

UDC 811.1'33

IRSTI 16.31.02

<https://doi.org/10.48371/PHILS.2025.4.79.003>

INTEGRATING SOFTWARE AND PROGRAMMING TOOLS IN LINGUISTIC AND SOCIAL SCIENCE RESEARCH IN KAZAKHSTAN

Akishev T.B.¹, *Kazybekova U.M.², de Lasala Porta S.³

^{1, *2,3}KIMEP University, Almaty, Kazakhstan

Abstract. When integrated into linguistic and social science research, digital data analysis tools help scholars explore linguistic and social science data, improving their understanding of complex theoretical concepts and their real-world applications. While these tools are commonly used by researchers at leading institutions abroad, Kazakhstani linguistic and social science studies are generally considered more theory-driven and less practice-oriented. The adoption of these tools by local researchers and institutions can help bridge the gap between theory and practice and provide them with useful, relevant, and competitive knowledge and skills. To help address these limitations, this paper presents an overview of a selection of digital tools that includes both software and programming resources. The tools presented have a scientific foundation, as they are generally used for digital data analysis and visualization in linguistics and social science research. Relying on this scientific basis, this study presents the general and specific characteristics of these digital tools and the ways in which they can be integrated into linguistics and social science-focused studies. In conclusion, the main contribution that this study makes to theory and practice is that it provides highly applicable and relevant suggestions for the improvement of these scientific components in a local context.

Keywords: research, linguistics, social sciences, philosophy, software, programming, tools, data analysis, visualization

Introduction

As the scientific study of language, linguistics examines the different aspects of language, including its sound system (phonetics and phonology), word formation patterns (morphology), word order and sentence structure (syntax), meaning (semantics), and contextual interpretation (pragmatics). The use of language in various social contexts and settings is also of interest to linguists. The acquisition of language is another important area of linguistic research. Finally, linguists seek to understand how language developed historically and how it is evolving in the present. These and many other topics are studied in terms of the various branches of linguistics. Closely connected with the humanities and social sciences, this field combines both theoretical analyses and practical approaches to understand the complex nature of language.

The social sciences include various types of disciplines: anthropology, sociology, psychology, economics, political science, geography, history, and communication studies. These various fields seek to explain the specificities

of human behavior, social hierarchies, and cultural processes. Like the social sciences, some branches of philosophy, such as philosophical anthropology, political philosophy, and ethics, focus on the study of human beings and society; however, while the social sciences tend to rely more on quantitative and scientific methods, philosophy emphasizes cultural interpretation, values and conceptual understanding. The goals of philosophy typically include the development of critical thinking skills, aesthetic appreciation, and historical understanding.

Research methods in linguistics and the social sciences differ widely. However, they all include both qualitative approaches, such as, for instance, ethnographic observation, interviewing, and discourse analysis, and quantitative techniques like experiments, surveys, and statistical analyses. Recently, these fields have started to incorporate more computational and data-driven methods. For example, both linguistics and the social sciences use tools such as the programming language Python for data analysis. In linguistics, software like *#LancsBox X* and *TreeForm* is commonly used to analyze language data, while social science research often employs *NVivo* and *MAXQDA* for qualitative analysis and *SPSS* for statistical analysis.

In Western higher education institutions, researchers typically engage with both theoretical constructs and their practical, real-world applications. It is common for linguistics and social science programs to introduce them to specialized, research-based software tools and programming languages to help them analyze and visualize data. The integration of these digital tools is crucial for ensuring that they acquire the most relevant, useful, modern, and competitive knowledge.

In Kazakhstan, linguistics and social science research and education tends to prioritize theory-related discussions. There is typically little opportunity for scholars to practice and develop analytical skills in terms of data analysis procedures focusing on the exploration of theoretical concepts in various practical contexts. Integrating theory with practice is crucial because it allows them to engage in active learning.

The main purpose of this study is to provide an overview of selected software and programming tools that can be adapted by Kazakhstani researchers interested in linguistics and the social sciences. Importantly, this paper focuses more on the benefits of applying these tools to enhance the analysis and understanding of linguistic and social science data than on a purely theoretical discussion that would draw insight from contemporary research methodologies.

Materials and methods

This paper contains a presentation of research-based digital tools related to linguistics and the social sciences. These tools, including both software and programming-based resources, are commonly used by researchers and institutions abroad. The linguistics disciplines discussed here include morphology and syntax, corpus linguistics, and computational linguistics. The social science disciplines include political science, psychology, sociology, economics, anthropology, geography, history, philosophy, and communication studies.

