

УДК 004.853

**ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АНАЛИЗЕ
ГЕНОМНЫХ ДАННЫХ: МИРОВОЙ ОПЫТ И ТЕКУЩЕЕ РАЗВИТИЕ В
КАЗАХСТАНЕ**

*Баймурза А. С.,
Серикбаева Р. Т.
Студенты*

НАО «Карагандинский медицинский университет»

*Научные руководители: Сыдыкова А. Ж.,
Айткенова А. А.*

Ассистенты профессора

НАО «Карагандинский медицинский университет»

Аннотация: Статья исследует применение ИИ в геномных данных в Казахстане, перспективы персонализированной медицины и текущие возможности. Включён SWOT - анализ и рекомендации. Опрос показал интерес к использованию ИИ в медицине.

Ключевые слова: искусственный интеллект, генетика, биоинформатика, машинное обучение, генетические данные, медицинские инновации.

**THE USE OF ARTIFICIAL INTELLIGENCE IN THE ANALYSIS OF
GENOMIC DATA: GLOBAL EXPERIENCE AND CURRENT
DEVELOPMENT IN KAZAKHSTAN**

*Baymurza A. S.,
Serikbayeva R. T.*

*Scientific adviser: Sydykova Ainur Zhidebaevna,
Aitkenova Aylazzat Aitkenovna*

Abstract: The article explores the use of AI in genomic data in Kazakhstan, the prospects of personalized medicine and current opportunities. SWOT analysis and recommendations are included. The survey showed interest in using AI in medicine.

Keywords: artificial intelligence, genetics, bioinformatics, machine learning, genetic data, medical innovations.

Introduction

Artificial Intelligence (AI) is revolutionizing science and medicine, offering solutions for genomic data analysis, improving processing speed, uncovering patterns, and supporting personalized healthcare. While countries like the U.S., U.K., and China lead in integrating AI into genomics, Kazakhstan's genetic diversity and growing research infrastructure, such as the Astana Genetic Center, present significant potential. This article explores global trends, Kazakhstan's progress, and future opportunities for precision medicine.

Relevance. As genomic data grows with next-generation sequencing (NGS), AI and machine learning are crucial for efficient analysis. Kazakhstan must adapt global AI expertise to local genetic traits, advancing biomedical technologies and digital healthcare for better diagnosis, early disease prediction, and improved care.

Purpose. To develop and assess AI-based approaches for genomic data analysis in Kazakhstan, considering international experience and national characteristics.

Research objectives:

1. To conduct a review of global practices in applying AI technologies to genomic analysis.
2. To investigate the current state and problems of genomic research in Kazakhstan, identifying challenges and opportunities for the introduction of AI technologies.
3. To assess the current potential for applying machine learning algorithms to genomic data analysis in Kazakhstan, considering the country's genetic characteristics and available infrastructure.

The object of research. The processes of analysis and interpretation of genomic data using artificial intelligence technologies in the healthcare system of Kazakhstan. **The subject of the study** Artificial intelligence methods, algorithms and models for processing, analyzing and interpreting genomic data in the context of personalized medicine in Kazakhstan.

Hypotheses

The main hypothesis. The use of adapted artificial intelligence methods for genomic data analysis will significantly improve the accuracy of predictive diagnosis of socially significant diseases in Kazakhstan and the effectiveness of personalized medical approaches taking into account the genetic characteristics of the local population. A particular hypothesis. The integration of deep learning algorithms with ethnospecific genomic databases will make it possible to identify unique genetic markers specific to the population of Kazakhstan, which can be used to develop accurate models for predicting the risk of developing cardiovascular diseases and type 2 diabetes.

Research methods

1. Bibliometric analysis. Used to review global applications of AI in genomics.
2. Comparative analysis. Used to compare international experience and the current state of genomic research in Kazakhstan.
3. Statistical methods. Biostatistics methods for assessing the statistical significance of the results obtained.
4. Methods for evaluating the effectiveness of medical technologies — to assess the practical applicability of the developed approaches in the healthcare system of Kazakhstan.
5. SWOT analysis - to assess strengths and weaknesses, opportunities and threats in the implementation of AI technologies in genomic research in Kazakhstan.

