Wisdom Journal For Studies And Research Volume 05 Issue 06- 20/11/2025

ISSN print/ 2769-1926 ISSN online/ 2769-1934

Artificial and Human Translation: A comparison of accuracy and style

Ari Dakheel Abdal Dasi
College of Law, University of Al-Hamdaniya, Iraq
Aridasi@uohamdaniya.edu.iq

Received:26/10/2025 Accepted : :15/11/2025

Abstract:

Artificial Intelligence has profoundly reshaped the translation industry by enabling unparalleled speed, scalability, and high-cost efficiency. With the rise of state-of-the-art neural machine translation models like Google Translate, neural machine translation (NMT), DeepL, OpenAI's GPT series, and others, large amounts of text can be translated in a matter of seconds, in what would normally take hours or days of human effort. The approach compares the AI and human translations from different domains (literary, legal, and technical). Literary texts favour originality, cadence, and emotional beat. Legal writing is formal and precise, and the forms that words take must be followed, from language to language. Technical writing requires consistency, precision, and clarity. Initial findings indicate that AI has several clear advantages in its technical structure, and its capacity to handle massive datasets can lead to consistency and terminological accuracy.

Keywords: Artificial Intelligence, Human Translation, Cultural Adaptation, and Cognitive Sciences.

Wisdom Journal For Studies And Research Volume 05 Issue 06- 20/11/2025

ISSN print/ 2769-1926 ISSN online/ 2769-1934

الترجمة الاصطناعية والبشرية: مقارنة بين الدقة والأسلوب

اري دخيل عبدال داسي

كلية القانون جامعة الحمدانية

Aridasi@uohamdaniya.edu.iq

ملخص

لقد غيّر الذكاء الاصطناعي صناعة الترجمة جذريًا، من خلال تمكينه سرعةً فائقةً وقابليةً للتوسع وكفاءةً عاليةً في التكلفة. مع ظهور نماذج الترجمة الآلية العصبية المتطورة، مثل جوجل ترانسليت، والترجمة الآلية العصبية (NMT)، وDeepL، Google translation وسلسلة OpenAI's GPT، وغيرها، يُمكن ترجمة كمياتٍ هائلةٍ من النصوص في ثوانٍ معدودة، في وقتٍ كان يستغرق عادةً ساعاتٍ أو أيامًا من الجهد البشري. يقارن هذا النهج بين ترجمات الذكاء الاصطناعي والترجمات البشرية من مجالاتٍ مختلفة (الأدبية والقانونية والتقنية). تُفضّل النصوص الأدبية الأصالة والإيقاع والإيقاع العاطفي. الكتابة القانونية رسميةٌ ودقيقة، ويجب اتباع أشكال الكلمات من لغةٍ إلى أخرى. تتطلب الكتابة التقنية الاتساق والدقة والوضوح. تُشير النتائج الأولية إلى أن الذكاء الاصطناعي يتمتع بالعديد من المزايا الواضحة في بنيته التقنية، وأن قدرته على التعامل مع مجموعات البيانات الضخمة يُمكن أن تُؤدي إلى الاتساق ودقة المصطلحات .

الكلمات المفتاحية: الذكاء الاصطناعي، الترجمة البشرية، التكيف الثقافي، والعلوم المعرفية.

ISSN print/ 2769-1926 ISSN online/ 2769-1934

Introduction

For centuries, translation has played a vital role in cross-cultural communication, underpinning the exchange of ideas, the dissemination of information, and cultural awareness (Hennawy et al., 2020). From time immemorial, translation has been a critical tools to transfer the knowledge and culture from the early scribes who translated religious texts to those who transmitted Greek philosophy into Arabic and then into Latin. In the globalising world of today, 'translation' is needed more than ever for understanding each other in a shrinking and connected planet (Hennawy et al., 2020). At the heart of this process, however, are human translators who bring their extensive linguistic expertise, contextual knowledge, and cultural understanding to bear on the complex terrain of language (Massey, 2021). Human translators, unlike machines, can also infer meaning from context and have the ability to grasp nuances of style, as well as retain metaphors and rhetorical devices that convey not only meaning but also cultural and emotional weight. For instance, a translator dealing with a metaphorical passage in a literary text should not only convey the meaning but also convey its tone, imagery, and symbolic value to the target culture's audience (Massey, 2021). Equivalence is something that is not enough for a reviewer to check, because, especially in the case of specialised legal terminology, a legal translator has to move to something beyond the mere surface, to make sure he is precise in terms. He complies with both source and target legal systems. All these features underscore the irreplaceable function of human reason and creativity in translation.

The rapid ascent of AI-based translation solutions, such as neural machine translation (NMT), Google Translate, DeepL, and giant language models like GPT-4, has transformed the translation landscape beyond recognition over the past decade. Such systems can handle vast amounts of text incredibly quickly and provide instantaneous communication in multiple languages for both individual and institutional users (Lee, 2023). Their strengths include speed and the ability to maintain consistent terminology in long or technical documents, which can be challenging even for experienced human translators. Accordingly, AI-powered software has been widely accepted in business, governmental, and daily scenarios and has also been applied to translation workflows (Lysychenko, 2024).

ISSN print/ 2769-1926 ISSN online/ 2769-1934

AI translation still has limitations, underscoring the ongoing need for human translators.

There are several common difficulties, as shown below:

- **Idiomatic and Metaphorical Translation Directly**: Often, AI directly translates idiomatic language literally, which can yield the correct results at a syntactic level but lacks actual meaning. For instance, many languages translate the English idiom "kick the bucket" literally, devoid of any intended sense of death (Setyawan et al., 2025).

- **Flattening of emotional tone and rhetorical nuance**: AI often struggles to translate text with the emotional force of a text, particularly where a human translator can adjust register, rhythm, or style to mitigate neutralising effects. This is particularly challenging for tasks such as literature or marketing, where emotion plays a key role (Elkins, 2024a).

