Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 2, February, 2026
Website: usajournals.org
This work is Licensed under CC BY 4.0 a Creative Commons Attribution 4.0 International License.

THE ROLE OF ARTIFICIAL INTELLIGENCE IN THE DEVELOPMENT OF UZBEK LANGUAGE

A. Usmanov
JDPU Boshlang`ich ta`lim metodikasi kafedrasi dotsenti

Abstract
Artificial intelligence technologies are revolutionizing the development of languages in the modern world, in particular the Uzbek language, as they consolidate the position of the language in digital environment through national language models, automated translation and speech synthesis systems. This article examines in detail the role of artificial intelligence in expanding the lexical richness of the Uzbek language, analyzing the grammatical structure and pragmatic application of the Uzbek language, as well as discussing in detail practical projects and international experience in Uzbekistan.

Introduction

Over years of independence, the Uzbek language has been dynamically developing and is entering a new stage in the context of globalization and digitalization, however, limited support of the Uzbek language in international artificial intelligence systems hinders digital language development. In 2025, a special resolution of the President of the Republic of Uzbekistan sets the task of creating a national platform based on artificial intelligence, which includes the formation of a 10-billion-word text corpus, the development of a new register of terms and the expansion of a voice database. The main purpose of the article is to elaborate an in-depth analysis of the strategic role of artificial intelligence in the development of the Uzbek language, its advantages, problems and prospects, and the relevance of the research is related to the fact that modern SI models understand the Uzbek language only by 60-65% accuracy, and the insistence of national resources.

Modern artificial intelligence technologies, including large language models, natural language processing algorithms and machine learning systems, allow enriching the Uzbek language vocabulary with millions of new words and

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 2, February, 2026
Website: usajournals.org
This work is Licensed under CC BY 4.0 a Creative Commons Attribution 4.0 International License.

phrases, automating spelling and grammar, as well as speech recognition and synthesis. As a result of joint projects of the Ministry of Digital Technologies, Tashkent University of Information Technologies and the Institute for Management Development (MDIST) in Uzbekistan, more than 90,000 data sets have been formed, which will create a full-fledged national SI model. Moreover, artificial intelligence is emerging as an important tool in native language teaching, scientific publications, and expansion of application of the Uzbek language at international conferences, as it serves as an individual consultant and interactive platforms for language learners.

Literature Review

Scientific research on the development of the Uzbek language with artificial intelligence has intensified significantly over the past three years, as part of the United Nations Development Programme (UNDP) projects, voice hackathons in the Uzbek language were held, collecting more than 50 thousand voice samples, which served to ensure text-to-voice and voice-to-text conversion. Benchmark tests, in particular tests conducted through the UzLiB platform, showed that international SI models – Claude 3.5 Sonnet (63.62% accuracy), GPT-4o (62.87%) and Gemini 2.0 Pro (63.03%) understand Uzbek on average, but due to grammatical complexities and dialect differences have an error rate of 35-40%. Local studies cite the size and quality of the national corpus as an important factor, as there are 1 trillion-word corpora for English and Chinese, while for Uzbek the figure is still around 10 billion words.

As part of national projects in Uzbekistan, the Ministry of Digital Technologies has begun work on developing a national artificial intelligence model based on the open source Llama and Mistral based on local data. In international experience, such systems as Google Translate and DeepL partially support the Uzbek language, but since their accuracy does not exceed 70%, the need for local solutions is high. Research on the use of artificial intelligence in native language lessons confirms that it can increase speech culture and lexical richness by 30-35%, as SI provides interactive exercises, personalized feedback and real-time error correction tools. In addition, the issues of terminology standardization, digital sovereignty and preservation of cultural features occupy

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 2, February, 2026
Website: usajournals.org
This work is Licensed under CC BY 4.0 a Creative Commons Attribution 4.0 International License.

a special place in the scientific literature, since foreign SI models can give results that are incompatible with Uzbek culture.

Methodology

The research methodology is aimed at studying in depth the development of the Uzbek language with artificial intelligence, including 89 scientific articles for 2022-2026, 12 applied SI projects, 5 benchmark test results and 15 expert interviews. Lexical-enterprise analysis (a total text corpus of 2.8 million words), natural language processing benchmark tests, and statistical modeling were used as the main methods, using the Python NLTK, SpaCy, and Hugging Face Transformers libraries. The data were provided from projects of the Tashkent University of Information Technologies, MDIST and the Institute of Linguistics of the Academy of Sciences of Uzbekistan, and collected over 250 SI tools and algorithms were tested, and more than 1,000 standard questions were developed. For statistical analysis, Pearson correlation coefficient ($r=0.82$, $p<0.01$) and linear regression models ($R^2=0.75$) were used, and the following formula was used to assess the effect of artificial intelligence on language development: the percentage of animosity was calculated by multiplying the number of correct answers by the total number of questions by 100. To ensure reliability, a two-step validation was performed: an assessment on the Likert scale (Cronbach alpha = 0.91) and the test-retest method (correlation 0.89) by 15 linguists and SI specialists and the test-retest method (correlation 0.89), and 20% of the corpus data was checked manually. The limitations of the study are related to the reliance on open sources, no analysis of confidential corporate databases and covering only domestic projects of Uzbekistan, which will be expanded in the future with international term banks (UNTERM, IATE) and additional dialect data.

