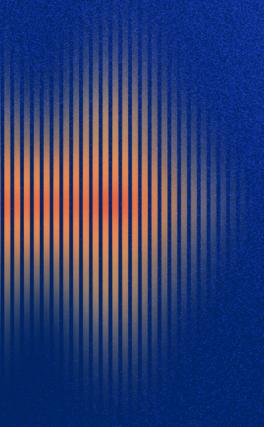


# Institute of Smart Systems and Artificial Intelligence



2024

**Annual Report** 

# **Table of Contents**



004	Introduction
<u>005</u>	Director's message
006	Appreciation
800	Team
<u>013</u>	ISSAI Kaz-LLM
<u>018</u>	Soyle App
<u>019</u>	Regional outreach
<u>022</u>	Summer Research Program 2024
<u>024</u>	Conferences
<u>032</u>	Publications
<u>037</u>	News Highlights

# **Table of Contents**



<u>055</u>	New Projects
<u>059</u>	Al Computational resources
<u>060</u>	Project Repository
<u>065</u>	YouTube
<u>066</u>	Ethical Principles
<u>067</u>	Memberships and Partners
<u>069</u>	Media
074	Contacts

#### Introduction



The Institute of Smart Systems and Artificial Intelligence (ISSAI) continues to contribute to AI development in Kazakhstan, supporting local researchers and sharing expertise and knowledge with regions. Our Annual Report 2024 highlights a year of significant projects and achievements, research activities, news and collaborations.

In this report, we delve into the work conducted by our talented team of researchers and students. From advancements in artificial intelligence to the development of cutting-edge smart systems, ISSAI has made substantial contributions to the development of AI for Kazakh language, computer vision, face recognition and AR-VR technologies.

We are proud to showcase our projects, including the ISSAI Kaz-LLM (Kazakh Large Language Model) and our first commercial Al product Soyle App.

We also delve into our commitment to education and training, fostering the next generation of Al experts, providing opportunities for internships as our Summer Research Program for school and university students.

As we look to the future, ISSAI remains dedicated to addressing challenges and driving positive change for our country. We invite you to explore the remarkable work being done at ISSAI in 2024.

# Founding Director's message



#### **Prof. Atakan Varol**

Founding Director of ISSAI

As we reflect on 2024, it is evident that the Institute of Smart Systems and Artificial Intelligence (ISSAI) stands at the forefront of groundbreaking innovation in Kazakhstan. Our dedication to advancing AI technologies tailored to Kazakhstan's unique linguistic and cultural diversity is embodied in projects such as the ISSAI KAZ-LLM and the Soyle App. These initiatives, presented to and commended by President Kassym-Jomart Tokayev, lay a robust foundation for the integration of AI into everyday life and industry. They also prove the the potential of the advanced AI workforce in the country.

This year, our notable achievement includes the creation of the ISSAI KAZ-LLM, which strengthens our capabilities in generative AI and enriches multilingual communication and technological independence in Kazakhstan through tailored AI tools. Moreover, we proudly introduced our cloud-based application, Soyle App, placing Kazakhstan among the select few that have developed a national text and speech translation service.

We also celebrate the transformation of ISSAI into a full-fledged research institute, marking a pivotal chapter in our journey. Our accomplishments were made possible by the unwavering support of Dr. Ilesanmi Adesida, who has been instrumental in ISSAI's evolution and establishment as an AI center of excellence at Nazarbayev University and in Kazakhstan. As Dr. Adesida retires, we express our deepest gratitude for his visionary leadership and steadfast commitment, which have significantly contributed to our success.

Looking ahead, 2025 promises to be a year of even greater ambition. We will embark on more challenging projects in generative AI aimed at positioning Kazakhstan as a frontrunner in AI research on the global stage. Our focus will be to drive innovation forward, nurturing local talent to continue our legacy of excellence and impact across various sectors. As the Founding Director of ISSAI, I am eager to collaborate with our talented team and esteemed partners to realize these aspirations, fostering sustainable growth and establishing Kazakhstan as a leader in AI research, ultimately enhancing the quality of life of its citizens and leaving a lasting impact.

# **Appreciation**

The success of the Institute of Smart Systems and Artificial Intelligence (ISSAI) is attributable to the invaluable encouragement and guidance of the President of Kazakhstan, Kassym Jomart Tokayev, who is an ardent supporter of science and research in Kazakhstan. In May 2024, following the directives of President Tokayev, ISSAI started its transformation into a full-fledged research Institute. This transformation has opened up wide opportunities for the development of the ISSAI.



We also thank Nazarbayev University Management and community for their continuous support of the ISSAI's initiatives and project.

We extend our sincere appreciation to our administrative, science, and media partners for their support and collaboration.

Additionally, we express our deepest gratitude to our dedicated team, researchers, collaborators, and students, whose tireless efforts, expertise, and passion for research have been crucial to our achievements and success.

# **Appreciation**



Farewell to Dr. Ilesanmi Adesida, Provost of Nazarbayev University

We extend our sincere gratitude to Dr. Ilesanmi Adesida, Provost of Nazarbayev University, for his exceptional leadership and significant contributions to the University's and our Institute's success over the past eight years.

Dr. Adesida's guidance and support have been instrumental in the establishment and development of the Institute of Smart Systems and Artificial Intelligence (ISSAI). His guidance has enabled ISSAI's transformation into a full-fledged research institute, driving innovation and research advancements within Kazakhstan and the wider Central Asian region.

Dr. Adesida's dedication to fostering a diverse and inclusive community has created an environment where the young generation of Kazakhstani researchers can thrive and reach their full potential.

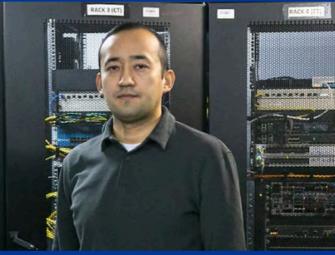
We express our appreciation for his wisdom and commitment to making a positive impact on AI development in our country.

















Dr. Huseyin Atakan Varol General Director



Yerbol Absalyamov **Executive Director** 



Madina Abdrakhmanova Deputy Director of Product and External Affairs, Senior Data Scientist



Dr. Azamat Yeshmukhametov Head of Advanced Robotics and Mechatronics laboratory



General HR Manager





General Manager







**Nurgul Shymyrbaeva** Budget Manager



**Grant Coordinator** 



**Madina Satybaldina** Administrative Manager



Aikerim Bissarinova IT Product Manager



Lead Data Scientist



Dr. Zhanat Makhataeva Senior Data Scientist

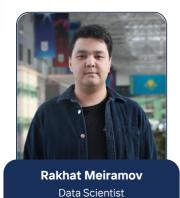


Saida Mussakhojayeva Senior Data Scientist



Data Scientist





**Ulzhan Bissarinova** Data Scientist





Mamyrbek Parakhat Data Scientist











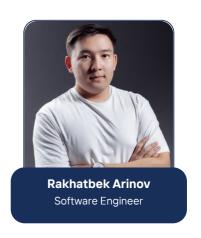
Aibota Sanatbyek Junior Data Scientist



Junior Data Scientist















Doctoral researcher







The year 2024 for the Institute of Smart Systems and Artificial Intelligence (ISSAI) was marked by the Kaz-LLM project. We have successfully developed the Kazakh Large Language Model (ISSAI KAZ-LLM) to benefit all strata of Kazakhstani society and economy by addressing local needs through customized AI technology.

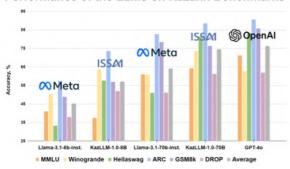
ISSAI KAZ-LLM is designed to generate content in Kazakhstan's three most relevant languages: Kazakh, Russian, and English, while also supporting Turkish. Through this effort, ISSAI KAZ-LLM shows how national AI projects can bridge linguistic gaps while contributing to the global AI landscape. The initiative also plays a significant role in preserving and promoting Kazakhstan's cultural heritage by embedding ideological perspectives, historical contexts, and specialized knowledge of the country's unique identity.

On December 10, 2024, we released the 8-billion and 70-billion parameter versions of ISSAI KAZ-LLM as open-source models for non-commercial purposes, under the CC-BY-NC (Attribution-NonCommercial) license. Both models are built on a variation of the Meta's Llama architecture and aligned with state-of-the-art standards.

The 70B ISSAI KAZ-LLM model shows superior performance compared to open-source models in Kazakh and also demonstrates strong results in Russian and English, approaching the benchmarks of OpenAI's models.

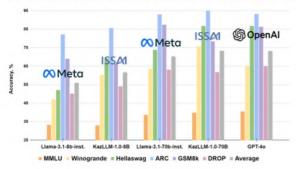
Six models are available on our public <u>Hugging Face</u> repository. We also created 4-bit quantized versions, significantly reducing the memory footprint and computational load, while still maintaining a relatively high level of accuracy. This makes these models particularly useful for deployment in resource-constrained environments, e.g., on notebook computers and workstations.

#### Performance of the LLMs on Kazakh Benchmarks



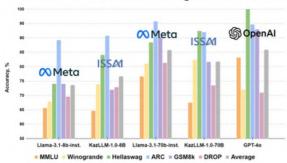
The final ISSAI KAZ-LLM training dataset comprises over 150 billion tokens across Kazakh, Russian, English, and Turkish, with 95% of the data collected and curated by ISSAI's team. Tokens were sourced from public domains, including Kazakh websites, news articles, and online libraries. Additionally, high-quality English content was translated into Kazakh, and data from various organizations were integrated. In addition, the ISSAI team mastered the art of synthetic data generation for creating supervised finetuning datasets.