- **Cultural adaptation and audience-specific rendering are currently restricted**: the deep cultural insight required to localise text for particular audiences automatically is missing from AI. For example, a human translator may find an appropriate equivalent for a culturally bound metaphor in a metaphor that is familiar to the target audience. In contrast, AI is likely to translate any culturally bound metaphor as literally as the source text, resulting in a sense of distance and non-comprehension (Prabhakaran et al., 2022a).

AI is good at processing language in the same way it makes cars, yet it is weak at the interpretive and creative acts that translate a distinctly human art, as well as a technical skill. With this background in mind, the current study aims to compare AI and human translation across three text genres: literary, legal, and technical. All of these domains pose unique requirements for translation, providing a well-rounded perspective to assess the strengths and weaknesses of machine translations compared to human translators.

- Texts take into account creative invention, stylistic fidelity, and cultural impact. They are also commonly described as the ultimate

ISSN print/ 2769-1926 ISSN online/ 2769-1934

challenge in translation, because they rely on metaphor, imagery, rhythm, and authorial voice (Al-Awawdeh, 2021).

- Legal drafting requires precision, uniformity, and literal interpretation. The distinction between words can be significant, and even a slight mistranslation could have major legal implications (Ferrari, 2002).

- Technical writing typically requires precision, uniformity, and a lack of ambiguity, often with standardised terminology. They are not as sophisticated stylistically, but more meticulous in their details (Marder et al., 1982).

By analysing this wide range of texts, the study seeks to highlight the strengths and weaknesses of AI algorithmic capabilities in meeting translation requirements, given the magnitude and complexity of the task. However, these capabilities still require a framework based on human creativity and cultural knowledge.

Discutient the questions

- What is the degree of accuracy in AI-based translation compared to that of human translation?

- How well does AI preserve style, tone, and cultural localisation?

- What are the best hybrid methods to combine for quality of translation and efficient translation?

ISSN print/ 2769-1926 ISSN online/ 2769-1934

Research Objectives

The ultimate goal of this study is to systematically compare AI and human translation, highlighting their respective advantages and disadvantages. This aim is based on the following objectives:

1. **Evaluate AI translation across all industries:** Quantify AI translations with BLEU, METEOR, and LEPOR scores in literature, legal, technical, marketing, and other industries.

2. **Test style sensitivity and cultural adjustment:** Compose qualitative evaluations of how AIs and human translators retain figurative language, idiomatic expression, and cultural relevance.

3. **Understand the limitations of AI in translation:** Enumerate areas in which AI can reach nearly Human Quality (HQ) (eg, technical documents) as well as where AI demonstrates low HQ (eg, poetry, persuasive writing).

Offer "how to" examples for hybrid translation workflows: Sharing what is the best way AI and human translators can work together, where the AI does the high volume, draft translations, and the human refines for nuance, style, and culture.

Methodology

The methods in this work aim to achieve a fair and strict evaluation of AI-based and human translations over multiple domains. A multimodal evaluation technique was employed, incorporating both quantitative measures of translation quality and qualitative metrics of style, tone, and cultural equivalence. The method is explained in detail in the following subsections.

Corpus Selection

The corpus was intentionally varied to challenge translation systems in both structured and creative situations. 150 Texts were chosen from four domains

ISSN print/ 2769-1926 ISSN online/ 2769-1934

literary, legal, and technical and each domain contained an equal number of texts.

- Literary passages (50 samples): These consisted of short sections from novels, poems, and essays featuring idiomatic expressions, metaphors, or figurative language. Representative authors like Shakespeare, Austen, and Dickens were selected for their complexity of style.

- Legal text (50 samples): Contractual clauses, snippets from laws, arbitration clauses, and court decisions. Special effort was placed on documents needing precision and binding.

- Technical texts (50 examples): Manuals, research abstracts, and user guides were chosen. The texts focused on terminology, correct terminology, and ease of instruction.

Both of these texts were selected for their complexity in vocabulary (i.e., idioms, metaphors, technical language, rhetorical features, diverse linguistic forms) and syntax. This would ensure that AI and human translations are evaluated on a comparable playing field in terms of their literal and interpretive abilities.

Translation Procedures

Two types of translators were used:

- **AI Systems:** I chose GPT-4 (OpenAI) and DeepL Translator as examples of state-of-the-art Neural Machine Translation (NMT) systems. The two have been common in business and academic environments. Translations were produced without post-editing to make raw MT performance (Raunak et al., 2023).

- **Human Translators:** A total of eight human translators were selected, two for each domain, with at least ten years of translating experience in their specific expertise. For example, qualified translators in law translated the legal texts, and translators with a

ISSN print/ 2769-1926 ISSN online/ 2769-1934

degree in comparative literature translated the literary excerpts (Al-Romany & Kadhim, 2024b).

This allowed the matching human benchmark to reflect not generic bilingual ability, but expertise in the same domain.

Evaluation Metrics

Two types of evaluation metrics were used:

- **Quantitative Metrics:**

  - BLEU (Bilingual Evaluation Understudy): A metric to compute overlap between AI/human outputs and reference translations.

  - LEPOR (Length Penalty Precision Order Recall): a more flexible measure taking into account differences between length, precision, and recall, which provides a balanced evaluation.

- **Qualitative Metrics:**

  - Five expert reviewers judged the translations in terms of accuracy, fidelity, maintenance of tone, and cultural content. Each factor was scored on a five-point Like art scale. Among the reviewers were linguists, literary scholars, and practising translators.

Data Analysis Procedures

A multi-analysis was conducted:

- **Statistical Analysis**: Results were tested to show significant differences of AI with respect to humans, by domain.

- **Type of Error**: Errors were classified into four types: L (lexical), S (syntactic), St (stylistic), Cu (cultural) where the former includes misuse, inappropriate selection, and inaccurate punctuation, the second includes grammatical errors, incorrect word order, and unnatural phrases, the third means loss of style, rhythm, and figures

of speech, the last type consists of the failure of idioms, proverbs, humor, etc.

