді табуды жүзеге асырды.

Көп тілді корпустық тәсілмен келесі жұмыс жасалды мұнда чех, орыс, поляк, неміс, испан және каталан тілдеріндегі референцияны табуда терең оқыту Алгоритмін қолданды.

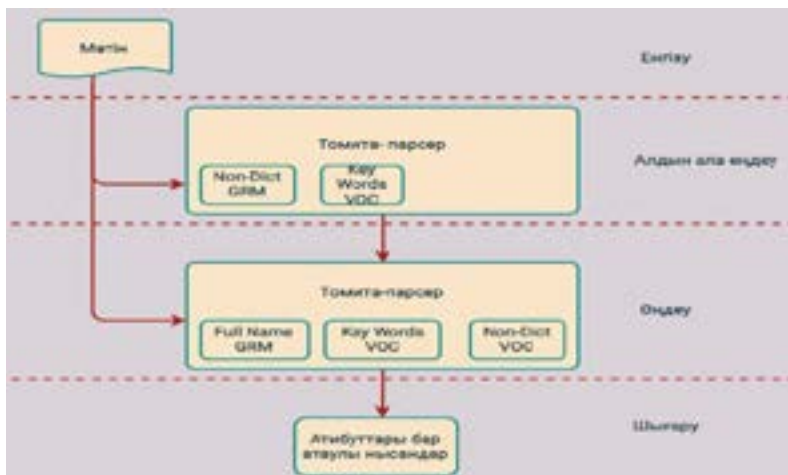
Осындай (Zhumabay et al., 2022: 27) көп тілді жүйеде қазақ орыс және ағылшын тілдеріндегі референцияны шешудің жүйесін құрды, мұнда кибриктің әдісін (Kibrik et al., 2013: 118) негізге ала отырып, антцедент пен анафордың референты болу жағдайын есептеу формуласын ұсынды.

Қазақ тілінде жасалған тағыда бір жұмысты атауға болады, мұнда машиналық оқыту әдін қолданып есімдік анафорасын шешу тәсілі ұсынылды нәтиже айтарлықтай жақсы болды. (Қалман т.б., 2022: 2709).

(Liyan et al., 2020: 8527) және (Yuval et al., 2021: 14) жұмыстары бойынша жұптық моделді кластерлік әдістермен біріктіру және декоддауды пайдаланып референциялық қатынасты шешті. Мұнда кластерлік әдіс өте тиімді және шешу жылдамдығы мен қарапайымдылығы басым, референтті жұптар дағы ықтимал антецедентті табуда ерекше роль атқарады, сонымен қатар бұл әдіске басқа моделдері қосу арқылы әдістің соңғы алатын нәтижесін арттыруға болатындығын дәлелдеді

Бізге осы бағытта зерттеуге негізгі түрткі болған мәселе бұл қазақ тілі үшін бұндай зерттеулердің жоқтығы бізге үлкен мотивация беруде, бұл жұмыстың тағы бір артықшылығы бұнда біз жұптық моделді кластерлеу әдісімен бірге қолданылуымыз.

Атаулы нысандарды шығару мақсаты бастапқы мәтіннен ерекшеліктері бар ескертулер жинағын алу болып табылады. Осы мақсатқа жету үшін біз Tomita талдаушысына екі грамматиканы жаздық. Біріншісі алдын ала өңдеу сатысында қолданылады, оның көмегімен біз сөздік емес атауларды шығарып, олармен уақытша сөздіктерді толтырамыз. Екінші грамматика құрастырылған сөздіктерді пайдалана отырып, атау бөліктерін біріктіріп, сәйкес белгілерді шығарады. Екі грамматика да тоқтату сөздерді, географиялық нысандар мен ұйымдарға арналған дескрипторларды, тақырыптарды және т.б. сипаттайтын қолдан жасалған кілт сөздік сөздіктерді пайдаланады. Атаулы нысандарды шығару моделі I-суретте көрсетілген.



Сур. 1. Атаулы нысандарды шығару моделі. GRM-грамматика, VOC-кілт сөздер сөздіктері (Fig. 1. Named object extraction model. GRM-grammar, VOC-keyword dictionaries)

Зерттеу әдістері

Қойылған міндеттерді шешу үшін атаулы нысандарды шығару әдістері, кластерлеу, кілт сөздер сөздіктері және жұптық модельді қолданылды.

Сөздік атауларды шығаруға арналған грамматика. Бұл кезеңнің негізгі проблемасы-жеке тұлғаларға қатысты атауларды басқа нысандар түрлеріне (мысалы, орындар немесе ұйымдар) қатысты атаулардан бөлу. Осы мақсат үшін біз белгілі бір нысан түріне жату ықтималдығы жоғары контекстті сипаттайтын ережелер жасаймыз. Мысалы, географиялық нысандар мен ұйымдарға арналған дескрипторларды қамтитын түйінді сөздіктердегі сөздер.

	A	B	C	D
1	Өлеңдерге қатысты атау ана, ана, тәте, қандайымы, ару, қашым, сүйсің, азаматтың, бөйжеткен.	Ерлерге қатысты атау мырза азамат әке, аға, іні іншек, бауыр,	Орын дескрипторлары сусжай, тау, қала, провинция на, сарай, аймақ, теңіз, көл, арал, облыс.	Ұйым дескрипторлары агенттік, одақ, жандатыс - біртектік ; Оанк, Оқар, Бөлім, Заруыт, институт

Сур. 2. кілттік сөздер сөздік үлгісі (Fig. 2. dictionary of keywords)

Толық атауды алуға арналған грамматика: Негізгі атаулардың сөздігін жасағаннан кейін олардың құрамдас бөліктерінен жасалған толық атауларды шығаруға болады: Аты (Аты), Фамилия (Тегі), Әкесінің аты, Сондай-ақ, біз келесідей гендерлік (Gender), Мекенжай (Address) және Дескриптор (Descriptor) сияқты қосымша атрибуттарды аламыз.

Мекенжай (Address) атрибутының мәні атаулардың сөздігінен кез келген сөзді қабылдай алады (2-сурет кілттік сөздер сөздік үлгісі).

Дескриптор (Descriptor) — бұл белгілі бір адамға қатысты сыныпты сипаттайтын атаулы топ (мысалы, кәсібі немесе туыстық байланысты мәртебе).

