HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

PROSPECTS FOR THE DEVELOPMENT OF THE UZBEK LANGUAGE AND ISSUES OF PRESERVING LANGUAGE NORMS IN THE ERA OF ARTIFICIAL INTELLIGENCE

To‘ychiyeva Gulsanam Umid qizi
Toshkent Practical Sciences University
+998934870260
umidovnagulsanam7@gmail.com
Doi: https://doi.org/10.5281/zenodo.21905837

Abstract. This article examines the prospects for the development of the Uzbek language in the era of artificial intelligence and the preservation of literary language norms in the digital environment. Particular attention is paid to language corpora, machine translation, speech technologies, terminology, spelling and grammatical standards. The study argues that the effective integration of Uzbek into artificial intelligence systems requires high-quality linguistic resources and systematic cooperation between linguists and information technology specialists.

Key words: Uzbek language, artificial intelligence, language norms, digital linguistics, natural language processing, language corpus, terminology.

Introduction

In the 21st century, the rapid development of artificial intelligence and digital technologies is directly impacting language and communication processes. Text generation systems, automatic translators, voice assistants, speech recognition software, and electronic dictionaries are transforming the ways humans interact with language. Consequently, the presence of languages in the digital space and their adaptability to modern technologies are becoming crucial factors determining their future development.

As the state language, Uzbek holds a special place in social, scientific, cultural, and educational life. Its legal status is enshrined in legislation, and measures aimed at enhancing the prestige and expanding the scope of the state language are being consistently implemented [1; 2]. In the context of digital transformation, ensuring the high-quality use of the Uzbek language across electronic platforms, information systems, and artificial intelligence products has become a pressing priority.

Artificial intelligence enables the generation, translation, and editing of Uzbek texts, as well as speech-to-text conversion and the analysis of vast amounts of linguistic data. However, if these systems are trained on low-quality or non-standardized data, they may replicate existing orthographic, stylistic, and terminological errors. Therefore, the integration of the Uzbek language into artificial intelligence must be approached not merely as a technical issue, but also as a linguistic and cultural process.

Literature review

Studying the grammatical structure and normative characteristics of the Uzbek literary language is a key area of national linguistics. Works by scholars such as H. Jamolxonov, A. Nurmonov, and N. Mahmudov elucidate the phonetic, lexical, grammatical, and stylistic systems of the literary language, as well as the interrelationships between linguistic units and their normative usage [4; 5]. These theoretical perspectives also serve as a foundation for assessing the status of the Uzbek language within the digital environment.

In computational linguistics, natural language processing is studied in the context of large text corpora, morphological and syntactic analysis, machine translation, speech technologies, and

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

generative models [6; 7]. The effectiveness of such systems depends on the volume, quality, and level of standardization of the data used. In Uzbek-language internet texts, the presence of abbreviations, loanwords, inconsistent apostrophe usage, and the mixing of Latin and Cyrillic scripts complicates the issue of stability in language models.

Metodology

The study was conducted using a qualitative approach. Employing analytical, descriptive, comparative, and systemic methods, the research analyzed the impact of artificial intelligence on the Uzbek language, observed orthographic and terminological trends in digital communication, and examined the relationship between literary language norms and emerging technological trends. Additionally, the study synthesized regulatory documents concerning the state language, scholarly sources on Uzbek linguistics, and literature on natural language processing theory. This approach made it possible to evaluate artificial intelligence as a complex factor influencing the development and standardization of the language.

Results

Analyses indicate that artificial intelligence technologies can expand the functional capabilities of the Uzbek language within the digital space. The first key area is the development of national electronic language corpora. AI systems require vast amounts of reliable, standardized data; therefore, it is essential that these corpora provide a balanced representation of literary, scientific, journalistic, official, and spoken language samples, with texts tagged according to genre and grammatical characteristics.

The second area concerns the digitization of Uzbek orthographic and grammatical norms. Standardized spelling dictionaries, morphological analyzers, grammar checkers, and automated editing tools enhance the quality of digital texts. Automated processing is particularly complicated by issues such as the use of various technical characters for the letters “o‘” and “g‘,” the mixed use of Latin and Cyrillic scripts, and the existence of multiple spelling variants for a single word. Consequently, technical standardization must become the digital embodiment of linguistic norms.

The third area involves the systematic development of modern terminology. New concepts are rapidly emerging in fields such as artificial intelligence, programming, cybersecurity, data analysis, and robotics. Certain terms are used in multiple forms within the Uzbek language or coexist alongside foreign equivalents. To ensure precision in scientific and official communication, it is essential to create an open, regularly updated electronic terminology database.

