Binhua Wang*

A Survey-Based Analysis About Consequential Risk of Errors and Ethical Complexities in the Use of AI-Powered Machine Translation in High-Stakes Situations

https://doi.org/10.1515/dsll-2025-0009
Received June 3, 2025; accepted June 30, 2025; published online August 1, 2025

Abstract: The availability of AI tools in recent years has led to increased use of automatic AI-powered machine translation (AI-MT) in various sectors. While AI-MT might be useful to support language diversity and enhance cross-lingual information access, the use of AI-MT involves consequential risk of errors and serious ethical issues, particularly when they are used in high-stakes situations such as government public services that matters to the rights, benefits and well-being of the public. Through analysis based on a survey about consequences in the use of AI-MT in government public services, this article reveals the hidden risk of errors and ethical complexities in using AI-MT, which have been overlooked by the non-professional users and the public. Through reflections on the status quo and by drawing upon relevant discussions, the article discusses the ethical issues in the use of AI-MT and proposed tentative principles for the formulation of guidelines on proper, responsible and ethical use of AI-MT.

Keywords: AI-powered machine translation; high-stakes situations; government public services; consequential risk of errors; ethical complexities

Article note: A draft of the article was presented as an invited keynote in the International Forum on Digital Studies in Language and Culture (Chongqing University, 31 May – 1 Jun 2025). An earlier draft of the article was presented as an invited keynote in the 15 National Conference & International Forum on Interpreting (The Chinese University of Hong Kong, Shenzhen, 9–10 November 2024). Special thanks go to the organisers of the conferences.

***Corresponding author: Binhua Wang**, Centre for Translation and Interpreting Studies, University of Leeds, Leeds, UK, E-mail: b.h.w.wang@leeds.ac.uk. https://orcid.org/0000-0003-2404-5214

1 Introduction

The evolution of machine translation from rule-based machine translation (RMT) in its early infancy to statistics-based machine translation (SMT) about one to two decades ago and to today's neural machine translation (NMT), which is labelled as AI-powered machine translation (AI-MT), has profoundly transformed how translation is practised, studied and perceived, which is catalysing heated debate about the various aspects as well as issues and challenges in the application of AI technology in translation (Mo et al. 2024). In this regard, digital humanities can and should be functional, not only in exploring how to leverage artificial intelligence (AI) technology to expand the boundaries of language, literature and translation studies and how to harness AI technology best for the public well-being of the humankind, but also in revealing implications of the wider and wider application of AI technology in human-AI interactions.

1.1 Issues in the Application of AI Technology: A Strand of Digital Humanities

As highlighted by the Preface of the translated Chinese version of a classical book on digital humanities – *Digital Humanities. Knowledge and Critique in a Digital Age*, current research in digital humanities covers the following three main strands (Ma 2019, vii): (a) using digital information technology to address existing problems in the humanities that traditional methods cannot solve, such as preserving cultural heritage digitally; (b) identifying and studying new humanistic issues brought about by digital information technology, such as digital capitalism, data sovereignty, data donation, etc.; (c) exploring the application and impact of new digital technologies, tools and platforms in the production, dissemination, and teaching of (humanistic) knowledge.

In translation studies, as summarised by a representative translation scholar, four facets of digital humanities have been applied to translation studies, including: studies about translation technologies, translation studies with big data, translation in digital spaces, translation studies with data visualization (Bowker 2021, 38).

This study aligns with the third strand of digital humanities, which concerns translation technologies, in particular, issues in the application and impact of new digital technologies such as AI in the production, dissemination, and reproduction of humanistic knowledge.

1.2 Increased Use of AI-MT Tools and Scarcity of Guidelines on Their Use

The availability of AI technology and tools in recent years, symbolised by the release of ChatGPT by the OpenAI on 30 Nov 2023, has led to increased use of AI-powered machine translation (AI-MT). AI-MT tools are becoming more and more widely used in various sectors, particularly by non-professional general users. While these tools have the advantage of convenience and high efficiency in output, which might be useful in enhancing language diversity and facilitating cross-lingual access to information, a major issue that has been overlooked is that the use of these tools involve consequential risk of errors and serious ethical issues, especially when they are used in government public services that matters to the rights, benefits and well-being of the public in multilingual communities. Despite the high stakes, the risk of errors and the ethical complexities involved in using AI-MT have been largely neglected by general users and the public.

In order to identify whether any guidelines have been provided on the use of AI-MT, I conducted two surveys based on Google searches with the query string of 'AI machine translation guideline' in March 2024 and May 2025. The survey conducted in March 2024 revealed that none of the results found in Google Search were about "guideline" for the use of AI-powered machine translation tools, most of which were 'guide' or 'guidance' provided by translation companies and agencies on how to use such tools. The content includes: Which AI-MT tools are available? How to choose AI-MT tools? How to use them? What is an AI-MT solution like? However, there was no mentioning of "guideline" on the use of AI machine translation tools. The survey in May 2025 indicated that there is still a glaring lack of guidelines on proper, responsible and ethical use of AI-MT.