Importantly, for the purposes of this study, we draw no strict demarcation

between software and programming tools. All the software tools mentioned in this paper are based on a programming language and incorporate one or more programming-based language analysis procedures. For us, the main difference lies in usability: software tools are user-friendly applications with graphical interfaces, while programming tools are procedures that experienced users can implement themselves, typically within a programming environment without relying on a simplified GUI. Furthermore, while software tools described here may be more accessible, the programming-based ones invite scholars to master certain aspects of a programming language in order to be able to perform more advanced analytical procedures and directly control their data analysis process.

Results and discussion

This section presents our findings along with their interpretation. We describe analytical tools used in linguistics and the social sciences. The linguistics tools discussed here were integrated into the first author's research practice. For the social sciences, we decided to include a more extensive collection of tools, without limiting the selection to those used in research by the co-authors of this paper.

1. Tools used in linguistic research

The procedures and tools described below were applied in research, including the following subfields of linguistics: *Morphology and Syntax*, *Corpus Linguistics*, and *Computational Linguistics*. Some of the procedures described below naturally overlap. This is because a lot of advanced analytical procedures are based on standard tasks performed in terms of morphosyntactic analysis. The discussion below presents each tool only once in order to avoid redundancy.

(1) Morphology and Syntax

In morphosyntactic analysis, a wide variety of tools are available. They focus on lemmatization, morphological tagging, recognition of inflectional and derivational patterns, part-of-speech tagging, dependency parsing, and other analytical procedures [1]. The following example illustrates the use of the Google Colab programming environment for one of these procedures, dependency parsing.

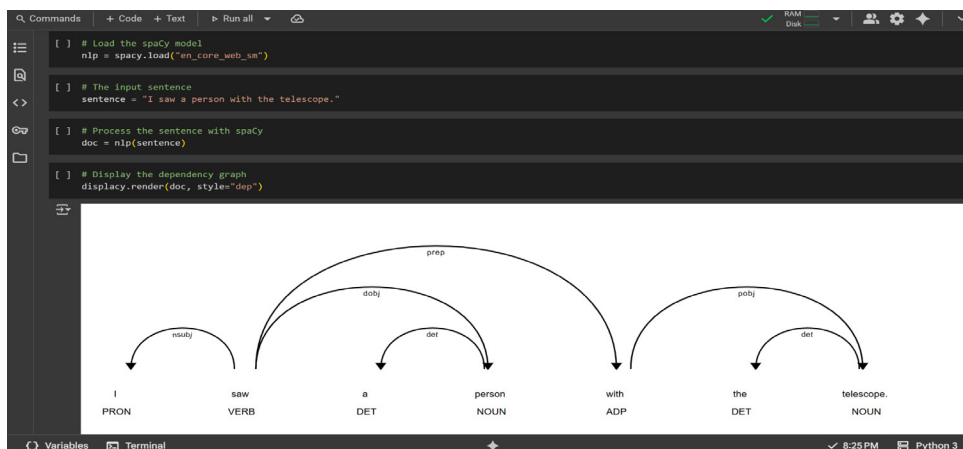


Figure 1 – Dependency parsing in *Google Colab*

In generative syntax, a popular software tool for data analysis and visualization is *TreeForm*, which helps researchers create syntax trees [2]. Online syntax tree generators are also available, offering various capabilities.