Gender атрибуты үш мәннің кез келгенін қабылдай алады: ”male”, ”female”, ”dual”.

Gender бірнеше жолмен анықталады. Біріншіден, егер қол жетімді болса, біз сөздік ақпаратты қолданамыз. Содан кейін синтаксистік байланысты етістіктердің, сын есімдердің грамматикалық түріне байланысты анықталады.

Толық атауды шығару үшін грамматикалық нәтиже:

Мысал: “Молекулярлық биология және биохимия институтының ғалымы [Бақытжан Әлжанұлы] геномды редакциялау көмегімен I типті қант диабетін емдеуге арналған бірегей жасушаларды ойлап тапты.”

StrictName

```
{
  Firstname = Бақытжан
  Lastname = Әлжанұлы
  Patronymic = None
  Gender = male
  Address = None
  Descriptor = ғалым
}
```

Белгілері бар ескертулер: {“Бақытжан”, “Әлжанұлы”, “None”, “male”, “None”, “ғалым” }

Кластерлеу

Әрбір D құжаты $M = \{m_1, \dots, m_n\}$ n атаулы нысандар жиыны ретінде ұсынылсын. Барлық атаулы нысандар $MP = \{(i, j) \mid 1 \leq i < j \leq N\}$ бірыңғай жұптарының жиынтығын құрайды, мұнда әр атаулы нысандар өз индексімен берілген. Әрбір MP атаулар жұбының өз салмағы $P_{(i,j) \in MP}$ болады, ол score (i, j) функциясының нәтижесі болып табылады, ол екі атаулы нысанның қаншалықты референты екенін көрсетеді:

$$\text{score}(i, j) = \begin{cases} \text{None} & \text{if } \text{Disjoint}(i, j) \\ \theta^T f(i, j), & \text{otherwise} \end{cases} \quad (1)$$

мұндағы $f(i, j)$ — атаулар үшін мүмкіндік векторы, θ атаулардың векторлық салмағы, Disjoint функциясы екі ескертпелер арасында олардың бір нысанға қатысты болуына кедергі болатын қайшылықтардың бар-жоғын көрсетеді.

Кореференциялық қатынасты шешу міндетін қою үшін, біз $n \times n$ логикалық үшбұрышты матрицаны C кластерлеу нәтижесі деп санаймыз, егер кореференты болса, онда $C_{ij} = 1$, кері жағдайда 0ге тең.

C кластерленуі транзитивті шектеулерге қолдау көрсетілетін жағдайда ғана жарамды болады: $(C_{ij} = 1 \wedge C_{jk} = 1) \Rightarrow (C_{ik} = 1) \forall 1 \leq i < j < k \leq n$

C[i]- і-ескертуге жататын кластер.

Жұптық модель. Кореференциялық модельдің екінші түрі екілік классификатор болып табылады. Идея - екі ескертпенің негізгі ықтималдығын табатын классификаторды дайындау. (Loic et al., 2022: 88) Ол сөйлемдегі әрбір екі сөзді салыстырады және олардың байланысу ықтималдығын тағайындайды (Gorka et al., 2019: 35). Идеал жағдай теріс мысалдар 0-ге жақын, ал оң мысалдар 1-ге жақын ықтималдық алады.

Жұптық моделді қолданудағы негізгі мақсат:

а) Ескертулердің әрбір жұбы үшін олардың бір негізгі референциялық тізбекте екенін анықтау.

б) Соңғы негізгі референциалдық тізбектерді құру үшін жіктелген жұптарды алу.

(Auliarachman et al., 2019) референциялық тізбекте ескерту жұптарын шығру және тізбектерді жіктеу әдістері келесі жұмыстан көруге болады (Shanu et al., 2019: 4179)

Біз ескертпелердің салмағын алу үшін логистикалық классификатормен аталған жұптық моделін оқыттық. Модель екі ескертудің бір кластерге жататынын болжайды. Кореференттік ықтималдық стандартты логистикалық формуланы қабылдайды:

$$P_{\theta}(i, j) = (1 + e^{-\theta^T f(i, j)})^{-1} \quad (2)$$

Мұндағы $f(i, j)$ m_i мен m_j үшін атаулар векторы, ал θ салмақ векторы θ салмақ векторы табу үшін, — жаттығу жиынындағы барлық ескертпелердің жиыны деп есептейміз, $T(j)$ — j алдындағы ескертпелердің индекстік жиыны болсын, осылайша, негізгі референциалды және $F(j)$ — референциалды емес. $T(j)$ және $F(j)$ жиындары ескертпесі жататын әрбір D құжаты үшін құрылған. $T(j)$ жиынындағы ықтималды үміткерлер саны $F(j)$ жиынындағы ықтималды үміткерлер санынан көп екендігін есептейтін салмақ векторы келесі формула арқылы есептеледі.

$$L(\theta) = -\sum_{m \in M} (\sum_{t \in T(m)} \log(p_{\theta}(t, m)) + \sum_{t \in T(m)} \log(1 - p_{\theta}(t, m))) + \min \|\theta\|_1 \quad (3)$$

Мүмкіндік векторы (Вектор признаков). Мүмкіндік векторы екі ескертпенің әртүрлі мүмкіндік мәндерінің арасындағы сәйкестіктерді көрсетеді.

First name, Last name және Patronymic белгілерінде сәйкестіктің 5 нұсқасы бар:

1. Strict Match (қағаз сәйкестік)-Белгінің мәндері бірдей.

2. Incomplete Match (толық емес сәйкестік) — Белгінің мәндері іс жүзінде бірдей.

Толық емес сәйкестікті тексеру үшін біз белгілердегі екі жол көрінісі арасындағы Левенштейн қашықтығын есептейміз. Левенштейн қашықтығы- бір сөзді екіншісіне айналдырмас бұрын қажет болатын бір әріптен тұратын түрлендірулердің (кірістіру, ауыстыру, жою) ең аз санын табуға арналған жол көрсеткіші.

Толық емес сәйкестікті тексеру қалыпқа келтіру қателері мен технологиялық қателер түзетілген бірдей атауларды табу үшін қажет. Егер Левенштейннің арақашықтығы қазақ тіліндегі зат есімнің аяқталуының 3 максимумынан аспаса, біз жолдарды бір - біріне жақын деп санаймыз.