Against this background, this article intends to reveal the hidden risk of errors and ethical complexities involved in using AI-MT, particularly when they are used in such high-stakes situations as government public services, which highlights the necessity of ethical considerations and guidelines for proper, responsible and ethical use of AI-MT.

The analysis and discussion will revolve around the following research questions:

1) What are the hidden risks and ethical complexities associated with the use of AI-powered machine translation (AI-MT) in high-stakes situations such as government public services, as revealed through survey-based analysis?
2) How do non-professional users and the general public perceive or overlook the consequences of AI-MT use in high-stakes public service contexts?

3) What tentative principles can be proposed for the formulation of guidelines on proper, responsible and ethical use of AI-MT in high-stakes situations such as government public services?

2 Consequences from the Use of AI-MT in High-Stake Situations: A Survey Based on News Reports About Government Public Services

In order to reveal possible consequences involved in the use of AI-powered machine translation in high-stakes situations, a survey was conducted about two high-stakes settings of government public services, i.e., governmental legal services and governmental healthcare services. The survey was done based on web-searches for relevant news reports published in public media in recent years on Google Search, which, with its mission "to organize the world's information and make it universally accessible and useful",[1] has possibly the most comprehensive information from around the world. As indicated by the top search results, it is found that consequences from the use of AI-MT in government public services have made breaking news which have been reported by major global news portals such as *The Guardian*, BBC News (British Broadcasting Corporation), and ABC News (Australian Broadcasting Corporation), etc. As revealed by the news reports, the consequences exist even in major developed countries, including US, UK and Australia.

2.1 The Use of AI-MT in Governmental Legal Services

The news report published by *The Guardian* on 7 Sep 2023 is titled "Lost in AI Translation". It is reported that the US immigration system growingly relies on language apps or AI-powered machine translation without grasping the limit of such tools, which jeopardizes asylum seekers' applications (Figure 1).

According to the news report, problems with machine translation tools occur throughout the asylum process, from border stations to detention centres to immigration courts. The CBP One app, which the Biden administration has mandated anyone seeking asylum to use to schedule an appointment with CBP before entering the country, is translated into only a handful of languages. And even in those translations, errors appear. The version of the FAQ section of the app in Haitian

1 https://www.google.com/intl/en_uk/search/howsearchworks/our-approach/.

Figure 1: Screenshot of the news report in *The Guardian*.

Creole, for instance, largely shows a string of letters with no spaces or the necessary accent marks (*The Guardian* 2023a).

As reported by BBC News on 27 Sep 2019, the US Citizenship and Immigration Services (USCIS) uses online machine translation tools such as Google Translate to vet refugees (Figure 2).

Figure 2: Screenshot of the news report in BBC News.

According to the news report, a staff manual of the USCIS said pasting text into tools such as Google Translate and Microsoft Bing was “the most efficient approach”. It acknowledges that “occasionally” the online translation services may not fully

understand "dialect" or colloquial language. It also says individual officers can decide whether to request human expert translation services (BBC News 2019).

As reported by *The Guardian* on 22 Feb 2023, the UK Home Office requires that the application questionnaire for refugees, which have more than 50 complicated questions, "must be completed in English" (Figure 3). It recommends using "online translation tools" if the refugee claimants have difficulty in completing the form in English, which is often the case. It also says that a failure to return the document

Figure 3: Screenshot of the news report in *The Guardian*.

within 20 working days "may result in an individual's asylum claim being withdrawn" (*The Guardian* 2023b).

2.2 The Use of AI-MT in Governmental Healthcare Services

As reported by BBC News on 21 Nov 2023, NHS interpreting service problems contributed to patient deaths (Figure 4).

According to the news report, the BBC has found that interpreting issues were a contributing factor in at least 80 babies dying or suffering serious brain injuries in England between 2018 and 2022. It has also found that some NHS staff have resorted to online translation tools to convey critical information to non-English speaking patients (BBC News 2023).

In the recent Awaab Ishak Inquiry that have been widely reported by the media in the UK, which was about the tragic death of a two-year-old boy whose mother is limited in English proficiency, the doctor admitted that she used Google Translate rather than a telephone translation service by human interpreters to convey the message that the boy should "go to hospital" if he struggled with breathing again, which later turned out to be a critical message that is a matter of life and death (Garden Court North Chambers 2022). It is stated in the coroner's verdict that "Clear instructions using an appropriate translator should have been provided to the family." (HM Court Ruling 2022)

As reported by ABC News on 18 Nov 2020, the Australian Federal Government used Google Translate for COVID-19 messaging aimed at multicultural communities (Figure 5).

