
[Book Review] Machine Learning in Translation

Peng Wang and David B. Sawyer (eds.). *Machine Learning in Translation*
Routledge, 2023. 218 pp. Paperback £ 39.99.
ISBN: 9781032323800

Hao Yin
(The Hong Kong Polytechnic University)
Jianwen Liu
(Hong Kong Shue Yan University)

Yin, H., & Liu, J. (2026). [Book Review] Machine learning in translation. *Asia Pacific Journal of Corpus Research*, 7(1), 25-28.

This review examines *Machine Learning in Translation*, edited by Peng Wang and David B. Sawyer, as an accessible, human-centered introduction to the integration of machine learning into translation theory, practice, and education. Organized into three parts, the book connects human and machine approaches to translation, explains key technologies and tasks, including neural machine translation, quality assessment, and natural language processing, and highlights the central role of language data in translator-computer interaction. Particular attention is given to translators' intentionality, model customization and personalization, machine learning literacy, and evolving professional competences. The book's main strength lies in presenting machine learning as a continuation of linguistic, communicative, and cognitive approaches rather than merely an engineering domain. However, its predominantly conceptual orientation, limited empirical evidence, emphasis on resource-rich languages, and brief treatment of ethics and implementation restrict its scope. Overall, the volume provides a valuable roadmap for translation practitioners, educators, and researchers navigating increasingly data-driven, human-machine translation environments and future interdisciplinary collaboration.

Keywords: Machine Learning, Translation Technology, Translator-Computer Interaction, Translator Education, Human-Centered Translation

The recent surge in machine learning has significantly impacted diverse interdisciplinary fields (Carleo et al., 2019; Greener et al., 2022), particularly influencing machine translation (Yang et al., 2020). *Machine Learning in Translation* is an insightful and comprehensive guide that merges human linguistic expertise with machine learning technologies in the field of translation. Adopting a human-centric approach to Machine Learning in Translation (MLinTrans), this book stands out for its accessibility and practicality. Tailored specifically for translation practitioners, linguists, and stakeholders in the translation field, it equips readers with relevant knowledge and skills essential for MLinTrans, offering a comprehensive view of the theoretical foundations, practical applications, and human-centered strategies needed to integrate machine learning into translation workflows effectively. This inclusive perspective encompasses machine learning applications in the fields of document analysis, language data management, and various translation-related Natural Language Processing (NLP) tasks.

The book presents a novel perspective on translation, moving away from conventional approaches

such as statistical machine translation (Koehn, 2009) and established paradigms in machine translation education (Bhattacharyya, 2015). It explores translation through the lens of machine learning. It situates MLinTrans not as an independent discipline approached solely from a computer programming and engineering standpoint but rather as a seamless continuation of linguistic, communicative, and cognitive approaches to translation. The book is organized into a carefully structured three-part framework, each consisting of three chapters. This structure effectively guides readers through foundational translation concepts, key machine learning techniques, and the crucial role of data in shaping modern translation workflows.

In the inaugural section (Part I) of the book, the authors systematically analyze both human and machine approaches to translation. Chapter 1 presents a compelling analysis of the intersection between traditional human translation practices and machine learning, highlighting their convergence from linguistic, communicative, and mathematical perspectives. By exploring how words can be represented as vectors in a continuous space, it explores how machine learning models, particularly neural networks, are revolutionizing translation. Ultimately, it sets the stage for a broader exploration of how these two approaches can be integrated to enhance translation processes, suggesting that while they differ in methodology, they share a fundamental goal: to bridge language gaps and facilitate communication.

Chapter 2 focuses on the role of linguistic structure and levels of analysis in both human and machine approaches to translation. It explores the importance of segmentation in translation tasks, highlighting the distinction between form, meaning, and context. The chapter explains how machine segmentation, which relies on both mathematical calculations and human perception of language, enables more effective language processing. By analyzing the different linguistic units and their context, the chapter outlines how various levels of analysis—such as word, sentence, and discourse—are necessary for both human translators and machine translation systems to process and understand language effectively.

Chapter 3 conducts a meticulous examination of predicative language models, showing how both humans and machines rely on prediction in language comprehension and production. By linking cognitive insights from translation and interpreting studies with computational approaches, the authors demonstrate how predictive modeling underpins key NLP tasks such as translation, summarization, and speech recognition. This chapter effectively illustrates how prediction serves as a bridge between human cognitive processes and machine learning techniques in translation.

Part II introduces crucial machine learning tasks in translation, spanning Chapters 4 to 6. Chapter 4 offers a lucid, holistic account of machine translation within the wider localisation ecosystem, mapping its evolution from rule-based machine translation (RBMT) through statistical machine translation (SMT) to neural machine translation (NMT). It clarifies NMT as a conditioned neural language model and relates encoder-decoder architectures (and later attention/Transformer advances) to interlingua-style representations. A notable strength is the integration of machine translation with CAT (Computer-Assisted Translation) workflows, addressing practical challenges such as APIs, post-editing, and terminology/TM integration, showing how human-machine interaction influences translation quality. The discussion is largely descriptive but insightful on practical trade-offs—data, sequence length, content type, and user input—underscoring machine translation as an instantiation of human translation processes rather than a replacement.