#### Performance of the LLMs on Russian Benchmarks



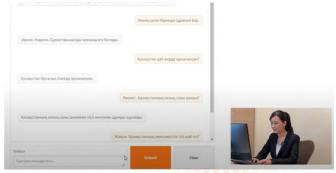
Looking ahead, we plan to extend our work into next-generation language-vision models, further advancing AI capabilities. In addition, we're exploring how to expand the model from its current capabilities in Kazakh and Turkish to other Turkic languages. By doing so, we aim to strengthen the ties between Turkic-speaking communities through technology and create opportunities for broader language inclusion.

#### Performance of the LLMs on English Benchmarks



We also aim to develop AI products and services that bring tangible benefits to the people of Kazakhstan and have a meaningful economic impact. By collaborating with partners, we seek to bridge the gap between academia and industry, driving innovation and accelerating the application of cutting-edge research to support growth and development in the local economy.





ISSAI expresses sincere gratitude to scientific partners for their invaluable contribution to the creation of the ISSAI KAZ-LLM model. Thanks for professionalism, exceptional support, and involvement in the project, the translation, and adaptation of datasets for benchmarks were carried out at the highest level, allowing the model to achieve high accuracy and relevance for the cultural and social contexts of Kazakhstan. Each partner made a unique contribution, providing a solid foundation for the implementation of this important project. The knowledge and experience of the partner organizations' specialists helped in the high-quality translation and adaptation of the tests, contributing to the expansion of the model's capabilities.

The contribution of our partners was crucial in the translation of questions and answer choices for two benchmarks: ARC (Al2 Reasoning Challenge) which assesses scientific thinking, and MMLU (Massive Multitask Language Understanding) that evaluates knowledge across 57 areas.

#### The translation of questions and answer choices for the ARC benchmark dataset involved:

- Institute of Mathematics and Mathematical Modeling (Mathematics)
- L. N. Gumilyov Eurasian National University (Physics and General Science)
- Al-Farabi Kazakh National University (Astrophysics, Astronomy, Biology)
- Karaganda State University named after Academician E. A. Buketov (Social Sciences, Logic and Reasoning)
- Abai Kazakh National Pedagogical University (Earth Science)
- Institute of Geography and Water Safety (Environmental Science)
- Almaty Technical University and Karaganda Technical University named after Abylkas Saginov (Engineering and Technology)

#### In the translation of questions and answer options for the MMLU benchmark:

- Institute of Information and Computing Technologies (Computer Science)
- Institute of Combustion Problems (High School and College Chemistry)
- Institute of Mathematics and Computing Systems (Abstract Algebra and Professional Mathematics)
- L. N. Gumilyov Eurasian National University (Physics and Conceptual Physics)
- M. A. Aitkhozhin Institute of Molecular Biology and Biochemistry (High School and College Biology)
- Al-Farabi Kazakh National University (Electrical Engineering, High School Computer Science, Computer Science, High School Statistics, Security Studies, Computer Security, Philosophy, Logical Fallacies, Moral Disputes, Moral Scenarios, Astronomy, Public Relations)
- Kazakhstan Medical University "VSHOZ" (Human Sexuality)
- Karaganda State University named after Academician E. A. Buketov (High School Government and Politics, Sociology, Formal Logic, Philosophy)
- Ch. Valikhanov Institute of History and Ethnology (History, High School U.S., European and World History, Prehistory)
- Abai Kazakh National Pedagogical University (Professional Psychology)
- Kazakh Ablai khan University of International Relations and World Languages (US Foreign Policy)
- South Kazakhstan Medical Academy (Professional Medicine)
- Karaganda Medical University (Clinical Knowledge)
- Astana Medical University (Anatomy and Nutrition)
- Semey Medical University (Virology)
- West Kazakhstan Medical University named after M. Ospanov (High School Psychology)
- Kazakhstan-Russian Medical University (College Medicine)
- Kazakh National Medical University named after S. D. Asfendiyarov (Human Aging)
- Narxoz University (High School Microeconomics, Professional Accounting, Marketing, Business Ethics, Econometrics, High School Macroeconomics, Management)
- "Bolashak" Academy (Jurisprudence)

Experts from the MIND (Maqsut Narikbayev Institute for Network and Development) created questionanswer pairs that reflect the political and social situation in Kazakhstan. This was an important step in ensuring the model's relevance to our region and current realities.

Specialists from the National Scientific and Practical Center "Til-Qazyna" named after Shaiysultan Shayahmetov supported us in cleaning the datasets for the ISSAI KAZ-LLM model.

Soyle App: ISSAI's first commercial AI product



In November 2024, the ISSAI team launched the Soyle App (soyle.nu.edu.kz), a multifunctional application for speech recognition, voice synthesis, text, and speech translation into four languages: Kazakh, Russian, English, and Turkish, serving as a universal communication tool in the country's rapidly evolving digital landscape. Soyle App became the first ISSAI commercial AI product and a result of a major milestone after five years of research by the Institute's team of Kazakhstani researchers and developers.

Powered by Kazakhstan's foundational speech model, Soyle App is hosted on NU's supercomputing cluster to guarantee the protection of user data, with no user data being used for training. Funded entirely by the NU and NIS Foundation, the app was created to promote the sustainable growth of ISSAI. Soyle App offers both free (250,000 characters) and paid versions (1 million characters for 9,000 KZT), with new features planned for future releases.

Potential impact of this product is significant. Early user reviews highlight the app's accuracy in translation, particularly for Kazakh, a language that has posed challenges for previous Al translation models. Additionally, the ease of use and clear interface can make Soyle App a popular tool for students, journalists, professional translators and travelers in Kazakhstan. The ISSAI team is actively gathering user feedback to make SoyleApp better for users, focusing on expanding language options and incorporating new functionalities like image and document translation.

Looking ahead, the Soyle App serves as a springboard for future ISSAI commercial products. It demonstrates the Institute's capabilities in developing and deploying AI solutions. ISSAI plans to develop Soyle App, ensuring the app remains at the forefront of AI translation technology in Kazakhstan.

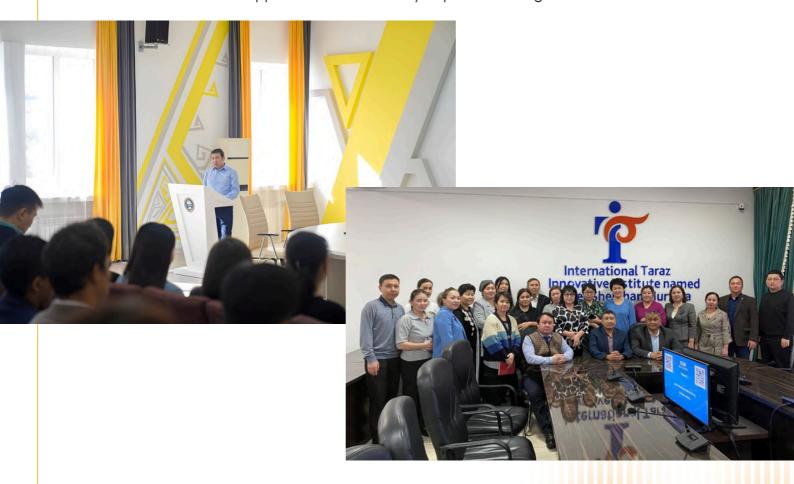


# Regional outreach:

ISSAI actively supports regional universities, ensuring that talent from all corners of the country and beyond has an opportunity to do research and gain practical experience in developing AI projects.

In the summer of 2022, under the initiative of the leadership of Nazarbayev University and the National Academy of Sciences of the Republic of Kazakhstan under the President of the Republic, ISSAI launched a project to hold seminars for all regional universities and cities of national significance in our country.

Over the past two years, ISSAI has been actively implementing this extensive initiative across the country's regions, within the framework of which it covers all the main universities in the country. This initiative spearheaded by Executive Director Yerbol Absalyamov has proven itself to be highly successful and strategically important. We were able to cover key educational institutions, providing comprehensive information about our projects, the resources used, as well as the resources and support that we are ready to provide to regional universities.



# Regional outreach:







As part of the seminars, we focused on our innovative projects, providing the audience with overviews of our methodologies and technological solutions. We also shared information on modern computing resources and opportunities for their use by regional universities.

Particular attention was paid to discussing the possibilities of providing additional assistance, consultation, summer internships, and resources to regional universities. These events have enabled the establishment of strong partnerships that facilitate the integration of scientific and educational standards at the national level. We are confident that our initiative will serve as a basis for further successful cooperation and the development of the country's research potential in artificial intelligence.

All these seminars in the main universities of Kazakhstan marked a significant milestone in ISSAI's efforts to expand its outreach to regional universities, creating a bridge between cutting-edge research and academic communities throughout Kazakhstan. Such collaborations are crucial steps toward boosting Kazakhstan's prominence in Al research and development, nurturing a robust ecosystem of innovation and expertise in artificial intelligence.



# Summer Research Program





In 2024 our Summer Research Program (SRP2024) attracted over 500 school and university students from local and international universities and schools. During the last three years, our summer internship program offers hands-on experience, mentorship from ISSAI data scientists and researchers, and access to cutting-edge technologies.

Through the Summer Research Program, we aim to inspire and support the young generation to become leaders in the field of artificial intelligence.

In 2024, 49 of 500 participants were selected and participated in the program for two months. This year we had participants from international and regional universities, including students from New Uzbekistan University, Korkyt Ata University, KAIST, University of Amsterdam, Astana IT University, Kazakhstan Branch of Moscow State University, and Tashkent International School.

