Егер бір жол екіншісінің бөлігі болса Firstname Match белгісінің мәні Incomplete Match тең болады. Мысал : ‘Анна Мария’ және ‘Анна’.

3) First Letter Match (бірінші әріп бойынша сәйкестік)- егер белгілердің кем дегенде біреуінің мәні бір әріпке дейін қысқартылса және бұл әріп басқа мәннің бірінші әрпімен сәйкес келсе, онда Firstname Match = First Letter Match болады.

4) None Field- Егер өрістердің кем дегенде біреуі “ None “ мәндерін қабылдаса

5) Errormatch -Белгілердің мәндері мүлдем өзгеше.

Рейтинг моделі. Рейтинг моделі жұптық және маңызды модельдер арасындағы аралық қадам болып саналады. Осы модельді қолданған кезде үміткердің бірнеше ескертулерін тексеруге болады, осылайша барлық үміткерлер арасындағы ықтималдығы ең жоғарғысы таңдалады.

Gender және Address белгілерінде тек 3 сәйкестік опциясы бар: Match, Errormatch және None Field. Gender және Address белгілері шектеулі сөздердің мағыналарын қабылдай алады.

Descriptor белгісі ескертулерден әрбір жолда тек дескриптордың болуын көрсетеді.

Жұптық салмақ. Жоғарыда сипатталғандай, құжаттағы барлық ескертулер көптеген бірегей жұптарды құрайды және олардың әрқайсысының өз салмағы бар, бұл ескертулердің бір кластерде қаншалықты жақын болуын көрсетеді, яғни кореллянтты болу ықтималдығы.

Кесте 1-Атаулы жұптарға арналған мүмкіндіктер

id	Sign	Meaning	Example
1	Firstname_match	Strict Match Incomplete Match First Letter Match None Field Errormatch	“Анар” “Анар” “Анар” “Анара” “Анар” “А” “Анар” “None” “Анар” “Жанат”
2	Lastname_match	Strict Match Incomplete Match First Letter Match None Field Errormatch	“Омарова” “Омарова” “Омарову” “Омарова” “Омарова” “О” “Омарова” “None” “Омарова” “Ержанов”

3	Patr_match	Strict Match Incomplete Match First Letter Match None Field Erromatch	“Серікұлы” “Серікұлы” “Серікұлы” “Серікұлын” “Серікұлы” “С” “Серікұлы” “None” “Серікұлы” “Ержанұлы”
4	Gender_match	Match None Field Erromatch	”male” ”male” ”male” ”None” ”male” ”female”
5	Addr_match	Match None Field Erromatch	“ханым” “ханым” “ханым” “None” “ханым” “мырза”
6	Descr_presence	Both true First is true Second is true Both false	“true” “true” “true” “false” “true” “true” “false” “false”

Егер ескертпелер сәйкес келмесе және нақты кореференциялық қатынас болмаса, жұптық салмақ функциясы “None” мәнін қайтара алады, әйтпесе функция ескертулер жұптары үшін өлшенген мүмкіндіктердің қосындысын қайтарады.

$$scor(i, j) = \begin{cases} None, & \text{if } Disjoint(i, j) = true \\ \theta^T f(i, j), & \text{otherwise} \end{cases}$$

Disjoint функциясы келесі көрініске ие:

$$Disjoint(i, j) = \begin{cases} true & \text{if } Firstname_{match}[Mismatch] = true \\ & \text{or } Lastname_{match}[mismatch] = true \\ & \text{or } Gender_{match}[mismatch] = true \\ & \text{or } \theta^T f(i, j) < 0 \\ false & \text{otherwise} \end{cases}$$

Нәтижесінде біз келесідей жиынтықты аламыз: $p = \{score(i, j) | (i, j) \in MP$

Кластерлеу алгоритмі: Кластерлеу жұп салмақтарды қолдану және транзиттілік шектеу арқылы жүзеге асады. Корференцияны шешу үшін кластерлеудің кең таралған тәсілі-best-first кластерлеу қолданылады (Juntao et al., 2020: 11). Әрбір ескерту үшін алгоритм ықтималдығы жоғары кореференцияны таңдайды.

Бұл тәсілдің кемшілігі бұл алгоритм субъект деңгейінде емес, жұптық деңгейде шешім жасауға арналған. Яғни кілттік сөздер мен атрибуттар негізінде жұптық модельге негізделеді.

Бұл мәселені болдырмау үшін біз кластерлерді біріктірудің осы кезеңінде мән деңгейін енгіземіз, жұптасқан салмақ функциясы екі кластердің кез-келген ескертулерінің кем дегенде біреуін қабылдамайтынын тексереміз. Кластерлерді біріктіру процесі агломердің көмегімен жүзеге асырылады:

әр ескерту өзінің жеке бір элементтік кластерінде басталады, содан кейін әр қадамда екі кластер біріктіріледі.

Біріншіден, біз барлық жұптарды олардың жұптық салмақ функциясына сүйене отырып сұрыптаймыз. Бұл оңай шешімдер қабылдауға және қол жетімді ақпарат санын көбейтуге мүмкіндік береді.

Екіншіден сұрыпталған тізім бойынша Итерация жасаймыз. Әр жұп үшін біз элементтері бар кластерлерді біріктіру туралы шешім қабылдаймыз. Шешім екі кластың кез-келген екі ескертуі арасындағы қайшылықтардың болуына байланысты. Алгоритм 1 кластерлеу процедураларын сипаттайды.

```

input: p
output: C
S ← Sortdesc(P(i, j))
score(i, j) /= None
for (i, j) ∈ S do
  if C[i] != C[j] then
    if CanMerge(C[i], C[j]) then Merge(C[i], C[j])
  
```

Сур. 3. Кластерлеу алгоритмы
(Fig. 3. Clustering algorithm)

Эксперимент және нәтижелер

Оқу тәжірибе деректеріне Tengrinews.kz- тен жаңалықтар топтамасын қолдандық. Жалпы қолданылған мәтіндер әр түрлі тақырыптардағы 300 ге жуық мәтін алдынды, оның 122-і оқу жинағы (development set) ретінде берілген, ал 132-і сынақ жинақ (test set) және бағалау үшін пайдаланылады (2-кесте)

Кесте -2: Жаттығулар мен сынақ жинақтары бойынша статистика

Құжаттар	development set	test set
Documents	145	165
Mentions	173	201
Entities	162	182
Average # mentions	7.04	12.5
Average # entities	6.07	5.03

Алгоритмдердің нәтижелері дәстүрлі бағалау метрикасының көмегімен бағаланды: Дәлдік (Precision), Recall (толықтық) және F-measure (F-өлшемі).

