PROBLEMS AND SOLUTIONS TO THE CREATION OF THE OLD UZBEK LANGUAGE AND ITS TAJVID

[1]Iskandarova S.N., [2]Begimkulova P.A., [3]Akbarova M.A., [4]Khaitov Kh.S.
[1,3,4]Tashkent University of Information Technologies named after Muhammad al-Khorazmiy
[2]Navoi state pedagogical institute
https://doi.org/10.5281/zenodo.11403399

Abstract*. Ancient Uzbek writing has served the peoples of Central Asia for a thousand years, and it is known that rich written cultural monuments have been created in it. The same can be said about the fact that after 1917 this inscription was reformed, and then switched to Latin and Slavic writing. Today, when the attitude to cultural heritage and to the study of ancient Uzbek writing, which has become the key to heritage, has changed dramatically, there have been controversial places in the issue of the name of this writing, which has become the national - historical writing of our people. Some felt it should be called "Arabic script" and others "Arabic - Uzbek."*

Keywords*: ancient, cultural monuments, cultural heritage, inscription.*

Ancient Uzbek writing has served the peoples of Central Asia for a thousand years, and it is known that rich written cultural monuments have been created in it. The same can be said about the fact that after 1917 this inscription was reformed, and then switched to Latin and Slavic writing. Today, when the attitude to cultural heritage and to the study of ancient Uzbek writing, which has become the key to heritage, has changed dramatically, there have been controversial places in the issue of the name of this writing, which has become the national - historical writing of our people. Some felt it should be called "Arabic script" and others "Arabic - Uzbek." Yes, writing has surpassed Arabic. " More than a thousand years have passed since then. Now the inscription has become not only Arabic, but also Iranian, Turkic and other peoples. Secondly, the inscription itself has changed, letters have been added to it, which were not in Arabic, its graphic capabilities have expanded and improved. In Arabic, writing served as a consonant. Among the Turkic peoples, including in the history of the Uzbek people, this writing (based on the ancient Turkic writing) was transformed into phonographic (evenly expressing vowels and consonants) writing, and its spelling basis changed - 3 www.ziyouz.com libraries. For these and similar reasons, ancient writing in the history of our people cannot be called "Arabic writing." Arabic writing - Persian, Turkish - could be called "Turkic." However, "ancient Turkic writing" refers to ancient Turkic writing, such as O'rhun - Enasai writing and Uyghur writing. In contrast to the concepts of "Old Turkic language," "Old Turkic script" is used the term "Old Turkic language" in relation to the history of the Uzbek language, according to which our great ancestors, such as Lutfi, Alisher Navoi, Babur, worked and left behind a tired inscription based on the Arabic alphabet Old Turkic script can be called "Old Uzbek". It may also differ in the fate of the Arabic script in comparison with the history of other peoples who created great spiritual values. Many of our historical sources are given in Old Uzbek. Using handwriting recognition programs, you can create the electronic form of these manuscripts. In addition, it is required to translate it into Latin or crillography in order to convey to modern readers.

aspect	Old Uzbek script	Modern language platforms
Recording system	Arabic script	Current Latin
Lexical content	old words	New words, technological and modern terminology
Grammar	Complex and consistent additions	Simplified grammar rules
Language platforms	Limited support	Broad support, e.g. Duolingo, Google Translate
Users	Mostly scientists and linguists	Wide audience, including youth and students

The table compares the capabilities of the existing language platform. Modernity, technology, and language grammar also differ from modern language platforms. Therefore, the creation of the text of the Old Uzbek script in krill or Latin script is one of the pressing problems.

Methodology

To translate the Old Uzbek language into Latin graphics, it is necessary to go through a number of methodological stages. This process can be applied to automatic translation systems or to language learning applications. Let's take a look at the main steps to create this platform:

New Database

Old texts Collect texts: Use of old texts, books, manuscripts and other sources.

Transliteration rules: Development of specific rules for the transition from Old Uzbek to Latin graphics. It is recommended that these rules be developed in conjunction with linguists and linguists.

Development of transliteration algorithms

Software development: creating software to convert old text into Latin graphics. This can be done in programming languages such as Python.

Automated transliteration: Create an algorithm to convert each letter or combination of letters into corresponding letters of Latin graphics.

Quality assessment

Testing: Testing the transliteration algorithm in various texts. Correcting errors found in this process. Experts Evaluation: Evaluation of transliteration quality involving linguists and historians. Processing based on their opinions.

User interface

User friendship: Create a simple and intuitive user experience. This can be for mobile and web applications.

User Assistance - Explain the transliteration process and platform capabilities to users.

Integration and extensibility

Integration with other platforms: for example, consideration of integration with educational platforms or translation systems.