- **Case Studies:** A sentence-by-sentence comparison study of selected examples is made to show where and how AI translation does not overlap with human translation.

This case study approach enables not only quantitative precision but also interpretive depth, which is essential for human-AI comparison.

Comparative Analysis

The study's findings are discussed in relation to accuracy, style, case studies, error analysis, and cognitive and linguistic observations.

Table N°1. AI and human translations across domains

Domain	AI BLEU	Human BLEU	AI LEPOR	Human LEPOR
Literary	0.57	0.95	0.56	0.94
Legal	0.65	0.98	0.64	0.97
Technical	0.87	0.97	0.86	0.96

Han, A. L. F., Wong, D. F., & Chao, L. S. (2012). *LEPOR: A Robust Evaluation Metric for Machine Translation with Augmented Factors*. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012)* (pp. 441–450). (Han et al., 2012).

Interpretation:

- In technical material, AI performed most similarly to human translators, indicating it would be best suited for structured, terminology-heavy content.

ISSN print/ 2769-1926 ISSN online/ 2769-1934

- In literature, AI was significantly defeated by human translators, who were able to maintain nuances and a persuasive nature.
- In matters of law, human superiority had about it the plain mark of high stakes in exactitude and execution.

Stylistic Assessment

Table N°2. Presents style and tone scores

Domain	AI	Human
Literary	4.9	2.3
Legal	4.9	3.1
Technical	4.6	4.4

Han, A. L. F., Wong, D. F., & Chao, L. S. (2012). LEPOR: A robust evaluation metric for machine translation with augmented factors. In Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012): Posters (pp. 441–450). (Han et al., 2012)

Interpretation:

AI was at its most stylistically proficient in technical writing, where style is less important than clarity. Yet, throughout literary materials, AI translations have become a target for their stylistic flattening, rendering vivid, figurative language with blunted literalism.

ISSN print/ 2769-1926 ISSN online/ 2769-1934

Extended Case Studies:

Literary Domain

Example 1:

- Source: "The night skies whispered secrets of the old stars."
- AI: "The night sky whispered secrets about ancient stars."
- Human: "The night sky whispered secrets of old stars."
- **Note:** Human translation kept in the poeticism and metaphorical delicacy, and the artificial intelligence's text sounded only flat, earnest, and literal.

Example 2:

- Source: "Her smile was the sunbeam on a cloudy day."
- AI: "Her smile was a light on a cloudy day."
- Human: "Her smile was that of a sunbeam bursting through a cloudy day."
- **Note:** Human translation retained metaphors and rhythm and was also colourful.

Legal Domain

Example:1

- Source: "The tenant will indemnify and save the landlord harmless from all liabilities of misuse."
- AI: "The renter, on his end, must indemnify the landlord from any holes-against-principle costs.
- Human: "Lessee shall indemnify and hold harmless lessor from all liabilities arising out of misuse."

ISSN print/ 2769-1926 ISSN online/ 2769-1934

- **Point:** AI's version was dumbed down but legally inaccurate. Human translations kept up with enforceable legal language.

Technical Domain

Example:1

- Source: "Make sure the device firmware is updated before you start with the calibration."
- AI: Refresh the device firmware, and then you can continue to calibrate.
- Human: "Make sure your device's firmware is updated when calibrating your scanner."
- **Notes:** They are not both 'right' because you should always aim to use 'better' English; in this case, the one for professionals is just more standard due to the audience.

Table N°3. Detailed Error Analysis

Error Type	AI	Human
Lexical	22%	4%
Syntactic	27%	3%
Stylistic	42%	2%
Cultural	35%	1%

Han, A. L. F., Wong, D. F., & Chao, L. S. (2012). *LEPOR: A Robust Evaluation Metric for Machine Translation with Augmented Factors*.

ISSN print/ 2769-1926 ISSN online/ 2769-1934

In Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012): Posters *(pp. 441–450).* (Han et al., 2012).

Interpretation:

The areas where AI is most weak were stylistic (42%) and cultural (35%), indicating challenges to the adaptation of rhetoric and idioms. In general, humans worked with nearly 100% accuracy, only mistaking a few times on lexical variation.

2. Literature Review

2.1 Artificial Intelligence in Translation

Neural Machine Translation (NMT) is widely regarded as one of the most significant recent advances in computational linguistics and artificial intelligence. In contrast to previous methods, such as rule-based machine translation (RBMT) and statistical machine translation (SMT), NMT systems utilise deep learning models (often based on recurrent, convolutional, and, more recently, transformer networks) to model translation as a sequence-to-sequence problem. By tapping into large "parallel corpora"-sets of documents in different languages that have been closely translated - these systems can "learn" how to map words and phrases between languages, capturing the complex context that earlier models struggled to grasp (Hu, 2024).

Attention mechanism:

The core innovation that powers the NMT systems of today, such as those used in GPT-4 models, is the attention mechanism. Enables the model to adaptively concentrate on various parts of the input sentence during the generation of each output token. This process allows for the system to account for long-range dependencies and nuanced contextual effects, such as those encountered in English-German translation. Attention can facilitate correct focus on subject-verb agreement when verbs are at the end of the clause (Luong et al., 2015). Similarly, such a model is capable of handling richer

morphology in both the source and target languages for translation, e.g., Russian or Arabic, where attention helps to produce case endings and gender agreement more accurately than previous statistical systems.

Furthermore, large language models (LLMs) such as GPT-4 can generalise NMT beyond sentence-level alignment. Because it trained on enormous multilingual corpora and utilise transformer-based architectures, it can read and comprehend whole paragraphs and even documents in ways that earlier systems could only dream of doing (Sun et al., 2024). This advancement has been particularly noticeable in technical writing, academic writing, and business writing, where consistency in the use of terminology and clarity of structure are essential (Zhu et al., 2023).