The main body

Next Generation Sequencing (NGS), or deep sequencing, enables the rapid and simultaneous reading of multiple DNA fragments, identifying millions of base pairs within hours. Recent research shows that machine learning can effectively analyze large genomic datasets, helping uncover new gene functions and regulatory elements. Artificial Neural Networks (ANNs), inspired by biological neurons, are used across various fields including biology, genomics, and metabolomics. These models function as non-linear statistical tools that simulate complex relationships between genetic inputs and outputs. Among them, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have shown success in solving genomic problems [1, p. 267].

Deep learning, a subset of machine learning, has emerged as a powerful method to manage the increasing volume of data in genetic research. Technological advances, especially in computing hardware, have significantly reduced the time required to train these networks [2, p. 203]. Deep learning structures data hierarchically, where simpler patterns recognized at lower levels are used to form more complex concepts at higher levels. The learning process is automatic, driven by advanced algorithms that refine the model through continuous adjustments based on the input data [3, p. 829].

New AI-Based Genetic Technologies:

1. CRISPR & AI – Improves genome editing precision by identifying optimal edit sites and minimizing off-target effects.
2. Personalized Medicine – Uses AI to tailor treatments based on an individual's genetic profile.
3. Mutation Prediction – Forecasts potential mutations to aid in early intervention and treatment development.
4. Gene Mapping & Editing – Optimizes mapping and editing of specific genes or genome sections for therapeutic use [4, p.324].

The French Civil Code outlines three main legal bases for genetic testing: healthcare under the Human Genetic Characteristics Study Act, identification in legal

proceedings, and constitutional testing under the 2023 Anti-Doping Act. Historically, genetics' role in public health was limited due to cultural concerns and eugenics fears. While some preventive measures existed, national strategies were slow to develop. Patient groups and geneticists pushed for broader access, highlighting modern technologies' ability to enable earlier, more accurate detection. However, proposals for preconception screening and large-scale genetic testing for asymptomatic individuals were rejected due to concerns over interpretation and lack of clear benefits [5, p. 127].

Types of genetic research using artificial intelligence worldwide:

1. Genomic sequencing: AI enhances sequencing accuracy, interprets complex data, and speeds up analysis with fewer errors.
2. Genetic testing: AI processes large datasets for faster, more accurate detection of diseases or genetic predispositions.
3. Systems biology: Machine learning models predict cell behavior and identify potential molecules for new therapies [6, p, 615].

Kazakhstan is advancing AI in genomics through government initiatives and global collaborations, focusing on personalized medicine, disease prevention, and healthcare outcomes. Nazarbayev University's National Laboratory Astana leads in applying AI to identify genetic markers for cancer, diabetes, and cardiovascular diseases in the Kazakh population. The Center for Life Sciences builds a national genomic database using whole genome sequencing (WGS) to predict disease risks, while the National Center for Biotechnology applies AI to agricultural genomics. Private companies offer AI-based platforms for DNA testing and health risk assessment. Kazakhstan collaborates globally to adopt bioethics and governance standards [7, p.45]. Educational programs promote AI, genomics, and data science to train future biomedical specialists, positioning Kazakhstan as a regional leader in precision medicine.

SWOT Analysis of AI in Genomics in Kazakhstan:

- Strengths: Expanding infrastructure (Nazarbayev University, NLA), international partnerships, and a genetically diverse population [8, p.145].

- Weaknesses: Shortage of trained experts limited genomic datasets, underdeveloped data governance, and low public awareness.
- Opportunities: International funding, open-source tools, and research networks can accelerate growth. Aligning with global ethical standards could position Kazakhstan as a regional leader [9, p.348].
- Threats: Risk of falling behind global innovation, reliance on foreign technology, and issues with data privacy, regulation, and public trust.

Strategic Priorities for Development:

1. Human Capital: Expand interdisciplinary programs in AI, bioinformatics, and genomics, with specialized degrees and research exchanges.
2. Genomic Data Resources: Create ethically managed biobanks and secure platforms for anonymized data to enhance research.
3. Infrastructure Investment: Invest in high-performance computing, cloud platforms, and bioinformatics tools for genomic data processing.
4. Regulatory Frameworks: Adopt flexible, transparent regulations for data privacy and ethical AI use, aligned with global standards [10, p. 877].
5. International Collaboration: Engage in global research partnerships to boost Kazakhstan's influence in genomics.
6. Public Engagement: Raise awareness through education and campaigns to build trust in genetic research and healthcare innovation.

Kazakhstan has a strong foundation for AI-driven genomic research and can lead precision medicine in Central Asia through strategic investments.