Төменде бұл көрсеткіштердің есептеу формуласы:

$$Precision = \frac{TP}{(TP+FP)} \quad (4)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (5)$$

$$F_1 = \frac{2*Recall*Precision}{(Recall+Precision)} \quad (6)$$

Барлық нәтижелер F-өлшемі бойынша сұрыпталған.

Томиита-парсер көмегімен жеке атауларды алу үшін кіріктірілген алгоритм ұсынады. Ол мәтіндегі ішкі сөздіктен атауларды табады, содан кейін тегі, саны және жағдайы бойынша қарама-қайшылықсыз байланыс аттарын біріктіреді. Біз томит-парсердің нәтижелерін негізгі деңгей ретінде қолдандық.

Сынақ жинағы бойынша салыстыру нәтижелері 3-кестеде келтірілген.

Кесте-3 Сынақ жинағы бойынша салыстыру нәтижелері

	Precision	Recall	
violet	0.8750	0.8690	0.8850
crimson	0.8640	0.8430	0.8345
black	0.8775	0.8685	0.3615
pink	0.8634	0.8473	0.8422
aquamarine	0.7625	0.7385	0.7485
grey	0.8201	0.8289	0.8109
Біздің метод	0.8809	0.8726	0.8786

Кластерлеу кезеңінің нәтижелері

Зерттеу жұмысымыздың екінші трегі олар сілтеме жасайтын нысандағы сілтемелерді біріктіруге арналды. Екінші трекке тек 3 команданы алдық.

Сынақ жиынтығындағы салыстыру нәтижелері 4-кестеде көрсетілген.

Кесте 4-. Сынақ жиынтығындағы кластерлеу нәтижелері

	Precision	Recall	
violet	0.9050	0.8802	0.8850
crimson	0.8940	0.8630	0.8945
black	0.8209	0.8103	0.203
Біздің метод	0.8193	0.8245	0.8159

Қортынды

Осы жұмыстың аясында қазақ тіліндегі мәтіндер бойынша кореференциялық қатынасты шешудің тәсілі сипатталған.

Кореференциялық қатынасты шешу екі кезеңнен тұрды: ескертулерді алу және оларды кластерлеу. Ескертулер Томиита -парсер үшін жазылған грамматика арқылы алынды. Алгоритм үшін субъектілер деңгейінде өтетін және жұптық белгілердің салмақ векторларын пайдалана отырып, агломеративті кластерлеу таңдалды.

Тәжірибе көрсеткендей, біз салыстырмалы нәтижелер алдық, ол Томиита — парсердың кірістірілген алгоритмдерінің базалық деңгейінен жоғары болды. Болашақта біз аталған экстракция сатысындағы дәлдікті жақсарту, қалыпқа келтіру қателерін түзету және мүмкіндіктер жиынтығын жақсарту арқылы нәтижелерді жақсартуды жоспарлап отырмыз.

ӘДЕБИЕТТЕР

Қалман Г., Самбетбаева М.А., Жұмабай Е.С., 2022 — *Қалман Г., Самбетбаева М.А., Жұмабай Е.С.* Қазақ тіліндегі есімдік анафорасын шешу алгоритімі, *Еуразия инновациялық университетінің Хабаршысы*, № 2 ISSN 2709-3077, 126 бет.

Қалман Г., Самбетбаева М., Актаева Д. & Илюбаев А., 2022 — *Қалман Г., Самбетбаева М., Актаева Д. & Илюбаев А.* Машиналық оқыту әдістеріне негізделген анафораны шешу моделі. *Известия НАН РК. Серия физико-математическая*, (4), 56–67. <https://doi.org/10.32014/2022.2518-1726.156>.

A. Bodrova and N. Grafeeva, 2016 — *A. Bodrova and N. Grafeeva* “Coreference resolution using clusterization,” *2016 International FRUCT Conference on Intelligence, Social Media and Web (ISMW FRUCT)*, 2016. Pp. 1–8, doi: 10.1109/FRUCT.2016.7584764.

Benjamin Radford, 2020 — *Benjamin Radford*. Seeing the Forest and the Trees: Detection and Cross-Document Coreference Resolution of Militarized Interstate Disputes. In *Proceedings of the Workshop on Automated Extraction of Socio-political Events from News 2020*. Pp 35–41.

Damrin Kim, Seongsik Park, Mirae Han, and Harksoo Kim, 2022 — *Damrin Kim, Seongsik Park, Mirae Han, and Harksoo Kim*. Pipeline Coreference Resolution Model for Anaphoric Identity in Dialogues. In *Proceedings of the CODI-CRAC 2022 Shared Task on Anaphora, Bridging, and Discourse Deixis in Dialogue*, Gyeongju, Republic of Korea. Pp 28–31.

Gorka Urbizu, Ander Soraluze, and Olatz Arregi, 2019 — *Gorka Urbizu, Ander Soraluze, and Olatz Arregi*. Deep Cross-Lingual Coreference Resolution for Less-Resourced Languages: The Case of Basque. In *Proceedings of the Second Workshop on Computational Models of Reference, Anaphora and Coreference*, Minneapolis. USA. Pp 35–41.

Juntao Yu, Nafise Sadat Moosavi, Silviu Paun, and Massimo Poesio, 2021 — *Juntao Yu, Nafise Sadat Moosavi, Silviu Paun, and Massimo Poesio*. Stay Together: A System for Single and Split-antecedent Anaphora Resolution. In *Proceedings of the 2021. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Pp. 4174–4184.

Juntao Yu, Sopan Khosla, Nafise Sadat Moosavi, Silviu Paun, Sameer Pradhan, and Massimo Poesio, 2022 — *Juntao Yu, Sopan Khosla, Nafise Sadat Moosavi, Silviu Paun, Sameer Pradhan, and Massimo Poesio*. The Universal Anaphora Scorer. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. Pp. 4873–4883.