According to the news report, the Department of Home Affairs said it used the Google Translate subscription service for its website in the pandemic to "ensure there was an easy-to-use repository of translated information available to multicultural communities as quickly as possible" (ABC News 2020). Critical public health messages by the Commonwealth government about the coronavirus pandemic were bungled as bureaucrats used Google Translate. As a result, "nonsensical" and "laughable" language translations of COVID-19 public health messages are being distributed to multicultural communities, prompting fears that migrants and refugees will lose trust in authorities' handling of the coronavirus pandemic. For example, one of those translations tweeted by the Australian Federal Government was a rendition of "Information in your language", which was supposed to tell Chinese speakers where to look for more details about the pandemic. However, it was translated to something ungrammatical: "Use your language supplied information" (Figure 6).

In summary, it is revealed by the survey analysis that AI-powered machine translation tools are becoming more widely used not only by private users but also by

Figure 4: Screenshot of the news report in BBC News.

governmental institutions in various countries, and they are used in high-stakes settings such as governmental legal services and governmental healthcare services. It is also identified that the use of AI-MT in high-stakes situations such as government

Figure 5: Screenshot of the news report in ABC News.

public services involves serious consequences, which not only matters to the protection of legitimate rights, benefits and well-being of the public, but also have potential life-and-death implications. It seems that, however, the risk of errors and ethical complexities involved in this are largely overlooked by government institutional users.

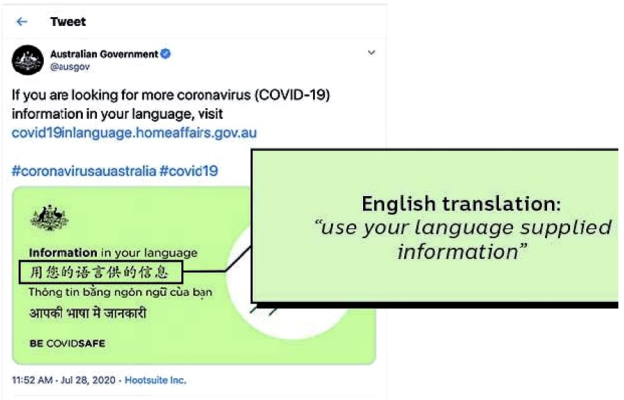

Figure 6: Screenshot of the Tweet message with translations in several languages.

3 Reflections on the Status Quo: Risk of Errors and Ethical Complexities in Contrast with the Scarcity of Regulation About the Wide Use of AI-MT

A possible reason for the wide use of AI-MT tools might be that machine translation has seen significant improvement in recent years thanks to the breakthrough of AI technology, particularly in grammatical correctness and structural coherence. However, it seems that general users of AI-MT tools are not aware of or have neglected several facts about current AI-powered machine translation, which can be elaborated as follows:

1) Even powered by AI technology, the current machine translation has not reached such a reliable quality level in terms of accuracy and completeness that it is free of errors. For example, when a Chinese university put on the poster on its newly opened dining hall: 欢迎新老师生前来就餐 (Literal meaning: Welcome new and old teachers and students to come and dine here), people found that it was translated by ChatGPT as: "Welcome new teachers to dine before they pass away". When it was translated by DeepL, a machine translation tool that is widely used by professional translators, the translation became: "New teachers are welcome to dine before they were born". Although with the human intervention of proper prompt engineering, some LLMs might be able to avoid this error due to wrong segmentation by the machine, no one can guarantee that AI-MT outputs are error-free.

2) It has to be noted that machine translation of speeches, which organisers of big multilingual meetings or conferences are tempted to use and is typically shown as live subtitles or transcripts on the screen to the audience, is particularly susceptible to risk of errors. Many errors originate from the automatic speech recognition (ASR) of the speech in the original language since speeches are varied in terms of pronunciation, intonation, pace of delivery and accent; errors also happen in machine translation of the machine-transcribed text. Also, it is very difficult for the machine, even powered by AI technology to capture the paraverbal and nonverbal elements of meaning in human communication, such as intonation, tone, facial expression, gesture and other body language (Wang 2023). Compounded by all these factors, the current machine translation of speeches is non-sensical for most of the time, which is actually not useable if cross-lingual communication needs to be achieved through translation. It is also important to note that although this type of machine translation is often labelled by technology companies as "simultaneous interpreting by the machine" or "AI machine interpreting", it is actually a wrong label because this is only speech-to-text translation while "interpreting" in its real sense should be speech-to-speech translation through which the audience can hear the speech in the target language to achieve instant on-situ cross-lingual understanding and communication.
3) For general users who are not capable of making professional evaluation of translation quality, the seemingly 'coherent and fluent' text produced by AI-powered machine translation nowadays actually entails an even higher risk for errors than before, which is due to the fact that the errors have become much less obvious and much more difficult to detect with its coherence and fluency in the surface structure. What is worse is that most general users are not even aware of the existence of errors and that there is the risk for errors in AI-MT. Many people have got so carried away by the convenience and efficiency of AI-MT that they would neglect the risk of errors, which can be very consequential.