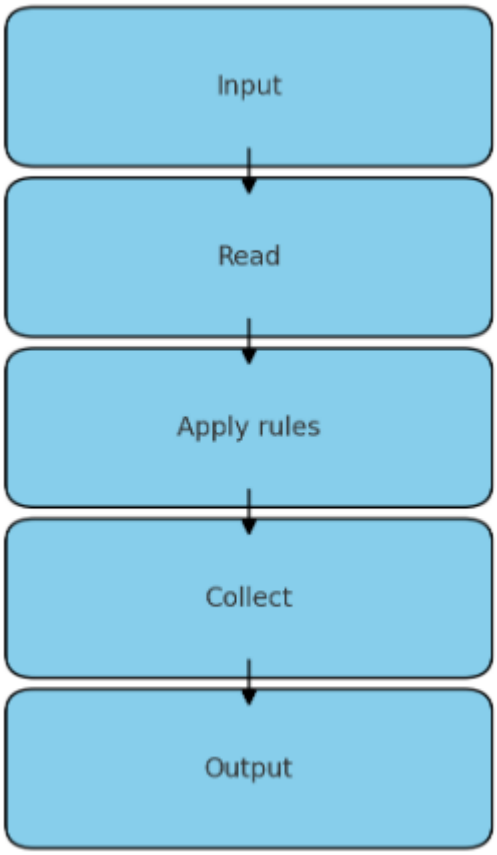
Language variants: Creating the ability to support various dialects and variants of the historical language.

Continuous updates and support

Updates: Regular updates to adapt to changes in language rules and graphics.

User support: continuous improvement of the platform by accepting user feedback.

Algorithm for translating Old Uzbek into Latin



The above block diagram is an algorithm for translating old Uzbek into Latin graphics. Each step includes:

Reading the opening text: The program accepts the Old Uzbek text.

Checking each character: The program will read each character in the text.

Applying Transliteration Rules: Change each old Uzbek letter to the corresponding Latin letter or leave punctuation marks unchanged.

Collect final text: The changed letters or characters will be added to the new text.

Output of the final text: The modified text will be displayed to the user or saved for further processing.

To translate Old Uzbek from Arabic into Latin graphics, clear rules must be developed. This involves converting each Arabic letter or combination according to the Latin letter during the transliteration process. Below I will give general rules for this process:

Basic letters:

ا	A	ر	ف	F
ب	B	ز	ق	Q
پ	P	ژ	ك,ک	K
ت	T	س	گ	G
ث	S	ش	ل	L
ج	J	ص	م	M
چ	Ch	ض	ن	N

ح	H	ط	و	V, O, U (depending on rational pronunciation)
خ	Kh	ظ	ه	H
د	D	ع		
ذ	Z	غ		

ى, ي - Y, I (depending on rational pronunciation)

Special Marks and Diacritical Mark

تاشديد (شدّة) - Overwrite letter twice (for example, الله -> Alloh)

فتحة - A

ضمّة - U

كسرة - I

سكون - Read aloud

Other rules

Capital words are capitalized if necessary.

Sentences and sentences are based on the grammatical principles of the Uzbek language.

Diacritics and punctuation marks (periods, commas) are used in accordance with the Latin alphabet.

Case studies

مكتب - school

درس - lesson

كتاب - book

علم - scientist

صحبت - Sahbat

These rules are intended to be translated from Old Uzbek into modern Latin. In practice, this process can be complex, as some letters and sounds are not the same in both writing systems. Based on these rules, we translate the text of the old Uzbek language into Latin script. It is changes in the language that do not lead to a full text value.

Using Natural Language Processing (NLP) technologies to translate letters into Latin graphics in old Uzbek and then understand the meaning of the text can be a very effective approach. This process involves several stages, and each stage is aimed at improving the efficiency of working with text. The following is a general outline of this process and its benefits.

Scheme for translating Old Uzbek into Latin and using NLP:

Transliteration:

Introduction: Text in old Uzbek.

Process: Translating each Arabic letter or combination into Latin.

Output: transliterated text based on Latin script.

Tokenization:

Login: Transliterated text.

Process: divide text into small fragments (tokens).

Output: Individual tokens (words and signs).

Lexical and grammatical analysis:

Introduction: Tokens.

Process: determining the lexical meaning and grammatical function of each token.

Output: Annotated tokens (e.g. horse, felt, quality).

Material analysis:

Introduction: Annotated tokens.

Process: context analysis, for example, recognition of context units, syntactic and semantic analysis.

Exit. Fully understandable text and related information.

To compare the results of the program, which includes the use of only the letter exchange program and the use of NLP in the old Uzbek language, you can create a table based on the following criteria. This table helps to compare the advantages and disadvantages of both approaches.

