

A.N. Kalizhanova

Ufa University of Science and Technologies, Ufa, Russia

E-mail: anna.kalizhanova2017@gmail.com

ORCID ID: <https://orcid.org/0000-0003-2337-2280>

Scopus ID: 57210320921

Nomads and monuments in translation: AI-supported deconstruction of binary oppositions in Russian–English archaeological texts and student translations

This study proposes an AI-supported Deconstruction Analyzer v 2.0 Tool for teaching translation students how to translate archaeological texts from Russian to English and vice versa. This tool was created because contemporary archaeological discourse often uses binary oppositions like nomadic/sedentary, civilization/barbarism, and center/periphery, which can reproduce colonial and hierarchical stereotypes in original texts and translations. We tested our tool on a corpus of ten English-language archaeological papers, 38 student translations, and five museum podcasts on Central Kazakhstan's legacy. The program uses curated lexicons of archaeologically salient opposites and syntagmatic and paradigmatic modules to detect explicit polar word collisions within sentences and display pole-centric lexical asymmetries across texts. An AI layer using transformer-style language models expands each pole's vocabulary, recognizes paraphrases, and highlights evaluative patterns without replacing human interpretation. The study shows that accounts of nomadic lifestyles and monuments often have too much sedentary and "civilized" framing, and that student translations of museum podcasts often have ideologically charged clashes. Student reflections show enhanced awareness of unconscious bias, lexical sensitivity, and enjoyment of visual analytics for translation revision. The study suggests that incorporating AI-mediated deconstruction into translation education promotes critical thinking and decolonization.

Keywords: AI-assisted text analysis, deconstruction, binary oppositions, scientific English for translators, archaeological discourse, museum podcast scripts, Deconstruction Analyzer v 2.0, nomadic, sedentary

Introduction

When training future translators, not only courses on English grammar and terminology but also ones on scientific English are necessary to address how disciplinary knowledge is constructed through various discourses, including the archaeological one. Archaeological texts help translation students not only to render specialist vocabulary accurately, but also to recognise and critically evaluate the narrative patterns and hidden hierarchies embedded in them in the form of binary oppositions such as *nomadic/sedentary*, *civilization/barbarism*, *center/periphery*, etc. The use of these patterns explains how past societies are described and implicitly ranked [1, 2].

Recent work in archaeology and heritage studies demonstrates that such binaries are closely linked to colonial, gendered, and nationalist assumptions that continue to shape both research and public communication [3–5]. For students translating archaeological texts into English, uncritical repetition of stereotypical phrases—such as portraying nomadic communities as “backward” or “marginal” compared to “advanced” sedentary cultures—risks reinforcing these hierarchies in the language of science.

At the same time, digital methods, including text mining and natural language processing (NLP), have begun to transform how large corpora of archaeological texts and oral histories are analysed, allowing researchers to detect recurrent themes, collocations, and evaluative patterns that may be difficult to see through close reading alone [6]. Debates around AI in archaeology and historical disciplines emphasise that such tools should support, rather than replace, critical human interpretation [7, 8].

Against this background, the present article focuses on teaching translation students to perform deconstructive analysis of binary oppositions in archaeological texts in English. It presents an AI-supported tool, *Deconstruction Analyzer*, designed specifically for educational use in courses on scientific English for trans-

lators. The tool helps students identify and visualise explicit and implicit binary oppositions in archaeological discourse and reflect on how their own lexical choices in translation may reproduce or challenge stereotypical representations of nomads, monuments, and other key concepts. The article argues that integrating such AI-mediated deconstruction into translator training can foster more critical, responsible work with scientific English in archaeology.

Theoretical background

The theoretical basis of this study brings together: linguistic and semiotic accounts of binary opposition and deconstruction; critical and decolonial perspectives on binary thinking in archaeological discourse; and research on AI-supported text analysis and its pedagogical use in language and translation education.