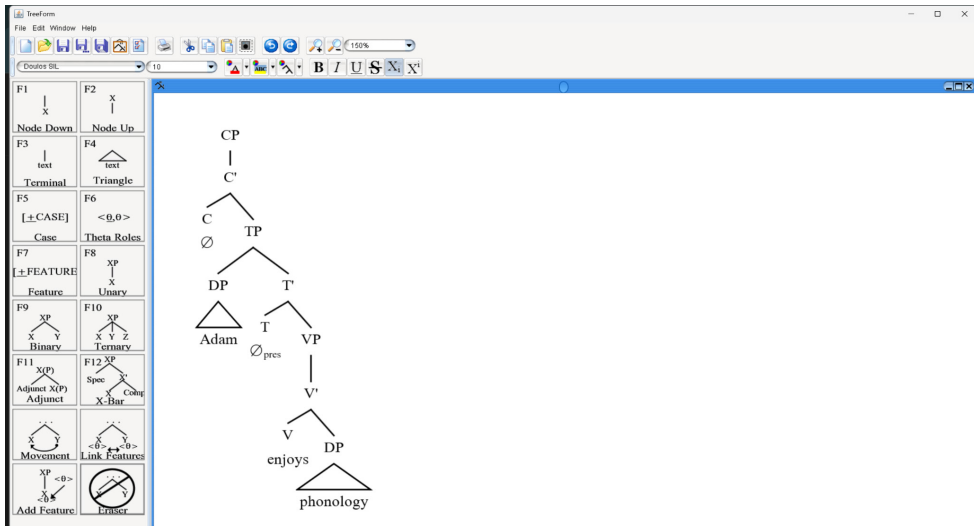


Figure 2 – Syntax tree created using *TreeForm*

(2) Corpus Linguistics

In corpus linguistics, tools like *#LancsBox X* tools are frequently used. Such tools allow researchers to do corpus compilation, querying, data elicitation, analysis, and visualization [3]. An example of such analytical procedures is given below, namely collocation analysis.

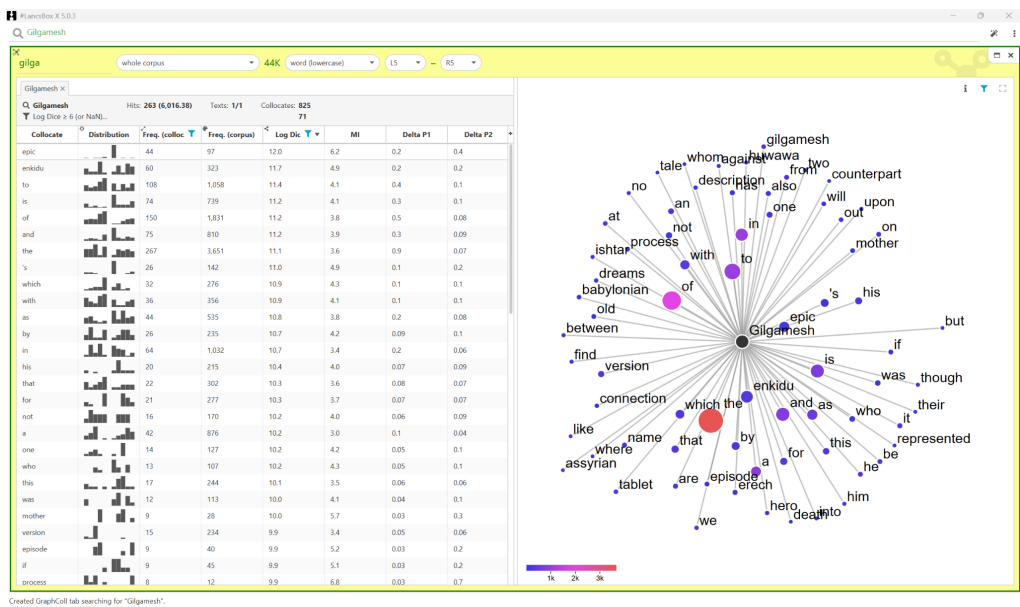
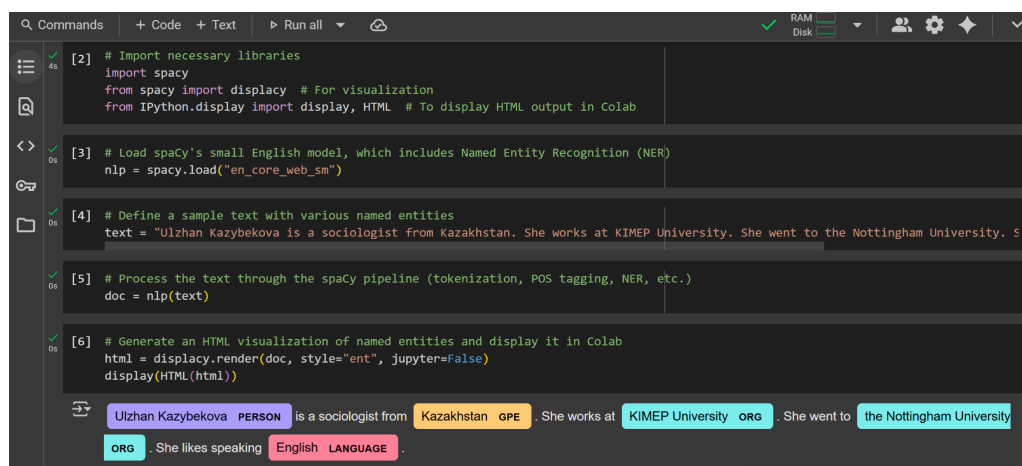