Juntao Yu, Alexandra Uma, and Massimo Poesio, 2020 — *Juntao Yu, Alexandra Uma, and Massimo Poesio*. A Cluster Ranking Model for Full Anaphora Resolution. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, Marseille, France. Pp 11–20.

Kibrik A.A., Dobrov G.B., Khudyakova M.V., Loukachevitch N.V., Pechenyj A., 2013 — *Kibrik A.A., Dobrov G.B., Khudyakova M.V., Loukachevitch N.V., Pechenyj A.* A corpus-based study of referential choice: Multiplicity of factors and machine learning techniques, *Text processing and cognitive technologies. Cognitive modeling in linguistics: Proceedings of the 13th International Conference*, Corfu. Pp. 118–126.

K. Wohiduzzaman and S. Ismail, 2018 — *K. Wohiduzzaman and S. Ismail*. “Recommendation System for Bangla News Article with Anaphora Resolution,” *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEICT)*. Pp. 467–472, doi: 10.1109/CEEICT.2018.8628075.

Lavie Alon, Tomita Masaru, 1993 — *Lavie Alon, Tomita Masaru*. GLR* - An Efficient Noise-skipping Parsing Algorithm for Context Free Grammars // In *Proceedings of the Third International Workshop on Parsing Technologies*. Pp. 123–134

Liyan Xu and Jinho D. Choi., 2020 — *Liyan Xu and Jinho D. Choi*. Revealing the Myth of Higher-Order Inference in Coreference Resolution. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Pp 8527–8533.

Loic De Langhe, Orphee De Clercq, and Veronique Hoste, 2022 — *Loic De Langhe, Orphee De Clercq, and Veronique Hoste*. Investigating Cross-Document Event Coreference for Dutch.

In Proceedings of the Fifth Workshop on Computational Models of Reference, Anaphora and Coreference, Gyeongju, Republic of Korea. Pp 88–98.

Ondřej Pražák, Miloslav Konopík, and Jakub Sido, 2021 — *Ondřej Pražák, Miloslav Konopík, and Jakub Sido*. Multilingual Coreference Resolution with Harmonized Annotations. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021). Pp 1119–1123.

Shany Barhom, Vered Shwartz, Alon Eirew, Michael Bugert, Nils Reimers, and Ido Dagan, 2019 — *Shany Barhom, Vered Shwartz, Alon Eirew, Michael Bugert, Nils Reimers, and Ido Dagan*. Revisiting Joint Modeling of Cross-document Entity and Event Coreference Resolution. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy. Pp 4179–4189.

Tomita parser. — Yandex technologies. — URL: <https://tech.yandex.ru/tomita/?ncrnd=2315>.

T. Auliarachman and A. Purwarianti, 2019 — *T. Auliarachman and A. Purwarianti*. “Coreference Resolution System for

Indonesian Text with Mention Pair Method and Singleton Exclusion using Convolutional Neural Network,” 2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA). Pp. 1–5, doi: 10.1109/ICAICTA.2019.8904261.

Yuval Kirstain, Ori Ram, and Omer Levy, 2021 — *Yuval Kirstain, Ori Ram, and Omer Levy*. Coreference Resolution without Span Representations. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Pp 14–19.

Zhumabay Y., Kalman G., Sambetbayeva M., Yerimbetova A., Ayapbergenova A. Bizhanova A., 2022 — *Zhumabay Y., Kalman G., Sambetbayeva M., Yerimbetova A., Ayapbergenova A. Bizhanova A.* Building a model for resolving referential relations in a multilingual system. Eastern-European Journal of Enterprise Technologies, V 2(2 (116)). Pp. 27–35.

REFERENCES:

A. Bodrova and N. Grafeeva, 2016 — *A. Bodrova and N. Grafeeva*. “Coreference resolution using clusterization,” 2016 International FRUCT Conference on Intelligence, Social Media and Web (ISMW FRUCT), 2016. Pp. 1–8, doi: 10.1109/FRUCT.2016.7584764. (in Eng.).

Benjamin Radford, 2020 — *Benjamin Radford*. Seeing the Forest and the Trees: Detection and Cross-Document Coreference Resolution of Militarized Interstate Disputes. In Proceedings of the Workshop on Automated Extraction of Socio-political Events from News 2020. Pp 35–41. (in Eng.).

Damrin Kim, Seongsik Park, Mirae Han, and Harksoo Kim, 2022 — *Damrin Kim, Seongsik Park, Mirae Han, and Harksoo Kim*. Pipeline Coreference Resolution Model for Anaphoric Identity in Dialogues. In Proceedings of the CODI-CRAC 2022 Shared Task on Anaphora, Bridging, and Discourse Deixis in Dialogue, Gyeongju, Republic of Korea. Pp 28–31. (in Eng.).

Gulzhamal K., Sambetbaeva M., Aktaeva D. & Ilyubaev A., 2022 — *Gulzhamal K., Sambetbaeva M., Aktaeva D. & Ilyubaev A.* ANAPHORA RESOLUTION MODEL BASED ON MACHINE LEARNING METHODS. News of NAS RK. Series physico-mathematical, (4), 56–67. <https://doi.org/10.32014/2022.2518-1726.156> (in Kazakh).

G. Kalman, M.A. Sambetbayeva, Y.S. Zhumabay, 2022 — *G. Kalman, M.A. Sambetbayeva, Y.S. Zhumabay*. Algorithm for solving the anaphora of a pronoun in the Kazakh language, bulletin of the innovative university of Eurasia, № 2. ISSN 2709-3077. P. 126 (in Kaz.).

Gorka Urbizu, Ander Soraluze, and Olatz Arregi, 2019 — *Gorka Urbizu, Ander Soraluze, and Olatz Arregi*. Deep Cross-Lingual Coreference Resolution for Less-Resourced Languages: The Case of Basque. In Proceedings of the Second Workshop on Computational Models of Reference, Anaphora and Coreference, Minneapolis, USA. Pp 35–41. (in Eng.).

Juntao Yu, Nafise Sadat Moosavi, Silviu Paun, and Massimo Poesio, 2021 — *Juntao Yu, Nafise Sadat Moosavi, Silviu Paun, and Massimo Poesio*. Stay Together: A System for Single and Split-antecedent Anaphora Resolution. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Pp. 4174–4184 (in Eng.).