Chapter 5 examines machine translation quality assessment, emphasizing both automated and human-centered approaches. It explains that relying solely on human evaluation is often slow, inconsistent, and costly, and thus insufficient for handling the scale of MT output. The chapter introduces widely used automated metrics, such as BLEU (Bilingual Evaluation Understudy) and TER/HTER (Translation Edit Rate / Human-targeted Translation Edit Rate), which quantify translation quality by comparing MT output with human references. In addition, it highlights the role of MQM (Multidimensional Quality Metrics) as a customizable framework for categorizing translation errors,

and discusses Quality Estimation (QE), a set of machine learning methods that evaluate MT output without requiring reference translations, by analyzing linguistic and statistical features of the source text and translation output. Together, these approaches underscore the importance of combining automation with human expertise to ensure both efficiency and accuracy in translation evaluation.

Chapter 6 examines the role of intentionality in translation and how it shapes the use of NLP tasks within machine learning–driven workflows. The chapter emphasizes that translators manage a range of NLP tasks rather than the core MT technology itself, highlighting two categories: facilitation tasks, which assist productivity, and intelligent tasks, in which translators not only use tools but also provide data (e.g., corrections or annotations) to retrain and improve models. It also explains that these tasks are situated within a continuous translation workflow, from content extraction to final publishing, underscoring the centrality of human intentionality. Ultimately, the chapter argues that while machines execute formal operations, intentionality—the purpose-driven direction of tasks—remains a uniquely human attribute, guiding how NLP tools are selected, customized, and integrated in translation practice.

In Part III (Chapters 7-9), the focus shifts to the indispensable role of data in both human learning and machine learning processes within translation. Chapter 7 explores translator–computer interaction (TCI) through the lens of language data, highlighting their dual role in human analysis and machine learning. It introduces *translation-driven language data*, identifying comparability as a defining feature (extended from translation-driven corpora). TCI is shown through three lenses—data design and acquisition, data annotation, and data use—demonstrating how translators’ expertise complements machine analytics. By stressing the two states of language data and the importance of meaningful human intervention, the chapter underscores translators’ central role in shaping ML-driven translation workflows with attention to fairness, robustness, and safety.

Chapter 8 highlights the parallels between human and machine learning, emphasizing shared mechanisms such as statistical inference and parameter adjustment. It introduces the concept of translators’ mental multilingual language models, linking individual cognitive processes with machine learning approaches. The chapter distinguishes between customization for collective purposes and personalization for individual learning, showing how translators can directly engage with data and models. Practical examples, including terminological schematic context, source-text-oriented comparable corpora, and intelligent tutoring systems, illustrate how personalized machine learning tools can simultaneously enhance translation efficiency and human learning.

Chapter 9 examines how machine learning reshapes translator education by redefining professional roles and competences. Translators increasingly shift from translation proper toward translation-related linguistic tasks, requiring new skills in data management, evaluation, and human–computer collaboration. The chapter highlights two overarching competences—ML literacy and continuous learning—alongside task-specific linguistic, cultural, and NLP competences. From an educational perspective, it contrasts top-down, educator-guided and bottom-up, learner-centered environments, emphasizing the importance of personalization, human–computer interaction, and the enduring value of educators in guiding meaningful human learning in an ML-driven context.

While *Machine Learning in Translation* provides a valuable and well-structured overview of the field, certain limitations should be noted. The book is largely conceptual in nature, with relatively few empirical studies or detailed case examples to illustrate its arguments. Its focus leans toward English and resource-rich languages, leaving low-resource and multimodal contexts less fully explored. Discussions of practical implementation issues—such as data management, workflow integration, and translator workload—remain somewhat limited. Ethical considerations like bias, fairness, and privacy are mentioned but not extensively developed. Similarly, the pedagogical recommendations could be further supported by empirical evidence from translator education practice.

Looking ahead, research on machine learning in translation would benefit from a stronger focus on interdisciplinary collaboration, bringing together linguists, computer scientists, and educators to

design more human-centered tools and workflows. Particular promise lies in developing personalized ML systems that align more closely with translators' cognitive processes, thereby enhancing both creativity and efficiency. Future work could also explore innovative pedagogical models that integrate machine learning into translator training, equipping students with the competences required for emerging professional roles. Finally, investigating new forms of human-machine interaction may open avenues for more adaptive and context-sensitive translation support.

In conclusion, *Machine Learning in Translation* offers valuable insights into the intersection of machine learning and translation, highlighting the evolving role of translators and the potential of personalized machine learning tools. It successfully bridges human learning and machine capabilities, providing a roadmap for both educators and researchers to navigate the future of translation technology.

References

- Bhattacharyya, P. (2015). *Machine Translation*. Boca Raton, FL: CRC Press.
- Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., & Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4), 045002.
- Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23(1), 40-55.
- Koehn, P. (2009). *Statistical Machine Translation*. Cambridge: Cambridge University Press.
- Yang, S., Wang, Y., & Chu, X. (2020). A survey of deep learning techniques for neural machine translation. *arXiv*. <https://arxiv.org/abs/2002.07526>

THE AUTHORS

Hao Yin (Anna) is a PhD student in the Department of Language Science and Technology at The Hong Kong Polytechnic University, Hong Kong, China.

Jianwen Liu (Kacey) is an Associate Professor in the Department of English Language and Literature at Hong Kong Shue Yan University, Hong Kong, China.

THE AUTHORS' ADDRESSES

First Author

Hao Yin (Anna)
PhD Student
The Department of Language Science and Technology
The Hong Kong Polytechnic University, Hong Kong, China
Email: annayinhao@gmail.com

Corresponding Author

Jianwen Liu (Kacey)
Associate Professor
Department of English Language and Literature
Hong Kong Shue Yan University, Hong Kong, China
Email: jwliu@hksyu.edu

Received: 13 March 2026

Received in Revised Form: 21 July 2026

Accepted: 20 August 2026