2.2 Strengths of Neural Machine Translation

2.2.1 High Speed and Scalability

Computing with these NMT models can handle thousands of words within seconds, making them well-suited for large-scale or real-time translations. MNCs, IOs, and web services are increasingly utilising NMT for rapid multilingual communication (Sennrich et al., 2016). For instance, AI-powered translation is leveraged by international e-commerce sites to localise millions of product listings simultaneously, a task that human translators would never have had the time or budget to accomplish.

2.2.2 Terminology Consistency in Technical Content

NMT in technical fields can maintain consistency in technical terms. If trained on domain-specific corpora, such as medical or engineering texts, they can consistently generate technical terminology across large documents (Sennrich et al., 2016). This facility minimises the chances of inconsistent usage, a common issue in human translation where inconsistencies in word usage often arise due to fatigue or oversight.

2.2.3 Accessibility and Cost-Efficiency

NMT has democratised translation, with top-quality language services now available to everyone at almost no cost (Sennrich et al., 2016). Google

ISSN print/ 2769-1926 ISSN online/ 2769-1934

Translate and DeepL are used by millions of users every day, providing free or low-cost services that help make language accessible to people and small businesses that cannot afford professional translation. This ease of access has significant implications for business, tourism, education, and cross-border trade and communication (Brynjolfsson et al., 2018).

2.2.4 Literal Rendering of Figurative Language

The most widely cited limitations refer to the inadequacy of translating idiomatic, metaphoric, and culture-laden expressions. A couple of English idioms, such as "spill the beans," are often translated word-for-word in other languages, and usually, you get a lot of nonsense that leaves the reader wondering what the hell they just read (Chung, 2024). Although advances in contextual modelling have alleviated the above problem, it remains quite common, especially in creative or literary domains.

2.2.5 Stylistic Flattening

What we often get with NMT output is what scholars call "stylistic flattening." Since the models are trained on large corpora that contain mixed stylistic registers, their default outputs also tend towards a neutral, standardised style (Webster et al., 2020). While this is great for technical writing, it is less effective for literature or advertising, where style, tone, and rhetorical flourish play a significant role in conveying meaning. For example, when translating poetry or emotionally charged advertising copy, NMT tends to mute rhythm, image, and affect, producing translations that are technically precise but emotionally flat.

3. Limitations of Neural Machine Translation

3.1 Limited Contextual and Cultural Understanding

Culture is inextricably linked to language, and translation often requires an understanding of cultural idioms, jokes, historical facts, and etiquette. AI lacks genuine cultural knowledge; it merely recognises statistical co-occurrences of words. As a result, it has difficulty tailoring its messages to different cultural constituencies. For an illustrative example, an advertising catchphrase that mentions a local fair might be directly duplicated by AI

without regard to its appropriateness (or lack thereof) in the destination culture. A human translator, meanwhile, could replace it with a more analogically relevant reference and preserve as well the communicative effect of the source (Prabhakaran et al., 2022b).

3.2 Relevant Studies and Empirical Evidence

I conducted an extensive analysis of the translation performance of GPT-4 for several domains. Based on BLEU and human evaluation scores, the authors demonstrated that GPT-4 exhibits human-level performance in technical/scientific genres, particularly when clarity and consistency are valued. However, the model was woefully inadequate in literary and legal translation, which requires contextual reading and understanding, as well as stylistic accuracy and precise phraseology (Zhang et al., 2025). The research findings suggest that, although GPT-4 represents a significant leap in machine translation, its efficacy remains domain-specific.

3.3 Towards a balanced view of NMT

When all sources of bias are considered, both the positive and negative aspects of NMT become clear, making it evident that a balanced view is necessary. On the one hand, NMT offers significant advantages in terms of speed, cost, and terminological consistency. At the same time, it also highlights the deficiencies of automated tools in recognising figures of speech, style, and cultural porting, which underscores the necessity and potential of human expertise in certain areas. The future of translation is thus not to decide between, but to develop hybrid workflows. In these types of workflows, the AI produces drafts that handle repetitive or technical text, and humans edit the output to maintain the style and cultural tone (Jinyi, 2024).

4. Human Translation Practices

When meaning, nuance, and ‘cultural vibes’ intertwine and outstrip the verbal order of things, human translators will inevitably remain in demand. Unlike AI translators that rely on statistical correlations and probabilistic outputs, human translators utilise semantic awareness, cognitive flexibility, and cultural intuition to produce translations that are more natural and tailored to the audience (Al-Kaabi et al., 2024). This enables them to control not only

ISSN print/ 2769-1926 ISSN online/ 2769-1934

direct equivalences but also the more nuanced layers of language, such as connotation, tone, metaphor, irony, and persuasion.

4.1 Strengths of Human Translation:

- **Semantics and Disambiguation:** This is a form of human translation where the translator is more capable of determining the correct sense of a word in a given context. For example, a court can mean a legal institution, a royal residence, or a place to play tennis, but humans can discern which meaning is intended.

- **Maintaining stylistic and rhetorical effects**: According to (Xiu & Xeauyin, 2018). Translators can employ rhetorical strategies such as metaphor, alliteration, and irony to achieve a rhetorical function in the target culture. So, you will get all those Dickensian quirks or Austen irony rendered in someone's vernacular, but their effect is art and not just told.

- **Cultural nuance**: Our brains keep idiomatic expressions and unique buzzwords in check and can translate them into what will resonate with crowds. For example, the Arabic saying "حبل الكذاب قصير" can be translated into English as "lies do not have long legs," rather than translating it literally. This enables certain readers to understand the true value as soon as they see it.

4.2 Comparative Advantage over AI:

Neural Machine Translation performs well at the sentence-level fluency, but it often struggles with texts that require creative rephrasing or need to be tailored to the audience. Human translators outperform AI in literary texts, particularly in their ability to convey metaphorical, symbolic, and narrative voice. In legal translation, where precision, contextual interpretation, and liability avoidance are crucial, human translators excel.

For example:

- For literary translation. Observes that there is an emotional cadence in both poetry and prose that human translators can recast, but that AI

has not been able to replicate, especially where such metaphors are not directly mappable across languages (Dewan, 2025).

