

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



**INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION
TECHNOLOGIES**

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

**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ
ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

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

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

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

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

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

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

EDITOR-IN-CHIEF:

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

DEPUTY EDITOR-IN-CHIEF:

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

EDITORIAL BOARD:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

MANAGING EDITOR

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

International Journal of Information and Communication Technologies

Periodicity: 4 times a year.

Languages: Kazakh, Russian, English

DOI prefix: 10.54309

ISSN 2708-2032 (print)

ISSN 2708-2040 (online)

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

Distribution: Materials are distributed under the Creative Commons Attribution 4.0

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

Owner: International Information Technology University JSC (Almaty).

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

РЕДАКЦИЯ

БАС РЕДАКТОР:

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

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

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

РЕДАКЦИЯЛЫҚ АЛҚА:

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

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

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

Микеле Пагано — PhD, Пиза Университетінің (Италия) профессоры

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

ЖАУАПТЫ РЕДАКТОР:

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

Халықаралық ақпараттық және коммуникациялық технологиялар журналы

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Префикс DOI: 10.54309

Мерзімділігі: жылына 4 рет.

Басылым тілі: қазақ, орыс, ағылшын.

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

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

Тарату: материалдар Creative Commons Attribution 4.0 лицензиясы бойынша таратылады

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

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

РЕДАКЦИЯ

ГЛАВНЫЙ РЕДАКТОР:

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

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

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

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

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

Микеле Пагано — PhD, профессор Университета Пизы (Италия)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

ОТВЕТСТВЕННЫЙ РЕДАКТОР:

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

Международный журнал информационных и коммуникационных технологий

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Префикс DOI: 10.54309

Периодичность: 4 выпусков в год.

Язык издания: казахский, русский, английский.

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

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

Распространение: материалы распространяются по лицензии Creative Commons Attribution 4.0

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

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

CONTENTS

DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF SOCIOECONOMIC SYSTEMS

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

INFORMATION TECHNOLOGY

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

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

INFORMATION SECURITY AND COMMUNICATION TECHNOLOGIES

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

МАЗМУНЫ

ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

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

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

АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

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

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

АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН

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

СОДЕРЖАНИЕ ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИ СОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ

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

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

М.Ж. Айтимов, Г.К. Муратова, Ж.К. Бисенбаева, И.М. Бапиев, М. Кассим

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

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

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

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

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Vol. 7. Is.1. Number 25 (2026). Pp. 247–268

Journal homepage: <https://journal.iitu.edu.kz><https://doi.org/10.54309/IJICT.2026.25.1.015>

FORMATION OF A UNIFIED DATA SET FOR ANALYZING THE RELATIONSHIP BETWEEN DOG BREEDS AND DISEASES

A. Kalykulova, K. Bakirov, A. Shoman, A. Nugumanova*

Astana IT University, Astana, Kazakhstan.

E-mail: aliyakalykulova01@gmail.com

Aliya Kalykulova — master student, Astana IT University, Astana, Kazakhstan

E-mail: aliyakalykulova01@gmail.com, <https://orcid.org/0009-0006-5641-3797>;

Kuanysh Bakirov — PhD, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0009-0002-8059-7240>;

Aruzhan Shoman — PhD, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0000-0002-7844-8601>;

Aliya Nugumanova — PhD, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0000-0001-5522-4421>.

© A. Kalykulova, K. Bakirov, A. Shoman, A. Nugumanova

Abstract. This article presents an approach to creating a dataset containing information about dog breeds, their classification in accordance with international standards, and hereditary diseases associated with these breeds. The main objective of the study was to integrate disparate sources of information and standardize breed and disease names, which enabled further comprehensive analysis. Official databases and specialized resources were used to form the final dataset: iDog, Canine Inherited Disorders Database (CIDD), The Kennel Club (UK), and American Kennel Club (AKC). The developed standardization algorithm, based on natural language processing methods, made it possible to eliminate inconsistencies in names, ensure data comparability, and form a structured dataset including 4,093 records. As part of the study, exploratory data analysis (EDA) was conducted to identify patterns in the distribution of breeds by size, classification groups, and disease types. The analysis showed that large and small breeds are characterized by a wider range of hereditary pathologies compared to medium-sized breeds. In addition, differences in the structure of diseases between The Kennel Club and American Kennel Club classifications were identified, which emphasizes the importance of taking national characteristics into account in international comparative studies. The results of the study demonstrate the significance of the developed dataset as a basis for building predictive models and decision support tools in veterinary medicine. The resulting dataset can be used to more accurately predict dogs' predisposition to



diseases and develop effective prevention and treatment strategies.

Keywords: Creating a dataset, data standardization, Exploratory Data Analysis, NLP

For citation: Kalykulova A., Bakirov K., Shoman A., Nugumanova A. (2026). Formation of a unified data set for analyzing the relationship between dog breeds and diseases // International journal of information and communication technologies. Vol. 7. No. 25. Pp. 247-268. <https://doi.org/10.54309/IJICT.2026.25.1.015>. (In Eng.).

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

Acknowledgements. *This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP23488904 Development of scientifically based formulation using unsupervised machine learning in the production of full canned wet feed for unproductive animals).*

ИТ ТҰҚЫМДАРЫ МЕН АУРУЛАРЫ АРАСЫНДАҒЫ БАЙЛАНЫСТЫ ТАЛДАУҒА АРНАЛҒАН БІРЫҢҒАЙ ДЕРЕКТЕР ЖИЫНТЫҒЫН ҚАЛЫПТАСТЫРУ

А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова*

Astana IT University, Астана, Қазақстан.

E-mail: aliyakalykulova01@gmail.com

Калыкулова Алия — магистрант, Astana IT University, Астана, Қазақстан
E-mail: aliyakalykulova01@gmail.com, <https://orcid.org/0009-0006-5641-3797>;

Бакиров Куаныш — PhD, Astana IT University, Астана, Қазақстан
<https://orcid.org/0009-0002-8059-7240>;

Шоман Аружан — PhD, Astana IT University, Астана, Қазақстан
<https://orcid.org/0000-0002-7844-8601>;

Нугуманова Алия — PhD, Astana IT University, Астана, Қазақстан
<https://orcid.org/0000-0001-5522-4421>.

© А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова

Аннотация. Бұл мақалада ит тұқымдары, олардың халықаралық стандарттарға сәйкес жіктелуі және осы тұқымдармен байланысты тұқым қуалайтын аурулар туралы ақпаратты қамтитын деректер жиынтығын құру тәсілі ұсынылады. Зерттеудің негізгі мақсаты ақпараттың әртүрлі дереккөздерінде шашыраңқы орналасқан мәліметтерді біріктіру және тұқымдар мен аурулар атауларын стандарттау болды. Бұл кейіннен деректерге кешенді талдау жүргізуге мүмкіндік берді. Қорытынды деректер жиынтығын қалыптастыру үшін ресми дерекқорлар мен мамандандырылған ресурстар: iDog, Canine Inherited Disorders Database (CIDD), The Kennel Club (UK) және American Kennel Club (AKC) пайдаланылды. Табиғи тілді өңдеу (NLP) әдістеріне негізделген әзірленген стандарттау алгоритмі атау-

лардағы сәйкессіздіктерді жоюға, деректердің салыстырмалылығын қамтамасыз етуге және 4 093 жазбадан тұратын құрылымдалған деректер жиынтығын қалыптастыруға мүмкіндік берді. Зерттеу аясында тұқымдардың мөлшері, жіктеу топтары және ауру түрлері бойынша таралу заңдылықтарын анықтау мақсатында деректерге барлау талдауы (Exploratory Data Analysis, EDA) жүргізілді. Талдау нәтижелері ірі және ұсақ ит тұқымдарында орташа мөлшердегі тұқымдармен салыстырғанда тұқым қуалайтын патологиялардың ауқымы кеңірек екенін көрсетті. Сонымен қатар, The Kennel Club және American Kennel Club жіктеулеріндегі аурулар құрылымының айырмашылықтары анықталды. Бұл халықаралық салыстырмалы зерттеулер жүргізу кезінде ұлттық ерекшеліктерді ескерудің маңыздылығын көрсетеді. Зерттеу нәтижелері әзірленген деректер жиынтығының ветеринариялық медицина саласында болжамдық модельдер мен шешім қабылдауды қолдау құралдарын құруға негіз бола алатынын көрсетеді. Алынған деректер жиынтығын иттердің ауруларға бейімділігін дәлірек болжау, сондай-ақ аурулардың алдын алу мен емдеудің тиімді стратегияларын әзірлеу үшін пайдалануға болады.