It is useful to do an overall assessment of the current status quo about the use of AI-MT by asking the following questions and seeing their answers:

(1) Are general users aware of the risk of errors and consequences in using AI-MT? Do they know how to avoid such risks?
As seen from the survey result presented in Section 2, the answer is 'No'.
(2) Who should be held accountable when there are consequences from using AI-MT?
Not specified. AI companies that developed these tools have made disclaimers.
(3) Are there considerations about ethics in using AI-MT?
As seen from the survey result presented in Section 2, the answer is 'No'.
(4) Are there guidelines about proper, responsible and ethical use of MT?

As seen from the survey result presented in Section 1, the answer is 'No'.

Also, in academic research, there is the scarcity of evidence-based study about risk of errors, consequences and ethical complexities in using AI-powered machine translation in high-stakes government public services. It was only very recently in Oct 2024 that a preliminary report from a survey was released on uses of AI translation in UK public service contexts on the website of the Chartered Institute of Linguists (CIOL) (Vieira 2024). According to that report, AI translation is used prevalently: "At least 33 % of 2,520 UK public service professionals surveyed had used machine translation in their work." (Vieira 2024) Technology is heavily involved in the uses of AI translation: "Most users accessed translation tools via openly available browser interfaces. Over half reported using personal devices for work-related translation tasks." (Vieira 2024) However, there is the serious lack of institutional guidelines or training about the proper use of AI translation: "For 72 % of respondents, machine translation had never been mentioned in workplace training", in spite of the fact that "15 % reported that machine translation use was recommended by their [institutional] employer." (Vieira 2024) The preliminary report identified the existence of three major risks involved in the use of AI-MT, including: (1) high possibility for miscommunication in critical public services; (2) serious concerns over privacy and information security due to use of personal devices and open web interfaces; (3) the risk of machine translation becoming common practice but without proper policy consideration (Vieira 2024).

As seen from the above reflections, while AI-powered machine translation tools are being used more and more widely including in such high-stakes situations as governmental public services, the risk of errors and ethical complexities in their use have been glaringly overlooked. There is the regulatory gap about the use of AI-MT though their use has caused fatal miscommunications and with a lack of accountability. Therefore, it is imperative now to promote the awareness about the risk of errors, consequences and ethical complexities involved in the use of AI-MT tools among their users, esp. government institutional users.

4 Major Ethical Issues Involved in the Use of AI-MT

According to the survey report on the use of AI-MT tools (Wang and Wang 2025), 76.12 % of the language service professionals in China have concerns about risk of errors and inaccuracies; nearly 70 % have concerns about data privacy and safety; 36.32 % have concerns about copyright issues; about 55 % have concerns about the lack of verifiability in AI-MT output.

The main ethical issues involved in using AI-MT can be summarised as follows:

4.1 Risk of Errors and Inaccuracies

Accuracy is the lifeline for translation, esp. of translations to be used in high-stakes situations. However, the accuracy of AI-MT depends on various factors, such as quality and quantity of training data, complexity of the language, and the algorithm used.

Errors and inaccuracies in AI-powered machine translation systems often stem from the nature and quality of the data used during AI training. These systems rely heavily on large-scale parallel corpora to learn how to translate between languages, but such data is not always representative or comprehensive. AI-MT frequently struggles to capture the grammatical, pragmatic and sociocultural nuances that differ significantly across languages, leading to mistranslations or contextually inappropriate outputs. This issue is particularly pronounced in the case of less commonly used or low-resource languages, where training data is often sparse, outdated or inconsistent. As a result, AI-MT systems trained on such limited datasets may produce translations that are not only inaccurate but also potentially misleading or culturally insensitive. These limitations highlight the need for more inclusive data collection practices and the integration of linguistic and cultural expertise in the development of AI-MT technologies.

As pointed out by the Institute of Translation & Interpreting (ITI) and the Chartered Institute of Linguists (CIOL): “Notwithstanding advances in AI and Machine Translation, even in English and major European languages, the frequency and severity of errors in machine translated texts is such that no public authority should contemplate its routine use, without human oversight” (The Institute of Translation and Interpreting 2023).