# ISSAI Junior data scientist attended the 7th International conference on AI and Cloud Computing in Tokyo, Japan





In December 2024, ISSAI Junior data scientist Vladimir Albrekht participated in the 7th International Conference on Artificial Intelligence and Cloud Computing (AICCC 2024) held at Waseda University in Tokyo, Japan. This conference brought together researchers and professionals from academia, industry, and research organizations worldwide to discuss advancements in artificial intelligence and cloud computing.

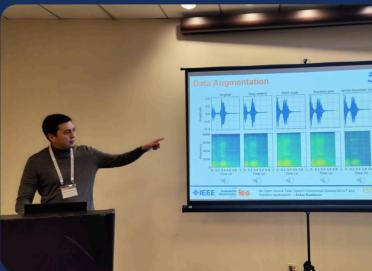
During the conference, Vladimir Albrekht listened to a keynote speech by Prof. Wenbing Zhao from Cleveland State University, USA, who addressed the advantages and disadvantages of blockchain research and development. Vladimir visited the "Software and Information System Development" session, where he explored research topics including Al for music recommendations, malware detection, and blockchain-based license management. He also attended the session on "Small-Data Lightweight Deep Learning for Al-Aided Diagnosis" by Prof. Kenji Suzuki from the Institute of Integrated Research at Tokyo Institute of Science.

Vladimir Albrekht participated in the "Next Generation Artificial Intelligence and Engineering Applications" session, which included presentations on prosocial microlending platforms, Al adoption in dairy farming, and network traffic prediction using ensemble methods. He attended the "Machine Vision and Virtual Technology Applications" session by Prof. Chien-Sing Lee from Sunway University, Malaysia, who showcased diverse applications of Al in cloud computing, such as multi-cloud management, Aldriven cloud security, and private time series forecasting using Transformers.

The conference provided a platform for Vladimir to present his research on Al-driven solutions for cloud security. It also offered opportunities to engage with experts from various fields, acquire valuable knowledge, and establish professional connections.

# Four ISSAI papers were presented at IECON 2024 in Chicago





In November, 2024 ISSAI Founding Director Dr. Atakan Varol, Lead Data Scientist Askat Kuzdeuov, Data Scientist Rakhat Meiramov, and Researcher Daniil Filimonov, attended the IEEE Industrial Electronics Society's 50th Annual Conference (IECON 2024) in Chicago.

ISSAl's four research works received high interest from the scientific community. "Field of View Invariant Object Recognition Using Non-Linear Transformation Augmentation" work was presented by researchers Muslim Alaran, Zarema Balgabekova. Additionally, another two projects were presented at the conference: "A Tensegrity Hybrid Mobile Robot with Mechanical Impact Resistance" by Daniil Filimonov, and "Enhancing Human Pose Estimation Accuracy Using Synthetic Data" by Rakhat Meiramov. Askat Kuzdeuov showcased his collaborative work with Rinat Gilmullin, Bulat Khamitov from the Tatarstan Academy of Sciences, and Dr. Varol, titled "An Open-Source Tatar Speech Commands Dataset for IoT and Robotics Applications".

The ISSAI team engaged with innovative discussions on industrial AI and attended keynote sessions from esteemed representatives of NASA, Intel, Tesla, and several national laboratories and universities.

# ISSAI Mini Symposium at the Nazarbayev University Annual Research Conference (NU ARC 2024)





In September 2024, members of the Institute of Smart Systems and Artificial Intelligence (ISSAI) participated in the NU Annual Research Conference (NU ARC) at Nazarbayev University.

They attended a plenary sessions and talks by Prof. Rashid Bashir, Grainger Distinguished Chair, Professor of Bioengineering, and Dean of The Grainger College of Engineering at the University of Illinois Urbana-Champaign. Next session by Prof. John Rogers, Founding Director of the Querrey-Simpson Institute of Bioelectronics at Northwestern University focused on Transient Electronics: From Bioelectronic Medicines to Environmental Monitors" providing an overview of his pioneering research in this field.

The Mini Symposium on AI was organized during the NU ARC 2024. The symposium included a plenary session titled "Harnessing AI for Societal Good" by Prof. Mohan Kankanhalli, Provost's Chair Professor, Deputy Executive Chairman of AI Singapore, and Director of the NUS AI Institute. During the symposium ISSAI team members presented their projects: Askat Kuzdeuov discussed "Current State of the Kaz-LLM Project: Data Preparation, Model Training, and Applications" detailing the process of data preparation and model training for the foundational Kaz-LLM model, Rustem Yeshpanov shared a presentation on "Kaz-LLM Leaderboard: A Benchmark Suite for Large Language Models in Kazakh" explaining differences between various LLMs and demonstrating benchmark examples with tricky sentence interpretations. Rakhat Meiramov talked about "Virtual Avatars", highlighting the digital human prototype Umay, and its potential applications in education and other fields. Data scientist Ulzhan Bissarinova introduced the Soyle product, demonstrating its text-to-speech and text-to-text translation features in Kazakh, English, Turkish, and Russian, and showcased the Soyle App to the audience. Madina Abdrakhmanova presented a work on "Rainfall-runoff Modeling via AI for Kazakhstani Watersheds". The ISSAI panel session attracted many attendees, including distinguished invited speakers of the NU Annual Research Conference, Prof. Andreas Cangellaris and Prof. Mohan Kankanhalli. Prof. Mohan Kankanhalli specifically shared his positive impressions of ISSAI's contribution to AI development in Kazakhstan.

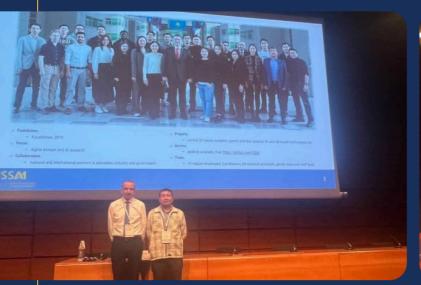
The IEEE International Conference on Automatic Face and Gesture Recognition: ISSAI paper selected as one of the best reviewed papers





In May, 2024 ISSAI team members participated in the premier international IEEE Automatic Face and Gesture Recognition (FG) conference, which was organized at the Istanbul Technical University in Istanbul, Turkey. ISSAI Senior data scientist, Askat Kuzdeuov, presented a paper entitled "OpenThermalPose: An Open-Source Annotated Thermal Human Pose Dataset and Initial YOLOv8-Pose Baselines" in both poster and oral sessions. This work, authored by Askat Kuzdeuov, Darya Taratynova, Alim Tleuliyev, and Prof. Huseyin Atakan Varol, was selected as one of the best reviewed papers by the Program Chairs.

# New ISSAI papers at the LREC-COLING conference in Torino, Italy





Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024) was organized by two major international key players in computational linguistics, the ELRA Language Resources Association (ELRA) and the International Committee on Computational Linguistics (ICCL) on 20-25 May, 2024 in Torino, Italy.

Four ISSAI papers were presented at the LREC-COLING 2024:

- 1) "KazEmoTTS: A Dataset for Kazakh Emotional Text-to-Speech Synthesis" authored by Adal Abilbekov, Saida Mussakhojayeva, Rustem Yeshpanov, and Huseyin Atakan Varol;
- 2) "KazParC: Kazakh Parallel Corpus for Machine Translation" authored by Rustem Yeshpanov, Alina Polonskaya, Huseyin Atakan Varol;
- 3) "KazQAD: Kazakh Open-Domain Question Answering Dataset" authored by Rustem Yeshpanov, Pavel Efimov, Leonid Boytsov, Ardak Shalkarbayuli, Pavel Braslavski;
- 4) "KazSAnDRA: Kazakh Sentiment Analysis Dataset of Reviews and Attitudes" authored by Rustem Yeshpanov, Huseyin Atakan Varol.

During the conference ISSAI data scientists Rustem Yeshpanov and Alina Polonskaya, along with deputy director Yerbol Absalyamov showcased the Institute's projects and accomplishments to the global research community. ELRA Chair Dr. Khalid Choukri expressed enthusiasm about ISSAI's active involvement in the natural language processing field and praised the Institute for being a consistent and high-quality contributor to LREC.

# Three ISSAI papers were presented at the ICAIIC 2024 in Osaka, Japan





The International Conference on Artificial Intelligence in Information and Communication (ICAIIC 2024) was organized at the Nakanoshima Center of Osaka University in Japan in February, 2024. Three ISSAI papers were presented at the conference:

- 1) "ChatGPT for Visually Impaired and Blind" authored by Askat Kuzdeuov, Olzhas Mukayev, Shakh-Izat Nurgaliyev, Alisher Kunbolsyn and Huseyin Atakan Varol;
- 2) "Noise-Robust Multilingual Speech Recognition and the Tatar Speech Corpus" authored by Saida Mussakhojayeva, Rinat Gilmullin, Bulat Khakimov, Mansur Galimov, Daniil Orel, Adal Adilbekov and Huseyin Atakan Varol;
- 3) "TatarTTS: An Open-Source Text-to-Speech Synthesis Dataset for the Tatar Language" authored by Daniil Orel, Askat Kuzdeuov, Rinat Gilmullin, Bulat Khakimov, Huseyin Atakan Varol.

Two of these papers were the results of collaborative work with the Institute of Applied Semiotics of Tatarstan Academy of Sciences. The conference was attended by founding director Dr. Atakan Varol, senior data scientist Askat Kuzdeuov, and undergraduate researcher Daniil Orel.

