ETHICS OF ARTIFICIAL INTELLIGENCE IN TRANSLATION AND INTERPRETING

A FIT Working Group Report

Preamble

Artificial Intelligence (AI) is transforming the way translation, and interpreting services are delivered across both the public and private sectors. AI technologies are increasingly used to support language professionals in translation, interpreting, terminology management, transcription, accessibility, and multilingual communication. While these technologies present significant opportunities for improving efficiency and expanding access to information, they also introduce complex ethical, professional and legal challenges.

This report has been prepared by the FIT Working Group on AI Ethics to examine the ethical implications of AI in the translation and interpreting profession. It identifies the core ethical principles that should guide the responsible development, deployment and use of AI technologies while recognising the continuing central role of professional translators and interpreters.

The report does not seek to discourage innovation. Rather, it promotes the responsible and ethical use of AI in ways that protect the public interest, preserve professional standards and maintain trust in multilingual communication.

The FIT AI Ethics Working Group recognises that, while AI can provide valuable assistance in supporting translation and interpreting workflows, it cannot replace the professional judgement, ethical responsibility and cultural competence of qualified translators and interpreters. Human oversight remains essential, particularly in high-stakes contexts such as law, the judiciary, healthcare, education, finance, diplomacy and other settings where inaccurate, incomplete or misleading outputs may have significant consequences for individuals, organisations and society.
The Working Group further recognises the important role of professional translation and interpreting associations in promoting the ethical use of AI and contributing to the development of appropriate governance frameworks. Through advocacy, professional guidance and engagement with

policymakers and other stakeholders, these organisations play a vital role in supporting the adoption of legal and regulatory measures, including governance mechanisms, transparency and labelling requirements, and broader regulatory frameworks, such as the European Union Artificial Intelligence Act[1], to minimise potential harm, protect the public interest and maintain confidence in multilingual communication.

Acknowledgments

This report is the result of a collaborative effort by members of the FIT AI Ethics Working Group. It reflects the collective knowledge, professional experience and thoughtful discussion of the Working Group, whose members contributed to the development, review and refinement of its content.

The Working Group extends its appreciation to all those who contributed their expertise, research, comments and constructive feedback throughout the drafting process.

Adriana Zúñiga Hernández (Chair) – ACOTIP, Costa Rica
Aleksandra Hasior – TEPIS, Poland
Andy Benzo – ATA, United States
Edita Jiráková – JTP, Czech Republic
Ginta Rudd – AIT, United Kingdom
Irina Sanders – AIT, United Kingdom
Judith Platter – OSDV, Austria
María Eugenia Poblete – COTICH, Chile

The Working Group also recognises the importance of ongoing dialogue within the translation and interpreting profession and hopes that this report will contribute to the continued development of ethical guidance for the responsible use of artificial intelligence.

Background

Artificial intelligence has developed rapidly in recent years and is now embedded in many aspects of professional practice. Machine translation, large language models, speech recognition systems and other AI-powered technologies are increasingly integrated into workflows used by translators, interpreters, organisations and public institutions.

These developments have created new opportunities but have also raised significant ethical concerns relating to professional responsibility, confidentiality, transparency, accountability, accessibility,

[1] European Union. *EU Artificial Intelligence Act*. Available at: https://artificialintelligenceact.eu/ (accessed 22 July 2026).

fairness and accuracy. As AI systems become more capable and increasingly influence professional decision-making, it is essential that ethical principles remain at the centre of their development and application.

Throughout its discussions, the Working Group identified six interrelated ethical principles that underpin the responsible use of AI within the translation and interpreting profession:

- Accessibility
- Accountability
- Accuracy
- Confidentiality
- Fairness
- Transparency

These principles form the basis of the report.

Purpose

The purpose of this report is to:

- identify the principal ethical considerations associated with the use of AI in translation and interpreting;
- encourage responsible and transparent use of AI technologies;
- support translators and interpreters in maintaining the highest professional standards;
- promote human oversight where AI is used in professional practice and especially in high-stakes contexts; and
- contribute to the wider international discussion on AI governance and ethics.

