ISSN:3030-332X
Impact factor: 9,493
International journal of scientific researchers

ARTIFICIAL INTELLIGENCE IN TRANSLATION: BENEFITS, CONVENIENCES, RISKS, AND LIMITATIONS — A CRITICAL REVIEW

Sultonova Nodirabonu Rustamjon qizi

FerSU teacher

Nodirabonusultonova07@gmail.com

Abstract: Artificial intelligence has transformed translation from a slow, specialised craft into an instantaneous, ubiquitous service embedded in everyday digital life. The shift from statistical to neural machine translation, and subsequently to large language models, has produced fluency that rivals professional human output on certain tasks while simultaneously introducing new categories of error and risk. This article offers a critical review of the scholarly evidence on AI-driven translation, organised around four dimensions: its measurable benefits, its practical conveniences, its harms and limitations, and the strategies proposed to manage them. Drawing on machine-translation evaluation studies, productivity research and work on algorithmic bias, the review argues that headline claims of human parity are methodologically fragile, that quality gains are unevenly distributed across languages and domains, and that fluency can mask serious failures of adequacy, bias and factual fidelity. It concludes that AI translation is most safely understood not as a replacement for human judgement but as a powerful instrument whose value depends on informed, supervised use.

Keywords: artificial intelligence; machine translation; neural machine translation; large language models; translation quality; algorithmic bias; post-editing

1. Introduction

Translation has long stood among the most demanding intellectual tasks, requiring not only command of two languages but sensitivity to context, register and culture. For most of the twentieth century, automating it remained an aspiration pursued in research laboratories with limited practical success. The situation changed decisively over the past decade. The introduction of neural architectures, followed by the public arrival of large language models, has pushed automatic translation to a level of fluency that would have seemed implausible a generation ago, and the technology is now consulted billions of times each day through browsers, mobile applications and office software.

The scale of this change has been accompanied by extraordinary claims. When Google deployed its neural translation system, its developers reported that the gap between machine and human translation had narrowed dramatically, with output approaching the accuracy of average bilingual translators on some test sets [1]. Shortly afterwards, a widely publicised study asserted that machine translation had reached human parity on a Chinese-to-English news task [2]. Such announcements captured public attention and fuelled expectations that professional translation might soon be rendered obsolete.

Yet the same period produced a more cautious counter-literature, questioning how parity was measured and documenting persistent weaknesses hidden beneath fluent surfaces. This article reviews both currents. Its aim is to provide a balanced account of what AI translation now does well, where it remains convenient but imperfect, and where its limitations carry real consequences. The discussion is organised around four questions: what benefits the technology delivers, what conveniences explain its mass adoption, what harms and limitations qualify its use, and how those risks can responsibly be managed.

The review is narrative rather than systematic in form, drawing selectively on three bodies of scholarship: machine-translation evaluation research, which measures output quality against human references; productivity studies, which assess the technology in professional workflows; and work on algorithmic bias and fairness, which examines its social effects. These literatures are rarely read together, yet they are complementary, because the value of a translation tool cannot be judged by fluency alone but only by the joint consideration of quality, usability and consequence. By synthesising them, the article

seeks to replace the binary framing of machines versus humans with a more useful question about the conditions under which automated translation can be trusted.

2. From Statistical to Neural and Generative Translation

The capabilities under discussion are the product of a rapid succession of technical paradigms. Statistical machine translation, dominant from the 1990s, assembled translations from probabilities learned over large bilingual corpora; it produced usable but often disjointed output. Neural machine translation, which became the field standard in the mid-2010s, models entire sentences through deep neural networks and generates markedly more idiomatic prose, outperforming earlier phrase-based approaches across many language pairs and shared tasks [3].