Түйінді сөздер: деректер жиынтығын құру, деректерді стандарттау, деректерге барлау талдауы, табиғи тілді өңдеу (NLP)

Дәйексөздер үшін: Калыкулова А., Бакиров К., Шоман А., Нугуманова А. (2026). Ит тұқымдары мен аурулары арасындағы байланысты талдауға арналған бірыңғай деректер жиынтығын қалыптастыру // Халықаралық ақпараттық және коммуникациялық технологиялар журналы. Т. 7. No 25. Б. 247–268. <https://doi.org/10.54309/IJICT.2026.25.1.015>. (Ағыл. тіл.).

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

Алғыс. Бұл зерттеу Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитетінің қаржылық қолдауымен жүзеге асырылды (№ AP23488904 грант «Өнімсіз жануарларға арналған толыққанды ылғалды консервіленген жем өндірісінде оқытусыз машиналық оқыту әдістерін қолдана отырып, ғылыми негізделген рецептураны әзірлеу»).

ФОРМИРОВАНИЕ ЕДИНОГО НАБОРА ДАННЫХ ДЛЯ АНАЛИЗА ВЗАИМОСВЯЗИ МЕЖДУ ПОРОДАМИ СОБАК И ЗАБОЛЕВАНИЯМИ

А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова*

Astana IT University, Астана, Казахстан.

E-mail: aliyakalykulova01@gmail.com

Калыкулова Алия — магистрант, Astana IT University, Астана, Казахстан
E-mail: aliyakalykulova01@gmail.com, <https://orcid.org/0009-0006-5641-3797>;

Бакиров Куаныш — PhD, Astana IT University, Астана, Казахстан
<https://orcid.org/0009-0002-8059-7240>;

Шоман Аружан — PhD, Astana IT University, Астана, Казахстан

<https://orcid.org/0000-0002-7844-8601>;

Нугуманова Алия — PhD, Astana IT University, Астана, Казахстан

<https://orcid.org/0000-0001-5522-4421>.

© А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова

Аннотация. В статье представлен подход к созданию набора данных, содержащего информацию о породах собак, их классификации в соответствии с международными стандартами, а также о наследственных заболеваниях, связанных с данными породами. Основной целью исследования являлась интеграция разрозненных источников информации и стандартизация названий пород и заболеваний, что обеспечило возможность их дальнейшего комплексного анализа. Для формирования итогового набора данных были использованы официальные базы данных и специализированные ресурсы: iDog, Canine Inherited Disorders Database (CIDD), The Kennel Club (UK) и American Kennel Club (AKC). Разработанный алгоритм стандартизации, основанный на методах обработки естественного языка, позволил устранить несоответствия в названиях, обеспечить сопоставимость данных и сформировать структурированный набор данных, включающий 4 093 записи. В рамках исследования был проведен разведочный анализ данных (Exploratory Data Analysis, EDA) для выявления закономерностей в распределении пород по размеру, классификационным группам и типам заболеваний. Анализ показал, что крупные и мелкие породы характеризуются более широким спектром наследственных патологий по сравнению с породами среднего размера. Кроме того, были выявлены различия в структуре заболеваний между классификациями The Kennel Club и American Kennel Club, что подчеркивает важность учета национальных особенностей при проведении международных сравнительных исследований. Результаты исследования демонстрируют значимость разработанного набора данных как основы для построения прогностических моделей и инструментов поддержки принятия решений в ветеринарной медицине. Полученный набор данных может использоваться для более точного прогнозирования предрасположенности собак к заболеваниям, а также для разработки эффективных стратегий профилактики и лечения.

Ключевые слова: создание набора данных, стандартизация данных, разведочный анализ данных, обработка естественного языка (NLP)

Для цитирования: Калыкулова А., Бакиров К., Шоман А., Нугуманова А. (2026). Формирование единого набора данных для анализа взаимосвязи между породами собак и заболеваниями // Международный журнал информационных и коммуникационных технологий. Т. 7. No. 25. Стр. 247-268. <https://doi.org/10.54309/IJICT.2026.25.1.015>. (На англ.).

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Благодарность. Данное исследование выполнено при финансовой поддерж-

ке Комитета науки Министерства науки и высшего образования Республики Казахстан (грант № AP23488904 «Разработка научно обоснованной рецептуры с использованием методов обучения без учителя при производстве полнораціонного влажного консервированного корма для непродуктивных животных»).

Introduction.

Modern information technologies enable the collection, processing, and analysis of large amounts of data, opening up new opportunities for veterinary medicine, particularly in studying the predisposition of animal breeds to various diseases. The problem is that information about breeds, breed diseases, types of disorders, predisposition levels, and breed sizes is often scattered across different sources, formatted in different ways, and even uses different terms and classifications. This fragmentation hinders the creation of large-scale, unified databases and, as a result, prevents the effective application of machine learning, predictive analytics, and other IT approaches to disease prevention and diagnosis.

The relevance of the study lies in the need for a systematic approach to collecting and analyzing data on dog breeds and related diseases for the application of modern information technologies in veterinary medicine. Currently, information on the types of disorders, predisposition levels, and breed characteristics is scattered, presented in different formats, and often uses inconsistent classifications, which makes it difficult to integrate and analyze. In a study by (Donner et al., 2018), based on the analysis of the DNA of more than 100,000 dogs, it was found that about 40 % of animals have at least one genetic variant associated with hereditary diseases, which indicates a high prevalence of disease predisposition and the need for comprehensive data analysis. Similarly, recent reviews in the field of veterinary informatics emphasize the importance of aggregating information from various sources, standardizing it, and improving its quality to create representative and reliable sets suitable for automated analysis and the implementation of digital solutions in veterinary practice (Lustgarten et al., 2020). These facts confirm that the development of approaches to data structuring and analysis is a key direction for improving the effectiveness of animal health diagnosis, prevention, and management.

The purpose of this study is to create and analyze a dataset containing information about dog breeds and related diseases, as well as to conduct an explanatory analysis to identify relationships between key characteristics and demonstrate the potential applications of modern data processing methods in veterinary informatics.

During the formation of the dataset, special attention was paid to the standardization of breed names, as different sources use different spellings and classifications. An additional complication was the heterogeneity of the information: for some breeds, there are officially approved groupings, while for others, no such classification exists. In addition, some breeds cannot be included in the dataset without official recognition by international or national cynological organizations, which limits the completeness of the information presented.

The novelty of the study lies in the presentation of a compact, unified dataset

created on the basis of data from various sources, with the names of breeds and diseases converted to a single format. This approach allows for preliminary explanatory analysis and demonstrates ways of working with disparate information in the field of veterinary informatics. The resulting dataset can serve as a starting point for further research aimed at expanding the volume of information and developing more comprehensive analytical tools.

Materials and methods.

In recent years, there has been rapid growth in research related to the use of big data in veterinary medicine, opening up new opportunities for animal health analysis, disease diagnosis, and the development of preventive measures. However, the process of integrating data on dog breeds and their hereditary diseases remains a challenging task due to the heterogeneity of sources, lack of standardization, and variability of data formats.