Core Ethical Principles

1. Accessibility

Accessibility is a fundamental right, as outlined in the UN Convention on the Rights of People with Disabilities[2]. It guarantees individuals with physical, intellectual and sensory impairments to fully access and exercise human rights, promoting dignity, inclusion and access to information. The

[2] United Nations, *Convention on the Rights of Persons with Disabilities (CRPD)*, adopted 13 December 2006, entered into force 3 May 2008, available at: https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html

European Accessibility Act[3] emphasizes universal benefits of accessible communication, which also supports older adults and the public. Services like alternative text formats enhance usability in various contexts, aligning with the social model of disability and fostering active societal participation. This goes along with the latest research in accessibility that states that anyone can face communication barriers under certain circumstances[4], just think about a noisy environment where subtitles for video content are crucial to understanding.

Translators and interpreters are essential to accessibility. As trained experts, they work with diverse formats and contents, bridging language and cultural gaps. They ensure communication is clear to the audience's needs in conferences, public services, legal or medical settings, etc. They do so because they understand the barriers their audience may face.[5]

AI tools, despite being marketed as accessible, often fail to deliver accessibility. They provide limited solutions like transcriptions instead of subtitles or avatars instead of sign language, placing the responsibility of activating, functioning and controlling the tool on the user. This can be particularly challenging for people with disabilities, cultural or language barriers, often resulting in frustration and "inaccessibility".

For true accessibility professional translators and interpreters must use AI tools responsibly, leveraging their skills to provide tailor-made information, prevent cognitive overload for the user and ensure added-value, not basic "tick-the box" solutions.

---

2. Accountability

Artificial intelligence is increasingly making decisions that affect our lives, from aiding legal, medical, and government professionals to influencing real-world outcomes. As these systems grow in power, so does the ethical concern of accountability. When AI makes a mistake, who is responsible?

AI companies routinely include legal disclaimers stating that they are not liable for damages caused by errors in their systems. They claim their tools are meant to assist, not replace human judgement. But in practice, many professionals use AI in ways that directly influence outcomes, and when those

---

[3] European Union, *Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services* (European Accessibility Act), available at: https://eur-lex.europa.eu/eli/dir/2019/882/oj

[4] Maaß, Christiane. 2020. Easy language - plain language - easy language plus: Balancing comprehensibility and acceptability; Berlin: Frank & Timme, p. 21

[5] Jonathan Downie, *Interpreters vs Machines: Can Interpreters Survive in an AI-Dominated World?* (London: Routledge, 2020), available at: https://www.routledge.com/Interpreters-vs-Machines-Can-Interpreters-survive-in-an-AI-dominated-world/Downie/p/book/9781138586437

outcomes are harmful, responsibility is often deflected, leaving users and the public to deal with the consequences[6].

This scenario creates an ethical vacuum. Imagine a lawyer using an AI tool to draft a legal brief only to find that citations were fabricated, or a doctor relying on an AI diagnostic suggestion that leads to a misdiagnosis. These are not fictional scenarios; they have already happened. Yet when harm is done, companies refer to their disclaimers, placing responsibility on the end user.

Ethically, this is untenable. Developers and companies are not passive observers: they choose data, training methods, disclaimers and how tools are marketed. To say users alone are to blame ignores the role and accountability[7] of those who create and profit from the technology[8].

Some governments are responding. The European Union's AI Act classifies systems by risk and imposes strict transparency, safety, and accountability requirements on high-risk applications, with penalties for non-compliance. In contrast, the United States[9] and many other countries have no binding AI regulation. Ethical frameworks[10] are non-binding, with no legal consequences for releasing harmful systems and no clear pathways for victims to seek redress.

Professional interpreters and translators may be required to use AI without sufficient training, but responsibility should not rest with them alone. Employers, institutions and professionals must share accountability through proper support and continuous development. In this shared system, accountability should be genuinely distributed, particularly because the public is rarely informed when AI is involved.

To move forward ethically, responsibility must be shared. Disclaimers alone are not enough. The question remains: when AI fails, who pays the price?