The most recent transition involves large language models, which perform translation as one ability among many rather than as a dedicated function. Because they are trained on vast multilingual data and can be prompted conversationally, these systems handle context, tone and clarification requests in ways earlier engines could not. A comprehensive benchmark of one such model against human translators found that, unlike traditional neural systems whose quality drops sharply for resource-poor language directions, the model maintained consistent performance across all evaluated pairs, achieving quality on a par with junior to mid-level human translators [6]. This versatility underlies both the promise and the unpredictability of the current generation of tools.

3. Benefits of AI-Driven Translation

3.1. A genuine leap in quality

The first and most consequential benefit is the sheer improvement in output quality. Neural and generative systems produce fluent, grammatically coherent text that, for many everyday purposes, conveys meaning accurately and naturally. On certain constrained tasks the quality has been judged comparable to that of professional translators [2], and large language models have narrowed the historic gap between high- and low-resource languages, delivering more even performance across diverse language directions than their predecessors [6]. For users who once had no access to translation at all, this represents a substantive gain rather than a marginal one.

3.2. Productivity and post-editing gains

A second, well-documented benefit lies in professional productivity. Rather than translating from scratch, many translators now post-edit machine output, correcting and refining a generated draft. Controlled experiments have repeatedly shown that this workflow increases throughput: studies with professional translators report productivity gains in the region of thirteen to twenty-five per cent compared with unaided translation [9]. A large cross-language study confirmed that neural post-editing drives significant productivity improvements across typologically diverse target languages, although the magnitude of the gain depends heavily on the language and the underlying system [10]. AI translation, in this sense, functions less as a replacement than as an accelerator of skilled human work.

3.3. Consistency, scale and tireless operation

AI systems also offer advantages that humans structurally cannot match. They operate continuously without fatigue, maintain terminological consistency across very large volumes, and scale to quantities of text no human team could process. A benchmarking study noted pointedly that, whatever its weaknesses, the machine translator does not suffer from the fatigue or hallucination-by-exhaustion that affects human translators over long tasks [6]. For organisations handling high-volume, repetitive or time-critical material, these properties translate directly into reduced cost and faster delivery.

3.4. Access, cost, and inclusion

A further benefit, easily overlooked in technical evaluations, is the radical lowering of cost and barriers to access. Professional human translation is expensive and comparatively scarce, placing it

beyond the reach of most individuals and many institutions. Free, instantaneous AI translation makes cross-lingual communication available to populations that previously had none: a speaker of a less widely studied language can read foreign scholarship, a migrant can interpret an official letter, and a small enterprise can correspond across borders without a translation budget. For speakers of languages that large models have begun to cover more evenly, the inclusion effect is significant, since the historic concentration of quality in a few high-resource pairs is gradually giving way to broader, if still uneven, coverage [6]. Viewed in this light, AI translation is not merely a convenience for the well-resourced but a partial democratisation of access to information.

4. Conveniences That Drive Mass Adoption

Beyond measurable performance, the everyday conveniences of AI translation explain its pervasive use. The service is instantaneous: a sentence, a document or a webpage can be rendered into another language in seconds, at any hour, without appointment or cost. It is multilingual at a scale no individual could rival, covering scores or hundreds of language pairs from a single interface. It is embedded directly into the tools people already use, from search engines and messaging apps to word processors, so that translation has become an ambient feature of digital life rather than a separate task.

Generative systems add a further layer of convenience by accepting natural-language instructions. A user can ask not merely for a translation but for a more formal version, a simpler one, or an explanation of why a particular phrase was chosen, turning a static tool into an interactive assistant. For travellers, learners, researchers reading foreign-language sources, and migrants navigating unfamiliar bureaucracies, this immediacy lowers barriers that were once formidable. The convenience is real, and it is precisely this frictionlessness that makes critical awareness of the technology's limits so necessary.