Comparison table

Criteria	Letter changer only	Program used in NLP
Accuracy	Low. Only letters, grammar and content are replaced without attention.	Yuqori. Kontekst va til qurilmasi hisobga olinadi.
Ma'noni Tushunish	Limited. The meaning of the text is difficult to understand.	Strong. Deeply understands the meaning of the text.
Language rules	No rules apply.	The grammatical and lexical rules of the language apply.
Multiple circumstances	There is only one transliteration method.	Adaptive for different situations and contexts.
User friendships	It usually works simply and quickly.	Harder, but the results are more useful and accurate.
Widespread use	Limited, mainly used for regular transliteration.	It is widely used, including in automated translation and language learning.
Development potential	Limited. It is difficult to add new properties.	Big. There is the possibility of using artificial intelligence and machine learning technologies.
Correcting errors	Errors are often corrected manually.	Errors can be detected and corrected automatically.

Accuracy: Programs that work with NLP take into account complex aspects of the language, so their accuracy is much higher than just letter exchange programs.

Understanding the meaning: NLP programs are more effective in understanding the meaning of text because they are able to analyze the semantics and context of a language.

Widespread use: The versatility of NLP technologies allows them to be used in translation, in language learning applications, and in many other areas.

Therefore, we chose the NLp approach to translate Old Uzbek into Latin graphics and obtained the following results. A table has been compiled showing to what extent the meaning of

the text can be understood by translating 10 texts of the Old Uzbek script into Latin script using NLP. Previously, software tools for these two methods were developed. Real tests were carried out and the following results were obtained.

Book number	Book title	Author	Understand the value by replacing the letter (%)	Understanding the topic through NLP (%)
1	"Divan-i Hikmat"	Alisher Navoi	30%	86%
2	"Baburname"	Babur	30%	85%
3	"Qutadgu Bilig"	Yusuf Has Hajib	30%	88%
4	"Hibat al-haqoyiq"	Ahmad Yugnaki	30%	85%
5	"Akhbor al-akhyar"	Abdurahmon Jomiy	30%	85%
6	"Mahbub ul-qulub"	Abdurahmon Jomiy	30%	89%
7	"Munojot"	Mahmud Koshgari	30%	85%
8	"Lison ut-Tayr"	Ahmad Fariduddin Attar	30%	85%
9	"Chahor Maqola"	Nizomiddin Shami	30%	87%
10	"Asrar al-Tawhid"	Mohammad Ibn Monavvar	30%	85%

Based on these results, it can be said that studying rich historical texts in NLp can increase the recognition percentage.

Conclusion

When translated from Old Uzbek to Latin script, using only the letter replacement method provides relatively low comprehension (30%). This method does not take into account complex aspects of the text, including grammar, vocabulary and semantics. The use of NLP (Natural Language Processing) technologies allows you to understand text values at a higher level (85%). It depends on NLP's ability to analyze language deeply. NLP analysis allows you to process texts more efficiently by determining the contextual meaning of the language and understanding grammatical structures. This can be useful in translation, training and other areas. NLP approaches can be used not only to translate old records into modern graphics, but also to process various language variants, including dialects and texts written according to various language rules. This expands the possibilities of working in various language environments. The use of NLP technologies makes it easier to understand old language materials and pass them on to modern learners. This is important for linguistics and historical research. With the development of NLP technologies, the quality and efficiency of language processing is improving, which opens up new opportunities for wider scientific and applied research.

REFERENCES

1. Eisenstein J. Introduction to Natural Language Processing. — Cambridge, MA: MIT Press, 2019. — 350 s.
2. Jurafsky D., Martin J.H. Speech and Language Processing. 3rd ed. — Pearson, 2019. — 1024 s.
3. Manning C.D., Schütze H. Foundations of Statistical Natural Language Processing. — Cambridge, MA: MIT Press, 1999. — 680 s.
4. Bird S., Klein E., Loper E. Natural Language Processing with Python. — Sebastopol, CA: O'Reilly Media, 2009. — 504 s.
5. Versteegh K. The Arabic Language. — Edinburgh: Edinburgh University Press, 2001. — 416 s.
6. Losensky P. (red.) Languages and Linguistics of Central Asia: A Comprehensive Guide. — Edinburgh: Edinburgh University Press, 2018. — 288 s.
7. Pedersen T.T. Transliteration of Non-Roman Scripts: A Comprehensive Approach. — Aalborg: Datakon, 2006. — 234 s.
8. Mertz D. Text Processing in Python. — Boston, MA: Addison-Wesley Professional, 2003. — 560 s.
9. Arnold D. va boshq. Machine Translation: An Introductory Guide. — Oxford: NCC Blackwell, 1994. — 300 s.
10. Campbell L. Historical Linguistics: An Introduction. 3rd ed. — Edinburgh: Edinburgh University Press, 2013. — 448 s.
11. Sayfullayeva D. THE ROLE OF MONOLOGUE AND DIALOGUE IN KHUDOYBERDI TOKHTABAEV'S NOVEL" MUNGLI KOZLAR" //Science and innovation. – 2023. – T. 2. – №. C10. – C. 154-158.