The ISSAI paper "ChatGPT for Visually Impaired and Blind" won the Excellent Paper Award at the ICAIIC 2024. The work presented an assistive mobile application with an intuitive user interface (UI) for visually impaired and blind people to interact with ChatGPT via natural conversation. The app employs a number of advanced technologies such as automatic speech recognition (ASR), text-to-speech (TTS), keyword spotting (KWS), voice activity detection (VAD), and a convenient UI to interact with ChatGPT effortlessly.

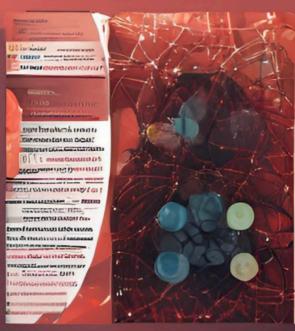
ISSAI Computer Engineer attended the International Conference for High Performance Computing, Networking, Storage, and Analysis in Atlanta, USA





From November 17 to 22, 2024, Makat Tlebaliyev, a Computer Engineer at the Institute of Smart Systems and Artificial Intelligence (ISSAI), visited the International Conference for High Performance Computing, Networking, Storage, and Analysis at the Georgia World Congress Center in Atlanta, Georgia, USA. This event is the largest global conference dedicated to high-performance computing (HPC), networking, storage, and analysis, attracting a diverse audience from academia, industry, and government sectors. Throughout the conference, Makat Tlebaliyev actively engaged in sessions, presentations, and exhibitions while attending talks by top expert speakers.

















Epile telios FIRE IA

ICATIONS





- A. Kuzdeuov, D. Taratynova, A. Tleuliyev and H. A. Varol, "OpenThermalPose: An Open-Source Annotated Thermal Human Pose Dataset and Initial YOLOv8-Pose Baselines," 2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG), Istanbul, Turkiye, 2024, pp. 1-8, doi: 10.1109/FG59268.2024.10581992
- 99 Rustem Yeshpanov, Alina Polonskaya, Huseyin Atakan Varol. KazParC: Kazakh Parallel Corpus for Machine Translation. LREC-COLING 2024, pages 9633–9644.

  https://doi.org/10.48550/arXiv.2403.19399
- Programment Progra
- Adal Abilbekov, Saida Mussakhojayeva, Rustem Yeshpanov, Huseyin Atakan Varol. KazEmoTTS: A Dataset for Kazakh Emotional Text-to-Speech Synthesis. LREC-COLING 2024, pages 9626–9632. <a href="https://doi.org/10.48550/arXiv.2404.01033">https://doi.org/10.48550/arXiv.2404.01033</a>
- Rustem Yeshpanov, Huseyin Atakan Varol. KazSAnDRA: Kazakh Sentiment Analysis Dataset of Reviews and Attitudes. LREC-COLING 2024, pages 9657–9667. <a href="https://doi.org/10.48550/arXiv.2403.19335">https://doi.org/10.48550/arXiv.2403.19335</a>
- S. Mussakhojayeva et al., "Noise-Robust Multilingual Speech Recognition and the Tatar Speech Corpus," 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Osaka, Japan, 2024, pp. 732-737, doi: 10.1109/ICAIIC60209.2024.10463419.
- D. Orel, A. Kuzdeuov, R. Gilmullin, B. Khakimov and H. A. Varol, "TatarTTS: An Open-Source Text-to-Speech Synthesis Dataset for the Tatar Language," 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Osaka, Japan, 2024, pp. 717-721, doi: 10.1109/ICAIIC60209.2024.10463261.
- A. Kuzdeuov, O. Mukayev, S. Nurgaliyev, A. Kunbolsyn and H. A. Varol, "ChatGPT for Visually Impaired and Blind," 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Osaka, Japan, 2024, pp. 722-727, doi: 10.1109/ICAIIC60209.2024.10463430.
- Akylbek Maxutov, Ayan Myrzakhmet, and Pavel Braslavski. 2024. Do LLMs Speak Kazakh? A Pilot Evaluation of Seven Models. In Proceedings of the First Workshop on Natural Language Processing for Turkic Languages (SIGTURK 2024), pages 81–91, Bangkok, Thailand and Online. Association for Computational Linguistics. <a href="https://aclanthology.org/2024.sigturk-1.8">https://aclanthology.org/2024.sigturk-1.8</a>



- G. Moger and H. A. Varol, "Design and Implementation of a Mobile Robot With Variable-Diameter Wheels," in IEEE/ASME Transactions on Mechatronics, doi: 10.1109/TMECH.2024.3457011.
- Rakhimzhanova, T.; Kuzdeuov, A.; Varol, H.A. AnyFace++: Deep Multi-Task, Multi-Domain Learning for Efficient Face Al. Sensors 2024, 24, 5993. https://doi.org/10.3390/s24185993
- Rakhimzhanova, T.; Kuzdeuov, A.; Varol, H.A. AnyFace++: Deep Multi-Task, Multi-Domain Learning for Efficient Face Al. Sensors 2024, 24, 5993. <a href="https://doi.org/10.3390/s24185993">https://doi.org/10.3390/s24185993</a>
- M. Abdrakhmanova, A. Yermekova, Y. Barko, V. Ryspayev, M. Jumadildayev and H. A. Varol, "One Model to Rule Them all: A Universal Transformer for Biometric Matching," in IEEE Access, vol. 12, pp. 96729-96739, 2024, doi: 10.1109/ACCESS.2024.3426602.
- 99 Bissarinova, U.; Rakhimzhanova, T.; Kenzhebalin, D.; Varol, H.A. Faces in Event Streams (FES): An Annotated Face Dataset for Event Cameras. Sensors 2024, 24, 1409. https://doi.org/10.3390/s24051409
- Bissarinova U, Tleuken A, Alimukhambetova S, Varol HA, Karaca F. DL-SLICER: Deep Learning for Satellite-Based Identification of Cities with Enhanced Resemblance. Buildings. 2024; 14(2):551. <a href="https://doi.org/10.3390/buildings14020551">https://doi.org/10.3390/buildings14020551</a>
- D. Zholtayev et al., "Smart Pipe Inspection Robot With In-Chassis Motor Actuation Design and Integrated Al-Powered Defect Detection System," in IEEE Access, vol. 12, pp. 119520-119534, 2024, doi: 10.1109/ACCESS.2024.3450502
- D. Zholtayev et al., "Smart Pipe Inspection Robot With In-Chassis Motor Actuation Design and Integrated Al-Powered Defect Detection System," in IEEE Access, vol. 12, pp. 119520-119534, 2024, doi: 10.1109/ACCESS.2024.3450502
- Hamad Hassan Awan, Marzhan Kabdrakhmanova, Huseyin Atakan Varol, Ferhat Karaca, A GIS-Based emergency response and management support framework for Earthquake crisis: A case study of Antakya and Kahramanmaras earthquake in Turkey, International Journal of Disaster Risk Reduction, Volume 111, 2024, 104706, ISSN 2212-4209, <a href="https://doi.org/10.1016/j.ijdrr.2024.104706">https://doi.org/10.1016/j.ijdrr.2024.104706</a>.



- D. Mukashev et al., "E-BTS: Event-Based Tactile Sensor for Haptic Teleoperation in Augmented Reality," in IEEE Transactions on Robotics, vol. 41, pp. 450-463, 2025, doi: 10.1109/TRO.2024.3502215
- R. Chibar et al., "Honeycomb-Inspired Metamaterial for Tactile Sensors with Variable Stiffness," in IEEE Sensors Journal, doi: 10.1109/JSEN.2024.3492498
- Mubarak, A., Sarsembayev, B., Serik, Y., Onabek, A., Kappassov, Z., Bakenov, Z., Tsuchiya, K. and Kalimuldina, G. (2024), Quenched PVDF/PMMA Porous Matrix for Triboelectric Energy Harvesting and Sensing. Energy Environ. Mater. e12808. https://doi.org/10.1002/eem2.12808
- Topp, S., Xiao, S., Duvernoy, B., Milroy, J., Kappassov, Z., Kabdyshev, N., ... Ziat, M. (2024). Mediated and non-mediated tactile fingerspelling: a comparative study. Assistive Technology, 1–10. <a href="https://doi.org/10.1080/10400435.2024.2369547">https://doi.org/10.1080/10400435.2024.2369547</a>
- Madina Yirgibay, M., Mussin, T., Seitzhan, S., Kenzhebek, D., Kappassov, Z., Soh, H., & Taunyazov, T. (2024). NUSense: Robust Soft Optical Tactile Sensor.ArXiv. https://arxiv.org/abs/2410.23516
- D. Mukashev et al., "Initial Exploration into Electrotactile Tongue Stimulation for Providing Force Feedback for Robot-Assisted Surgery," 2024 IEEE 20th International Conference on Body Sensor Networks (BSN), Chicago, IL, USA, 2024, pp. 1-4, doi: 10.1109/BSN63547.2024.10780705.
- 7. Mussin, Y. Kassym, N. Kabdyshev, M. Kaikanov and Z. Kappassov, "Action-Driven Tactile Object Exploration for Shape Reconstruction via Silver Nanowire Injected Sensors," 2024 IEEE/SICE International Symposium on System Integration (SII), Ha Long, Vietnam, 2024, pp. 532-538, doi: 10.1109/SII58957.2024.10417155.
- Ziat, M., Kabdsyhev, N., Topp, S., Duvernoy, B., Milroy, J., Kappassov, Z. (2025). Evaluation of HaptiComm-S for Replicating Tactile ASL Numbers: A Comparative Analysis of Direct and Mediated Modalities. In: Kajimoto, H., et al. Haptics: Understanding Touch; Technology and Systems; Applications and Interaction. EuroHaptics 2024. Lecture Notes in Computer Science, vol 14768. Springer, Cham. <a href="https://doi.org/10.1007/978-3-031-70058-3">https://doi.org/10.1007/978-3-031-70058-3</a> 36