5. Harms, Risks, and Limitations

5.1. Overstated parity and the fluency trap

The most basic risk is mistaking fluency for accuracy. Because modern output reads smoothly, errors are easy to overlook. When researchers re-examined the celebrated human-parity claims using more rigorous protocols, they found that human evaluators preferred genuine human translation over machine output once whole documents, rather than isolated sentences, were assessed, exposing failures of cohesion and consistency that sentence-level testing concealed [3]. A separate reassessment showed that parity findings were partly artefacts of flawed evaluation design, including the use of non-professional raters and degraded reference texts [4]. The lesson is that headline quality claims must be read with caution: surface fluency can mask substantive defects.

5.2. Uneven quality across languages and domains

Quality is also distributed unequally. Foundational analyses identified persistent challenges for neural translation, including poor performance on low-resource languages, difficulty with rare words, degradation on long sentences, and weak handling of out-of-domain text [5]. Even the strongest contemporary models retain characteristic faults: a detailed benchmark found that a leading large language model tended toward overly literal translation and lexical inconsistency, and remained comparatively weak on grammar and named entities [6]. For specialised material such as legal, medical or literary texts, where precision and nuance are decisive, these limitations can render automatic output unsuitable without expert revision.

5.3. Bias and the reproduction of stereotypes

A more insidious harm is the reproduction of social bias. Because these systems learn from human-generated data, they absorb and amplify the prejudices encoded within it. Examining translation between English and languages that rely on gender-neutral pronouns, researchers found that systems systematically defaulted to gender stereotypes, assigning occupations and social roles along conventional lines and frequently selecting masculine forms for prestigious or powerful positions [7]. A broader

survey of the field confirms that the problem has carried over into the large-language-model era: trained predominantly on English-centric data, contemporary systems continue to favour masculine defaults and to mishandle gender-neutral content, so that the appearance of improved fluency does not entail improved fairness [8]. Such bias is not a peripheral glitch but a structural feature of statistically trained translation, with real consequences for how individuals and groups are represented. The harm is amplified precisely because fluent output invites trust: a biased rendering that reads naturally is less likely to be questioned than an obviously clumsy one.

5.4. Factual infidelity: omission and hallucination

Closely related to the fluency trap is a class of failures in which the output departs from the source not through awkwardness but through silent distortion. Neural systems are known to omit portions of difficult text, to add material not present in the original, or to substitute a confident but incorrect rendering when they encounter ambiguity or unfamiliar content. The foundational survey of neural-translation challenges already flagged unreliable behaviour on out-of-domain and low-frequency input, where models may produce plausible-sounding output untethered from the source meaning [5]. Because such errors are fluent, they are particularly dangerous: a dropped negation, a reversed quantity or an invented clause can pass unnoticed by a reader who does not consult the original. In high-stakes contexts this transforms a convenience into a liability.

5.5. Over-reliance, confidentiality, and professional displacement

Finally, the convenience of AI translation generates risks of its own. Habitual reliance can erode users' own linguistic competence and critical attention, encouraging acceptance of output that is fluent but wrong. Feeding sensitive documents into cloud-based services raises genuine concerns about data confidentiality and privacy, particularly for legal, medical and commercial material, since text submitted to a public service may be stored, processed or used for further training in ways the user cannot control. And the economics of automation place pressure on the translation profession itself, reshaping skilled translation into lower-paid post-editing and raising questions about the long-term sustainability of the human expertise that the technology still depends upon for quality control. There is a quiet paradox here: the systems that promise to make human translators redundant continue to rely on human translation, both as the training data from which they learn and as the corrective layer that keeps their output reliable. Eroding that human capacity too quickly may therefore undermine the very resource on which the quality of automated translation ultimately rests.

6. Managing the Risks: Toward Responsible Use

None of these limitations argues for rejecting AI translation, but together they argue strongly for supervised, informed use. The dominant safeguard in professional practice is the human-in-the-loop model, in which machine output is treated as a draft to be verified rather than a finished product. Post-editing not only recovers productivity gains but interposes human judgement at the point where adequacy, bias and confidentiality risks are greatest [9; 10]. Importantly, the evidence shows that post-editing high-quality output raises productivity while editing poor output can be slower than translating afresh, so the value of the workflow depends on matching the tool to the task.