One of the key areas is the collection and systematization of data from various sources. (Paynter et al., 2021) review existing large databases in veterinary medicine, including clinical records, insurance data, and specially organized cohort studies. The authors highlight the problems of the lack of uniform standards for record keeping and the ambiguity of diagnostic terminology, which makes it difficult to integrate data and conduct comparative studies. These findings are confirmed by other studies, which show that electronic health records (EHRs) have significant potential for analysis but require careful processing and normalization. For example, (Davies et al., 2025) demonstrate how text processing and NLP methods can be applied to extract clinically relevant information, including the use of specialized dictionaries and regular expressions.

An important source of data is electronic medical records from primary veterinary care. (Summers et al., 2019) and (Jackson et al., 2024) showed that such data can be used to identify the most common diseases in dogs, analyze age-related characteristics, and study factors affecting health. However, they also note that data quality depends on the completeness of records, the level of detail, and the participation of clinics in data collection projects.

The issue of standardizing information on dog breeds is of fundamental importance for further analysis. (Mullen et al., 2025) proposed the development of an ontology of vertebrate breeds that would allow data from various sources, such as FAO DAD-IS and national registries, to be combined while preserving their provenance. Despite the promise of this approach, the researchers encountered difficulties related to conflicting sources and the need for manual data verification. Additionally, (Perez et al., 2025) consider computer vision methods for automatic breed identification from images, which could help fill gaps in the data, but emphasize that visual breed identification is difficult due to mixed lines and differences in classifications.

Analyzing the relationship between breeds and hereditary diseases requires high-quality data on disease incidence. A study by (Oberbauer et al., 2015) found that susceptibility to hereditary diseases varies not only between purebred and mixed-breed dogs, but also within individual breed groups. However, the authors note the limitations

of their data, which was based on a sample from a single veterinary clinic, which could have led to a bias in the results.

To solve the problem of integrating heterogeneous data, preliminary cleaning and structuring are important. (Boland et al., 2019) proposed a two-step approach to building a clinical repository of dog data: first, semantic cleaning is performed to unify breed and animal species names, followed by domain-specific cleaning to identify logical inconsistencies and errors in the records. This approach improves data quality and preserves the maximum amount of data for analysis.

Modern technologies such as knowledge graphs and multimodal models open up new opportunities for data analysis and supplementation. (Hoang et al., 2023) showed how electronic medical records can be used to build a knowledge graph for diagnosing diseases in domestic animals, taking into account textual descriptions of symptoms. In turn, (Yue et al., 2025) explore methods for supplementing knowledge graphs using large language models and multimodal data, which can be applied to provide a more complete representation of information about breeds and diseases.

In addition, blockchain and IoT technologies can be used to track pedigrees and monitor animal welfare conditions. For example, (Lisnevskiy et al., 2025) proposed an order logistics platform that can be adapted to track the origin and history of dog breeds. Similarly, (Harahap et al., 2022) consider smart feeding systems that allow real-time collection of data on dog nutrition and health.

In today's world, technologies aimed at improving comfort and quality of life for both humans and animals are actively developing. For example, (Neftissov et al., 2025) proposed the concept of a "smart office," which allows monitoring the indoor environment, including parameters such as lighting, air quality, temperature, and humidity. This data can be used to assess the impact of external factors on health and well-being, which opens up prospects for the implementation of similar solutions in animal care systems, for example, in premises where dogs are kept.

In addition, the diagnosis of diseases in dogs remains an important area of research. (Rathnayaka et al., 2022) developed an intelligent system that allows the diagnosis of skin diseases in dogs based on images. As part of their work, they used ontology to extract and systematize clinical information and created a web application that simplifies the diagnostic process. To train the model, a dataset was collected that included 300 images for three types of diseases, with preprocessing and augmentation to improve the quality of the results.

Creating a unified dataset for analyzing the relationships between dog breeds and their diseases requires a comprehensive approach that includes:

- standardization of data on breeds and diseases (Mullen et al., 2025; Perez et al., 2025);
- cleaning and structuring information to improve the quality of analysis (Boland et al., 2019);
- integration of various data sources, including clinical records and genetic information (Paynter et al., 2021; Davies et al., 2025);

- application of modern analysis methods, such as knowledge graphs and multimodal models (Hoang et al., 2023; Yue ET AL., 2025).

The combination of these approaches provides a basis for comprehensive analysis of hereditary diseases in dogs and identification of patterns that may be useful for both veterinary practice and scientific research.

The final dataset was compiled in stages and included the collection, integration, and preliminary processing of information obtained from various official sources. To this end, four separate tables were used, compiled on the basis of data collected using web scraping. In the next stage, the data was combined into a single structure, taking into account differences in the formats of information presentation and the classifications used. Particular attention was paid to ensuring data comparability, which required standardization of breed and disease names. This approach made it possible to create a comprehensive and structured dataset suitable for further analysis and visualization.

The main source of information on dog breeds is the specialized resource iDog, developed by the National Genomics Data Center (NGDC, China) (NGDC, 2025). This resource contains standardized information on breeds, their characteristics, diseases, and related genetic variations. An important advantage of iDog is the availability of unified breed identifiers (Breed Ontology ID), which ensures the comparability of data obtained from various sources, including the AKC, UKC, and other international cynological organizations. The study used standard breed names, their variations, and information about diseases, including the level of heritability or increased predisposition, as well as a list of diseases that may occur in the corresponding breeds. This dataset included 482 dog breeds.

The Canine Inherited Disorders Database (CIDD) (CIDD, 2025) was used to structure information on hereditary diseases, where diseases are classified by body system. This approach allowed for more effective grouping and analysis of the data. Information about the types of diseases and the breeds in which they occur was used. The final dataset contained 147 diseases, grouped into 11 categories, which were associated with 151 breeds. The total volume of the table was 1,649 records.

Additional information about breeds was collected from The Kennel Club's official website, which is the official source of breed data supported by the UK's largest dog registration and standards organization (The Kennel Club, 2025). This source provides information on breed names, their classification into groups (e.g., terriers, greyhounds, toy dogs), and size categories. The processing resulted in 223 unique breeds.

Data from the American Kennel Club (AKC) (Filho, 2025) was used to obtain information on the distribution of breeds by official groups, such as Herding, Sporting, Working, and others. This source is an independent American breed classification standard, which differs from the standards used in the UK or other countries. Using information from the AKC made it possible to supplement and expand the data on breeds, ensuring their comparability in an international context. A total of 277 dog breeds were collected.

After preliminary preparation, all datasets were integrated into a single table.

The main difficulty lay in discrepancies in the names of breeds and diseases, as well as incomplete data: for some breeds, information on disease types or heritability levels was missing. To solve these problems, a standardization methodology was applied, including preliminary data processing using natural language processing (NLP) tools. All data processing was performed using the Python language. Before integration, temporary tables were created to unify breed and disease names. Figure 1 shows a diagram of the data merging process using intermediate tables.

To ensure consistency, all names were converted to lowercase. There were often cases where an alternative name was added to the main name in parentheses. In such cases, both variants were analyzed separately. First, completely identical names were matched, which were then no longer taken into account in the subsequent classification, which reduced the analysis load and avoided duplication. The process of matching disease and breed names from two tables is described in more detail below.