Figure 3 – Collocation analysis in *#LancsBox X*

(3) Computational Linguistics

This subfield is where Python is more frequently used, and researchers work with it in a programming environment, such as, for example, Google Colab, PyCharm Community Edition, or Python IDLE. These programming environments offer various levels of accessibility and user-friendliness. Programming is used here to carry out advanced analytical procedures, such as sentiment analysis, named entity recognition, as well as those previously mentioned in relation to morphosyntax. Python libraries that are frequently used include spaCy and NLTK. They are popular tools for Natural Language Processing (NLP).



```
[2] # Import necessary libraries
import spacy
from spacy import displacy # For visualization
from IPython.display import display, HTML # To display HTML output in Colab

[3] # Load spaCy's small English model, which includes Named Entity Recognition (NER)
nlp = spacy.load("en_core_web_sm")

[4] # Define a sample text with various named entities
text = "Ulzhan Kazybekova is a sociologist from Kazakhstan. She works at KIMEP University. She went to the Nottingham University. She likes speaking English."

[5] # Process the text through the spaCy pipeline (tokenization, POS tagging, NER, etc.)
doc = nlp(text)

[6] # Generate an HTML visualization of named entities and display it in Colab
html = displacy.render(doc, style="ent", jupyter=False)
display(HTML(html))
```

Ulzhan Kazybekova PERSON is a sociologist from Kazakhstan GPE . She works at KIMEP University ORG . She went to the Nottingham University ORG . She likes speaking English LANGUAGE .

Figure 4 – Named entity recognition analysis using *spaCy* in Google Colab

In conclusion, the above-mentioned digital tools support a wide range of linguistic data analysis and visualization tasks. Mastering them equips researchers with important analytical skills and makes them competitive specialists in the field.

2. Tools used in social science research

In this section, instead of providing visualizations for the use of some of the tools below, the focus is on discussing them. As they are often employed in social science research for data analysis, these tools are recognized by researchers worldwide. They are used to illustrate how data from surveys and interviews could be analyzed and interpreted by scholars in their learning and research endeavours. The first example is *NVivo* which is a software tool that enables researchers to manage and analyze different types of data, like text, audio/video and photos. It is also used to organize, code and explore the data obtained from interviews and surveys and to visualize relationships within the data by thematic coding [4]. The second one is *MAXQDA*, which is a widely used software tool that helps to analyze all the data collected in the context of empirical social research. It provides tools for organizing, coding, analyzing and visualizing

data from various sources, e.g. interviews, focus groups, documents, video and audio recordings and is particularly helpful to explore social concepts [5]. And, the next example is *SPSS* which is widely used for statistical analysis in social sciences research that enables it to manage and analyze large amounts of datasets. Particularly, statistical analysis is useful because it allows scholars to conduct research, read and evaluate journal articles, and to enhance their analytical and critical thinking skills [6].

Notably, there are various data analysis techniques that are used in social sciences research. First, for qualitative analysis that focus on non-numeric data, such as interviews or surveys, software tools such as NVivo, MAXQDA, or Atlas.ti are utilized for coding, managing and analyzing data in order to obtain in-depth understanding about social concepts. Second, data mining techniques are used to get information from large datasets with software tools such as IBM SPSS Modeler, RapidMiner, or Weka that cluster, analyze and predict modeling algorithms about valuable patterns and information in social science research. Third, network analysis is one of the popular methods to study the relationships between networks by utilizing software tools, e.g. Gephi, Pajek, or UCINET that provide features to measure network centrality and analyze the dynamics of networks in the society. For multivariate analysis that is useful when analyzing data with multiple variables, statistical software tools, such as SPSS, SAS, R, STATA, and SmartPLS are used by social science researchers to conduct factor analysis and structural equation modeling effectively [7]. In addition, *OBSIDIAN* is a very useful tool for philosophical anthropology, political philosophy and epistemology in general, which organizes knowledge using plain text notes. It allows the scholar to view multiple notes at once, create mind maps, diagrams, and interactive graphs. It enables users to cross-reference information, clarify ideas, and search content based on specific criteria. It offers a visual and comprehensive way to connect and structure various types of information [8].