4.2 Bias of AI Systems

Bias of AI systems are evident in three aspects: (a) bias of AI training data, which occurs when training data used to develop AI models is incomplete, unrepresentative, or contains errors; (b) bias of AI algorithm, which occurs when an AI model replicates and amplifies existing societal biases, such as racial or gender biases; (c) bias of AI users, which occurs when users of AI systems exhibit biased behaviour, such as by selecting translations that align with their preconceived notions.

It is not uncommon that the datasets for AI training contain socially constructed biases, inaccuracies and inconsistencies that are then learned and reproduced by the AI models (Tomalin et al. 2021). For instance, gender bias is a well-documented issue in neural machine translation (NMT), where models may default to stereotypical gender roles based on historical data patterns – translating ‘doctor’ as male and

'nurse' as female. These biases are not just technical flaws but reflect deeper societal inequalities embedded in language use.

The question of whether AI system bias is inevitable is complex. While some level of bias may be inherent due to the nature of human language and the data used, it is not entirely unavoidable. Techniques such as domain adaptation – adjusting models after training rather than attempting to fully debias data beforehand – have shown promise in reducing bias without significantly compromising performance (Tomalin et al. 2021). However, if left unaddressed, these biases can lead to serious consequences, including reinforcing stereotypes, misrepresenting cultural contexts and undermining trust in public services that rely on AI-MT.

The bias of AI systems has serious ethical implications for the use of AI-powered machine translation. According to the ITI and the CIOL: "In world languages where there is less 'training data' (i.e. web content sucked into AI datasets) the frequency and severity of errors can be so high that professional translators would not waste time using the tools. The opposite problem can occur where there is more training data – the errors are harder to spot – but all the more problematic, including with numbers, dates and not registering the difference between a positive and a negative assertion. The risks here are very high" (The Institute of Translation and Interpreting 2023).

4.3 Data Privacy and Security of the Users

Data privacy and security is another critical concern in AI-MT, particularly when sensitive information is processed through cloud-based translation tools. Private data can be leaked after users enter them into AI-MT tools and servers, which might be used as training data and/or become accessible by other users using the AI-MT tools. Many free online translation services store user data to improve model performance, raising questions about data ownership, consent, and compliance with regulations like GDPR.

According to the ITI and the CIOL, serious ethical concerns are involved in data privacy and security (The Institute of Translation and Interpreting 2023):

> The UK Government encouraging people to use Google Translate or other online translation tools means that they are encouraging them to input and send their most personal and identifiable data on a trip around the world – to servers and data processors in the USA, and other countries, with very different data protection laws and potential interest in them as individuals than the UK's. This seems wholly inappropriate from a data and privacy perspective, for people who are by definition vulnerable.

Only with guidelines in place can these challenges be effectively addressed through best practices such as by using on-premises translation solutions, implementing strong encryption and anonymization protocols, and ensuring transparent data governance policies.

4.4 Transparency and Accountability

Unlike human translators, the working mechanism and workflow of AI-powered translation tools are difficult to track and understand. This lack of transparency can make it challenging to identify errors and biases in the translations produced by them.

Additionally, it is challenging to hold the developers of AI-powered translation tools accountable for any errors or biases as most AI companies have made disclaimers about accountability related to the use of AI tools. For example, OpenAI, the developer of ChatGPT, states the following explicitly in its Usage Policies (Open AI 2025):

> Don't perform or facilitate the following activities that may significantly impair the safety, wellbeing, or rights of others, including: Providing tailored legal, medical/health, or financial advice without review by a qualified professional and disclosure of the use of AI assistance and its potential limitations; Making high-stakes automated decisions in domains that affect an individual's safety, rights or well-being (e.g., law enforcement, migration, management of critical infrastructure, safety components of products, essential services, credit, employment, housing, education, social scoring, or insurance).

5 Establishing Tentative Principles for the Formulation of Guidelines on Proper, Responsible and Ethical Use of AI-MT

In Jun 2024, the SAFE-AI (Stakeholders Advocating for Fair and Ethical AI in Interpreting) Task Force (https://safeaitf.org/), a pioneer to establish, disseminate and promote industry-wide guidelines and best practices for accountable design and adoption of AI in interpreting, released its Guidance on AI and Interpreting Services (SAFE-AI 2024), which proposed four core ethical principles on the development and deployment of AI-based interpreting solutions.

In 2024, the International Association of Conference Interpreters (AIIC) also launched the "M.A.S.T.E.R. the Message" campaign (AIIC 2024) in response to the impact of AI on the interpreting profession, which outlined six key principles that distinguish human interpreters from AI algorithms.

On the basis of the above two documents (SAFE-AI 2024; AIIC 2024), the following tentative principles can be outlined for ethical use of AI-powered machine translation (SAFE-AI 2024; AIIC 2024).