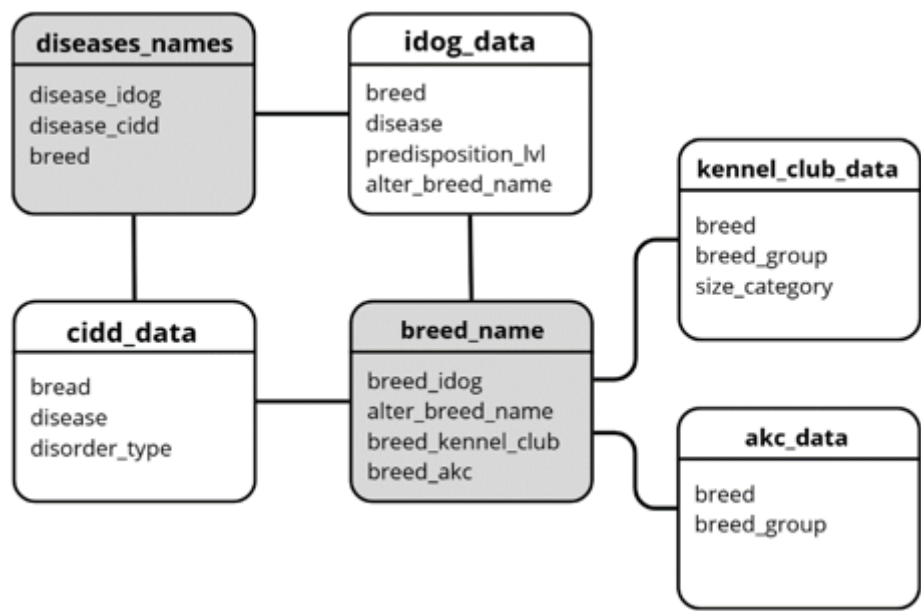


Fig. 1. Data sets and the process of merging using intermediate data.

When working with disease names from the iDog table and the CIDD table, the key task was to correctly match the corresponding entries. To do this, an approach based on tokenization and cleaning the text of punctuation marks was used. If the name contained an alternative form in parentheses, it was divided into two independent variants.

The formal condition for matching names was determined as follows:

$$(a_t \subseteq b_t) \vee (b_t \subseteq a_t) \vee (a_i \in b_j) \vee (b_i \in a_j); i, j \in \{1,2\} \tag{1}$$

where i and j are the first and second variants of the text name (if there is a second name in parentheses), a_i and b_i are the text names of diseases in the `diseases_idog` and `diseases_cidd` columns, and a_t and b_t are their tokenized and punctuation-cleaned forms.

Two records were considered to describe the same disease if the set of tokens of one record was completely contained in the set of tokens of the other, or if one of the text variants of the name was present in the other.

After successful matching of the names, a correspondence table was created, on the basis of which the data was merged. A new column containing information about the types of diseases was added to the final iDog table, if this information was present in the CIDD database. Thus, a unified and structured database was formed, providing further analysis.

Name of breeds

To process and match breed names, we used a method based on combining data from various sources with reference to the main `idog_data` table. This table contains a list of dog breeds, including data from several international organizations, and also provides alternative spellings and names for breeds. In particular, the `alter_breed_name` column in the `idog_data` table contains a list of such alternative names for each breed.

As part of the task, a set of names a_i is formed for each record in the `idog_data` table, including the main and all alternative names. If the name contains a variant in parentheses, it is split into two separate elements and added to the set. During the preprocessing stage, all punctuation marks are removed from the text. Based on the resulting set of names, a set of tokens a_t is formed, containing all unique tokens found in these names.

For data from other tables, each original breed name b_i is also split into two parts, if necessary, if it contains an alternative variant in parentheses. After that, tokenization is performed, resulting in a set of tokens b_{ti} being formed for each name variant. As in the case of the `idog_data` table, all punctuation marks are removed from the text.

Breeds are matched according to the following rule:

$$(b_{ti} \subseteq a_t) \vee (a_k \in b_i) \vee (b_i \in a_k); i \in \{1,2\}; k \in \{1,2 \dots n\} \quad (2)$$

where n is the number of variations of names of one breed from the `idog_data` table. The set a_n includes all breed names (main and alternative), a_t is the set of tokens obtained from these names, b_i is the text name of the breed from other tables, i is the name variant index (if it is divided into two with brackets), and b_{ti} is the tokenized form of the corresponding name.

Thus, breeds are considered to match if all tokens of one of the name variants b_{ti} are completely contained in the set of tokens a_t , or if the text of one of the name variants from one table is a substring of the name from another table. This allows for the correct consideration of differences in the spelling and structure of breed names, ensuring data matching even in the presence of alternative or partial names.

After creating intermediate tables with standardized names for breeds and

diseases, the data was integrated into a single structure. During the consolidation phase, records were compared between sources: if certain breeds were missing from the intermediate tables, they were added to the final set. If values were missing for some fields, they were left blank, which allowed the completeness of the information to be preserved without distorting the data.

Figure 2 shows a fragment of the final dataset, and Table 1 provides a brief description of it. The final dataset contains 4,093 records, each of which is a unique combination of the following characteristics: breed (Breed), type of disorder (Type_of_disorder), specific disorder (Disease), level of heritability or predisposition (Predisposition_to_disease), dog size (Size), The Kennel Club classification group (Group_UK), American Kennel Club classification group (Group_AKC), and alternative breed names (Alternative_breed_names).

If there are several rows in the dataset with the same breed but different diseases, this indicates that the breed may be predisposed to several types of pathologies.

The dataset includes two columns containing breed classifications according to the standards of various organizations: The Kennel Club (UK) – column Group_UK and American Kennel Club (AKC, USA) – column Group_AKC. Despite the general similarity of the classification systems, they do not completely coincide. For example, The Kennel Club distinguishes between the groups “gundog,” “toy,” “pastoral,” “utility,” “hound,” “working,” and “terrier,” while the American Kennel Club uses the categories “foundation stock service,” “toy,” “pastoral,” “miscellaneous class,” “non-sporting,” “hound,” “working,” “terrier,” and “sporting.” In some cases, the same breed may belong to different groups depending on the standard used.

Table 1. Brief description of the dataset obtained

Column Name	Count categories	Null
Breed	482	0
Type_of_disorder	11	1501
Disease	428	272
Predisposition_to_disease	4	2579
Size	6	557
Group_UK	7	557
Group_AKC	9	160
Alternative_breed_names	482	0

The result was a structured data set that allows for systematic analysis of the relationships between breeds, disease types, and classification groups according to various standards.

As part of this study, exploratory data analysis (EDA) was conducted to identify key patterns and characteristics that characterize the distribution of animal breeds, their sizes, classification groups, and types of diseases. Such analysis is necessary for a deeper understanding of the data structure and the formation of hypotheses about possible relationships between different characteristics.

	Breed	Type_of_disorder	Disease	Predisposition_to_disease	Size	Group_UK	Group_AEC	Alternative_breed_names
0	afghan hound	inherited eye disorders	cataract	have an increased incidence	large	hound	hound	afghan hound, tazi, tazi spay, kochyano spay...
1	afghan hound	inherited eye disorders	corneal dystrophy	have an increased incidence	large	hound	hound	afghan hound, tazi, tazi spay, kochyano spay...
2	afghan hound	inherited skin disorders	demodicosis	have an increased incidence	large	hound	hound	afghan hound, tazi, tazi spay, kochyano spay...
3	afghan hound	inherited eye disorders	glaucoma	may be inherited	large	hound	hound	afghan hound, tazi, tazi spay, kochyano spay...
4	afghan hound	inherited musculoskeletal disorders	hip dysplasia	have an increased incidence	large	hound	hound	afghan hound, tazi, tazi spay, kochyano spay...
...	...	...	...	...	...	...	...	...
4088	chinese imperial dog	inherited musculoskeletal disorders	patellar luxation		none small	utility	toy	chinese imperial dog, imperial shih-tzu, lion ...
4089	chinese imperial dog	inherited musculoskeletal disorders	intervertebral disk disease		none small	utility	toy	chinese imperial dog, imperial shih-tzu, lion ...
4090	chinese imperial dog	inherited nervous system disorders	intervertebral disk disease		none small	utility	toy	chinese imperial dog, imperial shih-tzu, lion ...
4091	paisley terrier	inherited musculoskeletal disorders	patellar luxation		none small	toy	toy	paisley terrier, clydesdale terrier, silky, si...
4092	soft-coated wheaten terrier	inherited endocrine disorders	hypoadrenocorticism (addison's disease)		none small	terrier	terrier	soft coated wheaten terrier, irish soft-coated...