- Keutayeva, A., Fakhrutdinov, N. & Abibullaev, B. Compact convolutional transformer for subject-independent motor imagery EEG-based BCIs. Sci Rep 14, 25775 (2024).
  <a href="https://doi.org/10.1038/s41598-024-73755-4">https://doi.org/10.1038/s41598-024-73755-4</a>
- A. Keutayeva and B. Abibullaev, "Data Constraints and Performance Optimization for Transformer-Based Models in EEG-Based Brain-Computer Interfaces: A Survey," in IEEE Access, vol. 12, pp. 62628-62647, 2024, doi: 10.1109/ACCESS.2024.3394696
- Meutayeva, A., Fakhrutdinov, N. & Abibullaev, B. Compact convolutional transformer for subject-independent motor imagery EEG-based BCIs. Sci Rep 14, 25775 (2024).
  <a href="https://doi.org/10.1038/s41598-024-73755-4">https://doi.org/10.1038/s41598-024-73755-4</a>
- Keutayeva, A., Zollanvari, A., Abibullaev, B. (2024). Evolving Trends and Future Prospects of Transformer Models in EEG-Based Motor-Imagery BCI Systems. In: Vinjamuri, R. (eds) Discovering the Frontiers of Human-Robot Interaction. Springer, Cham. <a href="https://doi.org/10.1007/978-3-031-66656-810.">https://doi.org/10.1007/978-3-031-66656-810.</a>



ISSAI team presented their projects to the Prime Minister of Kazakhstan and the Eurasian Economic Union delegation during the Digital Almaty 2024 forum





In February 2024, ISSAI team members, including Deputy Director Yerbol Absalyamov, Senior Data Scientist Askat Kuzdeuov, Data Scientist Madina Abdrakhmanova, Postdoctoral Scholar Dr. Tolegen Akhmetov, and Senior PR Manager Aliya Mukhidinova participated in the one of the largest digital forums in Eurasia, Digital Almaty 2024, which was themed "Industry X – Digital Evolution of the Future".

It convened statesmen from the Shanghai Cooperation Organisation (SCO) and Eurasian Economic Union (EAEU) countries, along with leading experts and entrepreneurs. The aim was to share experiences and ideas on ushering in a new era of digital industry.

During the forum, the ISSAI team showcased several new research projects, including Tilmash, Soyle, ISSAI Dauys App, and Kazakh text-to-speech 2, to the Prime Minister of Kazakhstan, Alikhan Smailov, and the Eurasian Economic Union delegation.

ISSAI postdoctoral researcher, Dr. Tolegen Akhmetov, demonstrated an augmented reality solution for industrial safety. This demonstration piqued the interest of industry representatives, students, researchers, and the general public.

Professor Yaacob Ibrahim of the Singapore Institute of Technology (SIT), who is also Singapore's former Minister-in-charge of Cybersecurity, met with Professor Atakan Varol, the Founding Director of ISSAI. They discussed the future prospects of AI development in Kazakhstan. Professor Varol also delivered a presentation during the Machine Learning panel session at the forum.

#### President of the National Academy of Sciences visited our Institute





In February 2024, the newly President of the National Academy of Sciences of the Republic of Kazakhstan Mr. Akhylbek Kurishbayev visited ISSAI.

ISSAI's postdoctoral researcher Dr. Azamat Yeshmukhametov presented his work on tensegrity robots. Dr Zhanat Makhataeva shared her work on a project involving Augmented Reality (AR) and Artificial Intelligence (AI), focusing on external object representation and the advancement of AR-assisted technology to aid people with memory impairments, such as Alzheimer's disease.

ISSAI Executive director Yerbol Absalyamov showcased the Institute's research initiatives to the visitors, emphasizing progress in AI for the Kazakh language, computer vision, facial recognition technology, and augmented reality innovations. Mr. Akhylbek Kurishbayev was impressed with the level of research projects at the Institute and thanked the team.

#### ISSAI at the ITU Regional Development Forum for CIS 2024





In March 2024, ISSAI Deputy director Yerbol Absalyamov attended the ITU Regional Development Forum for CIS (RDF-CIS) 2024, which was held in Astana. He participated in a round table dedicated to promoting innovation and partnerships in the introduction of new technologies such as Al. Yerbol presented Institute's NLP projects to the audience and also made several project demonstrations, including a digital prototype of Umay. ISSAI made its contribution to this forum by providing an initiative "ChatGPT for visually impaired and blind people", which was highly accessed by ITU team and audience.

To note, in 2019 ISSAI was the first institute from Kazakhstan among academic organizations, which became a member of the International Telecommunication Union (ITU). <u>ITU</u> is the United Nations specialized agency for information and communication technologies (ICTs) headquartered in Geneva, Switzerland. The Organization is made up of a membership of 193 Member States and more than 1000 companies, universities, and international and regional organizations.

## Memorandum of understanding was signed between NU and Ala-Too International University in Bishkek





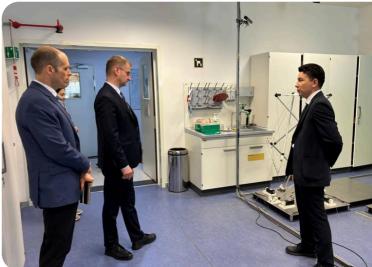
In April 2024, ISSAI Deputy director Yerbol Absalyamov and Senior data scientist Madina Abdrakhmanova as representatives of Nazarbayev University (NU) signed a Memorandum of Understanding between Ala-Too International University and NU in Bishkek, Kyrgyzstan.

ISSAI team members and Ala-Too International University's rector Dr. Zhyldyzbek Zhakshylykov, vice-rector for scientific work Dr. Salidin Kaldybayev, dean of Engineering and Informatics Dr. Ruslan Issayev, coordinator of Graduate programs in the Philology department Dr. Nurzat Kazakova and coordinator of Innovation development Center Mr. Radmir Gumerov discussed the Roadmap of creating the first Kyrgyz ASR. Kyrgyz colleagues expressed their interest in collaborating with Kazakhstani partners and working closely with ISSAI.

In 2021, ISSAI together with Computer Systems department of Tashkent University of Information Technology named after Muhammad Al-Khwarizmi created first Uzbek speech corpus and successfully created first Tatar ASR and TTS with Institute of Applied Semiotics under Tatarstan Academy of Sciences, Kazan, Russia.

#### Representatives of the US Embassy Astana visited ISSAI laboratories





In April 2024, representatives of the US Embassy in Astana, including Mr. Gregory Sadowski, economic assistant, and political officer Mr. Benjamin Overby, visited the ISSAI laboratories. During this visit, the esteemed guests were informed about the Institute's recent developments and achievements.

During the visit ISSAI postdoctoral researcher Azamat Yeshmukhametov showed the latest innovations in using artificial intelligence for robotics and highlighted their adaptability in various fields. Following this, postdoctoral researcher Zhanat Makhataeva shared her research on a project involving AR and AI, with a focus on external object representation and the advancement of AR-assisted technology to aid people with memory impairments, such as Alzheimer's disease.

ISSAI deputy director of operations Yerbol Absalyamov demonstrated the Institute's research advancements in AI for the Kazakh language, computer vision, facial recognition, and AR developments. Additionally, a separate meeting was organized with Prof. Atakan Varol, Mr. Yerbol Absalyamov, and data scientist Aknur Karabay. Representatives of the US Embassy expressed interest in the Institute's computational resources, future plans for development, project commercialization, and possible cooperation with universities and institutes in the United States.

#### ISSAI branch office opened its doors in Almaty





In April 2024, the Institute of Smart Systems and Artificial Intelligence (ISSAI) branch office was opened in Almaty. NU management with an aim to support our researchers in Almaty, and ensure a conducive environment for research activities, provided a space for ISSAI at the Nazarbayev University Almaty regional office located in the Forum business center.

Celebrating the launch of ISSAI's new branch office in Almaty, ISSAI founding director Prof. Atakan Varol gave a seminar for invited researchers and local partners. The seminar highlighted the Institute's recent initiatives and future plans, with a particular emphasis on developing KazLLM. Seminar was attended by 15 researchers and partners from KazNU Al-Farabi in particular.

#### ISSAI's transformation into a private entity





In May 2024, the Institute of Smart Systems and Artificial Intelligence (ISSAI) transformed into a full-fledged institute and celebrated this significant milestone with Nazarbayev University (NU) administration, professors, deans and vice-deans of NU schools, the Office of the Provost, NU research community, and members of various NU departments, such as law, finance, accountring, IT, PR, USM, and marketing. Representatives of the "National Information Technology" JSC and the Til-Qazyna center joined the celebration event and congratulated the ISSAI.

Acting President of NU, Dr. Ilesanmi Adesida congratulated the Institute's team, mentioned the establishment of the Institute, its achievements, and emphasized this crucial chapter in its history. He wished the team every success and pointed out that ISSAI must now become the fulcrum of AI at NU and in Kazakhstan. Dr. Aslan Sarinzhipov, Executive Vice President emphasized ISSAI's contributions to the development of AI in Kazakhstan and wished for further significant achievements.