5.1 End-User Autonomy of AI-MT Tools Must be Respected

AI-MT tools must respect the autonomy of its end users, including: end-users are involved in the design and evaluation process of such tools; informed consent is ensured to accept or decline AI use; opt-in/opt-out options are provided for data collection; clear mechanisms to switch between AI-MT and human translation are provided; transparency about AI involvement and data use is indicated.

5.2 Transparency of AI-MT Quality Must be Ensured

First, objective metrics about the functions and performance of the AI-MT tools must be published and explained.

Second, vendors of AI-MT tools must disclose AI capabilities, training data, dialect coverage, and performance metrics of the AI-MT tools. Mechanisms for feedback, error tracking, and bias auditing must be included as a component of the AI-MT tools.

Third, end-users of AI-MT tools must be informed about: The type and quality of AI-MT available; Their rights to accept/decline AI services; Data privacy and training requirements.

5.3 Accountability Structures and Regular Audits of AI-MT Must be in Place

Clear accountability structures must be in place, including: Developers, vendors, and purchasers of AI-MT tools must accept liability for AI performance; Contracts must define responsibilities and consequences for misuse or harm of AI-MT; AI-MT tools must be validated by qualified human translators before deployment.

Regular audits of AI-MT tools must be conducted to assess translation accuracy and user satisfaction, compliance of AI-MT tools with relevant laws and policies, correction of errors and mitigation of bias, and that performance metrics must be transparent and disaggregated by language.

5.4 Evidence of Improving Safety of AI-MT Must be Communicated Publicly

AI-MT tools should meet or exceed existing legal and ethical standards for translation. AI-MT tools should be evaluated against professional translation standards and legal requirements. The limitations of AI-MT tools should be clearly disclosed during the informed consent process. AI-MT developers must disclose their compliance with laws for ensuring equal language access to services.

5.5 Appropriate Human Involvement in Using AI-MT Must be Maintained

Maintaining appropriate human involvement in the use of AI-powered machine translation (AI-MT) is essential to ensure accuracy, cultural sensitivity, and ethical responsibility. While AI-MT systems have significantly advanced in speed and efficiency, they still require human oversight to function effectively and responsibly. Keeping a human-in-the-loop means that human translation experts are involved in setting up, tuning, and testing these systems, as well as reviewing and refining their outputs. This human role is crucial in improving decision-making and ensuring that translations are not only linguistically correct but also contextually appropriate and ethically sound.

Although the increasing availability of generative AI tools, including AI-MT, has led to more automated workflows, relying solely on AI for translation without meaningful human supervision can result in serious consequences. Machine-generated translations may contain inaccuracies, misinterpret cultural nuances, or even produce harmful content if left unchecked. Human intervention is necessary to assess the quality of AI-MT outputs, particularly machine translation post-editing by professional translators, ensuring that they are accurate, relevant, and aligned with societal and cultural values.

AI-powered machine translation also lacks the human qualities of empathy, discretion, and contextual understanding. Unlike humans, AI cannot adapt its translations based on individual intentions, the communicative situation or the sociocultural contexts of the source language and the target language. This limitation makes it unsuitable for high-stakes or sensitive communications, such as legal, medical, or governmental content, where the implications of mistranslation can be significant. In such cases, human translators must remain the final decision-makers, validating and cross-referencing AI-generated content before it is used or published.

To responsibly integrate AI-MT into workflows, organisations should evaluate whether its use is appropriate for the specific context and whether it serves a clear public or user benefit. They should understand the factors influencing AI outputs and maintain a human-in-the-loop for oversight, especially in high-stake situations. Safeguards must be developed to ensure that AI-MT supports rather than replaces human translator.

5.6 The Value of Human Professional Expertise in Translation Must be Highlighted

It has to be highlighted that translation is not simply 'language transfer', which is probably the biggest misconception about translation among technology developers of machine translation and among the public. If real communication needs to be achieved between people across languages and cultures through translation, that is not simply transferring the words from one language to another but cross-lingual and cross-cultural communication of meanings, which is much more complicated and nuanced because that involves not only different linguistic rules but also different cultural conventions, context- and culture-dependant attitudes, values and ideology.

As seen from the live machine translation of speeches, which is shown as bilingual subtitles or transcripts on the screen to the audience of conferences, the current machine translation, even powered by AI technology, is actually processing words rather than meanings though they are the real essence of human communication. In contrast, human translators and interpreters would extract meanings from words and convey the intended meanings by attending to the communicative context, and the cultural conventions, attitudes, values and ideology of both the source language and the target language. Therefore, it is typical that the machine would produce non-sensical text in its translation while it is rare that human translators and interpreters deliver non-sensical translation or interpretation. When the machine has problem in automatic speech recognition, it would spit out loads of 'rubbish' that do not make any sense; in contrast, when human interpreters have difficulty in rendering the content, they would at least convey the gist meaning by summarising and generalisation.