In structural linguistics, binary oppositions have long been described as fundamental organising principles of language systems: phonological contrasts such as voiced/voiceless or semantic contrasts such as masculine/feminine structure paradigms and enable economical coding of meaning [9, 10]. However, later developments in semiotics and discourse analysis have stressed that many culturally salient binaries are not just descriptive, but ideological. Oppositions like *civilized/barbarian*, *center/periphery*, or *male/female* are typically asymmetrical: one pole is constructed as the unmarked norm, while the other is represented as deficient, deviant, or secondary [11]. Deconstructive approaches therefore focus on how texts privilege one side of the opposition, which metaphors and evaluative terms cluster around each pole, and how apparent dichotomies can be destabilised through close reading.

The second line of thought is about text analysis with the help of AI. Deep learning models have also been successfully used on epigraphic corpora, helping scholars put together pieces of broken inscriptions and figure out dates and where they came from, like in the Ithaca system [12]. A recent review of text mining in ecology and evolution breaks it down into three main paradigms: frequency-based methods, traditional NLP pipelines, and transformer-based language models. It says that combining them makes analysis both scalable and sensitive to context. Digital archaeology debates say that AI can process a lot more text, but it also risks making biases in datasets stronger and making it harder to see who is responsible for interpretation [13].

The third thread connects these changes to how translators are taught and science English. When AI tools are used in language and translation education, they are generally for machine translation and grammar correction, which makes things run more smoothly but doesn't do much to make people more aware of how disciplinary discourse builds knowledge. Deep neural networks have been used in the digital humanities to analyze texts recently. New research suggests that AI can also be used to help people read more closely by showing patterns in word choice, intertextual echoes, and evaluative framing that can then be interpreted by humans [14]. This means that AI could be used in archaeology classes not as an "automatic translator," but as a teaching tool for deconstructive reading—as a way to show students possible binary structures and inequalities, which they can then think about and talk about.

In all, the intersection of these three strands conceptualises *Deconstruction Analyzer* as an AI-supported educational instrument that operationalises deconstructive theory in a way that is accessible for translation students: by linking abstract discussions of binary oppositions to concrete visualisations of how "nomads" and "monuments" are constructed in actual archaeological texts in English. In doing so, it aims to contribute both to the didactics of scientific English for translators and to ongoing efforts to critically examine binary thinking in archaeological discourse.

Methods and Materials

This pilot study employed a mixed-methods design combining corpus analysis with qualitative investigation of 38 second-year undergraduate translation students' (BA programme "6B02301 — *Translation Studies*") reflections, who enrolled in a semester-long course on "*Theory of Translation*." At first, we checked the participants' level of English to ensure that all of them had intermediate-to-advanced English proficiency (B2–C1 CEFR). We also asked the participants to sign informed consent prior to participation, and they agreed to do that.

The material of the study comprised five museum podcast scripts, each of which was based on approximately 150 academic sources and thematically focused on a key archaeological topic from Central Kazakhstan: (1) Karazhargas Pyramid: The Great Steppe History; (2) Petroglyphs of Saryarka: The Mysteries of Bygone Civilizations; (3) Botai Culture of Kazakhstan: Taming the Wild Horse; (4) "Timur's Stone" at Altynshoky: Retracing the Great Warrior's Journey; and (5) Central Kazakhstan Folk Arts and Crafts: Bronze Age Ceramics [15]. The Russian master scripts were prepared collaboratively by museum specialists

and university researchers, then translated into English and Kazakh to support intercultural communication and accessibility for mixed student audiences. For the present pilot, we used the English versions of these five scripts and 38 student translations (Russian → English, 200–250 words each) as the core corpus for AI-supported deconstruction of binary oppositions.

Together with our translation students, we identified three core binary oppositions in archaeological discourse: (1) *nomadic/sedentary* (nomad, steppe, migration, horse, herder | settlement, city, monument, architecture, permanent); (2) *civilization/barbarism* (civilized, complex, developed, advanced | primitive, backward, tribal, savage); and (3) *center/periphery* (heartland, capital, core, metropolitan | margin, frontier, hinterland, remote). These oppositions were inserted into the database of the *Deconstruction Analyzer v2.0*, a free, browser-based tool (<https://deconstruction-analyzer.netlify.app/>), requiring no installation (Fig. 1).