Three further principles follow from the literature. First, evaluation should be rigorous and document-level, since sentence-level metrics systematically overstate quality [3; 4]. Second, high-stakes domains demand mandatory expert review and, where confidentiality matters, secure or on-premises systems rather than public services. Third, addressing bias requires deliberate intervention, since left unattended these systems reproduce and amplify the inequities in their training data [7; 8]. Responsible use, in short, treats AI translation as an instrument that augments rather than replaces human competence.

These principles apply with particular force to specialised fields. In legal and medical translation, where a single mistranslated clause or dosage can have severe consequences, automatic output should be treated as a starting draft subject to obligatory verification by a qualified human, never as a final product. In literary and marketing translation, where voice, connotation and cultural resonance carry the meaning,

the literalness and lexical flattening characteristic of current systems make unedited output inadequate on its own [6]. Recognising that the appropriate level of automation varies by domain is itself a form of risk management: the same tool that is entirely sufficient for grasping the gist of a foreign news article may be wholly unsuitable for a contract or a poem. Matching the degree of human oversight to the stakes of the text is therefore the practical heart of responsible deployment.

7. Conclusion

Artificial intelligence has made translation faster, cheaper and more accessible than at any point in history, and these gains are genuine rather than illusory. AI systems produce fluent output across a remarkable range of languages, raise the productivity of professional translators, and operate at a scale and speed beyond human reach. Yet the same review reveals that claims of human parity rest on fragile evaluation, that quality remains uneven across languages and domains, and that fluency can conceal serious failures of adequacy, factual fidelity and fairness. The convenience that drives mass adoption is also what makes uncritical trust hazardous.

The balanced conclusion is neither rejection nor uncritical embrace. AI translation is best understood as a powerful but partial instrument whose benefits are fully realised only under informed human supervision. As the technology continues to advance, the central task is not to ask whether machines can translate, but to ensure that their growing capability is matched by a corresponding maturity in how, and where, we choose to rely on it. Future work should prioritise robust document-level evaluation, the mitigation of bias, and the protection of confidentiality, so that the undeniable advantages of automated translation are not undermined by its quieter risks.

References

1. Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., et al. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv:1609.08144.
2. Hassan, H., Aue, A., Chen, C., Chowdhary, V., Clark, J., et al. (2018). Achieving human parity on automatic Chinese to English news translation. arXiv:1803.05567.
3. Läubli, S., Sennrich, R., & Volk, M. (2018). Has machine translation achieved human parity? A case for document-level evaluation. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 4791–4796). Association for Computational Linguistics.
4. Toral, A., Castilho, S., Hu, K., & Way, A. (2018). Attaining the unattainable? Reassessing claims of human parity in neural machine translation. In Proceedings of the Third Conference on Machine Translation (WMT). Association for Computational Linguistics.
5. Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. In Proceedings of the First Workshop on Neural Machine Translation (pp. 28–39). Association for Computational Linguistics.
6. Yan, J., Yan, P., Chen, Y., Li, J., Zhu, X., & Zhang, Y. (2024). Benchmarking GPT-4 against human translators: A comprehensive evaluation across languages, domains, and expertise levels. arXiv:2411.13775.
7. Ghosh, S., & Caliskan, A. (2023). ChatGPT perpetuates gender bias in machine translation and ignores non-gendered pronouns: Findings across Bengali and five other low-resource languages. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (AIES). https://doi.org/10.1145/3600211.3604672
8. Vanmassenhove, E. (2024). Gender bias in machine translation and the era of large language models. arXiv:2401.10016.
9. Guerberof Arenas, A. (2009). Productivity and quality in the post-editing of outputs from translation memories and machine translation. In Proceedings of MT Summit XII. International Association for Machine Translation.