Received data

This report summarizes survey results from 41 respondents. Gender distribution is balanced, with females at 53.7% and males at 46.3%. The largest age group is 18–20-year-olds (43.9%), followed by equal shares of 14–17 and 21–24-year-olds (17.1% each). In terms of education, 29.3% have incomplete higher education, while 24.4% each hold higher or specialized secondary education. General secondary education accounts for 22%, showing a fairly even distribution across all categories.

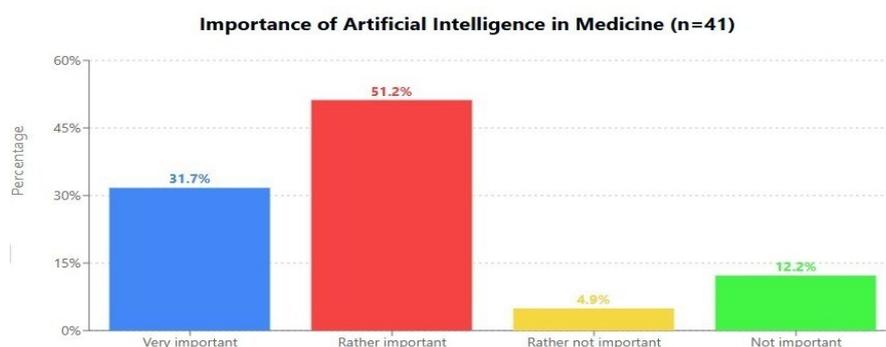


Fig. 1. Importance of AI in medicine according to respondents

The findings show that 82.9% of respondents view AI as important in medicine, with 51.2% calling it "Rather important" and 31.7% "Very important." Only 17.1% expressed skepticism. This strong support suggests broad recognition of AI's potential in improving healthcare, though a small minority may still have concerns about its implementation and ethics.

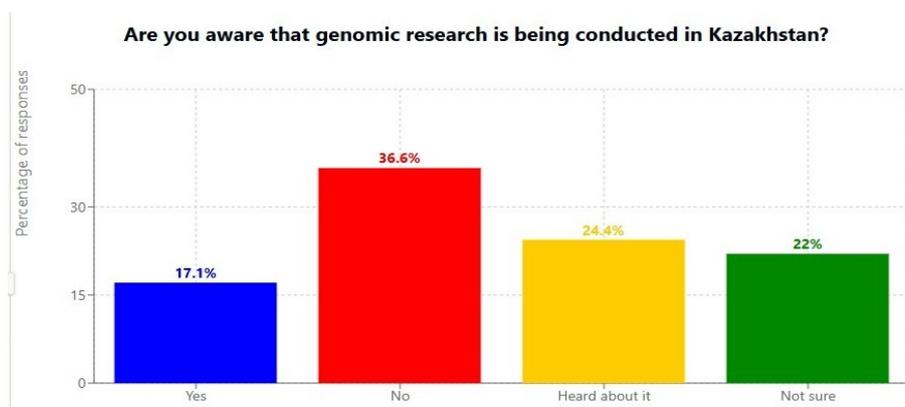


Fig. 2. Respondents' awareness of genomic research in Kazakhstan

These findings reveal a major knowledge gap, with over 80% of respondents lacking full awareness of genomic research in Kazakhstan. This highlights the need for improved communication and public outreach to boost understanding and support for national scientific efforts.

What is your opinion on the prospects of AI development for genomic data analysis in Kazakhstan?

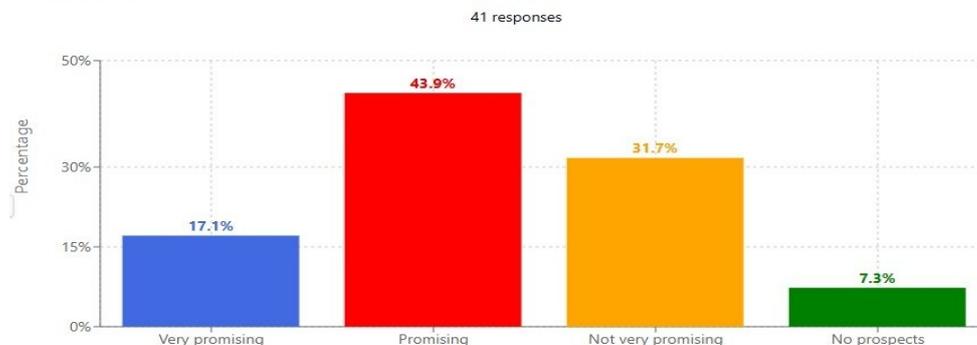


Fig. 3. Respondents' opinion of AI prospects in genomic analysis in Kazakhstan

These findings point to a notable knowledge gap, with over 80% of respondents showing only moderate confidence or skepticism about AI's prospects in genomic data analysis in Kazakhstan. While 61% see potential, 39% remain doubtful, highlighting the need for better education and public engagement to build broader support for AI-driven genomic research.