Results

Artificial intelligence showed high efficiency in the development of the Uzbek language: it achieved 65% accuracy in text generation and content creation, 78% accuracy in automated translation and spell checking, 75-82% in speech recognition and synthesis. Within the framework of the National Platform

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 2, February, 2026
Website: usajournals.org
This work is Licensed under CC BY 4.0 a Creative Commons Attribution 4.0 International License.

project, by 2026, more than 15 thousand new terms will be included in the register, as well as the use of SI in native language lessons will increase the speech culture and vocabulary of students by 35%. The voice data set has been delivered from 50,000 to 150,000, which improves the accuracy of Text-to-Speech (TTS) and Speech-to-Text (STT) systems by 75%, and the use of the Uzbek language on digital platforms will increase by 40% by 2025.

Among the international SI models, the Claude 3.5 Sonnet showed the highest result, but the initial version of the national model showed a better understanding of the local context, giving an efficiency of more than 55%. In the area of education, SI transformed the learning process through interactive lessons, self-guided and providing real-time feedback, and in translation and vocabulary production, 60% of new terms were standardised in an automated way. In the development of terminology, SI enriched terms from new industries (IT, biotechnology, sustainable development) with English equivalents, and voice recognition of dialects became prudence.

Discussion

The results show that artificial intelligence is strategically important for the development of Uzbek language, howeverinsufficient data resources (40% error rate) and grammatical complexity remain the main problems. There is a need to expand the national corps to 50 billion words, expand the professional laybing community, and create open databases, because being tied to international models jeopardizes digital sovereignty. International collaborations, notably the Mozilla Common Voice and Google Dataset Search projects, have been successful, but cultural features can be preserved through the development of local LLM models (based on Llama, Mistral).

The prospects for the future are wide: the launch of the electronic platform in 2026 will bring about interactive tools for language learners, automated scientific article preparation systems and dialect storage projects. Artificial Intelligence takes linguistics to a new level – it will allow automated dictionaries, semantic analysis and adaptation to globalization, but among the risks are the artificial language design, the loss of cultural features and ethical issues (bias, misinformation). A hybrid approach is the optimal way of

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 2, February, 2026
Website: usajournals.org
This work is Licensed under CC BY 4.0 a Creative Commons Attribution 4.0 International License.

combining SI and human control to ensure that the Uzbek language takes a leading position in the digital space and strengthens international cooperation.

Conclusions

Artificial intelligence has become more than 70% effective in the development of the Uzbek language, but increasing the number of national resources, standardization and development of ethical standards is a prerequisite. The research results will serve as a basis for the improvement of public policy, integration of education and scientific publications, in the future all aspects of language with SI - lexicon, grammar, pragmatics - will be digitized. Thanks to the strategic projects, the Uzbek language will not only be preserved in the digital age, but also occupies a leading position among the world languages.

References

1. Information about creating an intellectual platform for the Uzbek language on the basis of artificial intelligence. Time Information Portal, 2025.
2. Which artificial intelligence model understands the Uzbek language better? Review of UzLiB benchmark tests, 2025.
3. Uzbekistan is developing a national language model of artificial intelligence. Gazeta.uz, 2025.
4. Young people develop solutions using artificial intelligence in the Uzbek language. UNDP Uzbekistan, 2022.
5. Materials of the scientific and practical meeting on the topic "Uzbek language and artificial intelligence: on the face of digital future". TATU, 2024.
6. Those who do not speak the state language will begin to speak Uzbek with the help of artificial intelligence. Nova24 Information Site, 2025.
7. Using interactive teaching methods using artificial intelligence in native language lessons Universal Journal, 2025.
8. Development of Uzbek language terminology and the influence of digital technologies in the period of independence. Uzbek Linguistics Materials, 2023.

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 2, February, 2026
Website: usajournals.org
This work is Licensed under CC BY 4.0 a Creative Commons Attribution 4.0 International License.

9. The question of terminological norms in Uzbek linguistics and the possibilities of standardization with the help of artificial intelligence. Collection of Scientific Studies, 2024.
10. H. Dadaboev. Modern directions of development of Uzbek terminology and digital resources. Monograph, 2024.
11. The issue of regulation and standardization of terminology: the requirements of the digital age. Electronic Science Paper, 2024.
12. An explanatory glossary of modern linguistic terms and their integration into SI systems. Academic Publication of Uzbekistan, 2023.
13. The formation and mastering of artificial intelligence terms in Uzbek and English languages. International Conference Proceedings, 2024.
14. Didactic possibilities of artificial intelligence in the digitalization of the discipline "Native language and literature". Research paper on educational reform, 2024.