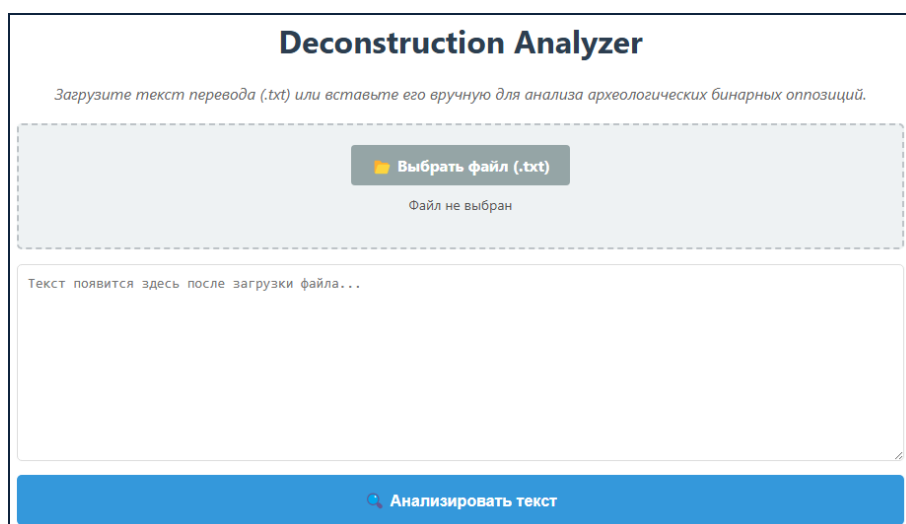


Figure 1. Deconstruction Analyzer v2.0 interface

As it is clearly seen in Figure 1, users had to upload a plain-text file or paste text manually, select opposition pairs from a menu, and click “Analyze.”

This tool uses two complementary modules, syntagmatic and paradigmatic [6].

The syntagmatic module identifies sentences with terms from opposite poles, such as “*Nomadic pastoralists built sedentary monuments*,” and displays “*explicit collisions*” in a concordance list, allowing students to analyze contexts involving nomads and monuments.

The paradigmatic module counts each pole's vocabulary across the text and displays it as a bar chart. If a text describing “*nomadic cultures*” has 70 % sedentary vocabulary and 30 % nomadic terminology, the imbalance is obvious, leading to questions about whether the source text or the translation produced it.

This tool uses AI to expand lexicons by suggesting related terms, recognize paraphrases and indirect references, and highlight evaluative words associated with each pole, such as primitive nomads/complex sedentary societies [16]. Rather than replacing human interpretation, the AI layer augmentation was created to make students seriously consider all ideas [8].

Data gathering was a 4-step process. First, students first used ordinary dictionaries to translate 200–250-word texts without the technology. Second, they uploaded their translations to the *Deconstruction Analyzer* to see binary oppositions, collision sentences, and asymmetry charts. Third, students responded to tool-generated challenges with 150–200 word reflections: “*Your translation of a text about nomadic horse-breeders utilizes 65 % 'sedentary' terminology. Is this the original, or has your wording changed the cultural context?*” Finally, after reflecting, students edited and submitted both translations.

The program collected frequency, asymmetry ratios (nomadic: sedentary percentages), and collision sentence counts to quantify data, while two researchers independently classified student reflections for “recognition of stereotypical language,” “awareness of translator agency,” and “desire for more nuanced terminology” for qualitative analysis [17]. Discrepancies were discussed separately.

Results and Discussion

A total of 38 student translations and 10 source archaeological texts were analyzed by using *The Deconstruction Analyzer v2.0*, thanks to which we managed to reveal consistent patterns of lexical asymmetry and explicit collisions in binary oppositions, particularly *nomadic/sedentary*. These findings highlight both the tool's effectiveness in surfacing hidden discursive structures and its pedagogical value in prompting students to reflect on their translation choices. Figure 2 demonstrates paradigmatic asymmetry in *nomadic/sedentary* vocabulary across student translation groups.