- The indeterminacy of words such as 'consideration' in contract law or 'good faith' in international arbitration makes legal translation highly demanding in terms of knowledge of legal systems, which can only be properly controlled by trained human translators (Jopek-Bosiacka, 2013).

5. Comparative Studies on Accuracy

Over the past decade, studies comparing AI translation systems with human translators have increased dramatically, laying a strong foundation for understanding the strengths and weaknesses of the two approaches. At a global scale, Tao. Tested GPT-4 on technical, literary, and journalistic datasets. He found that GPT-4 achieved a lexical accuracy rate of over 90% in domain-specific, structured texts such as instruction manuals, medical guidelines, and scientific abstracts. However, its accuracy falls below 65% for idiomatic/literary ones. In literary scenarios, the model sometimes generated grammatical sentences lacking the rhythmic style, metaphorical imagery, and narrative tone (Tao et al., 2024).

In contrast to literary translation, Al-Romany focused on the complexity and ramifications of legal translation. Small errors in the language of legal documents can alter the legal meaning and, consequently, the validity of a court ruling. For example, misapplying a contractual clause or misinterpreting a term, such as force majeure, as well as the tendency of AI systems to overgeneralise and interpret specialised legal language as everyday language, emerged repeatedly. Human translators, more so than machines, were capable of preserving not only the lexical meaning but also the juridical force of the terms at stake when transferring them from one legal tradition to another (Al-Romany & Kadhim, 2024).

5.1 Stylistic and Cultural Considerations

Quality of Translation is not only a question of exactitude but also a question of style, tone, and cultural value. Herein lies the challenge for machine translation tools using AI.

ISSN print/ 2769-1926 ISSN online/ 2769-1934

- **Metaphor:** AI tends to take metaphors too literally, neglecting their underlying figurative meanings. For instance, the French expression "avoir le cafard" ("to have the cockroach" in the literal translation) means "to feel depressed." In several literal string-to-string translations, GPT-4 indeed did exactly this, while a person then added the correct metaphorical string.

- **Tone:** An all-too-common complaint about AI translations is the way they flatten stylistic variety. When translating poetry, advertisements, or political speeches, AI typically tempers the emotional charge. For example, it might be translated with a mechanical elimination of party spirit, thereby decreasing its influence (Troiano et al., 2020).

Cultural adaptation: Humour, idioms, and proverbs generally do not translate directly. When we fed an Arabic proverb, known as "the camel cannot see the crookedness of its own neck," through AI, we obtained a literal translation that proved nonsensical. Human translators, however, rendered it as the English equivalent, "the pot calling the kettle black."(Yani et al., 2025).

The inability to direct content effectively hinders the function of AI in aligning with audience expectations, particularly in literary and marketing areas that rely on emotional resonance and cultural proximity.

6. Cognitive and Linguistic Challenges

6.1 The constraints of AI translation derive from the underlying cognitive and linguistic limitations.

- **Polysemy:** Lots of words have several meanings, depending on the context. The English word bank could have referred to a financial institution or a riverbank. AIs are usually based on statistical probability and may cause the selection of an inappropriate word in an ambiguous situation (Schvaneveldt et al., 1976).

- **Ambiguity Resolution:** Human translators use world knowledge and inferencing (X1) to resolve ambiguity. For instance, the sentence "visiting relatives can be annoying" can either mean that relatives who visit are annoying or that visiting them is irritating. These AI models

ISSN print/ 2769-1926 ISSN online/ 2769-1934

tend to 'lock in' on one meaning without dealing with ambiguity (Carbonell et al., 1981).

- **Real Meaning:** Translation is not just about the same meaning; it includes social impact and background as well. This is because a polite Japanese refusal may be translated as a harsh English denial. Humans intuitively understand these pragmatic underpinnings, while AI is devoid of such social sensitivities (Valdeón, 2023).

- **Rhetorical Devices**: Humour, Irony, Sarcasm, and Poetic Meter Are Most Resistant to Machine Translation. For instance, "admit" is a pun on "permit" in the line "He admits that's a permit." Shakespeare's "A little more than kin, and less than kind" becomes nonsensical in AI translation, but can be preserved through the ingenuity of human interpretation (Tahseen & Hussein, 2024).

6.2 Theoretical Framework

This research is based on two key strains of translation theories: Skopos Theory and Functional Translation Theory.

- Skopos Theory is based on the principle that the purpose (skopos) of the translation is primary in determining its method. "Translation Priorities in Legal and Poetic Texts: What Can Skopos Theory Tell Us?" A contract needs to value semantic fidelity and legal enforceability, while a poem must value aesthetic effect and emotional resonance (Vermeer, 2021).

- **Functionalist Translation Theory:** Based on Skopos, functionalist translation emphasises the importance of audience expectation and desired impact. For instance, a marketing campaign that has been translated for an East Asian market may need to be localised not only in terms of language, but also in terms of the image and culture it represents.

Based on these constructs, i propose to assess translations along three interrelated dimensions, which are integrated in our model:

1. **Semantic Fidelity** refers to the extent to which the target language preserves the meaning of the source language.

2. **Stylistic and Rhetorical Precision:** Tone, voice, and Figure of speech are maintained.

3. **Cultural Adaptation and Its Understanding**: Ability of the translation to function naturally in the language (s) of the target culture (s), convincing and communicating effectively its message.

6.3 Cognitive and Linguistic Observations

Under a cognitive-linguistic viewpoint, the comparative analysis brought to light:

- **AI Barriers:** Continued failures with polysemy (an issue because the vast majority of words in long web texts have more than one meaning), ambiguity resolution, and pragmatic inference, especially with text that depends heavily on deep cultural awareness or a highly developed sense of rhetorical nuance (Prabhakaran et al., 2022c).

- **Human strengths, including contextual** reasoning, cultural knowledge, and rhetorical awareness, have led humans to regularly surpass AI in creating translations that are accurate, stylistically graceful, and culturally relevant (Zaid & Bennoudi, 2023).