Fig. 2. Example of the structure of the final dataset.

At the initial stage, attention was focused on studying the distribution of breeds by size categories and classification groups, which are determined on the basis of international standards. Next, the distribution of diseases by type of pathology was analyzed, which made it possible to assess their diversity. After the general patterns of breed and disease distributions were examined, the next stage of the analysis focused on identifying potential relationships between these characteristics. For this purpose, only records with a specified value in the Predisposition_to_disease column were considered.

To determine whether statistically significant associations exist between disease types, breed size, and classification groups, a Chi-square test of independence (X2) was applied. This method assesses whether deviations between observed and expected frequencies arise from random variation or indicate genuine interdependence between categorical variables. The Chi-square statistic is computed as:

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (3)$$

where r is the number of rows (categories of the first variable), c is the number of columns (categories of the second variable), O_{ij} represents the observed frequency in cell (i, j) , E_{ij} - the expected frequency in cell (i, j) , calculated as:

$$E_{ij} = \frac{(\text{rowtotal})_i * (\text{columntotal})_j}{n} \quad (4)$$

where n is the total number of observations. The null hypothesis (H_0) assumes that the two variables are independent, meaning that deviations between O_{ij} and E_{ij} are due to random variation. A statistically significant χ^2 value ($p < 0.05$) indicates that the variables are not independent, implying an association between them.

The strength of association was further evaluated using Cramer's V, which provides a standardized effect size for categorical relationships, ranging from 0 (no association) to 1 (perfect association). It is calculated as:

$$V = \sqrt{\frac{X^2}{n*(k-1)}} \quad (5)$$

where n is the total sample size, k is the smaller of (r, c), i.e. the number of rows or columns in the contingency table.

Together, the χ^2 statistic and Cramer's V enable both the detection of statistically significant relationships between categorical variables and the evaluation of their practical strength.

To ensure the accuracy of comparisons between category groups, a contingency table was constructed, representing the distribution of observations by disease type and breed characteristics. However, the initial data exhibited a marked imbalance in category frequencies, which could bias the interpretation by overemphasizing categories with larger sample sizes. To mitigate this effect, Z-normalization was applied across the rows of the contingency table, allowing each value to be interpreted relative to its mean and standard deviation within the corresponding row. This normalization procedure reduced the influence of scale differences and ensured a more reliable comparison of categorical relationships.

To eliminate this effect, Z-normalization of the values in the contingency table was performed across rows. This approach allows values within each row to be compared relative to its mean and standard deviation, eliminating the influence of differences in category scales.

After normalization, all values within each row are converted to a comparable scale, where zero corresponds to the mean level. Positive values indicate above-average values, while negative values indicate below-average values. In cases where the standard deviation is zero, which means there is no variation within the row, all values in that row were replaced with zero to avoid division by zero.

The result of normalization was a matrix whose values reflect relative deviations from the average level for each combination of characteristics. This matrix was used to construct a heat map visualizing the relationships between animal size and disease types. The visualization made it possible to identify which size categories had a higher or lower association with certain diseases, eliminating the influence of the initial uneven distribution of data. A similar approach was used to analyze the relationships between

disease types and breed classification groups, which provided a more objective and accurate identification of structural patterns in the data.

Let x_{ij} denote the initial value in the table cell, where i is the row index (disease type) and j is the column index (animal size category). For each row, the row mean (μ_i) and row standard deviation (σ_i) are calculated. The normalized value Z_{ij} is then calculated as the difference between the original value and the row mean, divided by the row standard deviation:

$$Z_{ij} = \frac{x_{ij} - \mu_i}{\sigma_i} \quad (6)$$

$$\mu_i = \frac{1}{n} \sum_{j=1}^n x_{ij} \quad (7)$$

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (x_{ij} - \mu_i)^2} \quad (8)$$

where μ_i is the mean value for the row, σ_i is the standard deviation for the row, and n is the number of animal size categories in the row.

Thus, the exploratory data analysis not only allowed us to describe the overall structure of the sample, but also provided information for further stages of the study. Data normalization was an important step in ensuring the accuracy of subsequent analysis and the reliability of the interpretation of the results obtained.

Results and discussion.

As a result of the explanatory data analysis, key patterns were identified that reflect the relationships between dog size, breed classification according to international standards, and the diversity of hereditary diseases.

Figure 3 shows the distribution of breeds by size category. Large breeds account for the largest share (39.1 %), followed by small breeds (30.1 %) and medium breeds (24.3 %). The intermediate categories “small-medium” and “medium-large” are insignificant, accounting for 2.5 % each. The smallest group is “extra large,” which accounts for only 1.4 % of the total number of breeds. This distribution reflects the predominance of medium and large dogs in the sample compared to the extreme categories.

Figures 4 and 5 show the distribution of breeds by group according to the classifications of The Kennel Club (UK) and the American Kennel Club (AKC). In the UK classification, there is a relatively even distribution of breeds between groups, with

the largest number of varieties falling into the “hound” and “gundog” groups. In turn, the AKC classification is dominated by the foundation stock service group, followed by hound, working, and sporting. The smallest number of breeds are represented in the toy and miscellaneous class groups. These differences reflect the characteristics of national classification systems, which is important to consider in comparative studies.

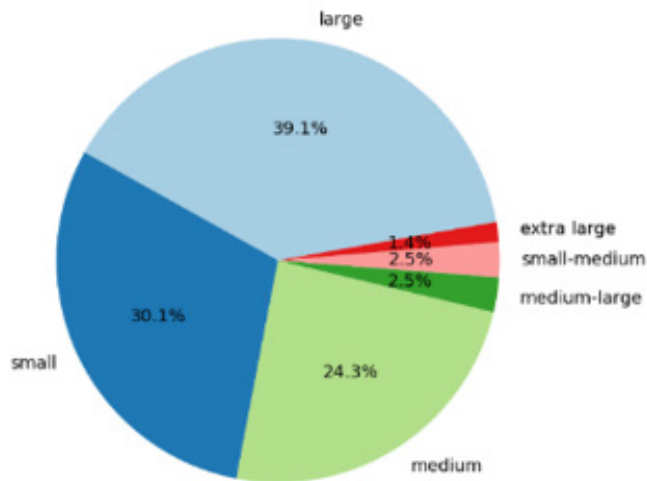


Fig. 3. Distribution of dog breeds by size categories.

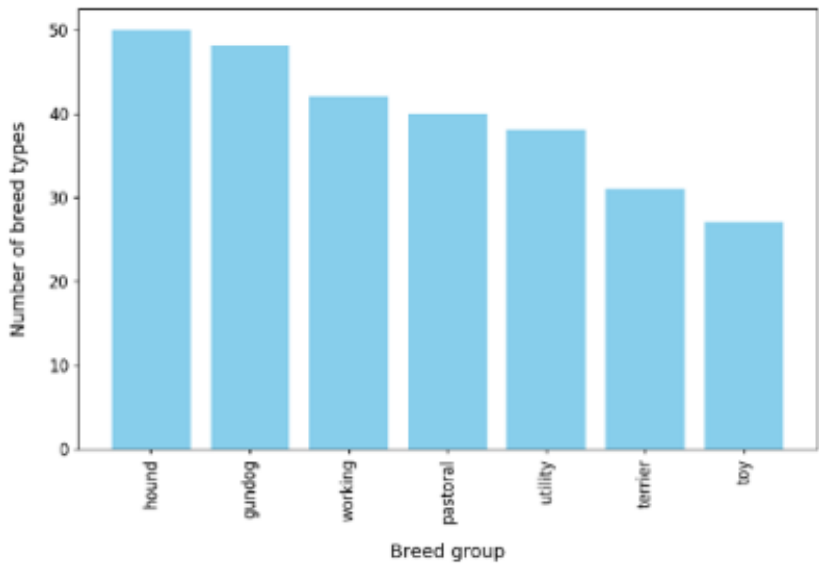


Fig. 4. Distribution of dog breeds by group (UK).