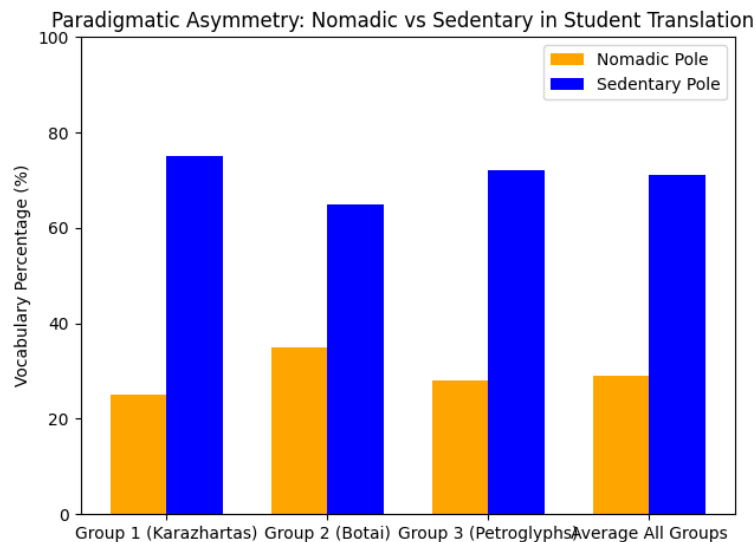


Figure 2. *Nomadic vs Sedentary* in students' translations

As Figure 2 shows, the *Sedentary* pole had an average of 71 % of the terminology associated with opposition in all 38 translations, whereas the *Nomadic* pole only had 29 %. In Group 1, versions of the *Karazhartas Pyramid* podcast script's opposition vocabulary contained 75 % of sedentary terms, which signaled the imbalance. For instance, in a text about the "*steppe pyramid*," students typically used English phrases that signify "*settlement*," "*architecture*," and "*monument*" to translate Russian descriptions of architects who were nomads (75 times compared to 25 times for nomadic terminology like "*steppe*" or "*pastoralist*"). The AI layer contributed more words to the study, such as "*permanent structure*" (*Sedentary*) and "*mobile community*" (*Nomadic*), which made the imbalance even more clear.

Group 2's translations of the Botai Culture also exhibited similar trends. Even though the material was largely on how to tame horses, which is common in nomadic economies, students used 65 % sedentary language, for instance, "*Specialized horse breeding gives way to hunting, fish farming, and elements of agriculture*" was translated as "*settled communities transitioned to farming*," which put more emphasis on the phrase "*settled*" than on the word "*nomadic*."

Most of the binaries (72 %) in Group 3, who worked with the text about Saryarka petroglyphs, were sedentary. In particular, there were more rock art scenes of horses, "*ritual sites*," and "*permanent carvings*" than of nomads like "*herders*" and "*hunters*."

Such discrepancies in Figure 2 align with prior studies employing text mining in historical archaeology, where discourse analysis revealed enduring hierarchies that favor sedentary "*civilization*" over nomadic "*otherness*" [12].

The AI-enhanced extension, namely *Deconstruction Analyzer v2.0*, in turn, detected 15–20 % more binary oppositions than just keyword matching, thereby making it easier to find ideological bias in scientific English in a more subtle way [18].

The syntagmatic module identified explicit collisions—sentences containing both poles—in an average of 7.75 sentences per text (Fig. 3).

