BUILDING AN EXPLANATORY STRUCTURED DICTIONARY OF THE UZBEK LANGUAGE USING ARTIFICIAL INTELLIGENCE

Daminova Barno Esanovna,
Associate Professor, Department of Algorithms and Programming Technologies, Karshi State University
https://orcid.org/0009-0001-4211-6082

Sunnatova Jasmin Humoyun qizi
Student of Karshi State University
https://orcid.org/0009-0000-2731-8068

Annotation. This article explores the development of an explanatory and structured dictionary of the Uzbek language through the application of artificial intelligence (AI) technologies. The research focuses on how natural language processing (NLP), machine learning, and semantic modeling can be used to systematically organize lexical units, generate definitions, identify relationships between words, and improve accessibility for users.

Keywords. artificial intelligence, Uzbek language, explanatory dictionary, natural language processing, lexicography, semantic structure, machine learning, linguistic resources.

Annotatsiya. Ushbu maqola sun'iy intellekt (AI) texnologiyalarini qo'llash orqali o'zbek tilining tushuntirishli va tuzilgan lug'atini ishlab chiqishni o'rganadi. Tadqiqot tabiiy tilni qayta ishlash (NLP), mashinani o'rganish va semantik modellashtirishdan leksik birliklarni tizimli ravishda tashkil qilish, ta'riflarni yaratish, so'zlar o'rtasidagi munosabatlarni aniqlash va foydalanuvchilar uchun qulaylikni oshirish uchun qanday foydalanish mumkinligiga qaratilgan.

Kalit so'zlar: sun'iy intellekt, o'zbek tili, tushuntirishli lug'at, tabiiy tilni qayta ishlash, leksikografiya, semantik tuzilma, mashinani o'rganish, lingvistik resurslar.

Аннотация. В данной статье рассматривается разработка пояснительного и структурированного словаря узбекского языка с применением технологий искусственного интеллекта (ИИ). Исследование сосредоточено на том, как обработка естественного языка (ОБЯ), машинное обучение и семантическое моделирование могут быть использованы для систематической организации лексических единиц, генерации определений, выявления связей между словами и повышения доступности для пользователей.

Ключевые слова: искусственный интеллект, узбекский язык, пояснительный словарь, обработка естественного языка, лексикография, семантическая структура, машинное обучение, лингвистические ресурсы.

This article presents an in-depth exploration of constructing a comprehensive explanatory and structured dictionary of the Uzbek language using artificial intelligence (AI). It examines the integration of natural language processing (NLP), deep learning, and semantic technologies in automating lexicographic processes. The study proposes a multi-layered framework for lexical representation, including morphological, syntactic, and semantic levels. Special attention is given to the challenges of processing an agglutinative language like Uzbek, including word formation complexity, polysemy, and dialect diversity. The paper also highlights the role of large-scale corpora, annotation strategies, and hybrid human-AI collaboration. The resulting model aims to produce a dynamic, scalable, and intelligent dictionary system capable of supporting education, research, and modern language technologies.

The Uzbek language, belonging to the Turkic language family, is characterized by rich morphology and flexible syntax. Despite its widespread use, there is still a lack of comprehensive, structured, and machine-readable dictionaries that fully capture its lexical and semantic richness. This article aims

to propose a detailed framework for building such a resource using AI technologies.

An explanatory dictionary differs from a simple word list by providing detailed meanings, usage examples, grammatical information, and semantic relationships. For Uzbek, creating such a resource is crucial for: Supporting language learners and educators, Preserving linguistic richness and cultural heritage, Enhancing computational applications such as machine translation and speech recognition, Facilitating academic research in linguistics.

Modern lexicography has evolved from static printed dictionaries to dynamic, interactive, and corpus-based systems. Digital dictionaries are expected to provide not only definitions but also semantic relations, contextual usage, frequency information, and multimedia support.

AI enables the automation of linguistic analysis through techniques such as: Natural Language Processing (NLP), Deep Learning, Knowledge Representation, Semantic Web Technologies.

Annotated corpora are essential for supervised learning. Annotation includes: Part-of-speech tagging, Morphological labeling, Semantic tagging.

AI models identify frequently used words and phrases, prioritizing them for dictionary inclusion.

Techniques such as word embeddings and transformer-based models capture semantic relationships between words.

The application of artificial intelligence in Uzbek lexicography offers unprecedented opportunities for innovation. By combining computational techniques with linguistic expertise, it is possible to create a comprehensive, structured, and intelligent explanatory dictionary. Such a resource will not only preserve the richness of the Uzbek language but also ensure its active participation in the global digital ecosystem.

References:

1. Daminova B. E. et al. SUN'IY INTELLEKTNING RIVOJLANISH TENDENSIYASI //Экономика и социум. – 2025. – №. 5-1 (132). – С. 221-225.

2. Daminova B. E. et al. SUN'IY INTELLEKTNING KASBLARGA TAHDIDI YOKI YORDAMCHILIGI //Экономика и социум. – 2025. – №. 5-1 (132). – С. 231-234.

3. Omonova M. NOMINATIVE-DEFINITIVE FUNCTIONS OF COMPONENTS OF AMELIORATIVE TERMS IN ENGLISH AND UZBEK LANGUAGES //Theoretical & Applied Science. – 2021. – №. 4. – С. 84-86.

4. Omonova M. K. Comparative analysis of semantical features of meliorative terms in English and Uzbek //Experientia est optima magistra. – 2021. – С. 269-272.

5. Omonova M. Innovative ways of teaching vocabulary in ESL and EFL classrooms //Science and Education. – 2020. – Т. 1. – №. 7. – С. 229-233.

6. Mallayev A. ALISHER NAVOIY VA BOBUR YOKI MILLATIMIZNING IKKI BUYUK QUYOSHI //Conferences. – 2025. – Т. 1. – №. 1. – С. 5-12.

7. QUVATOVA D. O 'QUVCHILARDA MANTIQIY-MUHOKAMAVIY FIKRLASH FAOLIYAT TUSHUNCHASINING NAZARIY TAHLILI //«ACTA NUUz». – 2025. – Т. 1. – №. 11. – С. 114-116.

8. Raximov N. et al. As a mechanism that achieves the goal of decision management //2021 International Conference on Information Science and Communications Technologies (ICISCT). – IEEE, 2021. – С. 1-4.

9. Kaynarov F. Z. THEORETICAL FOUNDATIONS FOR THE CREATION OF ELECTRONIC TEXTBOOKS FOR DISTANCE EDUCATION //Экономика и социум. – 2024. – №. 2-2 (117). – С. 169-175.

10. Кайнаров Ф. З. ИННОВАЦИОННЫЕ МЕТОДЫ ПРЕПОДАВАНИЯ ПРИКЛАДНОЙ МАТЕМАТИКИ //Экономика и социум. – 2023. – №. 1-2 (104). – С. 619-622.

11. Kaynarov F. APPLICATION OF MODERN INFORMATION TECHNOLOGIES IN MEDICINE //International Scientific and Practical Conference on Algorithms and Current Problems of Programming. – 2023.