BRIDGE
Trends & Traditions in Translation & Interpreting Studies

Introduction: Toward human-centred AI in translation studies

Marián Kabát
Comenius University in Bratislava
marian.kabat@uniba.sk

Kabát, Marián. 2025. Introduction: Toward human-centred AI in translation studies. In: *Bridge: Trends and Traditions in Translation and Interpreting Studies*. Vol. 6, No. 1: pp. 1-7.

Artificial intelligence (AI) is now woven into translation in ways that were scarcely imaginable a decade ago. Neural machine translation (NMT) built on the Transformer architecture (Vaswani et al., 2017) has altered expectations of speed, coverage and quality, while increasingly multimodal, massively multilingual systems promise "any-to-any" translation across text and speech. Yet if translation studies is to harness these advances responsibly, we need to place people (translators, interpreters, localizers, educators, communities, and end-users) at the centre of system design, policy, and evaluation. This is the core proposition of human-centred AI (HCAI): design AI that amplifies human agency, mastery, and accountability, rather than displacing them (Shneiderman, 2022, for HCAI in translation studies see Jiménez-Crespo, 2025).

HCAI is not a vague ideal. In human–computer interaction (HCI), concrete guidance exists for how AI systems should behave in practice (e.g., the "Guidelines for Human-AI Interaction," which articulate principles like keeping users in control, supporting efficient correction, and exposing system confidence). These guidelines are directly relevant to translation tools, from interactive MT to quality estimation dashboards and post-editing environments. At the same time, broader regulatory frameworks now impose obligations on providers and deployers of AI. In the European Union, for example, the AI Act entered into force on 1 August 2024 with phased applicability through 2025–2027, including near-term prohibitions and obligations for general-purpose AI and later rules for high-risk systems –

developments with clear implications for translation workflows and vendors.

This issue on Bridge takes up the challenge of human-centred AI in translation by foregrounding translators' expertise and lived practices, domain-specific quality, equity for under-resourced languages and modalities, and educational and regulatory transitions. Below is a sketch of the technological backdrop, and a map of key human-centred concerns.

1. The technological backdrop: from sequence models to multimodal, multilingual platforms

The last decade of MT has been defined by three intertwined developments. First, end-to-end neural MT (e.g., Google's GNMT) demonstrated large quality gains over phrase-based systems and catalysed industrial deployment at scale (Wu et al., 2016). Second, subword modelling (e.g., byte-pair encoding) tackled open-vocabulary translation and proved essential for robust generalization (Sennrich et al., 2016). Third, the Transformer made parallelizable attention-only architectures the default for translation and, later, for large language models (LLMs) that now power many "MT-like" experiences.

Recent work has pushed breadth (hundreds of languages) and modality (text ↔ speech ↔ speech). Meta's No Language Left Behind (NLLB-200) and SeamlessM4T/Seamless families, for instance, combine extensive data mining, multilingual training, and safety evaluation (toxicity, bias) to deliver cross-domain coverage across ~100–200 languages, including many that are traditionally under-served. At the same time, LLMs (e.g., GPT-4) increasingly perform translation as part of broader text generation and understanding, raising new questions about evaluation, authorship, and control.

For translation studies, the point is not merely that systems are better. It is that the design space has expanded: interactive and adaptive MT, predictive typing, terminology conditioning, streaming speech-to-speech, and integrated post-editing environments are now realistic. That space must be navigated with human roles and requirements in view.

2. Human roles, work practices, and standards

Human-centred AI starts with an accurate account of what translators actually do, under what constraints, and with what forms of craft knowledge. Post-editing (PE) is a prime example. A large body of research reports productivity gains for PE over from-scratch translation, but findings vary with domain, MT quality, translator profile, and task setup. A recent multi-year, multi-pair study shows that PE is "usually but not always" faster than human translation, warns that averages can be misleading, and cautions against using simple edit distance as a proxy for effort, which are precisely the kinds of nuances human-centred evaluation should surface (Terribile, 2024).