Figures 4 and 5 show the distribution of breeds by group according to the classifications of The Kennel Club (UK) and the American Kennel Club (AKC). In the UK classification, there is a relatively even distribution of breeds between groups, with the largest number of varieties falling into the “hound” and “gundog” groups. In turn, the AKC classification is dominated by the foundation stock service group, followed by

hound, working, and sporting. The smallest number of breeds are represented in the toy and miscellaneous class groups. These differences reflect the characteristics of national classification systems, which is important to consider in comparative studies.

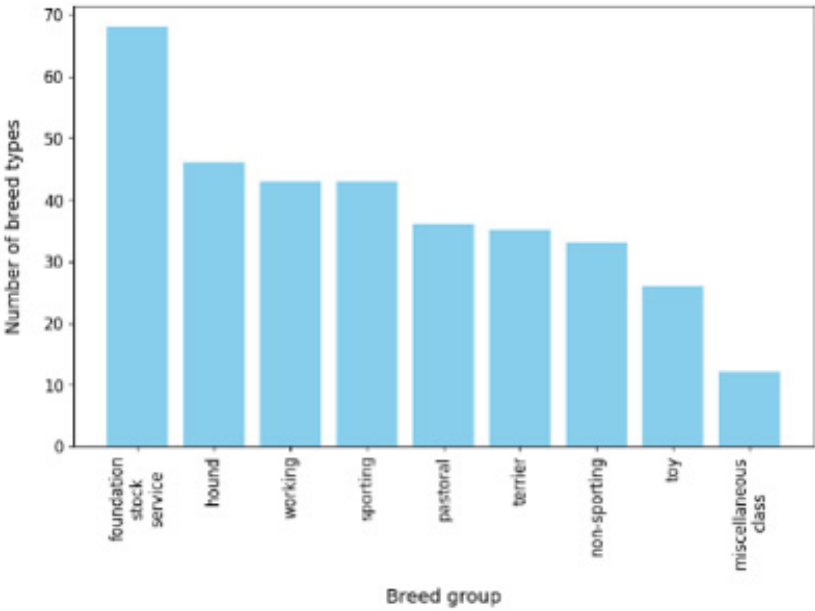


Fig. 5. Distribution of dog breeds by group (AKC).

An analysis of hereditary diseases has shown that the greatest variety of pathologies is observed among skin diseases – about 40 different types. A significant number of diseases are also associated with disorders of the nervous system and pathologies of the organs of vision. These data are presented in Figure 6 and indicate that these categories require special attention in prevention and diagnosis.

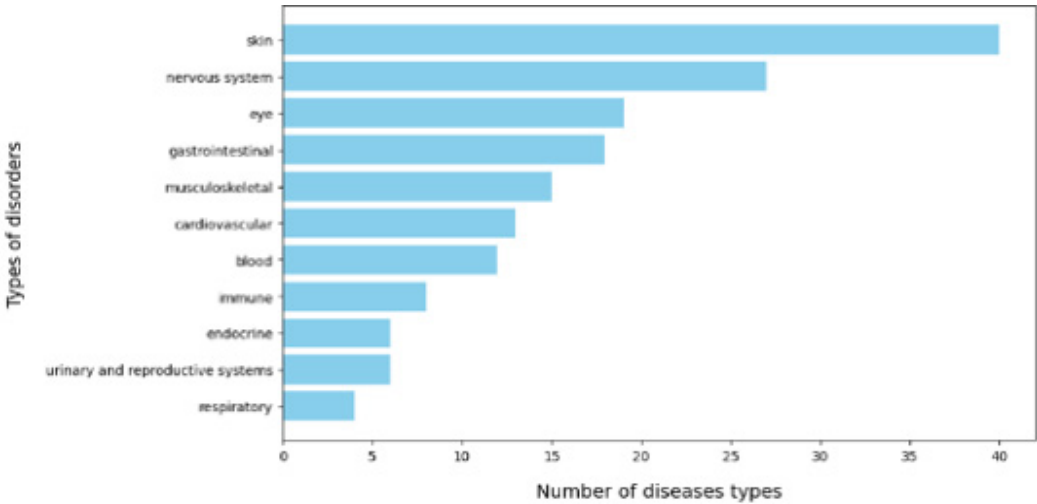


Fig. 6. Number of types of hereditary diseases by type of disorder.

The results of the Chi-square test of independence demonstrated that the distribution of disease types across breed size categories and breed classification groups is not random. The results are summarized in Table 2. In all cases, the p-values were below the 0.05 threshold, which leads to the rejection of the null hypothesis of independence between variables. This finding indicates the presence of statistically significant associations between disease types and breed characteristics. However, according to the values of Cramer’s V, the strength of these associations can be interpreted as weak, suggesting that the relationships between variables are partly influenced by other contributing factors.

Table 2. Results of the Chi-square test and Cramer’s V coefficient

Data	Chi-square statistic	Degrees of freedom	p-value	Check	Cramer’s V
Disease type × Breed size	66.801	20	6.01e-07	p-value < 0.05	0.159
Disease type × Breed group (UK)	88.779	60	9.27e-03	p-value < 0.05	0.106
Disease type × Breed group (AKC)	138.393	80	5.57e-05	p-value < 0.05	0.114

The following analysis presents a more detailed examination of the relationships between disease types and breed characteristics. Figure 7 shows the distribution of disease types according to dog size. Three main categories were identified for analysis: large, medium, and small breeds. The results indicate that large and small breeds have a wider range of diseases compared to medium-sized breeds.

Large breeds are more prone to skin and endocrine system pathologies, as well as musculoskeletal disorders. In addition, they have a higher level of diversity in diseases related to the circulatory system (blood disorders), gastrointestinal tract (gastrointestinal disorders), and cardiovascular system (cardiovascular disorders).

Small breeds, on the other hand, show an increased variety of urinary and reproductive system disorders, as well as eye, nervous system, and respiratory system disorders. For medium-sized breeds, the greatest variety is found among immune system diseases, while they show a more balanced profile in other categories.

Figures 8 and 9 show the results of the analysis of the distribution of disease types by breed groups in the UK and AKC classifications.

The UK system reveals marked differences between groups: the toy group is characterized by the greatest variety of cardiovascular diseases, terriers are more prone to eye diseases, breeds in the pastoral group are dominated by immune system pathologies, and the hound group is prone to gastrointestinal diseases. Gundog breeds are more prone to blood diseases, while in the utility group, diseases are more evenly distributed between the respiratory and endocrine systems and skin pathologies. The working group is characterized by a moderate variety of diseases, without any clearly dominant categories, but endocrine, musculoskeletal, and cardiovascular disorders are more common.

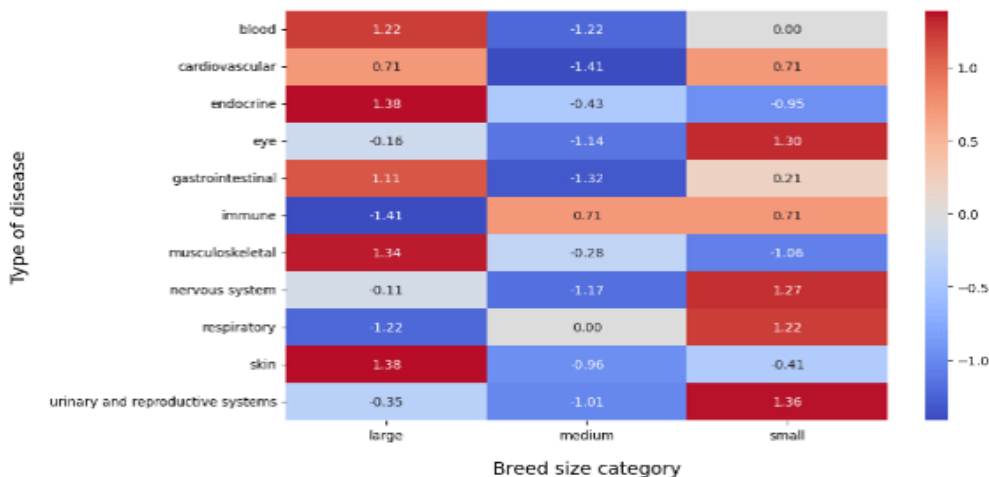


Fig. 7. Distribution of disease types by dog size categories.