To conclude, the above-mentioned digital tools are commonly used to manage and analyze various types of data in social sciences. Being able to handle such software tools enables scholars to enhance their practical, analytical and problem-solving skills, which are very crucial in today's age of advanced technologies.

Software and programming tools can be used for various purposes in research. Today's world is characterized by the availability of a variety of content, which means that traditional learning methods are no longer applicable. Specific actions must be taken to enhance educational activities that will replace ineffective methods and instead to use the new and advanced ways of learning methods in educational institutions. [9].

Pedagogical methodologies related to the use of software and programming tools are of paramount importance in the educational process. Modern pedagogical software tools enhance effectiveness and boost interest in learning in several ways [10]. First, visualization capabilities give the opportunity to use animations,

graphics and simulations to explain complex concepts in an easier way. Second, interactivity promotes active involvement and makes individual work possible. Third, adaptability permits presenting learning materials that match the needs and skills of each scholar. And, finally, it saves time for researchers during the preparation and carrying out their studies [10].

Also, the types of pedagogical software tools can be grouped by multimedia tools to explain the content of the lesson via Google Slides or PowerPoint; interactive educational platforms, like Moodle, Google Classroom, Edmodo effectively aid to manage the educational process; virtual and augmented reality enables the interactive view of the lesson content; software models and simulations with special programs designed to teach specific science subjects and use of online learning tools, e.g. YouTube tutorials, electronic books and distance learning platforms in the learning process [10].

Technology contributes significantly to innovation growth in learning, since digital tools allow researchers to implement various methods [11]. In addition, advanced technologies such as AI and virtual reality strengthen engagement by making learning experiences fascinating and unique. So, the role of Artificial Intelligence (AI) is very important in social science research, since AI tools are used for analysis, simulations and interactive learning, e.g. demographic data; social dynamics enable scholars to learn about social changes [12]. AI is also a useful tool in translation and translation studies [13].

Furthermore, with the AI tool of virtual and augmented reality researchers can get familiar with their desired content based on the visual scenarios, which make their learning and analysis more concrete. As a result, using AI technologies not only brings positive influence on their analytical skills development but also promotes critical thinking skills [12, p.112]. In relation to social sciences, using AI tools provides a positive impact due to opportunities to analyze complex and dynamic social concepts and problem-solving skills. However, the huge challenges for AI in the research process are infrastructure availability, the privacy of data security and the replacement of researchers by increased role of the AI tools [12, p.120].

Conclusion

Our findings highlight the importance of integrating innovative practice-oriented software and programming tools in research focused on linguistics and the sciences. If integrated by researchers, these tools can contribute to developing their analytical skills, helping them apply their theoretical knowledge to solving practical problems.

This study contributes to the development of linguistics and social science research methodologies in Kazakhstan by demonstrating how software and programming tools can enhance the learning and scientific components of these disciplines.

The main limitation of this study is its focus on a limited set of software

and programming tools that we have found most useful and usable for linguistics and social science research. Our selection of these tools was informed by recent research in these fields, recommendations from international colleagues, and our own experience of using these tools in academic work.

In the future, a wider range of tools can be explored and tested across different scientific and educational contexts to evaluate their applicability and effectiveness more systematically. The findings of this study can inform further research in this direction focusing on the application of digital data analysis tools in linguistics, the social sciences, and other related disciplines.