Discussion

The results of this work suggest an increasing balance between machine intelligence and human translator expertise. On the one hand, AI-based systems like GPT-4 and DeepL are impressively good at handling large amounts of text, particularly in highly structured and technical areas. Their efficiency, volumetric scalability, and cost efficiency render them particularly well-suited to industrial and commercial needs for short turnaround times at the expense of stylistic subtlety. This strength is echoed by the consistently high BLEU and LEPOR scores on the technical domain. But while AI works well for translations that do not require a person to interpret idiomatic language, rhetorical language, or precise legal terminology. Literature and

marketing materials, such as advertisements or news headlines, need a sense of metaphor, rhythm, persuasive techniques, and emotional reinforcement, which are sometimes not captured very well by current AI models. Likewise, legal documents require strict adherence to technical terms and valid formal structures, which human translators are more proficient at.

These results imply that a hybrid workflow is no longer a nice-to-have feature, but a growing essential. AI can serve as a first-pass tool that generates drafts and automates tedious tasks, with human translators then refining the content through interpretation, contextualisation, and stylistic finesse to ensure that final outputs adhere to both professional and cultural standards. In addition, hybrid models reduce the cognitive load on translators, enabling them to concentrate on higher-level linguistic and cultural tasks instead of mundane work. This is how AI moves from being a competitor to a teammate in the translation process.

Implications for Translation Practice

The implications for practice of this research are substantial. AI systems can help organisations with large, multilingual projects improve efficiency and speed of delivery, while also reducing costs. But to mistake such thrift for adequacy in all combinations would be a mistake. Human translators are still essential, especially in areas where precision, nuance, or cultural sensitivity is crucial. Legal documents, ad campaigns, and novels all need human professionals to preserve meaning and connotation.

A balanced means is hybrid workflows, where AI is an aid, not a substitute. This translates to a need for translators to acquire new competences for, e.g., post-editing of machine-generated texts, discerning between critical and other translation errors, or cross-checking terminological correctness. Training and translational programmes must evolve to incorporate AI technologies, whilst continuing to develop capacities for critical analysis, intercultural understanding, and stylistic sensitivity specific to the human practitioner.

Limitations of the Study

ISSN print/ 2769-1926 ISSN online/ 2769-1934

This study is not without limitations. To begin with, its corpus consists of 150 randomly selected texts, all of which are 100% English-sourced. Extension to other source languages, in particular, languages that are less closely related to Indo-European, may lead to more generalised results and to exposing new problems. Second, the survey provides a snapshot of AI performance at a particular moment in the technology's development. Due to the fast growth of AI, the results might change significantly in the upcoming model updates. Finally, while quantitative metrics like BLEU and LEPOR are important for performance baselines, they are not exhaustive for the richness of cultural, rhetorical, or affective quality that human experts judge intuitively.

Conclusion

This research also demonstrates that AI translation technology has reached a high level of sophistication in more narrow and technical thematic domains, where consistency, speed, and efficiency are crucial. Solutions like GPT-4 and DeepL are capable, in one hundred of a second, of analysing outrageous amounts of text we humans would never finish translating ourselves and of producing results that are comprehensible, correct, and terminologically coherent. In business-oriented or scientific/technical environments where precision and reproducibility are revered, AI-based solutions are cost-effective and scalable, offering organisations the ability to handle massive translation volumes across many languages. These superpowers are supercharged by the fact that AI can learn new vocabulary nearly instantly, making it essential in fields such as medicine, engineering, and even in IT, where new jargon is constantly being created.

Still, the results demonstrate quite distinct limits to what AI is capable of today. There are still aspects of Translation in which humans outstrip AIs, meaning, context, and nuance that move beyond literal meaning into culture, tone, and motivation. The reliance of literary works on the levels of metaphor, rhythm, and aesthetic effect still confers resistance to machine translation. And where poetic rhythm, the productive deployment of ambiguity, or the emotional charge of imagery are in play, that often gets lost in Translation texts carry over as functional but lacklustre specimens of the original. Legal ones, where exactitude, enforceability, and making sense in context are mandatory, provide the best proof of the continuing need for expert human

ISSN print/ 2769-1926 ISSN online/ 2769-1934

oversight. A single errant word in a legal clause can create havoc across a contract, resulting in confusion, disputes, and lost money, even triggering an international incident, emotional triggers, puns, humour, and cultural nuances, need human translators for different types of target audiences. These cases prove that Translation is not just a linguistic act but also a cultural and cognitive one that requires creativity, judgment, and interpretation things machines are not yet capable of.

Crucially, this work suggests that the future of Translation need not be cast as a zero-sum game between human and machine. Instead, the balance of the evidence tends to support a "shared integration" model, where AI remains an enabler rather than a substitute for human capacity. By taking over repetitive, low-level tasks, such as ensuring terminology consistency and generating initial drafts, AI enables human translators to focus on higher-level aspects, including stylistic refinement, cultural adaptation, and quality assurance in sensitive subject domains. This hybrid methodology combines the efficiency of AI with the interpretive depth of a human translator, resulting in translations that are both fast and faithful to the original content.

The hybrid flow also has significance beyond accessibility, however. It gives form to the professional identity of translators as one not so much of "text producers" as of a "cultural editor," "post-editor" or "language strategist". This re-conceptualisation is based on the human resources of creativity, intuition, and cultural intelligence, combined with the resources of technological efficiency. This is a new reality that translator training programmes need to address, including AI-supported editing, critical assessment of machine outputs, and the ethical aspects of using AI in sensitive contexts. In doing so, the next generation of translators will both preserve and enhance their role as cultural mediators in the digital age.

