Proceedings of the 8th International
Scientific and Practical Conference
«Progressive Science and Achievements»
(September 6-8, 2026).
Dallas, USA

No
308

LITERARY STUDIES

Peculiarities and Adaptation of Big Data Technologies in Contemporary Translation Studies under High-turbulence Environment and Systemic Risk

Svitlana Krasniuk[1]

[1] Senior Lecturer, Department of philology and translation
Kyiv National University of Technologies and Design; Ukraine

Abstract. The article explores the specifics, features and directions of adaptation of Big Data technologies in modern translation studies in conditions of uncertainty, turbulence and systemic risks of the external environment. The possibilities of using big linguistic data for the analysis of corpora, terminology, translation quality and improvement of machine translation are determined. The prospects of integrating Big Data with artificial intelligence, machine learning and expert technologies for the formation of adaptive hybrid systems of modern translation studies are substantiated.
Keywords: *modern linguistics, computational linguistics, Big Data, translation studies, digitalization, big language data, machine translation, artificial intelligence, machine learning, parallel corpuses, text mining, hybrid technologies, uncertainty, turbulence, systemic risks.*

INTRODUCTION.

Modern humanities are forced to adapt to complex environmental conditions [1], characterized by the dynamics of uncertainty, turbulence and systemic risks [2] (which conceptually corresponds to the SHIVA-world). Digitalization [3], the growth of information flows and the complication of socio-cultural processes require the updating of research tools and the transition to adaptive, intellectual and data-driven approaches.

Modern philology (as an important component of the humanities) is also currently developing in conditions of rapid complication of the information space [4], the growth of digital text resources [5], the acceleration of interlingual communication and constant environmental change. For translation studies, these processes are of particular

importance, since modern translation is increasingly carried out in a digital environment [6], which is characterized by scale, dynamism, multi-channel and high heterogeneity of language data [7]. At the same time, uncertainty, turbulence and systemic risks of the external environment directly affect the maintenance, organization and technological tools of translation activities.

One of the most promising areas of technological adaptation of modern translation studies is Big Data technologies, which provide collection, storage, processing, integration and intellectual analysis of significant arrays of various types of data [8]. Their application allows us to move from mainly local analysis of individual texts and corpora to systematic research of large-scale dynamic arrays of translated and untranslated texts, terminological resources, parallel corpora, data, metadata, machine translation results and other digital language objects [9].

The relevance of the study is due to the fact that classical approaches to translation studies are not always able to work effectively with constantly increasing volumes of dynamic data. In these conditions, Big Data acts not simply as a set of technical tools, but as the basis for the formation of a new analytical contour of modern translation studies, combining corpus linguistics, computational linguistics, artificial intelligence, machine learning, neural network technology and intelligent natural language processing [10]. Considering the above, the purpose of this article is to systematically determine the specifics, features and main directions of adapting Big Data technologies in modern translation studies, taking into account the complex and complex conditions of the dynamics of uncertainty, turbulence and systemic risks of the external environment of modern philology.

MAIN PART.

A). Transformation of modern translation studies under the influence of digitalization and Big Data.

The digital transformation of translation studies involves a qualitative transition from a predominantly text-centric research model to a data-driven approach, within which a significant part of scientific and practical solutions are based on the processing of large-scale digital data sets. While traditional translation studies focused mainly on individual texts, language pairs, authorial strategies and comparative analysis of translations, the modern digital

LITERARY STUDIES

paradigm allows us to study the patterns of translation activity based on millions of text and metatext units.

Big Data expands the analytical capabilities of translation studies by combining various sources: parallel and comparative corpora, electronic libraries, translation databases, term banks, translation memories, machine translation results, subtitles, multimedia resources, digital communications and other language arrays.

Thus, Big Data becomes a factor in the methodological transformation of translation studies, ensuring the transition from single observation to large-scale detection of patterns, from static analysis to dynamic monitoring and mainly expert interpretation to a combination of expert and algorithmic analysis.

B). Specificity of linguistic Big Data in modern translation studies.

Linguistic Big Data has an important specificity compared to data in traditional applied areas. They are characterized not only by large volume, but also by high variability, heterogeneity, context dependence and constant dynamics.

For translation studies, the classical characteristics of Big Data - Volume, Velocity, Variety, Veracity and Value - are especially significant. A large volume of data is associated with the scale of modern digital corpora; high speed - with the continuous emergence of new texts and updating of language resources; diversity - with the coexistence of artistic, scientific, technical, legal, media texts and digital communications; reliability - with the need for quality control and representativeness of data; value - with the possibility of extracting new linguistic and translation knowledge from data sets.