---

3. Accuracy

Artificial intelligence is increasingly relied upon to generate information, support decision-making, and automate tasks across professional and public domains. As AI systems become more capable

---

[6] Examples include *Mata v. Avianca Inc.* (2023), in which lawyers were sanctioned for submitting AI-generated fictitious case citations, and *Moffatt v. Air Canada* (2024 BCCRT 149), where Air Canada unsuccessfully argued that its chatbot was responsible for its own incorrect advice before the tribunal held the airline liable.

[7] See *The Guardian*, "AI errors: why companies must take responsibility," 24 June 2026, https://www.theguardian.com/commentisfree/2026/jun/24/ai-errors-companies-responsibility

[8] M. Coeckelbergh, "AI Accountability: Responsibility and the Challenges of Artificial Intelligence," arXiv:2209.09780 (2022), https://arxiv.org/abs/2209.09780, on the principle of distributed accountability in AI systems.

[9] VerifyWise, "State of AI Governance: Regulations in the United States (2026)," available at: https://verifywise.ai/blog/state-of-ai-governance-regulations-united-states-2026

[10] National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023, available at: https://www.nist.gov/itl/ai-risk-management-framework

and widespread, accuracy emerges as a central ethical concern. When AI produces incorrect, incomplete, or misleading outputs, the consequences can extend far beyond technical error.

AI systems are often perceived as reliable because they operate at scale and with apparent consistency. Yet accuracy in AI is probabilistic, not absolute. Outputs are generated based on patterns in training data, not on understanding, judgement, or verification. As a result, AI systems may produce plausible but false information, hallucinations, omit critical context, or replicate inaccuracies embedded in their source data. These risks are particularly acute in high-stakes settings such as law, healthcare, public administration, and education, where errors can affect rights, safety and trust.

The ethical challenge lies not only in the existence of error, but in how it is managed. When AI-generated content is presented without any revision conducted by a qualified professional, users may assume a level of precision that the system cannot guarantee. Overreliance on automated outputs can displace critical human review, turning assistance into substitution. In such contexts, inaccuracy is not merely a technical flaw; it becomes an ethical failure.

Ensuring accuracy requires shared responsibility. Developers must rigorously test systems, disclose limitations, and avoid overstating reliability. Organisations deploying AI must establish oversight mechanisms, define appropriate use cases, and require human verification where consequences are significant such as in high-stakes contexts. Professionals using AI tools must remain accountable for outcomes and resist treating automation as a substitute for expertise.

Regulatory approaches are beginning to address these concerns through requirements for transparency, documentation, and risk management, particularly for high-impact AI systems. Accuracy cannot be guaranteed, but it can be governed.

Accuracy in AI is not solely a technical objective; it is a moral obligation grounded in responsibility, trust and the protection of those affected by automated decisions. An ethical approach to AI recognizes that errors are unavoidable, but harm is not.

---

4. Confidentiality

Confidentiality is fundamentally the duty not to share information with individuals who are not authorised to receive it. As interpreters and translators, we often have access to significant volumes of sensitive information provided by clients. This places upon us both a commitment and responsibility to protect the data we work with.

However, current work conditions increasingly require translators and interpreters to rely on AI systems to enhance efficiency and productivity. As a result, the use of cloud computing and machine translation and interpreting tools has become essential. This, in turn, involves using off-site systems to store, manage, process or transmit such data.

Despite robust cloud security measures that significantly reduce the risk of cyber-attacks, data remains vulnerable to leaks or unauthorised access. Additionally, Large Language Models (LLMs) often collect and retain the information entered into them, supposedly for training purposes. Yet, their data usage policies are often unclear and inconsistent, making confidentiality an ongoing concern in the context of AI in general, and particularly generative AI.

To mitigate these risks, certain best practices should be followed:

- Use services with two-factor authentication.
- Implement strong passwords and data encryption.
- Before submitting texts to AI platforms, remove personally identifiable information, trade secrets and any other sensitive content.

Importantly, security is not a one-time product, but a continuous process that must evolve in response to emerging threats. Both FIT and national professional associations of interpreters and translators must regularly assess and disseminate updates on cybersecurity trends and best practices. This ensures their members are informed and able to make ethically sound decisions regarding the use of AI in their professional work.