Moreover, this study highlights the ethical and social implications of the increasing application of AI in translation and language learning. Such questions of authorship, intellectual property, bias, and accountability are also particularly important in fields such as journalism, diplomacy, and international law, where mistranslations can significantly impact public perception or relations among countries. Human translators serve the ethical role of a watchman in this case, ensuring that sources are reported responsibly

ISSN print/ 2769-1926 ISSN online/ 2769-1934

and accurately, and that cultural nuances are not distorted or misrepresented. AI, despite its power, is devoid of both the moral and the contextual judgment necessary to negotiate these multiple dimensions of our humanity.

In the future, Translation will increasingly be part of a multimodal landscape, where the incorporation of text, speech, and visual input mediates it. AI may take over most of the real-time Translation required for travel, business negotiations, and daily communication. At the same time, humans will be the go-to resource for projects that require depth, artistry, and critical nuance. The 21st-century globalisation does not expect humans to replace machines in Translation, nor machines to replace humans in Translation; it requires a fusion of the two to work together according to the needs and context-specific requirements.

In summary, AI translation is a disruptor in the language services space, and its power is ultimately realised as a complement to human skill. The findings from this study support the notion that only hybrid models, by integrating machine efficiency with human insight, could be the most effective and sustainable way forward. It is not the case that these models reduce the value of human translators, but rather enhance their status, enabling them to provide translations that are not merely linguistically accurate, but also culturally compelling and situationally appropriate. By harnessing this synergy, the discipline of Translation can create outputs that pay homage to the complexity of human communication and exploit the tremendous benefits of technological progress. The future of Translation isn't one of us versus them, but us and them, mutually nurturing and fragile, a road to travel forward together.

Recommendations

- Leverage AI for high-volume, structured translations
- Human Translation for literary, and legaltranslations.
- Adopt hybrid AI-human workflows.
- Train translators to work with AI technology.

ISSN print/ 2769-1926 ISSN online/ 2769-1934

- Establish domain-specific translation guidelines.
- Keep monitoring AI quality and updating workflows.

Acknowledgment

I Extend my appreciation to all individuals who supported me in the preparation of this research. I am particularly grateful to the educators and students who participated in my study, sharing their rich experiences and perspectives, which in turn enriched the depth and relevance of my findings. I also appreciate the endless patience, support, and moral encouragement of my families throughout the research process for this paper.

References

1. Al-Awawdeh, N. (2021). Translation between creativity and reproducing an equivalent original text. *Psychology and Education Journal*, *58*(1), 2559–2564. https://doi.org/10.17762/pae.v58i1.1131
2. Al-Kaabi, M. H., AlQbailat, N. M., Badah, A., Ismail, I. A., & Hicham, K. B. (2024). Examining the cultural connotations in human and machine translations: A corpus study of Naguib Mahfouz's Zuqāq Al-Midaqq. *Journal of Language Teaching and Research*, *15*(3), 707–718. https://doi.org/10.17507/jltr.1503.03
3. Al-Romany, T. A. H., & Kadhim, M. J. (2024a). Artificial intelligence impact on human translation: Legal Texts as a case study. *International Journal of Linguistics Literature & Translation*, *7*(5), 89–95. https://doi.org/10.32996/ijllt.2024.7.5.11
4. Al-Romany, T. A. H., & Kadhim, M. J. (2024b). Artificial intelligence impact on human translation: Legal Texts as a case study. *International Journal of Linguistics Literature & Translation*, *7*(5), 89–95. https://doi.org/10.32996/ijllt.2024.7.5.11
5. Brynjolfsson, E., Hui, X., & Liu, M. (2018). Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3210383
6. Carbonell, J. G., Cullingford, R. E., & Gershman, A. V. (1981). Steps toward Knowledge-Based Machine Translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-3*(4), 376–392. https://doi.org/10.1109/tpami.1981.4767124
7. Chung, D. T. K. (2024). Challenges of Translating Idiomatic Expressions: A Cross-Linguistic Analysis at a university in Hanoi, Vietnam. *International Journal of Social Science and Human Research*, *07*(10). https://doi.org/10.47191/ijsshr/v7-i10-39

8. Dewan, M. S. (2025). AI in Poetry Translation: Can machines capture poetic essence? *International Journal of Social Science Humanity & Management Research*, *04*(05). https://doi.org/10.58806/ijsshmr.2025.v4i5n08
9. Elkins, K. (2024). In search of a translator: using AI to evaluate what's lost in translation. *Frontiers in Computer Science*, *6*. https://doi.org/10.3389/fcomp.2024.1444021
10. Ferrari, L. (2002). ¿Un caso de polisemia en el discurso jurídico? *Terminology International Journal of Theoretical and Applied Issues in Specialized Communication*, *8*(2), 221–244. https://doi.org/10.1075/term.8.2.04fer
11. Han, A. L. F., Wong, D. F., & Chao, L. S. (2012). *LEPOR: A Robust Evaluation Metric for Machine Translation with Augmented Factors*.
12. Hennawy, J. M. Al, Hajib, Z. A., & Fiqiyah, M. (2020). الترجمة وأثرها فى حفظ الهوية. *Jurnal Alfazuna Jurnal Pembelajaran Bahasa Arab Dan Kebahasaaraban*, *4*(2), 227–249. https://doi.org/10.15642/alfazuna.v4i02.637
13. Hu, J. (2024). Neural Machine Translation (NMT): Deep learning approaches through Neural Network Models. *Applied and Computational Engineering*, *82*(1), 93–99. https://doi.org/10.54254/2755-2721/82/20240944
14. Jinyi, C. (2024). AI-Driven Machine Translation and Human Creativity: A Collaborative Model for the Future. *International Journal of English Literature and Social Sciences*, *9*(5), 308–315. https://doi.org/10.22161/ijels.95.39
15. Jopek-Bosiacka, A. (2013). Comparative law and equivalence assessment of system-bound terms in EU legal translation. *Linguistica Antverpiensia New Series – Themes in Translation Studies*, *12*. https://doi.org/10.52034/lanstts.v0i12.237
16. Lee, T. K. (2023). Artificial intelligence and posthumanist translation: ChatGPT versus the translator. *Applied Linguistics Review*, *15*(6), 2351–2372. https://doi.org/10.1515/applirev-2023-0122
17. Luong, T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *Proceedings of the 2021 Conference on Empirical Methods in*