REFERENCES

- [1] Holgado C. G., & Vergez-Couret M. Empowering low-resource regional languages with lexicons: A comparative study of NLP tools for morphosyntactic analysis // Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation. – 2024. – P. 5747-5756.
- [2] Derrick D., & Archambault D. TreeForm: Explaining and exploring grammar through syntax trees // Literary and linguistic computing. –2010. – №25(1). – P. 53-66.
- [3] Brezina, V. & Platt, W. #LancsBox X [software]. Lancaster University. – 2025. lancsbox.lancs.ac.uk.
- [4] Mortelmans D. Doing Qualitative Data Analysis with NVivo. Springer. – 2025. <https://library.oapen.org/handle/20.500.12657/96073>
- [5] Kuckartz U., & Rädiker S. Analyzing Qualitative Data with MAXQDA Text, Audio, and Video. Springer. – 2019. – 290 p.
- [6] Karamurugan S., & Govindarajan B. “Statistical Package for the Social Science” // International Journal of Business and Economics Research (IJBER). – 2022. – № 8(4). – P. 616-618.
- [7] William F. K. A. My Data Are Ready, How Do I Analyze Them: Navigating Data Analysis in Social Science Research // International Journal of Scientific Research and Management (IJSRM). – 2022. – № 12(03). – P. 1730-1741.
- [8] Fraga F., Poggi M., Casanova M., & Leme L. Creating automatic connections for personal knowledge management // SN Computer Science. – 2024. – № 5. doi.org/10.1007/s42979-024-02876-4
- [9] Azad N. H. An Introduction to the Use of Software and Its Impact in Educational Systems // International Journal of Advanced Studies in Humanities and Social Science. – 2024. –№13(4). – Pp. 264-272.
- [10] Khalikova N. M. Effective Use of Modern Pedagogical Software Tools in the course of the lesson // Web of Discoveries: Journal of Analysis and Inventions. – 2024. – № 2(12). – Pp. 114-116.
- [11] Malgazhdarova D. A., & Kenzhetaeva G. K. Teaching foreign language grammar through the communicative game-based project «Aengime» among

students. Bulletin of Ablai Khan KazUIR and WL. Series: Philological Sciences. – 2024. – №4(75). doi.org/10.48371/PHILS.2024.4.75.025

[12] Prasetya S.P. Artificial Intelligence in Social Sciences Education Presents New Challenges and Opportunities. Proceedings of the 4th International Conference on Social Sciences and Law (ICSSL 2024) // Advances in Social Science, Education and Humanities Research 87. –2024. doi.org/10.2991/978-2-38476-303-0_12

[13] Smagulova A. S., Muratbek N. M., & Rakhimbaeva R. M. The impact of artificial intelligence on translation: Current state and future prospects // Bulletin of Ablai Khan KazUIR and WL. Series: Philological Sciences. – 2025. – №76(1). doi.org/10.48371/PHILS.2025.1.76.029

ҚАЗАҚСТАНДА ТІЛ БІЛІМІ ЖӘНЕ ӘЛЕУМЕТТІК ҒЫЛЫМДАР САЛАСЫНДАҒЫ ЗЕРТТЕУЛЕРГЕ БАҒДАРЛАМАЛЫҚ ЖАСАҚТАМА МЕН БАҒДАРЛАМАЛАУ ҚҰРАЛДАРЫН ИНТЕГРАЦИЯЛАУ

Акишев Т.Б.¹, *Қазыбекова Ұ.М.², де Ласала Порта С.³

^{1, *,2,3}КИМЭП Университеті, Алматы, Қазақстан

Аңдатпа. Тіл білімі мен әлеуметтік ғылымдар саласындағы зерттеулерге цифрлық деректерді талдау құралдарын енгізу зерттеушілерге күрделі теориялық ұғымдарды және олардың шынайы өмірдегі қолданысын тереңірек түсінуге көмектеседі. Бұл құралдар шетелдің алдыңғы қатарлы оқу орындарындағы зерттеушілер тарапынан кеңінен қолданылып жүргенімен, Қазақстандағы лингвистика және әлеуметтік ғылымдар саласындағы зерттеулер әдетте теорияға негізделген, ал практикалық бағыт әлсіз деп бағаланады. Жергілікті зерттеушілер мен мекемелердің бұл құралдарды қолдануы теория мен практиканың арасындағы алшақтықты азайтып, бәсекеге қабілетті әрі өзекті білім мен дағдыларды ұсына алады. Осы шектеулерді жою мақсатында бұл мақалада бағдарламалық жасақтама мен бағдарламалау ресурстарын қамтитын бірқатар цифрлық құралдарға шолу ұсынылады. Ұсынылған құралдар лингвистика мен әлеуметтік ғылымдардағы цифрлық деректерді талдау мен визуализациялау үшін ғылыми негізде жасалған. Осы ғылыми негізге сүйене отырып, зерттеу бұл құралдардың жалпы және нақты сипаттамаларын, сондай-ақ оларды лингвистика мен әлеуметтік ғылымдарға бағытталған зерттеулерге енгізу жолдарын ұсынады. Қорытындысында, бұл зерттеудің теория мен тәжірибеге қосқан негізгі үлесі — жергілікті контексте бұл ғылыми құрамдастарды жетілдіруге арналған қолданбалы және өзекті ұсыныстар беруі.