Juntao Yu, Sopan Khosla, Nafise Sadat Moosavi, Silviu Paun, Sameer Pradhan, and Massimo Poesio, 2022 — *Juntao Yu, Sopan Khosla, Nafise Sadat Moosavi, Silviu Paun, Sameer Pradhan, and Massimo Poesio*. The Universal Anaphora Scorer. In Proceedings of the Thirteenth Language Resources and Evaluation Conference. Pp 4873–4883, (in Eng.).

Juntao Yu, Alexandra Uma, and Massimo Poesio, 2020 — *Juntao Yu, Alexandra Uma, and Massimo Poesio*. A Cluster Ranking Model for Full Anaphora Resolution. In Proceedings of the Twelfth Language Resources and Evaluation Conference, Marseille, France. Pp 11–20. (in Eng.).

Kibrik A.A., Dobrov G.B., Khudyakova M.V., Loukachevitch N.V., Pechenyj A., 2013 — *Kibrik A.A., Dobrov G.B., Khudyakova M.V., Loukachevitch N.V., Pechenyj A.* A corpus-based study of referential choice: Multiplicity of factors and machine learning techniques, Text processing and cognitive technologies. Cognitive modeling in linguistics: Proceedings of the 13th International Conference, Corfu. Pp. 118–126. (in Eng.).

K. Wohiduzzaman and S. Ismail, 2018 — *K. Wohiduzzaman and S. Ismail*. “Recommendation System for Bangla News Article with Anaphora Resolution,” 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (*iCEEICT*). Pp. 467-472, doi: 10.1109/CEEICT.2018.8628075. (in Eng.).

Lavie Alon, Tomita Masaru, 1993 — *Lavie Alon, Tomita Masaru*. GLR* - An Efficient Noise-skipping Parsing Algorithm For Context Free Grammars // In Proceedings of the Third International Workshop on Parsing Technologies. Pp. 123–134 (in Eng.).

Liyan Xu and Jinho D. Choi, 2020 — *Liyan Xu and Jinho D. Choi*. Revealing the Myth of Higher-Order Inference in Coreference Resolution. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Pp. 8527–8533. (in Eng.).

Loic De Langhe, Orphee De Clercq, and Veronique Hoste, 2022 — *Loic De Langhe, Orphee De Clercq, and Veronique Hoste*. Investigating Cross-Document Event Coreference for Dutch. In Proceedings of the Fifth Workshop on Computational Models of Reference, Anaphora and Coreference, Gyeongju, Republic of Korea. Pp 88–98. (in Eng.).

Ondřej Pražák, Miloslav Konopík, and Jakub Sido, 2021 — *Ondřej Pražák, Miloslav Konopík, and Jakub Sido*. Multilingual Coreference Resolution with Harmonized Annotations. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021). Pp 1119–1123. (in Eng.).

Shany Barhom, Vered Shwartz, Alon Eirew, Michael Bugert, Nils Reimers, and Ido Dagan, 2019 — *Shany Barhom, Vered Shwartz, Alon Eirew, Michael Bugert, Nils Reimers, and Ido Dagan*. Revisiting Joint Modeling of Cross-document Entity and Event Coreference Resolution. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy. Pp 4179–4189. (in Eng.).

Tomita parser. — Yandex technologies. — URL: <https://tech.yandex.ru/tomita/?ncrnd=2315>.

T. Auliarachman and A. Purwarianti, 2019 — *T. Auliarachman and A. Purwarianti*. “Coreference Resolution System for Indonesian Text with Mention Pair Method and Singleton Exclusion using Convolutional Neural Network,” 2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA). Pp. 1–5, doi: 10.1109/ICAICTA.2019.8904261. (in Eng.).

Yuval Kirstain, Ori Ram, and Omer Levy, 2021 — *Yuval Kirstain, Ori Ram, and Omer Levy*. Coreference Resolution without Span Representations. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Pp 14–19. (in Eng.).

Zhumabay Y., Kalman G., Sambetbayeva M., Yerimbetova A., Ayapbergenova A. Bizhanova A., 2022 — *Zhumabay Y., Kalman G., Sambetbayeva M., Yerimbetova A., Ayapbergenova A. Bizhanova A.* Building a model for resolving referential relations in a multilingual system. Eastern-European Journal of Enterprise Technologies, V 2(2) (116). Pp. 27–35. (in Eng.).

МАЗМҰНЫ

Ж.К. Абдугулова, Г.А. Ускенбаева, М.Н. Тлеген, А.К. Шукирова КҰБЫР ЖАБДЫҒЫНДА МАЙДЫ ҚЫЗДЫРУДЫҢ ТЕХНОЛОГИЯЛЫҚ ПРОЦЕСІН АВТОМАТТАНДЫРУ.....	5
Ж.С. Авкурова, С. Гнатюк, Л.М. Кыдыралина, Н.К. Курмангалиева АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ЖҮЙЕЛЕРДЕ ҚҰҚЫҚ БҰЗУШЫНЫ ЕРТЕ АНЫҚТАУ ЖӘНЕ СӘЙКЕСТЕНДІРУДІҢ ИНТЕЛЛЕКТУАЛДЫ ӘДІСІ.....	22
А. Бекарыстанкызы, Ө. Ж. Мамырбаев АГГЛЮТИНАТИВТІ ТІЛДЕРГЕ АРНАЛҒАН СӨЙЛЕУДІ АВТОМАТТЫ ТҮРДЕ ТАҢУ ЖҮЙЕСІ.....	37
А.С. Еримбетова, Э.Н. Дайырбаева, Л. Черикбаева БИКУБТЫҚ ИНТЕРПОЛЯЦИЯҒА НЕГІЗІНДЕ СУРЕТТЕРГЕ ЖАСЫРЫН АҚПАРАТТЫ ЕНГІЗУ.....	50
М.Б. Есенова, Г.Б. Абдикеримова, А. Толстой, Ж.Б. Ламашева, А.А. Некесова БИДАЙДАҒЫ АРАМШӨПТЕР ОШАҒЫН АНЫҚТАУ ҮШІН ТЕКСТУРАЛЫҚ БЕЛГІЛЕР ӘДІСТЕРІН ҚОЛДАНУ.....	64
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Р.К. Сенгирбаева НАҚТЫ УАҚЫТ РЕЖИМІНДЕ МЕДИАРИПЕ ЖӘНЕ SVM АРҚЫЛЫ ҚАЗАҚ ҰМ ТІЛІН ТАҢУ.....	82
Ж.С. Иксебаева, К. Жетписов, А.Б. Медешова, И.М. Бапиев, Ж.Ж. Багисов ҒАЛЫМДАРДЫҢ ҒЫЛЫМИ ЖОБАЛАР БОЙЫНША ГРАНТТЫҚ ҚАРЖЫЛАНДЫРУҒА ҚАТЫСУҒА ӨТІНІМДЕРІН ДАЙЫНДАУДЫҢ АҚПАРАТТЫҚ ЖҮЙЕСІ.....	94
А.А. Иманберді, Р.Н. Молдашева ӘЛЕУМЕТТІК МЕДИА ТАРАТУ ҮЛГІЛЕРІНЕ ШОЛУ.....	107
Г. Қалман, М.Ғ. Есмағанбет, М.М. Жаманкарин, А.И. Габдулина, Д.В. Плескачев КЛАСТЕРЛЕУ ӘДІСІН ҚОЛДАНЫП КОРЕФЕРЕНЦИЯН ШЕШУ.....	121