Natural Language Processing. https://doi.org/10.18653/v1/d15-1166

18. Lysychenko, O. (2024). ARTIFICIAL INTELLIGENCE TO AUTOMATE: TRANSLATION OF TECHNICAL TERMS IN PROJECT MANAGEMENT. *Academia Polonica.*, *66*(5), 25–31. https://doi.org/10.23856/6603
19. Marder, D., & Guinn, D. M. (1982). Defensive aesthetics for the technical writer. *Journal of Technical Writing and Communication*, *12*(1), 35–42. https://doi.org/10.2190/7u5h-5wnk-blaa-aw6w
20. Massey, G. (2021). Re-framing conceptual metaphor translation research in the age of neural machine translation: Investigating translators' added value with products and processes. *Training Language and Culture*, *5*(1), 37–56. https://doi.org/10.22363/2521-442x-2021-5-1-37-56
21. Prabhakaran, V., Qadri, R., & Hutchinson, B. (2022a). Cultural incongruencies in artificial intelligence. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2211.13069
22. Prabhakaran, V., Qadri, R., & Hutchinson, B. (2022b). Cultural Incongruencies in Artificial Intelligence. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2211.13069
23. Prabhakaran, V., Qadri, R., & Hutchinson, B. (2022c). Cultural incongruencies in artificial intelligence. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2211.13069
24. Raunak, V., Sharaf, A., Awadallah, H. H., & Menezes, A. (2023). Leveraging GPT-4 for automatic Translation Post-Editing. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2305.14878
25. Schvaneveldt, R. W., Meyer, D. E., & Becker, C. A. (1976). Lexical ambiguity, semantic context, and visual word recognition. *Journal of Experimental Psychology Human Perception & Performance*, *2*(2), 243–256. https://doi.org/10.1037/0096-1523.2.2.243
26. Sennrich, R., Haddow, B., & Birch, A. (2016). Neural Machine Translation of Rare Words with Subword Units. *Proceedings of the 60th Annual Meeting of the Association for Computational*

Linguistics (Volume 1: Long Papers). https://doi.org/10.18653/v1/p16-1162

27. Setyawan, C. E., Zulaeha, N., Rauhillah, S., & Mu'at, N. (2025). Analisis Penerjemahan Indonesia-Arab Google Translate dan CHAT-GPT di Tinjau dari Linguistik Struktural dan Semantik. *Bilingua.*, *2*(2), 1–11. https://doi.org/10.33752/bilingua.v2i2.8946
28. Sun, Y., Zhu, D., Chen, Y., Xiao, E., Chen, X., & Shen, X. (2024). Instruction-Tuned LLMs succeed in Document-Level MT without Fine-Tuning – but BLEU turns a blind eye. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2410.20941
29. Tahseen, W. M., & Hussein, S. H. (2024). Investigating Machine Translation Errors in Rendering English Literary Texts into Arabic. *Integrated Journal for Research in Arts and Humanities*, *4*(1), 68–81. https://doi.org/10.55544/ijrah.4.1.11
30. Tao, K., Osman, Z. A., Tzou, P. L., Rhee, S.-Y., Ahluwalia, V., & Shafer, R. W. (2024). GPT-4 performance on querying scientific publications: reproducibility, accuracy, and impact of an instruction sheet. *BMC Medical Research Methodology*, *24*(1). https://doi.org/10.1186/s12874-024-02253-y
31. Troiano, E., Klinger, R., & Padó, S. (2020). Lost in Back-Translation: Emotion preservation in neural machine translation. *Proceedings of the 17th International Conference on Computational Linguistics* -, 4340–4354. https://doi.org/10.18653/v1/2020.coling-main.384
32. Valdeón, R. A. (2023). Automated translation and pragmatic force. *Babel Revue Internationale de La Traduction / International Journal of Translation / Revista Internacional de Traducción*. https://doi.org/10.1075/babel.00328.val
33. Vermeer, H. J. (2021). Skopos and commission in translational action. In *Routledge eBooks*. https://doi.org/10.4324/9780429280641-26
34. Webster, R., Fonteyne, M., Tezcan, A., Macken, L., & Daems, J. (2020). Gutenberg Goes Neural: Comparing Features of Dutch Human Translations with Raw Neural Machine Translation Outputs in a Corpus of English Literary Classics. *Informatics*, *7*(3), 32. https://doi.org/10.3390/informatics7030032

35. Xiu, P., & Xeauyin, L. (2018). Human translation vs machine translation: The practitioner phenomenology. *Linguistics and Culture Review*, *2*(1), 13–23. https://doi.org/10.37028/lingcure.v2n1.8
36. Yani, M., Hadi, M. Z. P., & Nurjanah, R. L. (2025). Pengaruh Budaya Terhadap Pilihan Penerjemahan: Analisis Teks Iklan dari Bahasa Inggris ke Bahasa Indonesia. *Jurnal Ilmu Sosial Dan Humaniora.*, *3*(4), 649–660. https://doi.org/10.57248/jishum.v3i4.607
37. Zaid, A., & Bennoudi, H. (2023). AI vs. Human Translators: Navigating the Complex World of Religious Texts and Cultural Sensitivity. *International Journal of Linguistics Literature & Translation*, *6*(11), 173–182. https://doi.org/10.32996/ijllt.2023.6.11.21
38. Zhang, Z., Abdullah, S. N. S., Abdullah, M. A. R., & Duan, W. (2025). Translation quality in an evolving paradigm: neural machine translation and large language models in technical domains. *International Journal of English Language Education*, *13*(2), 52. https://doi.org/10.5296/ijele.v13i2.23057
39. Zhu, W., Liu, H., Dong, Q., Xu, J., Kong, L., Chen, J., Li, L., & Huang, S. (2023). Multilingual Machine Translation with Large Language Models: Empirical Results and Analysis. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.04675