Additional specificity is created by multilingualism of data. Translation Big Data must take into account differences in language systems, cultures, genres, registers, terminological systems and communicative norms. Therefore, simple scaling of data processing methods does not guarantee an increase in translation quality: algorithms must take into account the linguistic and extralinguistic context.

C). Main areas of application of Big Data in translation studies.

One of the most important areas is the formation and intellectual analysis of parallel and comparable corpora. Large-scale corpora allow to identify stable translation correspondences, transformations, patterns of choice of

LITERARY STUDIES

lexical and syntactic constructions, features of terminology transfer and differences between professional and machine translations.

The second area is related to terminological analysis. Big Data allows to automatically detect new terms, track changes in their use, establish contextual correspondences and form dynamic terminological bases. This is especially important for rapidly developing scientific and technical fields, in which terminology changes faster than traditional dictionaries and reference resources.

The third direction is the analysis of translation quality. Based on large arrays of translated texts, it is possible to identify typical errors, post-editing patterns, differences between translation options, and factors affecting the quality of the result.

The fourth direction is related to machine and neural network translation. Large multilingual data sets provide training and improvement of machine translation models, and also allow them to be tested in various genre, subject, and language conditions.

The fifth direction is the study of translation behavior. The analysis of digital traces of translation activity can be used to study the temporal parameters of work, the repeatability of translation decisions, the nature of editorial interventions, the use of translation memory, and human interaction with intelligent systems.

D). Big Data as a tool for adapting translation studies to uncertainty and turbulence.

In conditions of high environmental uncertainty, the importance of the translation system's ability to quickly adapt to changes in linguistic, technological, socio-cultural and professional conditions increases fundamentally. Big Data creates an information basis.

Dynamic monitoring of large language arrays allows for timely recording of the emergence of new terms, changes in word meanings, transformation of discursive practices, the spread of new communicative models and changes in translation requirements. Thus, translation studies gets the opportunity to move from a reactive model to a more proactive one.

In conditions of turbulence, the combination of Big Data with machine learning and artificial intelligence technologies becomes especially important. Such systems are able to detect weak signals and hidden patterns that are difficult to detect using traditional analysis. However, the

LITERARY STUDIES

final interpretation of the detected patterns must retain an expert and humanitarian component, since statistical correlation does not always correspond to linguistic or translation causality.

E). Systemic risks and limitations of using Big Data in translation studies.

Despite significant advantages, the adaptation of Big Data creates new risks. The key problem is the quality of the source data. A large amount of low-quality, erroneous or unrepresentative data can lead to systematic distortions of the analysis results.

No less important are the risks of algorithmic bias, when the models used reproduce errors and stereotypes contained in the training data. For multilingual translation studies, this is especially significant, since languages are represented in the digital space extremely unevenly.

Additional problems are related to confidentiality, copyright, data security, transparency of algorithms, reproducibility of results and the dependence of research processes on a specific digital infrastructure.

Therefore, Big Data should not be considered as a universal replacement for professional translation and philological analysis. It is advisable to understand them as an intellectual and technological expansion of the capabilities of the researcher and translator.

F). Adaptive model of Big Data use in modern translation studies.

Rational adaptation of Big Data involves the formation of a multi-level system: data collection → cleaning and structuring → integration → intelligent analysis → interpretation → expert verification → making a translation decision → feedback and updating data.

Of particular importance is the closed nature of such a system. The obtained translation result should become a source of new data for further analysis. Thanks to this, an adaptive circuit is formed in which the translation system is able to constantly learn from new materials and correct its own models.

A hybrid architecture that combines Big Data, corpus technologies, machine learning, neural network models, intelligent text analysis and expert knowledge of the translator is promising. Such an architecture allows combining the computational scalability of digital technologies with the contextual, cultural and interpretive

competence of a person.

G). Prospects for hybridization of Big Data and intelligent technologies in translation studies.

The most promising direction of further development is the formation of hybrid intelligent technologies of translation studies, in which Big Data is integrated with artificial intelligence, machine learning, neural network models, semantic analysis, knowledge graph technologies and expert systems [11].

Such hybridization can ensure the creation of adaptive translation platforms that can take into account the context, subject area, genre, terminology, communicative situation and individual characteristics of the translator. At the same time, what is promising is not the displacement of humans by algorithms, but the formation of a Human-AI collaboration model, in which computing technologies provide large-scale analysis and decision support, and humans retain the functions of interpretation, critical assessment and final choice.