Қ.Т. Қырғызбай, Е.Х. Какимжанов ГАЗ ТЕХНОЛОГИЯЛАРЫ НЕГІЗІНДЕ АЛМАТЫ ОБЛЫСЫНЫҢ ГЕОДЕРЕКТЕР БАЗАСЫН ҚҰРУ ВІТСОІН ЖЕЛІСІНДЕГІ КҮДІКТІ ТРАНЗАКЦИЯЛАРДЫ АНЫҚТАУ.....	136
Ш.Ж. Мусиралиева, М.Ж. Шайзат, А.К. Бекетова, Е. Абайұлы, А.Б. Манасова ВІТСОІН ЖЕЛІСІНДЕГІ КҮДІКТІ ТРАНЗАКЦИЯЛАРДЫ АНЫҚТАУ.....	154
А.Ұ. Мұхиядин, Ұ.Т. Махажанова, М.У. Мукашева, А.А. Муханова АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР ТӨТЕНШЕ ЖАҒДАЙДА ҚАШЫҚТАН ОҚЫТУДА ЭКСПЕРИМЕНТТЫҚ ДЕРЕКТЕРДІ ТАЛДАУ ҚҰРАЛЫ РЕТІНДЕ.....	170
А.Б. Тоқтарова, Б.С. Омаров, Г.Н. Казбекова, С.А. Мамиков, Ф.Е. Темірбекова ӘЛЕУМЕТТІК ЖЕЛІДЕГІ ҚАЗАҚ ТІЛДІ БЕЙӘДЕП СӨЗДЕР ҚОРЫН МАШИНАЛЫҚ ОҚЫТУДА ЖИНАҚТАУ.....	191
А.Ә. Шекербек, Г.Б. Абдикеримова, Ж.Б. Ламашева, М.Г. Байбулова, А.К. Токкулиева ТЕРЕҢ ОҚЫТУ АЛГОРИТМІМЕН РЕНТГЕНДІК КЕСКІННІҢ КЛАССИФИКАЦИЯСЫ.....	204
Э.Э. Эльдарова JPEG2000 ҚЫСУЫНАН KEЙІН ЦИФРЛІК БЕЙНЕЛЕРДІҢ ВИЗУАЛДЫ САПАСЫН ЖАҚСАРТУ.....	228

СОДЕРЖАНИЕ

Ж.К. Абдугулова, Г.А. Ускенбаева, М.Н. Глеген, А.К. Шукирова АВТОМАТИЗАЦИЯ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА ПОДОГРЕВА НЕФТИ НА ТРУБОПРОВОДНОМ ОБОРУДОВАНИИ.....	5
Ж.С. Авкурова, С.А. Гнатюк, Л.М. Кыдыралина, Н.К. Курмангалиева ИНТЕЛЛЕКТУАЛИЗИРОВАННЫЙ МЕТОД РАННЕГО ВЫЯВЛЕНИЯ ИДЕНТИФИКАЦИИ НАРУШИТЕЛЯ В ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ СИСТЕМАХ.....	22
А. Бекарыстанқызы, О. Ж. Мамырбаев ИНТЕГРАЛЬНАЯ СИСТЕМА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ СЛИТНОЙ РЕЧИ ДЛЯ АГГЛЮТИНАТИВНЫХ ЯЗЫКОВ.....	37
А.С. Еримбетова, Э.Н. Дайырбаева, Л. Черикбаева ВНЕДРЕНИЕ СКРЫТОЙ ИНФОРМАЦИИ В ИЗОБРАЖЕНИИ НА ОСНОВЕ БИКУБИЧЕСКОЙ ИНТЕРПОЛЯЦИИ.....	50
М.Б. Есенова, Г.Б. Абдикеримова, А. Толстой, Ж.Б. Ламашева, А.А. Некесова ПРИМЕНИМОСТЬ МЕТОДОВ АНАЛИЗА ТЕКСТУРНЫХ ИЗОБРАЖЕНИЙ ДЛЯ ВЫЯВЛЕНИЯ ОЧАГОВ СОРНЫХ ТРАВ ПШЕНИЦЫ.....	64
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Р.К. Сенгирбаева РАСПОЗНАВАНИЕ КАЗАХСКОГО ЖЕСТОВОГО ЯЗЫКА В РЕАЛЬНОМ ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ MEDIAPIPE и SVM.....	82
Ж.С. Иксебаева, К. Жетписов, А.Б. Медешова, И.М. Бапиев, Ж.Ж. Багисов ИНФОРМАЦИОННАЯ СИСТЕМА ПОДГОТОВКИ ЗАЯВОК ДЛЯ УЧАСТИЯ В ГРАНТОВОМ ФИНАНСИРОВАНИИ УЧЕНЫХ ПО НАУЧНЫМ ПРОЕКТАМ.....	94
А.А. Иманберді, Р.Н. Молдашева ОБЗОР МОДЕЛЕЙ РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ.....	107