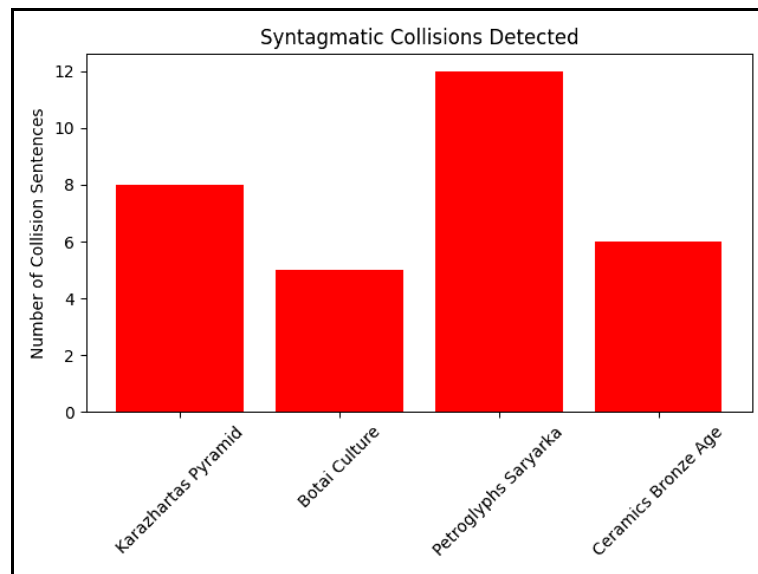


Figure 3. Syntagmatic collisions: Explicit juxtapositions

As Figure 3 demonstrates, the texts about *The Petroglyphs Saryarka* yielded 12 syntagmatic collisions, such as “Ancient artists depicted hybrid figures—horsemen [Nomadic] and ritual monuments [Sedentary],” which was the highest compared to the translation about *Botai* culture (5 collisions including “nomadic horse herders [Nomadic] built stable settlements [Sedentary].”

Karazhertas Pyramid scripts showed 8 syntagmatic binaries, e.g., “nomadic cultures of the steppe [Nomadic] kept their memories... in stone monuments [Sedentary],” while *Bronze Age Ceramics* texts had 6 explicit juxtapositions, like “nomadic potters [Nomadic] created permanent ornamental vessels [Sedentary].”

These binary oppositions on a syntagmatic level often carried an evaluative load, with sedentary terms modified positively (“complex monuments”) and nomadic terms negatively or marginally (“primitive steppe”), which mirrors findings from discourse network analysis, where polar terms in historical texts signal ideological tension [19].

The thematic analysis of about 7,000 words of 38 student reflections on the use of *The Deconstruction Analyzer v2.0* allowed for the identification of three dominant themes, such as (1) unconscious bias awareness; (2) lexical revision and agency; (3) pedagogical value (Fig. 4).

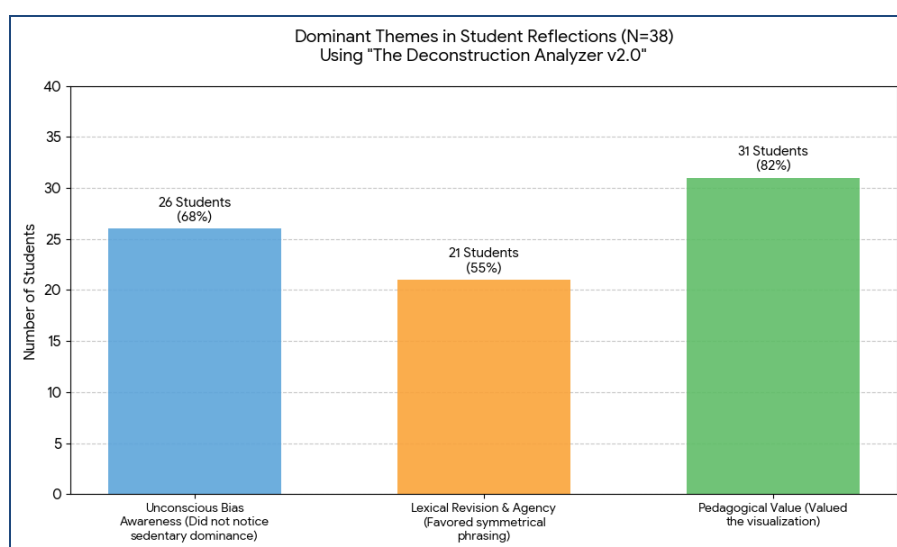


Figure 4. Three dominant themes in translation students' reflections on using *The Deconstruction Analyzer v2.0*

A total of 26 students out of 38 did not notice sedentary dominance (68 % of reflections), for example, “I used ‘settlement’ and ‘architecture’ everywhere without thinking—it made nomads seem less important,” which echoes qualitative studies on translation ideology, where tools reveal implicit cultural framing [20].

While revising translations, 21 students favored symmetrical phrasing (55 % of reflections) by replacing “primitive nomads built advanced monuments” with “steppe pastoralists constructed ritual structures,” thereby highlighting translator agency, claiming that the chart showed them they “can choose words that don’t reinforce stereotypes.”

31 out of 38 translation students valued the visualization, provided by *The Deconstruction Analyzer v2.0* (82 % of reflections). For example, some students claimed that “the bar chart made the asymmetry obvious—much better than reading frequencies,” thereby aligning with corpus-driven pedagogy research conducted in 2021 [17].