The fourth area of focus is the development of speech technologies for the Uzbek language. Automatic speech recognition and speech synthesis systems will expand the scope of Uzbek language usage in education, public services, media, and everyday communication. This requires high-quality audio databases that encompass not only standard pronunciation but also a variety of age groups and speech contexts. The fifth area involves monitoring texts generated by artificial intelligence for spelling, grammar, terminology, and stylistic accuracy.

Discussion

The relationship between artificial intelligence and the Uzbek language is two-way. On the one hand, new technologies expand the digital capabilities of the language; on the other hand, uncontrolled use can lead to the destabilization of literary language norms. One of the most important issues is the quality of information. Since the language model learns the regularities of the trained material, there is a possibility that spelling errors, inappropriate abbreviations and incorrect translations will be repeated in the system's responses. Therefore, in the formation of Uzbek digital resources, quality should be a priority criterion along with quantity. Another issue is the balance between literary language and everyday digital speech. Abbreviations, dialect units and foreign words are naturally used in social

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

networks. However, not all of them can be accepted as literary norms. Artificial intelligence should distinguish between official, scientific, journalistic and informal styles according to the context. At the same time, it is important that the national-cultural characteristics of the language, communication etiquette, expressions of respect and phraseological units are also correctly interpreted by the models.

Terminological stability also requires special attention. Due to the speed of technological development, the process of creating Uzbek alternatives to new terms sometimes lags behind. As a result, one concept is expressed by different names. To reduce this situation, it is advisable to develop electronic terminological platforms in cooperation with linguists, industry experts and programmers. Artificial intelligence can identify new terminological trends from large volumes of texts, and experts can determine their normative status.

A new approach is also needed in the education system. Since students and teachers use generative systems to prepare and translate texts, they need to develop the skills to critically check the answers given by artificial intelligence, to see spelling and stylistic errors. Rather than banning technology, it is more effective to develop a culture of responsible use, source verification, and adherence to literary language norms.

A single digital linguistic infrastructure is necessary for the sustainable development of the Uzbek language in artificial intelligence systems. This infrastructure should include a national corpus, explanatory and spelling dictionaries, a terminological database, morphological analysis tools, and speech resources in an interconnected manner. If such a system is created, different programmers can use the same normative base, which can reduce spelling and terminological differences that occur on separate platforms. Also, open and regularly updated resources will accelerate scientific research and the creation of local technological products.

Human expertise remains central to maintaining language norms. Artificial intelligence can quickly identify recurring errors, new words, or new usage cases from large volumes of texts, but the question of which form to recommend as a literary norm requires scientific assessment. Therefore, constant expert cooperation between linguists, terminologists, editors, and information technology specialists is desirable. Such cooperation will allow us to direct new technologies not against the natural development of the language, but, on the contrary, to strengthen the richness, accuracy, and modern functions of the Uzbek language.

Conclusion

The results of the study showed that artificial intelligence creates great opportunities for the digital development of the Uzbek language. With its help, it is possible to develop automatic text processing, translation, editing, voice services and educational resources. However, the effectiveness of such opportunities directly depends on the availability of high-quality language corpora, normative dictionaries, terminological databases and linguistic expertise.

The future digital competitiveness of the Uzbek language is also determined by the extent to which it is correctly recognized and processed by modern technologies. Therefore, expanding the national language corpus, making normative dictionaries machine-readable, improving automatic spelling and grammar tools, developing speech technologies and systematically standardizing new terms should be priority tasks.

In general, it is appropriate to view artificial intelligence not as a threat to the Uzbek language, but as a new development tool that should be managed on a scientific basis. If linguists, programmers, educational institutions, and government organizations collaborate, artificial intelligence will serve to strengthen the position of the Uzbek language in the digital space, preserve language standards, and continue the national cultural heritage in new generation technologies.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

The list of used literature

1. Law of the Republic of Uzbekistan “On the State Language”. October 21, 1989 (new edition December 21, 1995).
2. Decree of the President of the Republic of Uzbekistan “On measures to radically increase the prestige and status of the Uzbek language as the state language” No. PF-5850. October 21, 2019.
3. Regulatory and legal acts of the President of the Republic of Uzbekistan on the development of artificial intelligence technologies.
4. Jamolkhonov H. The current Uzbek literary language. – Tashkent: O’qituvati, 2012. – 256 p.
5. Nurmonov A., Mahmudov N. Theoretical grammar of the Uzbek language. – Tashkent: Fan, 2006. – 320 p.
6. Jurafsky D., Martin J. H. Speech and Language Processing. – Pearson, 2009.
7. Vaswani A. and others. Attention Is All You Need // Advances in Neural Information Processing Systems. - 2017.
8. UNESCO. Recommendation on the Ethics of Artificial Intelligence. – Paris: UNESCO, 2021.