In the AKC classification, the greatest variety of diseases is recorded in the “foundation stock service” group, where diseases of the nervous and endocrine systems, as well as the musculoskeletal system, predominate. Skin pathologies are more common in the “non-sporting” group, gastrointestinal diseases in the ‘hound’ group, and immune system diseases in the “pastoral” group.

Sporting breeds are characterized by a wide range of blood diseases, while terriers are more evenly represented by diseases of the gastrointestinal tract, genitourinary system, and blood. For the toy group, diseases of the respiratory and cardiovascular systems, as well as pathologies of the genitourinary system, are most typical. Cardiovascular diseases dominate in the working group, but the overall diversity of pathologies is lower compared to other groups.

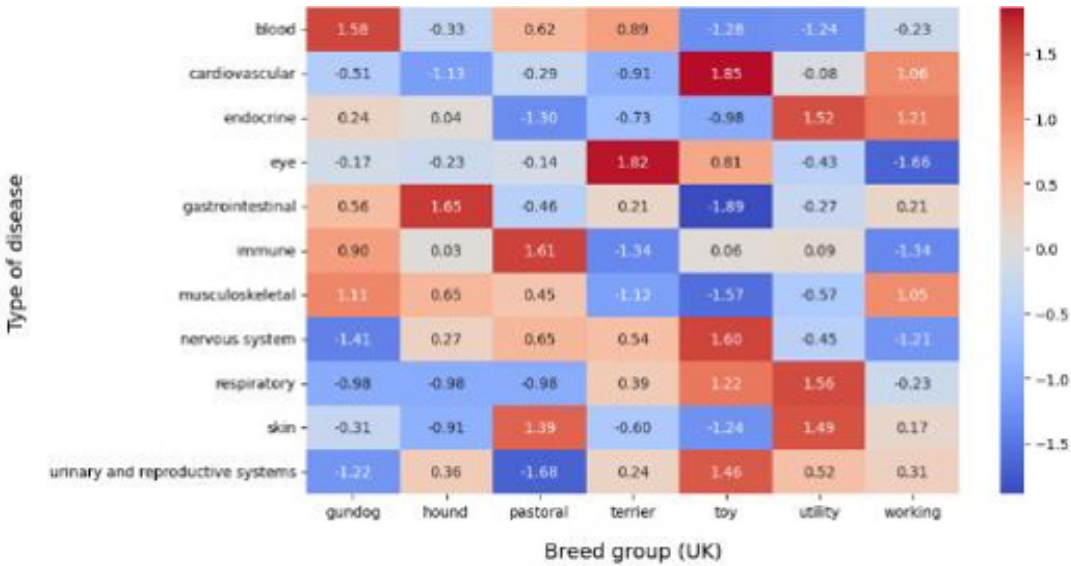


Fig. 8. Distribution of disease types by dog breed group (UK).

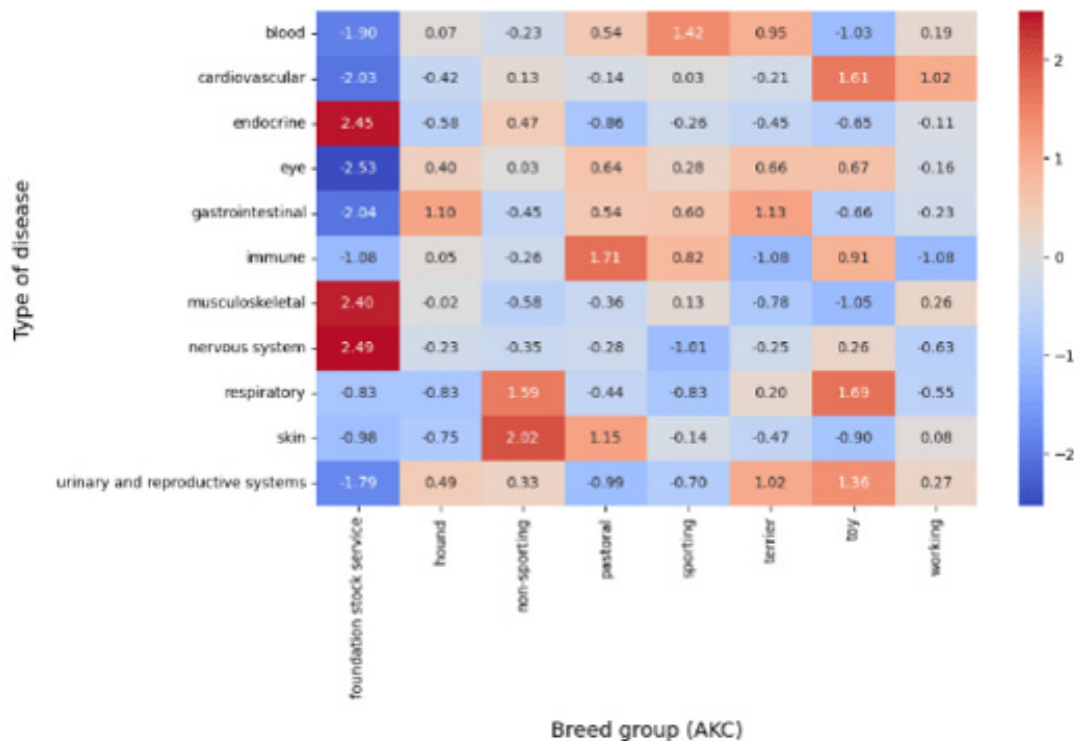


Fig. 9. Distribution of disease types by dog breed group (AKC).

Discussion.

The results of the study confirmed that the developed unified data set is an effective tool for analyzing the predisposition of dog breeds to various diseases and identifying patterns related to breed size, classification, and types of pathologies. The data obtained made it possible to systematize scattered information previously presented in various sources and create a basis for further research in the field of veterinary analytics and machine learning.

The identified relationship between dog size and the spectrum of diseases is consistent with known veterinary observations: large breeds are more prone to pathologies associated with the musculoskeletal, cardiovascular, and endocrine systems, which is due to both their physiological characteristics and the higher load on the body. Small breeds, on the contrary, show an increased incidence of diseases of the genitourinary system, eyes, and respiratory tract. Medium-sized breeds have a relatively balanced profile of pathologies, which may indicate a lower incidence of breed predispositions. These results are important for veterinarians and breeders, as they allow them to predict potential risks and develop targeted prevention strategies.

A comparison of breed classifications according to The Kennel Club (UK) and American Kennel Club (AKC) standards revealed both similarities and differences between the systems. Despite similarities in most categories, a number of breeds were classified into different groups depending on the standard used. These differences can affect the statistical interpretation of data and highlight the need to consider multiple

classification systems in analytical studies.

For example, the AKC system has a high concentration of breeds in the “foundation stock service” group, reflecting an emphasis on breeds that are in the process of being recognized or rare varieties, while the UK classification is more uniform. This difference may be due to the historical characteristics of the development of cynology in each country.