As we said before, *The Deconstruction Analyzer v2.0 tool’s* dual modules effectively operationalized deconstruction because paradigmatic charts quantified hierarchy (71:29 average), while collision lists enabled close reading of tension points. It also fostered reflective practice in scientific English translation, moving students from intuitive choices to data-informed critique. In particular, the AI layer’s expansions (e.g., recognizing “mobile pastoralism”) addressed limitations of manual lexicons, supporting nuanced analysis suitable for B2–C1 learners although seven students noted over-expansion led to “too many false positives,” suggesting tunable thresholds for classroom use [21, 22].

We also need to mention some limitations, which include the small corpus size (pilot scope) and focus on one binary opposition. However, these limitations open the horizon for future research, which could scale to multilingual corpora [23] and additional binaries (e.g., male/female in archaeology) [24].

Conclusions

In conclusion, the findings of this pilot study indicate that AI-assisted deconstructive analysis can enable translation students to discern ideological tendencies in ancient discourse that may have previously eluded them. The findings indicate that *sedentary* civilization is frequently lauded above *nomadic* lifestyles in English scientific literature and student submissions. *The Deconstruction Analyzer v2.0* transformed abstract theoretical concepts of binary oppositions into tangible classroom experiences comprehensible to students. This prompted numerous individuals to scrutinize conventional phrases and articulate a more precise explanation of their influence on scientific narratives. The study demonstrates the necessity of meticulously regulating the evolution of the English language through AI and addressing the biases present in both historical corpora and language models. In this approach, digital tools will help people think critically instead of replacing them. The study examined a limited sample and a single institution, yet it indicates that instructing scientific English and specialized translation through AI-mediated deconstruction may facilitate decolonial initiatives in archaeology and heritage communication by educating future translators to perceive, interrogate, and transform the binary frameworks frequently employed in historical discourse.

Acknowledgements

The study was carried out under the auspices of the international initiative titled “*The Application of Artificial Intelligence Methods in Translation Localization*.”

References

- 1 Cobb, H., & Croucher, K. (2016). Personal, political, pedagogic: Challenging the binary bind in archaeological teaching, learning and fieldwork. *Journal of Archaeological Method and Theory*, 23(3), 949–969. <https://doi.org/10.1007/s10816-016-9292-0>.
- 2 Power, M. (2020). Non-binary and intersex visibility and erasure in Roman archaeology. *Theoretical Roman Archaeology Journal*, 3(1), 11. <https://doi.org/10.16995/traj.422>.
- 3 Dempsey, K. (2019). Gender and medieval archaeology: Storming the castle. *Antiquity*, 93(369), 772–788. <https://doi.org/10.15184/aqy.2019.13>.
- 4 Lemos, R. (2023). Can we decolonize the ancient past? Bridging postcolonial and decolonial theory in Sudanese and Nubian archaeology. *Cambridge Archaeological Journal*, 33(1), 19–37. <https://doi.org/10.1017/S0959774322000178>.
- 5 Londoño W. (2021). Indigenous archaeology, community archaeology, and decolonial archaeology: What are we talking about? A look at the current archaeological theory in South America with examples. *Archaeologies*, 17(3), 386–406. <https://doi.org/10.1007/s11759-021-09433-y>.