ISSAI founding director, Dr. Atakan Varol, first thanked President Kassym-Jomart Tokayev for his continuous support in the development of AI in Kazakhstan and his instructions to transform ISSAI into a full-fledged research institute. Dr. Varol noted the significant assistance and guidance during the transformation from the NU administration, including the contributions of Dr. Adesida, Dr. Sarinzhipov, and Chief Operating Officer Maksat Mamashev, as well as various departments and university staff. Dr. Atakan Varol thanked partners from NIT JSC, the Til-Qazyna center, and partners from the Ministry of Digitalization, Astana Hub, and A. Baitursynuly Linguistic Institute, AI-Farabi National University.

#### ISSAI started its own Introduction to AI course





In June 2024, ISSAI started its own Introduction to AI course with esteemed members of Nazarbayev University management as the first cohort.

The course "Introduction to AI" comprised a total of 14 intensive modules designed to share foundational knowledge and key principles of AI. First lecture of this course was delivered by General Director of ISSAI, Prof. Atakan Varol, who presented two lectures on AI history.

The course's agenda covered essential areas such as AI infrastructure, providing a robust framework for understanding the technological component of AI systems. Participants also explored large language models, gaining insights into how these models are trained and utilized to process and generate text. In addition, the course included an introduction to the fundamentals of prompt engineering, which is crucial for AI responses and interactions. The lectures also extended to the realms of generative media, encompassing the creation of both videos and images using AI technologies. Such modules were included in the course program with an aim to share knowledge to appreciate the creative potential of AI in various multimedia formats.

The content of this course was crafted and delivered by the ISSAI team, including the Founding director Prof. Atakan Varol and data scientists Askat Kuzdeuov, Saida Mussakhojayeva, Alina Polonskaya, Madina Abdrakhmanova, Zhanat Makhataeva, and Aigerim Abdenbayeva. ISSAI members shared their expertise, ensuring that participants received a high quality learning experience.

# Members of Mazhilis of the Parliament of the Republic of Kazakhstan were introduced to the latest developments of ISSAI





The delegation of members of Mazhilis of the Parliament of the Republic of Kazakhstan visited the Institute of Smart Systems and Artificial Intelligence (ISSAI) in June 2024.

The high-profile guests, including vice speaker of the Mazhilis Dania Yespayeva, Chairman of the Committee for Social and Cultural Development Askhat Aimagambetov, Chairman of the Committee for Finance and Budget Tatyana Savelyeva, members of Mazhilis Kazbek Isa, Nurlan Auesbayev, Berik Beisengaliev, Nurgul Tau, Aina Musralimova, Guldara Nurumova, Zhuldiz Suleimenova, Rinat Zaitov, Bolat Kerimbek, Tansaule Serikov, Nartay Sarsengaliev, Daulet Mukaev, and Tilektes Adambekov.

The esteemed delegation was introduced with the Institute's latest projects and achievements. The ISSAI team presented projects for Kazakh language, computer vision, and face Al. They demonstrated the Soyle model, created by ISSAI, which can automatically recognize Kazakh speech and UMAY, ISSAI's digital human prototype. The guests were highly impressed by the Institute's progress and scientific achievements, highlighting its impact on the Kazakh language. The Mazhilis members wished the ISSAI team success and expressed a strong commitment to supporting the development of Al research Institutes such as ISSAI. Additionally, they highly appreciated ISSAI's initiative to promote the Kazakh language through the use of Al.

#### Ambassador of The Gambia to the UN Office at Geneva visited ISSAI





In September 2024, Ambassador and Permanent Representative of The Gambia to the United Nations Office in Geneva, Prof. Muhammadou M.O. Kah visited the ISSAI.

ISSAI General Director Prof. Atakan Varol welcomed Prof. Kah and introduced ISSAI Executive Director Yerbol Absalyamov, and Deputy Director of Product and External Affairs, Senior Data Scientist Madina Abdrakhmanova.

During the visit, Madina Abdrakhmanova presented the Institute's key projects and developments. She demonstrated ISSAI projects for the Kazakh language, computer vision, facial recognition and a collaborative project on nutrition AI with Nazarbayev University School of Medicine. Madina also shared that the ISSAI team works on a foundational Kazakh Large Language Model (Kaz-LLM), which plays a significant role in AI development in Kazakhstan.

Prof. Kah highly appreciated the level of ISSAI research projects. As a United Nations Office representative, Prof. Kah expressed his readiness for support and assistance to the Institute. ISSAI representatives thanked Prof.Kah for the visit, and both sides agreed to continue further communication.

### Distinguished guests of the NU Annual Research Conference visited ISSAI





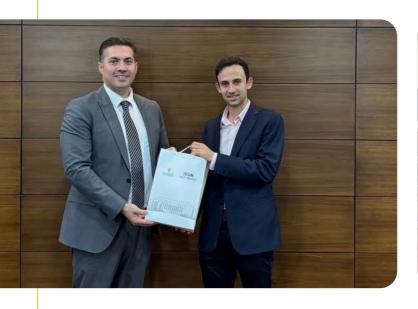
In September 2024, the ISSAI had the honor of hosting distinguished guests and invited speakers of the <u>NU Annual Research Conference</u>, dedicated to celebrating the 125th anniversary of Kanysh Satpayev, the visionary founder of the Kazakh National Academy of Sciences, organized at Nazarbayev University.

The delegation of esteemed guests accompanied by Professor Ilesanmi Adesida, Provost of Nazarbayev University, included Prof. Andreas Cangellaris, President of NEOM University, Prof. Rashid Bashir, Professor of Bioengineering and Dean of Grainger College of Engineering at the University of Illinois Urbana-Champaign, Prof. Mohan Kankanhalli, Provost's Chair Professor of Computer Science and Director of the NUS AI Institute from the National University of Singapore, Dr. Veena Sahajwalla, Australian Research Council Laureate Professor from the University of New South Wales, and Prof. Sarkyt Kudaibergenov from the Institute of Polymer Materials and Technology.

The Founding Director of ISSAI, Prof. Atakan Varol warmly welcomed distinguished guests, introducing them to the Institute's leadership team, data scientists, and researchers. As part of the visit, Prof. Varol conducted a guided tour through ISSAI's facilities, highlighting the cutting-edge research being carried out within the Institute. He presented an overview of ISSAI's key projects, highlighting the Institute's impact on AI development in Kazakhstan, and commitment to solving real social and economic problems in the country through AI and smart technologies.

During the visit, Junior data scientist Aibota Sanatbek presented her research work on a food recognition project. The guests were then introduced to ISSAI's digital human prototype, called Aibike. Data scientists Dr. Zhanat Makhataeva and Rakhat Meiramov demonstrated how this system works. The demonstration highlighted the potential for this technology to drive areas such as education, healthcare, and customer service. Following this, ISSAI Deputy Director of Product and External Affairs Madina Abdrakhmanova provided insights into the Institute's collaborations with governmental organizations and the private sector. She discussed how these partnerships are important to ISSAI's mission of bridging the gap between academia and industry, promoting the application of AI technologies across various sectors. Lead data scientist Askat Kuzdeuov and data scientists Rustem Yeshpanov and Alina Polonskaya shared details of their research projects, emphasizing ISSAI's contribution to AI development for Kazakh language.

#### Stanford University Lecturer visited the ISSAI





In October 2024, Stanford University Lecturer of Computer Science, Mr. Younes Bensouda Mourri visited the Institute of Smart Systems and Artificial Intelligence (ISSAI). The visit began with a welcome speech from the Founding Director of ISSAI, Prof. Atakan Varol, who introduced members of the institute including Executive Director Yerbol Absalyamov and Deputy Director of Product and External Affairs, Senior Data Scientist Madina Abdrakhmanova.

Madina Abdrakhmanova started with an overview presentation about the Institute. She showcased ISSAI's key projects and recent developments in AI, highlighting achievements in NLP, Computer vision, and Augmented Reality, and ISSAI's crucial role in AI development in Kazakhstan.

Mr. Bensouda Mourri was highly interested in ISSAI's project and expressed his admiration for the Institute's progress in the AI field. As a Founder of the LiveTech.AI startup, Mr. Mourri presented details and main goals of this platform, which helps to build AI tools for educators. He provided suggestions for potential partnership opportunities that could be explored further. The meeting concluded with an engaging and productive discussion, laying the foundation for potential collaboration.

# ISSAI members delivered lectures at the Ahmet Yassawi University in Turkistan as part of the Deans' School





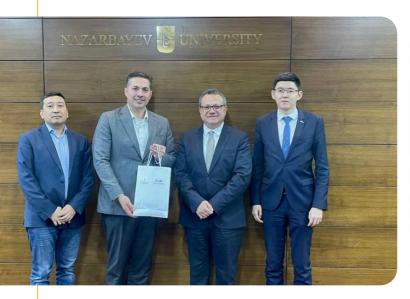
In October 2024, Senior Data Scientist of the Institute of Smart Systems and Artificial Intelligence (ISSAI) Dr. Zhanat Makhataeva and postdoctoral researcher Dr. Azamat Yeshmukhametov participated in the third module of "Digitalisation of Education" during a professional development program for deans of Kazakhstani universities, titled the "Institute of Academic Leadership Deans' School" at Akhmet Yassawi International Kazakh-Turkish University (AYU) in Turkestan.

Dr. Zhanat Makhataeva presented a lecture with two parts: Al in Learning and Teaching and Avatars and Mixed Reality in Education. During the presentation, Dr. Zhanat shared international experiences of implementing augmented reality, Al, and virtual avatars in human-computer interaction, as well as advancements in augmented reality technology. She also emphasized ISSAl's projects and developments in Al and computer vision and the Institute's role in the development of Al for Kazakh.