For modern philology in general and translation studies in particular, this means a gradual transition to an intellectual-analytical model based on the constant interaction of humans, data and algorithms. In conditions of dynamic uncertainty, such a model can increase the adaptability of translation research and practice, reduce the processing time of large information arrays, and simultaneously expand the possibilities of identifying new linguistic and translation patterns.

CONCLUSIONS.

The analysis shows that Big Data technologies are becoming one of the key factors in the digital transformation of modern translation studies. Their significance is defined as the ability to process huge amounts of information, but above all, the ability to form a fundamentally new data-driven contour of research and organization of translation activities.

The specificity of Big Data in translation studies is determined by multilingualism, heterogeneity, dynamism, context dependence and varying degrees of quality of language data. Therefore, effective adaptation of these technologies requires not a mechanical transfer of universal algorithms, but their linguistic, cognitive and professional translation adaptation.

The main areas of application of Big Data are the formation and analysis of parallel corpora, intellectual work

LITERARY STUDIES

with terminology, assessment of translation quality, improvement of machine translation, identification of translation patterns and research of digital models of professional behavior of translators.

In difficult conditions of uncertainty, turbulence and systemic risks, Big Data acquires additional value as a tool for constant monitoring of the dynamics of the language environment. They allow for timely detection of new linguistic phenomena, changes in terminology, transformation of discourses and the emergence of new communicative practices, thereby ensuring increased adaptability of translation studies.

At the same time, the use of Big Data is accompanied by risks of low data quality, algorithmic bias, violation of confidentiality, copyright issues, insufficient interpretability of models and a possible decrease in the role of professional expertise. Therefore, the optimal model is not a technologically deterministic one, but a hybrid model that combines Big Data, artificial intelligence, machine learning and expert knowledge of the translator.

The prospects for further development are associated with the formation of adaptive intelligent translation systems capable of functioning in a mode of continuous learning and updating. As a result, Big Data should be considered not only as a technological resource of modern translation studies, but also as the basis for its transition to a more dynamic, adaptive and intellectual-analytical paradigm that meets the new conditions for the development of modern philology.

References:

[1] De Luca, E. W., Fallucchi, F., Ghattas, B., & Spielhaus, R. (2024). The digital transformation processes for supporting digital humanities researchers in text analysis. Journal of Documentation, 80(2), 378-391. https://doi.org/10.1108/JD-07-2022-0143

[2] Zhang, Z., Song, W., & Liu, P. (2024). Making and interpreting: Digital humanities as embodied action. Humanities and Social Sciences Communications, 11, Article 13. https://doi.org/10.1057/s41599-023-02548-3

[3] Krasnyuk, M., Kulynych, Y., Tuhaienko, V., & Krasniuk, S. (2022). E-business and ecommerce technologies as an important factor for economic efficiency and stability in the modern conditions of the digital economy. Grail of Science, 2022, №. 17, pp. 69-81. https://archive.journal-grail.science/index.php/2710-3056/article/view/412

LITERARY STUDIES

[4] Blasi, D. E., & Roberts, S. G. (2018). Studying language evolution in the age of big data. Journal of Language Evolution, 3(2), 94-129. https://doi.org/10.1093/jole/lzy004
[5] Goncharenko S. (2025). BIG Data in modern literary studies. Science: Development and Factors its Influence : Proceedings of the 6th International Scientific and Practical Conference (October 6-8, 2025; Amsterdam, Netherlands). - Amsterdam: Scientific Collection «InterConf», 2025. - P. 82-85.
[6] Kaplan, F. (2015). A map for Big Data research in digital humanities. Frontiers in Digital Humanities, 2, Article 1. https://doi.org/10.3389/fdigh.2015.00001
[7] Goncharenko S. (2025). Big semi-structured data & deep ANN in computational linguistics. Science and Global Challenges in the Modern World : Proceedings of the 2nd International Scientific Conference (Leicester, United Kingdom, 5 October 2025). - Lulu Press, Inc., 2025. - pp. 133-136.
[8] Krasnyuk, M., Elishis, D. (2024). Perspectives and problems of big data analysis & analytics for effective marketing of tourism industry. Science and Technology Today, 4(32). https://doi.org/10.52058/2786-6025- 2024-4(32)-833-857
[9] Ganascia, J.-G. (2015). The logic of the Big Data turn in digital literary studies. Frontiers in Digital Humanities, 2, Article 7. https://doi.org/10.3389/fdigh.2015.00007
[10] Baayen, R. H. (2024). The wompom. Corpus Linguistics and Linguistic Theory, 20(3), 615-648. https://doi.org/10.1515/cllt-2024-0053