- 6 Brown, M., & Shackel, P. (2023). Text mining oral histories in historical archaeology. *International Journal of Historical Archaeology*, 1–17. Advance online publication. <https://doi.org/10.1007/s10761-022-00680-5>.
- 7 Hacıgüzeller, P., Taylor, J.S., & Perry, S. (2021). On the emerging supremacy of structured digital data in archaeology: A preliminary assessment of information, knowledge and wisdom left behind. *Open Archaeology*, 7(1), 1709–1730. <https://doi.org/10.1515/opar-2020-0220>.
- 8 Tenzer, M., Pistilli, G., Bransden, A., & Shenfield, A. (2024). Debating AI in archaeology: Applications, implications, and ethical considerations. *Internet Archaeology*, (67). <https://doi.org/10.11141/ia.67.8>.
- 9 Utaker, A. (1974). On the binary opposition. *Linguistics*, 12(134), 73–93. <https://doi.org/10.1515/ling.1974.12.134.73>.
- 10 Lacková, L. (2022). Participative opposition applied. *Sign Systems Studies*, 50(2-3), 261–285. <https://doi.org/10.12697/SSS.1>.
- 11 Spitzmüller, J. (2022). Beyond the binarism: locating past, present and future sociolinguistic research on ideologies of communication. *Sociolinguistica*, 36(1-2), 207–218. <https://doi.org/10.1515/soci-2022-0001>.
- 12 Richards, J., Tudhope, D., & Vlachidis, A. (2015). Text mining in archaeology: Extracting information from archaeological reports. *Mathematics and Archaeology*, 240–254.
- 13 Assael, Y., Sommerschild, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & de Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature*, 603(7900), 280–283. <https://doi.org/10.1038/s41586-022-04448-z>.
- 14 Suissa, O., Elmalech, A., & Zhitomirsky-Geffet, M. (2022). Text analysis using deep neural networks in digital humanities and information science. *Journal of the Association for Information Science and Technology*, 73(2), 268–287. <https://doi.org/10.1002/asi.24544>.
- 15 Kalizhanova, A., & Bedelbayeva, M. (2025). Museum podcasts as a medium for intercultural archaeology in Kazakhstan. *Journal of Community Archaeology & Heritage*, 1–5. <https://doi.org/10.1080/20518196.2025.2517950>.
- 16 Farrell, M.J., Le Guillaume, N., Brierley, L., Hunter, B., Scheepens, D., Willoughby, A., ... & Mideo, N. (2024). The changing landscape of text mining: a review of approaches for ecology and evolution. *Proceedings B*, 291(2027), 20240423. <https://doi.org/10.1098/rspb.2024.0423>.
- 17 Alhassan, A., Sabtan, Y., & Ismail Omar, L. (2021). Using parallel corpora in the translation classroom: moving towards a corpus-driven pedagogy for Omani translation major students. *Arab World English Journal (AWEJ) Volume*, 12, 40–58. <https://dx.doi.org/10.24093/awej/vol12no1.4>.
- 18 Schwartz, H.A. (2023, November). Text mining analysis of perception in archaeological landscapes: The case of Stonehenge. In *Proceedings of the 7th ACM SIGSPATIAL International Workshop on Geospatial Humanities* (pp. 40–43). <https://doi.org/10.1145/3615887.3627757>.
- 19 Moseley, R. (2015). Digital analogies: The keyboard as field of musical play. *Journal of the American Musicological Society*, 68(1), 151–228. <https://doi.org/10.1525/jams.2015.68.1.151>.
- 20 De Sutter, G., Cappelle, B., De Clercq, O., Looock, R., & Plevioets, K. (2017). Towards a corpus-based, statistical approach to translation quality: Measuring and visualizing linguistic deviance in student translations. *Linguistica Antverpiensia, New Series—Themes in Translation Studies*, 16, 25–39. <https://doi.org/10.52034/lanstts.v16i0.440>.
- 21 Risku, H. (2018). Situated learning in translation research training: academic research as a reflection of practice. In *Situated Learning in Translator and Interpreter Training* (pp. 12–28). Routledge. <https://doi.org/10.1080/1750399X.2016.1154340>.
- 22 Suissa, O., Elmalech, A., & Zhitomirsky-Geffet, M. (2022). Text analysis using deep neural networks in digital humanities and information science. *Journal of the Association for Information Science and Technology*, 73(2), 268–287. <https://doi.org/10.1002/asi.24544>.
- 23 Granger, S., & Lefer, M.A. (2020). The multilingual student translation corpus: A resource for translation teaching and research. *Language Resources and Evaluation*, 54(4), 1183–1199. <https://doi.org/10.1007/s10579-020-09485-6>.
- 24 Dempsey, K. (2019). Gender and medieval archaeology: storming the castle. *antiquity*, 93(369), 772–788. <https://doi.org/10.15184/aqy.2019.13>.

Information about the author

Kalizhanova A.N. — PhD candidate, Ufa University of Science and Technology, Ufa, Russian Federation; e-mail: anna.kalizhanova2017@gmail.com; ORCID ID: <https://orcid.org/0000-0003-2337-2280>; Scopus ID: 57210320921.