Dr. Azamat Yeshmukhametov showcased his research project on tensegrity robots, explaining their structure, key features, benefits, and potential applications in industry and manufacturing. He shared details about a collaborative project on innovative sensor technologies, developed with Prof. Gulnur Kalimuldina's Nanomaterials Laboratory at Nazarbayev University. Prof. Kalimuldina developed a TENG sensor, which was used to create smart insoles. Dr. Azamat Yeshmukhametov and PhD student Aisulu Tileukulova developed a new Stewart platform architecture. These two products will be combined to create a solution aimed at improving the diagnosis and rehabilitation of children with autism and other motor disorders.

The feedback from the audience was positive, as the deans were particularly impressed with ISSAI's projects and expressed interest in exploring potential collaborations.

## Visit of IOM Delegation to ISSAI: Exploring prospects for partnership





In November 2024, a delegation from the International Organization for Migration (IOM), a United Nations-related organization specializing in migration, visited the ISSAI. The delegation was composed of Regional Coordinator Mr. Serkan Oztoprak and Senior Project Assistant Mr. Alpamys Asabayev.

Prof. Atakan Varol, General Director of ISSAI, warmly welcomed Mr. Serkan Oztoprak. Following the welcome, Yerbol Absalyamov, Executive Director of ISSAI, presented the Institute's key projects, including Soyle App, Aibike, and Kaz-LLM".

The presentation piqued significant interest and was met with positive feedback from the IOM delegation, who recognized the value of ISSAI's contributions and conveyed their best wishes for the Institute's future endeavors. Following the presentation, participants explored potential areas for collaboration, such as developing efficient systems for public services, with a focus on security, safety, and data management. The IOM delegation highlighted their Global Data Institute division, which conducts research and data collection on migration and human mobility worldwide, as a potential area for future collaboration.

### ISSAI's First GenAl Hackathon: Fostering innovation and advancing the Kazakh language through Al





In November 2024, ISSAI co-organized its first GenAI hackathon with the NU Association for Computing Machinery, and the SIT Foundation, offering early access to its generative AI models, Soyle App, and ISSAI KAZ-LLM, via APIs. Seventy students from Nazarbayev University participated, using these locally developed AI tools to create innovative projects. The hackathon aimed to empower students and the broader IT community to explore AI's potential while promoting the use of the Kazakh language.

Participants integrated the Soyle App and KAZ-LLM into their projects, with teams evaluated not only on innovation and real-world applicability but also on their feedback to ISSAI's APIs and models. Projects ranged from a mobile avatar app for teaching Kazakh history, language learning apps, and an AI assistant for legal matters, to edge computing for tourism, a chatbot bridging language barriers, and a note-taking app. Fifteen teams presented their solutions, with 13 integrating both models.

The event allowed students to join ISSAI research teams or launch their startups. This hackathon is part of ISSAI's mission to foster innovation, promote the Kazakh language, and apply AI to real-world industry challenges, with more events planned for the future.

#### ISSAI team recognized at the NU Town Hall Event





Nazarbayev University organized the final Town Hall meeting of the year with university staff in December, 2024. The event was marked by a series of significant announcements and recognitions. Founding Director of the Institute of Smart Systems and Artificial Intelligence (ISSAI), Professor Atakan Varol was honored with a special state award for his "Contribution to the Development of Science". This prestigious recognition by Kazakhstan's Minister of Science and Higher Education, Sayasat Nurbek.

In addition to Prof. Atakan Varol's recognition, ISSAI Executive Director Yerbol Absalyamov, Deputy Director for Product and External Affairs, Senior data scientist Madina Abdrakhmanova, Lead data scientist Askat Kuzdeuov, Senior data scientist Saida Mussakhojayeva, and Data scientist Rustem Yeshpanov were also acknowledged with letters of appreciation from Prof. Waqar Ahmad.

Prof. Waqar Ahmad emphasized the crucial role of ISSAI for Nazarbayev University and Institute's impact for Al development in Kazakhstan.





## **New Projects**

#### **Soyle App**

Soyle App is a multifunctional application for speech recognition, voice synthesis, text, and speech translation into four languages. Powered by Kazakhstan's foundational speech model, Soyle App is a groundbreaking Al product developed to enhance multilingual communication for Kazakhstani audiences and beyond.

#### **Key Functions**

- Speech recognition, speech synthesis, and translation: Core features of Soyle App include accurate speech transcription, natural-sounding voice synthesis, and instant translation of text and speech between supported languages.
- Intuitive user interface: The application is designed with the needs of the Kazakhstani audience in mind, offering a user-friendly and intuitive interface for communication in the native or preferred language.
- Access modes: Soyle App offers both interactive in-app features and an API, enabling seamless integration of its capabilities for individual users and businesses.

Soyle App's translation quality, measured by BLEU scores on the FLoRes benchmark dataset, outperforms comparable tools like Google Translate, Yandex Translate, and GPT-4o, reflecting its superior accuracy for Kazakh, Russian, English, and Turkish translations.



### **New Projects**

# **Aibike - ISSAl's New Digital Human Prototype with KAZ-LLM Inside**

To overcome UMAY's limitations, such as reliance on external servers and the absence of cultural context, we have developed Aibike, our new digital human prototype as part of the KAZ-LLM project. Aibike is powered by our foundational Kazakh large language model (ISSAI Kaz-LLM), enabling direct understanding of the Kazakh language without the need for translation.



### **New Projects**

#### ISSAI Dauys: the first Kazakh text-tospeech application

In February, 2024 ISSAI Dauys App, the first Kazakh text-to-speech application, was launched during the Digital Almaty 2024 forum in Almaty.

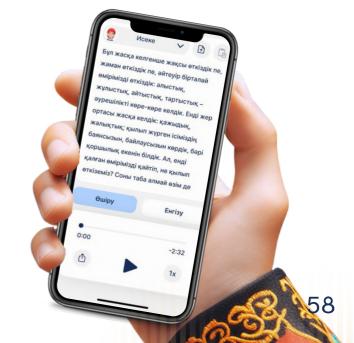
The developers Olzhas Mukayev, Ruslan Tasmukhanov, Nurali Rakhay, Yelaman Fazyl, Alisher Kunbolsyn are students of Nazarbayev University, and participants of the Summer Research Program 2023, organized by ISSAI. Students were supervised under ISSAI Senior Data scientist Askat Kuzdeuov.

ISSAI Dauys App is developed based on the large-scale dataset <u>KazakhTTS2</u>. The dataset features 271 hours of high-quality transcribed audio from five professional speakers.

ISSAI Dauys App allows users to instantly convert any Kazakh text to high quality speech with just one click! To convert text into audio, users need to input the text into the interface, choose a preferred voice, and get instant results. The first of its kind for the Kazakh language, ISSAI Dauys App offers offline access, no advertisements, free usage, and five high-quality voices.

The application is available for <u>download</u> for Android users on Google

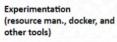
Play. The iOS version will launch soon.



# AI Computational resources

Our AI Computing cluster consists of many interconnected computing elements (nodes). The nodes in each cluster operate in parallel with each other, reaching higher processing power to train deep learning models. The current AI computing resources at ISSAI consist of 7 computing nodes.







DGX 1
GPUs: 8 x NVIDIA® Tesia® V100
GPU Memory 256 GB
CPU Dual 20-Core Intel Xeon E5-2698 v4 2.2
GHz
System Memory 512 GB, 2,133 MHz DDR4
RDIMM
Storage 4X 1.92 TB SSD RAID 0

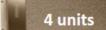
Delivered March 2018 1x1PF = 1 PF E.g.: NU employees, students, ISSAI RAs, and interns



DGX 2 (01, 02)
GPUs: 16 x NVIDIA<sup>9</sup> Tesla<sup>9</sup> V100
GPU: Memory 512 GB total
CPU: Dual Intel Xeon Platinum 8168, 2.7
GHz, 24-core
GHz, 24-core
System Memory: 1:5TB DDR4 RDIMM
Storage: 2 x 950GB NVME SSDs
Internal Storage: 30TB (8X 3.84TB) NVME

ximum Power Usage: 10kW

Delivered February 2020 2x2PF = 4 PF E.g.: ISSAI team (collaborating faculty, data scientists, postdocs, RA, computer engineer, etc.)



DGX A100 (01,02,03,04)
GPUs: 8 x NVIDIA A100 40 G8 GPUs
GPU: Memory 320 G8 total
CPU: Dual AMD Rome 7742, 128 cores
total, 2.25 GHz (base), 3.4 GHz (max
boost)
System Memory: 1T8 DDR4 RDIMM

boost) System Memory: 1TB DDR4 RDIMM Storage: 2 x 1.92TB M.2 NVME drives Internal Storage: 15 TB (4x 3.84 TB) U.2 NVMe drives Maximum Power Usage: 6.5 kW

Delivered December 2021 4x5PF = 20 PF AI

We are committed to share our know-how and resources with the Kazakhstani research community. Here, we provide information on how to interact with us to use these resources.



ISSAI shares the pre-trained AI models of its projects on the ISSAI GitHub page.

The models are available for free download to anyone:
<a href="https://github.com/IS2Al">https://github.com/IS2Al</a>



**59** 

Repositories



#### TOP REPOSITORIES:

- Kazakh TTS
- SpeakingFaces
- TurkicASR
- TurkicTTS
- ISSAI SAIDA Kazakh ASR
- thermal-facial-landmarks-detection

#### TOP LANGUAGES:

- Python
- Jupyter Notebook
- Shell
- C#
- Svelte

All research datasets and models can be freely downloaded from the official website: <a href="https://issai.nu.edu.kz/issai-datasets/">https://issai.nu.edu.kz/issai-datasets/</a>



#### **KAZAKH TEXT-TO-SPEECH 2**

Expanded version of the previously released Kazakh text-to-speech (KazakhTTS) synthesis corpus. In the new KazakhTTS2 corpus, the overall size has increased from 93 hours to 271 hours, the number of speakers has increased from two to five (three females and two males), and the topic coverage has been diversified with the help of new sources, including a book and Wikipedia articles.