The value of human professional expertise in translation and interpreting is highlighted explicitly by the AIIC with its "M.A.S.T.E.R. Principles" (AIIC 2024), as follows:

Meaning: Human interpreters go beyond literal translation to capture the true intent and cultural context of a message. They understand nuance, tone, and subtext – something AI, which relies on data and algorithms, cannot fully replicate.

Awareness: Interpreting requires cultural sensitivity, ethical judgment, and emotional intelligence. Humans can challenge biases and build trust – skills that AI lacks due to its inability to understand human dynamics.

Security: Human interpreters can be held accountable and trusted with confidential information, offering a level of data protection that AI systems, which may be vulnerable to breaches, cannot guarantee.

Tone: The emotional and rhetorical tone of speech is crucial. Human interpreters can convey humour, urgency, or formality accurately, while AI often misses these subtleties, leading to miscommunication.

Engagement: Interpreting is interactive. Humans can clarify, adapt, and connect with speakers and audiences in real time, fostering engagement and rapport – capabilities AI lacks.

Reliability: Human interpreters are consistent, adaptable, and accurate, especially in high-stakes or unpredictable situations. AI may offer convenience but often requires oversight and correction.

5.7 Particular Principles for Government Institutions

It is particularly important for government institutions to put in place guidelines and policies about proper, responsible and ethical use of AI-MT in public services. The following principles proposed by "the Generative AI framework for HM Government" of the UK can be employed as the basis in formulating relevant guidelines and policies about the use of AI-MT, as follows (The Central Digital and Data Office of the HM Government of the UK 2024):

First, when using of AI-MT in public services, government institutional users must know what AI-MT is and what its limitations are. Second, government institutional users must use AI-MT lawfully, ethically and responsibly. Third, government institutional users must know how to keep AI-MT tools secure. Fourth, government institutional users should have sufficient knowledge in choosing the right AI-MT tool for the job, if not, professional translation experts should be consulted. Fifth, government institutional users must understand how to manage the full AI-MT lifecycle, if not, professional experts should be consulted.

In order to ensure compliance of the above principles, it is necessary for government institutions to provide relevant training for government agencies and officers who need to use AI-MT.

6 Conclusion

As revealed by the survey based on news reports published by public media in different countries, AI-powered machine translation tools are being used by government institutions in high-stakes public services, which involves risk of errors and serious consequences as well as ethical complexities. However, as revealed by the surveys based on searches conducted in Google Search, there is the glaring lack of guidelines about proper, responsible and ethical use of AI-MT. Through the analysis about the consequences in the use of AI-MT in government public services, this article has revealed the hidden risk of errors and ethical complexities in using AI-MT, which have been overlooked by general users and the public. Through reflections on the status quo and by drawing upon relevant discussions, the article has discussed the ethical issues involved in the use of AI-MT and proposed tentative principles for the formulation ofguidelines on proper, responsible and ethical use of AI-MT.

This study is useful in raising awareness among general users about proper, responsible and ethical use of AI-MT tools to ensure accuracy, fairness, data privacy, transparency, and accountability. It can also serve as the basis for the formulation of guidelines, protocols and policies for the use of AI-MT tools by government institutions, AI-MT developers and professional translation associations.

In terms of limitation, as the survey results about the status quo and ethical consequences of using AI-MT in high-stakes situations are mainly about government public service in UK, US and Australia, the findings from this paper are not intended to be generalised to global contexts. In consideration of the AI technological disparities in developing countries and the factor of low-resource languages in contrast with global languages such as English as a lingua franca, there might be the possibility that the risk of errors is higher with more serious consequences and ethical complexities in AI-MT of those languages in developing countries.

Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: The author has accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: None declared.
Conflict of interest: The author states no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

ABC News. 2020. Federal Government used Google Translate for COVID-19 Messaging Aimed at Multicultural Communities. https://www.abc.net.au/news/2020-11-19/government-used-google-translate-for-nonsensical-covid-19-tweet/12897200.

AIIC. 2024. "M.A.S.T.E.R. the Message" Campaign. https://masterthemessage.aiic.org/?utm_content=buffer37b0f&utm_medium=social&utm_source=twitter&utm_campaign=master-the-message-2024-aiic-campaign.

BBC News. 2019. US Uses Google Translate to Vet Refugees. https://www.bbc.co.uk/news/technology-49852308.

BBC News. 2023. NHS Interpreting Service Problems Contributed to Patient Deaths. https://www.bbc.co.uk/news/uk-england-bristol-66605536.amp.