Human-centred design also means aligning tools and processes with professional standards. Two ISO standards are particularly relevant: ISO 17100 (translation services requirements) and ISO 18587 (post-editing of MT output). These specify role definitions, competence, and process requirements (e.g., revision), historically with a human translator "in the loop", and they continue to be touchstones for procurement and quality assurance as AI support expands. Complementary human-centred design guidance is formalized in ISO 9241-210, which codifies user-centred processes across the lifecycle, e.g., requirements analysis, participatory design, and iterative evaluation principles that translation tool builders should adopt explicitly.

Finally, the education pipeline is adapting. The European Master's in Translation (EMT) Competence Framework 2022 updates expectations for digital, technological and domain competences, providing a common reference for programs training translators to work in human-AI workflows.

3. Interaction design: from post-editing to mixed-initiative tools

Human-centred AI is, at heart, interaction design. Translation research has long explored interactive MT and mixed-initiative systems that share control between human and model. Beyond classic post-editing, systems such as predictive translation memory and interactive MT interfaces reduce cognitive friction by offering context-aware completions, signals of confidence, or quick ways to enforce terminology and style. Experimental comparisons of post-editing versus interactive MT illuminate trade-offs between human effort and model learnability, which are insights that speak directly to the "keep the user in control" and "support efficient correction" guidelines in HCI (Green et al., 2014).

To be truly human-centred, such tools should be co-designed with translators, surface transparent rationales (what changed and why), make model uncertainty explicit, and respect ergonomics (e.g., reducing unnecessary mode-switching and visual clutter). These are not luxuries: they are conditions for sustainable adoption and well-being in everyday work.

4. Quality, evaluation, and the meaning of "parity"

A human-centred perspective reframes translation quality as situated and purpose-relative. Industry-standard frameworks such as MQM (Multidimensional Quality Metrics) and TAUS DQF formalize error categories and link evaluation methods to communicative goals, enabling fine-grained, task-appropriate judgments. At the same time, Quality Estimation (QE), i.e., predicting quality without references, now complements human review and routing decisions in production.

Debates about "human parity" underscore why evaluation must be human-centred. When evaluation shifts from isolated sentences to document-level judgments that account for discourse coherence and pragmatics, human translations remain preferred – a reminder that context matters in both measurement and design.

A practical implication is that human-centred evaluation should: (i) use task-appropriate metrics (e.g., MQM categories aligned to content risk), (ii) integrate translator-centric measures (effort, cognitive load, satisfaction), and (iii) exploit QE judiciously to triage content, while validating model scores against human outcomes rather than treating them as ground truth (Burchardt et al., 2013).

5. Fairness, inclusion, and the long tail of languages

Human-centred AI must address bias and inclusion. Gender bias studies (e.g., WinoMT) exposed systematic errors in MT systems for gender-inflected languages and catalysed mitigation work; such audits are essential for any deployment touching identity or safety-critical domains. At a macro level, massively multilingual efforts like NLLB explicitly target low-resource languages, pairing model/dataset scaling with human evaluation and toxicity checks; this is a concrete step toward equitable access but not a solution by itself. Sustained investment in community-driven data governance, participatory evaluation, and localization infrastructures remain necessary.

6. Governance and accountability

Regulation is moving rapidly. The EU AI Act establishes a risk-based regime with phased obligations, from prohibited practices and GPAI transparency in 2025 to high-risk system requirements by 2026–2027. For translation, consequences include documentation and risk management for AI-enhanced CAT/PE tools (e.g., data provenance, bias/robustness testing, human oversight mechanisms), and clarity about provider versus deployer responsibilities in supply chains spanning LSPs, platform vendors, and clients (European Commission, 2024; European Parliament Research Service, 2025).

Governance also has a workforce dimension: confidentiality in client data used for model adaptation, authorship and remuneration in AI-assisted deliverables, and labour conditions on platformised marketplaces. Human-centred AI requires institutional arrangements, e.g., contracts, audit trails, and recourse mechanisms, that keep expertise and accountability visible.