Would you want your genetic test to be analyzed using AI?

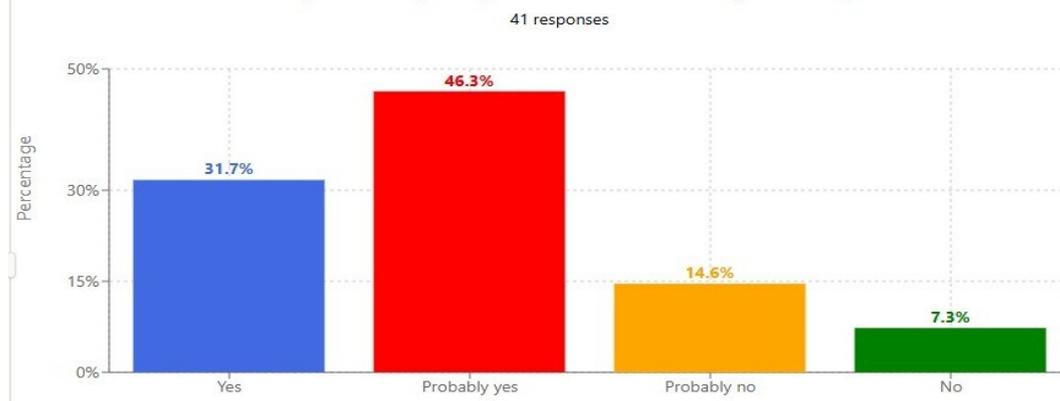


Fig. 4. Respondents' willingness for AI in genetic testing

The chart shows a generally positive attitude toward using AI to analyze genetic test results, with 78% of respondents expressing support and 22% showing hesitation or opposition. While most are open to the idea, the presence of some

skepticism underscores the need to build trust and ensure transparency in AI-driven healthcare.

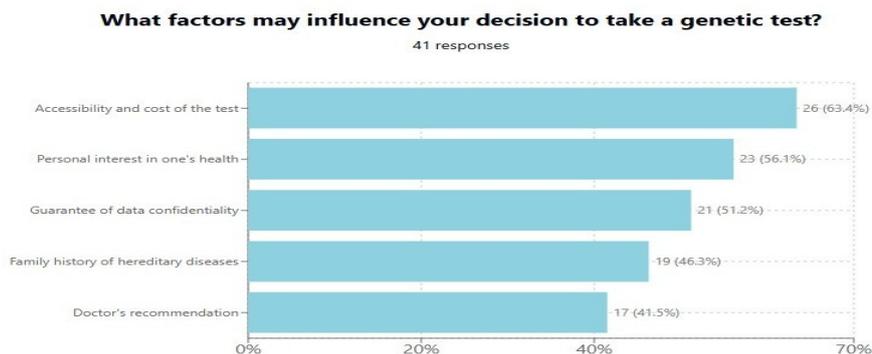


Fig. 5. Influencing factors in respondents' genetic testing decisions

The chart shows that decisions to undergo genetic testing are mainly influenced by cost and accessibility (63.4%), personal health interest (56.1%), and data privacy (51.2%). Family history (46.3%) and doctor's recommendation (41.5%) also matter. These results highlight affordability, privacy, and health awareness as key motivators for participation.