Г. Қалман, М.Ғ. Есмағанбет, М.М. Жаманқарин, А.Г. Габдулина, Д.В. Плескачев РЕШЕНИЕ КОРЕФЕРЕНЦИИ С ПОМОЩЬЮ МЕТОДА КЛАСТЕРИЗАЦИИ.....	121
Қ.Т. Қырғызбай, Е.Х. Какимжанов СОЗДАНИЕ БАЗЫ ГЕОДАНЫХ АЛМАТИНСКОЙ ОБЛАСТИ НА ОСНОВЕ ГИС-ТЕХНОЛОГИЙ О МЕТОДЕ ИДЕНТИФИКАЦИИ ПОДОЗРИТЕЛЬНЫХ ТРАНЗАКЦИЙ В БИТКОИН СЕТИ.....	136
Ш.Ж. Мусиралиева, М.Ж. Шайзат, А.К. Бекетова, Е. Абайұл, А.Б. Манасова О МЕТОДЕ ИДЕНТИФИКАЦИИ ПОДОЗРИТЕЛЬНЫХ ТРАНЗАКЦИЙ В БИТКОИН СЕТИ.....	154
А.Ұ. Мұхиядин, У.Т. Махажанова, М.У. Мукашева, А.А. Муханова ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ КАК СРЕДСТВО АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ ПРИ ЭКСТРЕННОМ ДИСТАНЦИОННОМ ОБУЧЕНИИ.....	170
А.Б. Токтарова, Б.С. Омаров, Г.Н. Казбекова, С.А. Мамиков, Ф.Е. Темирбекова СБОР БАЗЫ ДАННЫХ О ЯЗЫКЕ НЕНАВИСТИ В СОЦИАЛЬНОЙ СЕТИ НА КАЗАХСКОМ ЯЗЫКЕ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ.....	191
А.А. Шекербек, Г.Б. Абдикеримова, Ж.Б. Ламашева, М.Г. Байбулова, А.К. Токкулиева КЛАССИФИКАЦИЯ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ АЛГОРИТМА ГЛУБОКОГО ОБУЧЕНИЯ.....	204
Э.Э. Эльдарова УЛУЧШЕНИЕ ВИЗУАЛЬНОГО КАЧЕСТВА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ ПОСЛЕ СЖАТИЕ JPEG2000.....	228

CONTENTS

J.K. Abdugulova, G.A. Uskenbayeva, M.N. Tlegen, A.K. Shukirova AUTOMATION OF THE TECHNOLOGICAL PROCESS OF HEATING OIL PIPELINE EQUIPMENT.....	5
Z. Avkurova, S. Gnatyuk, L. Kydyralina, N. Kurmangaliev THE INTELLECTUALIZED METHOD OF EARLY DETECTION AND IDENTIFICATION OF THE VIOLATOR IN INFORMATION AND COMMUNICATION SYSTEMS.....	22
A. Bekarystankyzy, O. Zh. Mamyrbayev INTEGRATED AUTOMATIC SPEECH RECOGNITION SYSTEM FOR AGGLUTINATIVE LANGUAGES.....	37
A. Yerimbetova, E. Daiyrbayeva, L. Cherikbayeva EMBEDDING HIDDEN INFORMATION IN IMAGES BASED ON BICUBIC INTERPOLATION.....	50
M. Yessenova, G. Abdikerimova, A. Tolstoy, Zh. Lamasheva, A. Nekessova APPLICABILITY OF TEXTURE IMAGE ANALYSIS METHODS FOR DETECTION OF WHEAT WEED POCKS.....	64
L. Zholshiyeva, T. Zhukabayeva, Sh. Turaev, M. Berdieva, R. Sengirbayeva REAL-TIME KAZAKH SIGN LANGUAGE RECOGNITION USING MEDIAPIPE AND SVM.....	82
Zh.S. Ixebayeva, K. Jetpisov, A.B. Medeshova, I.M. Bapiyev , Zh.Zh. Bagisov AN INFORMATION SYSTEM FOR THE PREPARATION OF APPLICATIONS FOR PARTICIPATION IN GRANT FUNDING OF SCIENTISTS IN SCIENTIFIC PROJECTS.....	94
A. Imanberdi, R. Moldasheva REVIEW OF MODELS OF DISSEMINATION OF INFORMATION IN SOCIAL NETWORKS.....	107
G. Kalman, M.G. Esmaganbet, M.M. Zhamankarin, A.I. Gabdulina, D.V. Pleskachev COREFERENCE SOLUTION USING THE CLUSTERING METHOD.....	121

K. Kyrgyzbay, E. Kakimzhanov CREATION OF A GEODATABASE OF ALMATY REGION BASED ON GIS TECHNOLOGIES.....	136
Sh. Mussiraliyeva, M. Shaizat, A. Beketova, Y. Abayuly, A. Manassova IDENTIFICATION OF SUSPICIOUS TRANSACTIONS IN THE BITCOIN NETWORK.....	154
A. Mukhiyadin, U. Makhazhanova, M. Mukasheva, A. Mukhanova INFORMATION TECHNOLOGIES AS A MEANS OF EXPERIMENTAL DATA ANALYSIS IN EMERGENCY DISTANCE LEARNING.....	170
A.B. Toktarova, B.S. Omarov, G.N. Kazbekova, S.A. Mamikov, F.E. Temirbekova COLLECTING HATE SPEECH DATABASE ON SOCIAL NETWORK IN KAZAKH LANGUAGE BY USING MACHINE LEARNING.....	191
A. Shekerbek, G. Abdikerimova, Zh. Lamasheva, M. Baibulova, A. Tokkuliyeva CLASSIFICATION OF X-RAY IMAGES USING THE DEEP LEARNING ALGORITHM.....	204
E.E. Eldarova IMPROVING THE VISUAL QUALITY OF DIGITAL IMAGES AFTER JPEG2000 COMPRESSION.....	228

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

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

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

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

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

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

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

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

Правила оформления статьи для публикации в журнале смотреть на сайтах:

www.nauka-nanrk.kz

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

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

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

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

Верстка на компьютере *Г.Д. Жадыранова*

Подписано в печать 30.03.2023.

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

15,5 п.л. Тираж 300. Заказ 1.