7. A research agenda for human-centred AI in translation

There are at least five strands where translation studies can lead:

- Participatory, value-sensitive toolmaking. Involve translators and stakeholders throughout requirements gathering, prototyping, and evaluation; apply user-centred standards (ISO 9241-210) and HAI guidelines explicitly; document design rationales and trade-offs.
- Ergonomics and well-being at scale. Measure cognitive effort beyond speed: eye-tracking, keystroke logging, subjective workload, satisfaction; study long-horizon effects of PE/interactive MT on quality, fatigue, and professional identity.
- Context-aware evaluation. Advance document-level human evaluation aligned with MQM/DQF; triangulate with QE to route content and decide where human attention is most valuable.
- Equity and inclusion. Co-develop datasets, benchmarks, and interfaces with speakers of under-represented languages and modalities; operationalize fairness auditing (e.g., gender, dialect) as part of deployment checklists.
- Education and transition. Use competence frameworks (e.g., EMT 2022) to redesign curricula around human-AI collaboration: prompt-aware terminology work, critical QE literacy, post-editing strategies, and ethical/data governance.

8. Conclusion

Human-centred AI reframes the role of technology in translation: not as an autonomous replacement for expert labour, but as a set of augmentations that can make translators more effective, keep quality within human-defined bounds, and expand access across languages and modalities. Achieving this requires tight coupling between technical progress (multilingual and multimodal modelling, robust interaction design), professional norms (standards, translator training), evaluation practices that respect human judgments and context, and governance that safeguards rights and accountability. The articles gathered in this issue speak to different pieces of that puzzle. Taken together, they make a case for a translation field that not only uses AI, but helps to shape it around people.

References:

Burchardt, Arle; Lommel, Aljoscha and Popović, Maja, et al. 2013. Multidimensional quality metrics: A flexible system for declaring translation quality metrics. In: *Translating and the Computer.* 35.

European Commission. 2024. *AI Act enters into force.* https://commission.europa.eu/news-and-media/news/ai-act-enters-force-2024-08-01_en?utm_source=chatgpt.com. Accessed on: 15 December 2025.

European Parliament Research Service. 2025. AI Act implementation timeline. https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/772906/EPRS_ATA%282025%29772906_EN.pdf?utm_source=chatgpt.com. Accessed on: 15 December 2025.

Green, Spence; Wang, Sida; Chuang, Jason; Heer, Jeffrey; Schuster, Sebastian and Manning, Christopher D. 2014. Human Effort and Machine Learnability in Computer-Aided Translation. In: *EMNLP.* pp. 1225-1236.

ISO 17100:2015. *Translation services – Requirements for translation services.* International Organization for Standardization.

ISO 18587:2017. *Translation services – Post-editing of machine translation output – Requirements.* International Organization for Standardization.

ISO 9241-210:2019. *Ergonomics of human-system interaction – Human-centred design for interactive systems.* International Organization for Standardization.

Jiménez-Crespo, Miguel A. 2025. Human-centeredness in Translation: Advancing Translation Studies in a human-centered AI era. In: *INContext: Studies in Translation and Interculturalism.* 5(1). https://doi.org/10.54754/incontext.v5i1.132.

Meta AI (NLLB Team). 2022. No Language Left Behind: Scaling Human-Centered Machine Translation. In: *arXiv*:2207.04672.

Meta AI. 2023–2024. SeamlessM4T / Seamless: Massively Multilingual & Multimodal Machine Translation. In: *arXiv*:2308.11596; *arXiv*:2312.05187.

Sennrich, Rico; Haddow, Barry and Birch, Alexandra. 2016. *Neural Machine Translation of Rare Words with Subword Units.* ACL.

Shneiderman, Ben. 2022. *Human-Centered AI.* Oxford University Press.

TAUS. 2025. *Dynamic Quality Framework (DQF) resources and error typologies.*

Vaswani, A., et al. 2017. Attention Is All You Need. In: *NeurIPS.*

Wu, Y. et al. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. In: *arXiv*:1609.08144.

European Commission, DGT/EMT. 2022. *EMT Competence Framework 2022.*

Terribile, Silvia. 2024. Is post-editing really faster than human translation? In: *Translation Spaces.* 13(2). DOI: https://doi.org/10.1075/ts.22044.ter.