Тірек сөздер: зерттеулер, лингвистика, әлеуметтік ғылымдар, философия, бағдарламалық жасақтама, бағдарламалау, құралдар, деректерді талдау, визуализация

ИНТЕГРАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНСТРУМЕНТОВ ПРОГРАММИРОВАНИЯ В ЛИНГВИСТИЧЕСКИЕ И СОЦИАЛЬНО-ГУМАНИТАРНЫЕ ИССЛЕДОВАНИЯ В КАЗАХСТАНЕ

Акишев Т.Б.¹, *Казыбекова У.М.², де Ласала Порта С.³

^{1, *2,3} Университет КИМЭП, Алматы, Казахстан

Аннотация. Интеграция цифровых инструментов анализа данных в лингвистические и социально-гуманитарные исследования помогает ученым глубже изучать языковые и социологические данные, способствуя лучшему пониманию сложных теоретических концепций и их практического применения. Несмотря на то, что такие инструменты активно используются исследователями в ведущих зарубежных учреждениях, казахстанская лингвистика и социальные науки, как правило, остаются более теоретически ориентированными и менее практико-ориентированными. Применение этих инструментов местными исследователями и учреждениями может помочь сократить разрыв между теорией и практикой, предоставив актуальные, полезные и конкурентоспособные знания и навыки. Для устранения этих ограничений данная статья предлагает обзор ряда цифровых инструментов, включающих как программное обеспечение, так и ресурсы для программирования. Представленные инструменты имеют научную основу, поскольку широко используются для анализа и визуализации данных в лингвистике и социальных науках. Основываясь на этой научной базе, в исследовании описываются как общие, так и специфические характеристики этих цифровых инструментов, а также способы их интеграции в исследования, ориентированные на лингвистику и социальные науки. В заключение, основным вкладом этого исследования в теорию и практику является предложение практически применимых и актуальных рекомендаций по совершенствованию этих научных компонентов в локальном контексте.

Ключевые слова: исследования, лингвистика, социальные науки, философия, программное обеспечение, программирование, инструменты, анализ данных, визуализация

Мақала түсті / Статья поступила / Received: 26.08.2025.

Жариялауға қабылданды / Принята к публикации / Accepted: 25.12.2025.

Information about the authors:

Timur Akishev – Assistant Professor, Doctor of Philosophy (PhD), KIMEP University, Almaty, Kazakhstan, <https://orcid.org/0009-3684-5500>, e-mail: t.akishev@kimep.kz

Ulzhan Kazybekova – Assistant Professor, Doctor of Philosophy (PhD),

KIMEP University, Almaty, Kazakhstan, <https://orcid.org/0009-0008-8698-4659>, e-mail: u.kazybekova@kimep.kz

Santiago de Lasala Porta – Assistant Professor, Doctor of Philosophy (PhD), KIMEP University, Almaty, Kazakhstan, <https://orcid.org/0009-0001-0308-1083>, e-mail: s.porta@kimep.kz

Авторлар туралы мәлімет:

Акишев Тимур Ботаұлы – ассистент-профессор, философия докторы (PhD), КИМЭП Университеті, Алматы, Қазақстан, <https://orcid.org/0009-3684-5500>, e-mail: t.akishev@kimep.kz

Қазыбекова Ұлжан Мұратқызы – ассистент-профессор, философия докторы (PhD), КИМЭП Университеті, Алматы, Қазақстан, <https://orcid.org/0009-0008-8698-4659>, e-mail: u.kazybekova@kimep.kz

Сантьяго де Ласала Порта – ассистент-профессор, философия докторы (PhD), КИМЭП Университеті, Алматы, Қазақстан, <https://orcid.org/0009-0001-0308-1083>, e-mail: s.porta@kimep.kz

Информация об авторах:

Акишев Тимур Ботаевич – ассистент-профессор, доктор философии (PhD), Университет КИМЭП, Алматы, Казахстан, <https://orcid.org/0009-3684-5500>, e-mail: t.akishev@kimep.kz

Қазыбекова Ұлжан Муратовна – ассистент-профессор, доктор философии (PhD), Университет КИМЭП, Алматы, Казахстан, <https://orcid.org/0009-0008-8698-4659>, e-mail: u.kazybekova@kimep.kz

Сантьяго де Ласала Порта – ассистент-профессор, доктор философии (PhD), КИМЭП Университеті, Алматы, Қазақстан, <https://orcid.org/0009-0001-0308-1083>, e-mail: s.porta@kimep.kz