An analysis of the distribution of hereditary diseases showed that the greatest variety of pathologies is associated with skin diseases, nervous system disorders, and diseases of the organs of vision. These categories may require special attention when developing diagnostic and preventive measures, as they are most prevalent among various breed groups.

The methodological aspect of this study is particularly important. In the process of standardizing breed and disease names, we managed to overcome the problem of duplication and inconsistencies characteristic of the original sources. The use of natural language processing (NLP) methods made it possible to unify text data and ensure correct matching of records, which is an important step in preparing data for machine learning. This approach can be used in other areas of bioinformatics and veterinary analytics, where the problem of data heterogeneity also occurs.

It should be noted that the results of the study have certain limitations. First, data on rare breeds and diseases may be incomplete, which affects the completeness of the analysis. Second, the use of different international classifications creates the risk of ambiguity in interpretation when comparing groups.

Conclusion.

The study solved the problem of creating a unified and structured dataset containing information about dog breeds, their classification according to international standards, and related hereditary diseases. The developed approach to standardizing breed and disease names made it possible to integrate disparate sources of information into a single dataset containing 4,093 records and covering a wide range of characteristics, making it a valuable resource for further analytical and prognostic research in veterinary informatics.

The results of the exploratory data analysis (EDA) showed the presence of pronounced patterns reflecting the relationship between dog size, classification group membership, and the variety of hereditary diseases. It was found that large and small breeds are characterized by a wider range of diseases compared to medium-sized breeds, confirming the need for a differentiated approach to prevention and diagnosis. In addition, differences in the structure of diseases between classification groups according to The Kennel Club and American Kennel Club standards were identified, indicating the importance of taking national classification systems into account in international comparative studies.

The developed dataset represents a foundation for continued analytical studies. Its structure allows integration into decision-support systems and breed-specific disease risk assessment tools aimed at improving preventive strategies in veterinary medicine.

Future research will focus on expanding the dataset by incorporating environmental and nutritional information, which will enhance its relevance for both practical and research applications.

REFERENCES

- Boland M.R., Casal M.L., Kraus M.S., Gelzer A.R. (2019). Applied veterinary informatics: development of a semantic and domain-specific method to construct a canine data repository // *Scientific Reports*. — Vol. 9. — Pp. 55035. DOI:10.1038/s41598-019-55035-8. [in Eng.].
- Canine Inherited Disorders Database (CIDD). (2025). Disorder types, CIDD. URL: <https://cidd.discoveryspace.ca/disorder-types.html>. [in Eng.].
- Davies H., Noble P., Fins I.S., Pinchbeck G., Singleton D., Pirmohamed M., Killick D. (2025). Developing electronic health records as a source of real-world data for veterinary pharmacoepidemiology, *Frontiers in Veterinary Science*. — Vol. 12. — Article 1550468. DOI:10.3389/fvets.2025.1550468. [in Eng.].
- Donner J., Anderson H., Davison S., Hughes A.M., Bouirmane J., Lindqvist J., Lytle K.M., Ganesan B., Ottka C., Ruotanen P., Kaukonen M., Forman O.P., Fretwell N., Cole C.A., Lohi H. (2018). Frequency and distribution of 152 genetic disease variants in over 100,000 mixed breed and purebred dogs, *PLoS Genetics*. — Vol. 14. — Article e1007361. DOI:10.1371/journal.pgen.1007361. [in Eng.].
- Filho T.M. (2025). AKC data, GitHub Pages. URL: <https://tmfilho.github.io/akcdata/>. [in Eng.].
- Harahap R.K., Wibowo E.P., Nur'ainingsih D., Wijaya A.K., Widyastuti N., R.A.S.C. (2022). Anindya, Dogs feed smart system with food scales indicator IoT based, 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS). Pp. 1–7. DOI:10.1109/icoris56080.2022.10031344. [in Eng.].
- Lisnevskiy R., Biloshchytska S., Kushkimbayeva D., Kryshchyn P., Lisnevskiy V. (2025). Development of a blockchain platform for logistics order management using technologies, *CEUR Workshop Proceedings*. Pp. 3966. [in Eng.].
- Lustgarten J.L., Zehnder A., Shipman W., Gancher E., Webb T.L. (2020). Veterinary informatics: forging the future between veterinary medicine, human medicine, and One Health initiatives—a joint paper by the Association for Veterinary Informatics (AVI) and the CTSA One Health Alliance (COHA). *JAMIA Open*. — Vol. 3. — Pp. 306–317. DOI:10.1093/jamiaopen/ooaa005. [in Eng.].
- Mullen K.R., Tammen I., Matentzoglou N.A., Mather M., Balhoff J.P., Esdaile E., Leroy G., Park C.A., Rando H.M., Saklou N.T., Webb T.L., Vasilevsky N.A., Mungall C.J., Haendel M.A., Nicholas F.W., Toro S. (2025). The Vertebrate Breed Ontology: Towards effective breed data standardization. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11177956/#S8>. [in Eng.].
- National Genomics Data Center, Dog breed information, NGDC, 2025. URL: <https://ngdc.cncb.ac.cn/idog/dogph/breed/getBreedByA2Z.action?initial=A>. [in Eng.].
- Neftissov A., Sarinova A., Kazambaev I., Kirichenko L., Lisnevskiy R., Rzayeva L. (2023). Development of the smart office concept, 2021 // *IEEE International Conference on Smart Information Systems and Technologies (SIST)*. Pp. 382–386. DOI:10.1109/sist58284.2023.10223512. [in Eng.].
- The Kennel Club, Breeds A to Z, The Kennel Club (2025). URL: <https://www.thekennelclub.org.uk/search/breeds-a-to-z/>. [in Eng.].
- Oberbauer A.M., Belanger J.M., Bellumori T., Bannasch D.L., Famula T.R. (2015). Ten inherited disorders in purebred dogs by functional breed groupings, *Canine Genetics and Epidemiology*. — Vol. 2. — No. 21. DOI:10.1186/s40575-015-0021-x. [in Eng.].
- Paynter A.N., Dunbar M.D., Creevy K.E., Ruple A. (2021). Veterinary big data: when data goes to the dogs. *Animals*. — Vol. 11. — No. 1872. DOI:10.3390/ani11071872. [in Eng.].
- Perez G., He Y., Lyu Z., Chen Y., Howe N.R., Rando H.M. (2025). Standardizing canine breed data in veterinary records is challenging, but computer vision offers an alternative perspective on breed assignment // *American Journal of Veterinary Research*. Pp. 1–8. DOI:10.2460/ajvr.24.10.0315. [in Eng.].
- Rathnayaka R., Anuththara K., Wickramasinghe R., Gimhana P., Weerasinghe L., Wimalaratne G. (2022). Intelligent system for skin disease detection of dogs with ontology based clinical information extraction, 2022 *IEEE 13th Annual Ubiquitous Computing / Electronics & Mobile Communication Conference (UEMCON)*. Pp. 59–66. DOI:10.1109/uemcon54665.2022.9965696. [in Eng.].
- Summers J.F., O'Neill D.G., Church D., Collins L., Sargan D., Brodbelt D.C. (2019). Health-related welfare prioritisation of canine disorders using electronic health records in primary care practice in the UK // *BMC Veterinary Research*. — Vol. 15. — Pp. 1902. DOI:10.1186/s12917-019-1902-0. [in Eng.].
- Jackson J., Radford A.D., Belshaw Z., Wallis L.J., Kubinyi E., German A.J., Westgarth C. (2024). Using

veterinary health records at scale to investigate ageing dogs and their common issues in primary care // Journal of Small Animal Practice. DOI:10.1111/jsap.13809. [in Eng.].

Yue P., Tang H., Li W., Zhang W., Yan B. (2025). MLKGC: large language models for knowledge graph completion under multimodal augmentation, Mathematics. — Vol. 13. — Pp. 1463. DOI:10.3390/math13091463. [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 30.03.2026.

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