#### **KAZAKH SPEECH CORPUS 2**

The KSC2 is the first industrial-scale open-source Kazakh speech corpus, which subsumes the previously introduced two corpora: Kazakh speech corpus and Kazakh Text-To-Speech 2, and supplements additional data from other sources like tv programs, radio, senate, and podcasts. In total, KSC2 contains around 1.2k hours of high-quality transcribed data comprising over 600k utterances.



#### **UZBEK SPEECH CORPUS (USC)**

The USC is an open-source speech corpus that has been developed in collaboration between ISSAI and the Image and Speech Processing Laboratory in the Department of Computer Systems of the Tashkent University of Information Technologies (https://tuit.uz/en/kompyuter-tizimlari). The USC comprises 958 different speakers with a total of 105 hours of transcribed audio recordings.



#### **SPEAKING FACES**

A large-scale publicly-available dataset designed to encourage research in the general areas of user authentication, facial recognition, speech recognition and human-computer interaction.



#### **TURKISH SPEECH CORPUS (USC)**

The corpus contains 218.2 hours of transcribed speech with 186,171 utterances and is the largest publicly available Turkish dataset of its kind. The datasets and codes used to train the models are available for download at TurkicASR.



#### TATAR SPEECH CORPUS ASR

A large-scale publicly-available dataset designed to encourage research in the general areas of user authentication, facial recognition, speech recognition and human-computer interaction.



#### **IMUWIFINE**

A finer-level sequential dataset of IMU and WiFi received signal strengths (RSS). The dataset contains 120 trajectories covering an aggregate distance of over 14 kilometers. The dataset was collected across 3 floors of the C4 building of Nazarbayev University.





#### SF-TL54: A THERMAL FACIAL LANDMARK DATASET WITH VISUAL PAIRS

A thermal face dataset with manually annotated bounding boxes and facial landmarks. The dataset was constructed using our large-scale SpeakingFaces dataset (https://issai.nu.edu.kz/speaking-faces/). In total, the dataset contains 2,556 thermal-visual image pairs of 142 subjects, where each subject has 18 thermal-visual image pairs (2 trial x 9 positions).



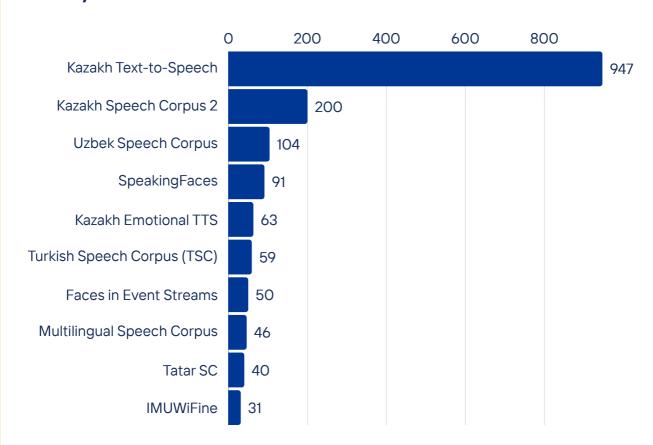
#### TFW: ANNOTATED THERMAL FACES IN THE WILD DATASET

The dataset contains thermal images acquired in indoor (controlled) and outdoor (uncontrolled) environments. The indoor dataset was constructed using our previously published SpeakingFaces dataset. The outdoor dataset was collected using the same FLIR T540 thermal camera with a resolution of 464x348 pixels, a wave-band of 7.5–14  $\mu$ m, the field of view 24, and an iron color palette. The dataset was manually annotated with face bounding boxes and five-point facial landmarks.



More than 100 local and international companies and educational organizations have downloaded ISSAI datasets. The list includes Nvidia, Google, Amazon, Halyk Bank, Otbasy Bank, Veon AdTech, Stanford University, Johns Hopkins University, Carnegie Mellon University, Columbia University, MIT Lincoln Laboratory, Darmstadt University of Applied Sciences, Staffordshire University, Strasbourg University, Tsinghua University, National University of Singapore, Samarkand State University, Mongolian National University, Alatoo International University, Kazakh British Technical University, SDU, nFactorial Incubator etc.

#### The number of downloads for each dataset January-December 2024:



# Youtube Analytics



1,133
Subscribers

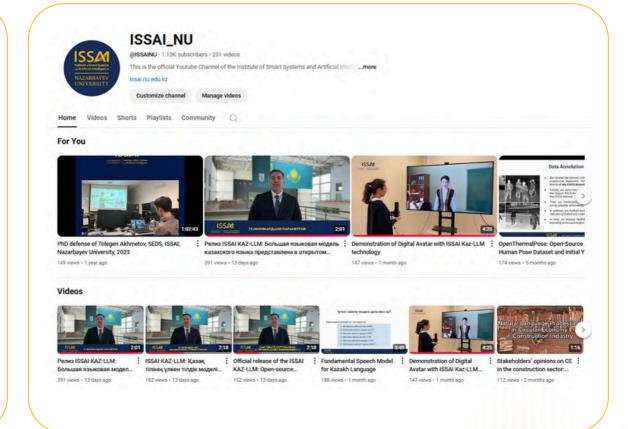


**23** 656 Views



247.5K Impressions







### **ISSAI Ethical Principles**

ISSAI operates in accordance with the following ethical principles:



#### **Societal Well-Being**

All systems should prioritize the benefits for humanity and the stewardship of the environment, emphasizing sustainability, and observing the Hippocratic credo of "first, do no harm".



#### **Human-Centered Values**

Al systems should respect human rights, the rule of law, and democratic values of freedom and dignity. Al systems should respect the privacy and anonymity of people, incorporating data protection, and observing values of equality, non-discrimination, diversity, social justice, and internationally recognized labor rights.



#### Transparency

Al systems utilize algorithms and learning methodologies that can be inscrutable, thus it is imperative to ensure responsible disclosure of a system's design, methodologies, capabilities, limitations, and risks such that humans can understand and challenge the outcomes.



#### **Technical Resilience and Robustness**

All systems must operate in a safe and secure manner, with engineered fault-tolerance and the capacity to detect risks and avoid harm in the event of an error or system failure.



#### Accountability

Organizations and humans developing, using and/or operating AI systems should be accountable for their proper functioning in line with the above principles.

### Subscriptions

#### ISSAI has subscriptions to services listed below:



**NVIDIA** invents the Graphic Processing Units and drives advances in AI, HPC, gaming, creative design, autonomous vehicles, and robotics. ISSAI holds 6 Nvidia DGX servers.



**Overleaf** is a collaborative cloud-based LaTeX editor used for writing, editing and publishing scientific documents. It partners with a wide range of scientific publishers to provide official journal LaTeX templates, and direct submission links.



ISSAI researchers are members of the **Midjourney** - an independent research lab, which explores new mediums of thought and expanding the imaginative powers of the human species.



ISSAI researchers are members of **OpenAI's GPT-4** - the most advanced system, producing safer and more useful responses for users.

### Memberships

#### ISSAI is a member of international organizations listed below:



The International Telecommunication Union (ITU) is the specialized United Nations Agency for ICT services and technologies promotion, collaboration, and standardization.



DataCite is a leading global non-profit organization that provides persistent identifiers (DOIs) for research outputs. With the help of DataCite, organizations assign DOIs to their research works. Since August 2020, ISSAI has been a certified repository with the right to assign up to 100 DOIs.



ISSAI researchers are members of the Institute of Electrical and Electronics Engineers (IEEE), the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.



#### Media About Us

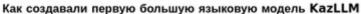










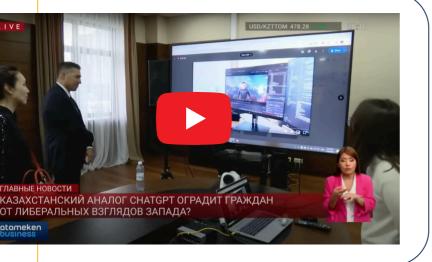


В Казахстане завершена работа над первой версией большой языковой модели на казахском языке (KazLLM). Kazinform побеседовал с разработчиками до того, как их детище накануне было презентовано Главе...





#### Media about us



#### atameken business





Исследователи Nazarbayev University представили уникальную языковую модель KAZ-LLM

Исследователи Института интеллектуальных систем и искусственного интеллекта (ISSAI) Nazarbayev University (NU) представили Главе государству Касым-Жомарту Токаеву большую языковую модель казахского...

🍸 Главные новости Казахстана/ Dec 12, 2024





#### Media about us



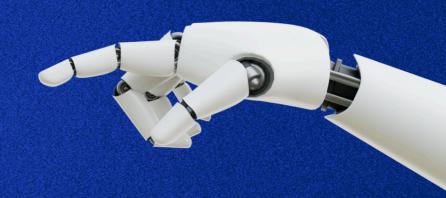












FIND MORE INFORMATION AND NEWS ABOUT ISSAI ON THE OFFICIAL WEBSITE: issai.nu.edu.kz

#### Contact us









Institute of Smart Systems and Artificial Intelligence



Email

issai@nu.edu.kz

Address

53, Kabanbay Batyr Avenue, Block 1, Nazarbayev University, Astana 010000, Kazakhstan

Website

issai.nu.edu.kz