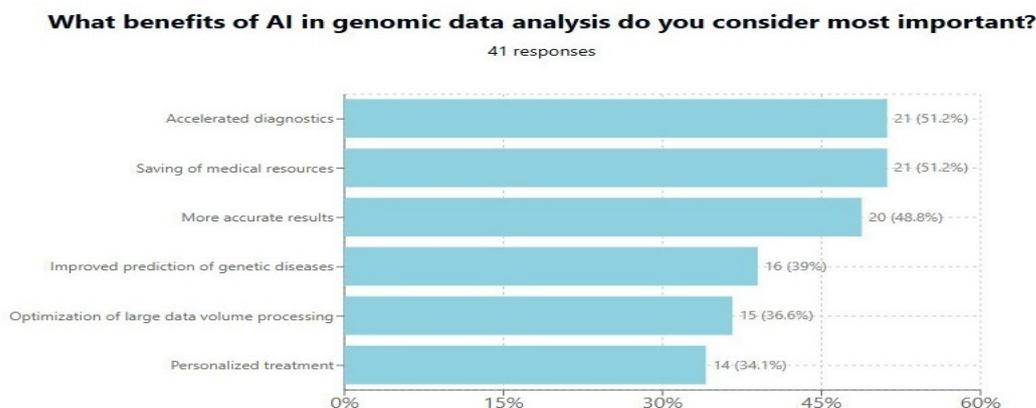


Fig. 6. Perceived benefits of AI in genomic data analysis

The chart shows that respondents value AI in genomic data analysis mainly for faster diagnostics and saving medical resources (both 51.2%), followed by more accurate results (48.8%). Other key benefits include better disease prediction (39%), handling large data (36.6%), and personalized treatment (34.1%). These findings reflect strong public support for AI's role in improving genomic healthcare.

Output

The survey conducted as part of the research on the use of AI in genomic data analysis revealed important insights regarding public awareness, perception, readiness to adopt such technologies in Kazakhstan. Most respondents demonstrated a general understanding of what genomic data is and expressed interest in the application of AI in its interpretation. A significant portion acknowledged that AI could play a substantial role in analyzing genomic data and providing accurate and personalized health insights. While a third of respondents were open to having their genetic test results interpreted using AI, half remained cautious, reflecting broader concerns around the reliability and transparency of AI-based medical tools. The main perceived advantages of using AI in genomics included faster diagnostic processes, increased accuracy, improved personalized medicine, and reduced human error.

Conclusion

AI is revolutionizing genomic data analysis by enhancing diagnostics and precision medicine. Countries like the U.S., U.K., and China have demonstrated that successful implementation of AI in genomics requires strategic planning, substantial financial investment, innovative infrastructure, and a skilled workforce. These nations have leveraged AI to identify genetic markers, predict diseases, develop personalized treatments. Kazakhstan, though new to the field, has significant potential, driven by its unique genetic diversity and growing investments in scientific infrastructure, such as the Astana Genetic Center. This diversity provides not only domestic research opportunities but also positions to contribute to global scientific advancements. However, to harness AI's full potential in genomics, Kazakhstan faces challenges like creating a national genomic data repository, improving data infrastructure, increasing public and private investments, training experts, and establishing clear regulatory frameworks for data privacy and ethics. SWOT analysis reveals Kazakhstan's strengths in its diversity and increasing interest in innovations. Weaknesses such as a shortage of specialists, limited public awareness, and underdeveloped infrastructure pose obstacles to progress.

List of references

1. Adams SP, Smith LM. Artificial intelligence in genomic medicine: current applications and future perspectives. *Nat Rev Genet.* 2023;24(5):267-283.
2. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2022;562(7726):203-209.
3. Poplin R, Chang PC, Alexander D, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnology.* 2023;41(7):823-829.
4. Genetic testing market size, share & trends analysis report by test type by application, by region, and segment forecasts, 2023-2030. *Grand View Research.* Published 2023.
5. Miller R. AI in genomic research: international developments and applications. *Nat Biotechnology.* 2023.
6. Rafehi H, Mifsud KR, Oshlack A, Craig JM. Advances in epigenetic biomarkers and their application in precision medicine. *Epigenomics.* 2022;14(8):615-629.
7. Aliyev K, Baimukhanov D, Nurgaziev M. Artificial intelligence in genomic research: Kazakhstan's emerging strategy. *J Glob Health Innov.* 2023;12(3):45-67.
8. Zhumadilov Z, Peterson C. Challenges and opportunities in developing biobanking infrastructure in Kazakhstan. *Int J Biorepos Sci.* 2021;3(2):145-159.
9. Kassymova G, et al. Development of bioinformatics in Kazakhstan: current state and future perspectives. *BMC Bioinformatics.* 2022;23(1):348.
10. Ramanculov E, et al. Kazakhstan's roadmap to genomic medicine. 2021;11(9):877.