Bowker, L. 2021. "Digital Humanities and Translation Studies." In *Handbook of Translation Studies*, Vol. 5, edited by Y. Gambier, and L. van Doorslaer, 37–44. Amsterdam: John Benjamins.

Garden Court North Chambers. 2022. Awaab Ishak Inquest. https://gcnchambers.co.uk/awaab-ishak-inquest-prolonged-exposure-to-mould-led-to-death/.

HM Court Ruling. 2022. Inquest Touching the Death of Awaab Ishak. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://nearlylegal.co.uk/wp-content/uploads/2022/11/HMC-RULING-download-from-170443.docx&ved=2ahUKEwj7jJH44K-NAxV9U0EAHareF8MQFnoECBoQAQ&usg=AOvVaw2iEJ0Ir8ojU22YMxB64iWl.

Ma, F. C., 2019. Preface. In *Digital Humanities: Knowledge and Critique in a Digital Age*, edited by Wang, X. G., et al. (trans.). Berry, D. M., and A. Fagerjord, vi–xi. Dalian: Dongbei University of Finance and Economics Press.

Mo, Q. Y., S. F. Li, A. D. Pionke, and J. Wang. 2024. "Editorial." *Digital Studies in Language and Literature* 1 (1–2): 1–8.

Open AI. 2025. Usage Policies. https://openai.com/en-GB/policies/usage-policies/.

SAFE-AI. 2024. *Guidance on AI and Interpreting Services.* https://safeaitf.org/guidance/.

The Central Digital and Data Office of the HM Government of the UK. 2024. Generative AI Framework for HM Government. https://www.gov.uk/government/publications/generative-ai-framework-for-hmg.

The Guardian. 2023a. US Immigration System Relies on AI-Powered Translation, which Jeopardizes some Asylum Applications. https://www.theguardian.com/us-news/2023/sep/07/asylum-seekers-ai-translation-apps.

The Guardian. 2023b. Home Office to Tell Refugees to Complete Questionnaire in English or Risk Refusal. https://www.theguardian.com/uk-news/2023/feb/22/home-office-plans-to-use-questionnaires-to-clear-asylum-backlog.

The Institute of Translation & Interpreting. 2023. ITI and CIOL Express Our Concerns to the Home Office. https://www.iti.org.uk/resource/iti-and-ciol-express-our-concerns-to-the-home-office.html.

Tomalin, M., B. Byrne, S. Concannon, D. Saunders, and S. Ullmann. 2021. "The Practical Ethics of Bias Reduction in Machine Translation: Why Domain Adaptation Is Better Than Data Debiasing." *Ethics and Information Technology* 23: 419–433.

Vieira, L. N. 2024. Uses of AI Translation in UK Public Service Contexts. https://www.ciol.org.uk/ai-translation-uk-public-services.

Wang, B. H. 2023. "Exploring Information Processing as a New Research Orientation Beyond Cognitive Operations and Their Management in Interpreting Studies: Taking Stock and Looking Forward." *Perspectives* 31 (6): 996–1013.

Wang, W. W., and B. H. Wang 2025. "A Survey Report on the Current Status of Using AI-Assisted MT Tools in the Chinese Language Service Industry." *Technology-Enhanced Foreign Language Teaching* (外语电化教学), 2025(2): 1–7.

Bionote

Binhua Wang
Centre for Translation and Interpreting Studies, University of Leeds, Leeds, UK
b.h.w.wang@leeds.ac.uk
https://orcid.org/0000-0003-2404-5214

Binhua Wang is Chair/Professor of interpreting and translation studies at University of Leeds in the UK, where he served as Director of the Centre for Translation Studies and Programme Manager of the MA programmes in interpreting. His research interests include interpreting and translation studies, digital humanities, AI-human interactions in T&I, and intercultural studies of Chinese literature and culture, in which he has published many articles in refereed SSCI/A&HCI/CSSCI journals and in edited volumes published by Routledge, John Benjamins, Springer, Brill, Palgrave and Bloomsbury. His recent articles have appeared in *Across Languages and Cultures*, *Digital Scholarship in Humanities*, *Perspectives*, *Babel*, *Target*, *The Translator*, and in *The Routledge Handbook of Conference Interpreting*. He has authored the monographs *New Orientations in Interpreting Studies and Interpreting Education* (Routledge forthcoming), *Theorising Interpreting Studies* (2019) and *A Descriptive Study of Norms in Interpreting* (2013), and edited with Jeremy Munday *Advances in Discourse Analysis of Translation and Interpreting* (Routledge). He serves as co-editor of *Interpreting and Society* and on the editorial boards of *Babel*, *ITT*, *Forum*, *Foreign Language Teaching & Research*, *Chinese Translators Journal*, *Translation Quarterly* and *The Linguist*.