---

5. Fairness

Artificial intelligence is often portrayed as objective, data-driven, and neutral. Yet AI systems do not exist outside human influence. They are trained on data created, selected, and structured by people, and as a result, they can reproduce and amplify existing biases. This highlights another key ethical issue: fairness.

Bias in AI can appear in many forms. Training data may reflect historical inequalities, social stereotypes, or institutional imbalances. When such data is used without critical oversight, AI systems may produce outcomes that disadvantage certain groups based on language, race, gender, socioeconomic status, or geographic location. These risks are especially serious when AI is used in areas such as healthcare, law, finance, education, and public services.

The ethical challenge is compounded by the perception that AI decisions are inherently impartial. Because algorithmic outputs are often framed as technical or mathematical, their results may be accepted without scrutiny. This illusion of neutrality can make biased outcomes harder to detect and easier to justify. When decisions are automated, responsibility becomes diffused, and affected individuals may have little opportunity to understand or challenge the process.

Addressing bias in AI requires more than technical fixes. It demands transparency about how systems are trained, what data they rely on and what limitations they carry. It also requires human oversight, diverse perspectives in system design and ongoing evaluation of real-world impacts. Without these safeguards, AI risks reinforcing the very inequalities it is often claimed to eliminate.

Regulatory approaches are beginning to acknowledge these concerns. Risk-based frameworks, such as those emerging in the European Union, emphasize fairness, accountability and oversight for high-impact AI systems. In contrast, lack of a clear framework, such as UNESCO's[11], leaves ethical responsibility uneven and often unenforced.

An ethical approach to AI recognizes that bias is not a flaw that can be fully eliminated, but a risk that must be managed. Fairness in AI is not automatic; it is a deliberate choice that requires vigilance, accountability and a commitment to human values.

---

6. Transparency

Artificial intelligence has transformed the field of translation. AI-powered systems can instantly translate text between languages, making international communication faster, cheaper, and more available. However, as AI translation becomes increasingly embedded in digital platforms and public services, a critical ethical issue arises: transparency.

Transparency is essential in an age when machine-generated language can mimic human expression with increasing fluency. Users should be informed when they read content that is entirely generated by an AI engine without human oversight or verified by a professional translator and/or interpreter. Unfortunately, many platforms deliver translations and/or interpreting without labels or clarification, leaving people unaware of how such information is generated and/or processed.

The lack of transparency gives rise to significant issues because AI translations and/or interpreting are not always reliable. They still struggle with idioms, legal nuances, cultural references, and tone. In high-stakes contexts, such as medical instructions or asylum applications, errors can lead to confusion, harm or violation of rights. Without knowing that a translation and/or interpreting was generated by a machine, a user may place full trust in its accuracy without challenging it.

One of the simplest ways to promote ethical transparency in AI translation is clear, visible labelling. Translations and/or interpreting should indicate their origin: whether they were fully human-generated, fully machine-generated, edited or reviewed by a professional translator and/or interpreter.

The labelling requirement fits within broader international initiatives to regulate AI. For instance, the European Union's AI Act enforces strict transparency rules for high-risk AI systems. It mandates that users be clearly informed of AI-generated content through a label. The labelling will be enforced as of August 2026. The underlying principle is the same: if people are reading a text produced or

---

[11] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, adopted by the General Conference at its 41st session, Paris, 23 November 2021, available at: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

modified by a machine, they have the right to be informed. In contrast, the United States and many other countries worldwide currently lack binding governmental regulations on AI transparency.

Moreover, labelling supports accountability. When an AI translation and/or interpreting causes harm, a lack of transparency makes it impossible to determine who is accountable. If users knew upfront that a translation and/or interpreting was unverified machine output, they could double-check it or seek human support by professionals.

Informing users whether they are reading human-made, human-reviewed, or raw machine content is not an option, it is an essential ethical obligation.

Conclusion

Artificial intelligence is expected to continue transforming the translation and interpreting profession. Ethical governance will therefore remain essential to ensuring that technological innovation serves both professionals and society responsibly.

The Working Group recognises that AI can provide valuable support for translators and interpreters when used appropriately. However, AI should complement rather than replace professional expertise, particularly in high-stakes legal, judicial, medical, educational, financial, diplomatic and other contexts where inaccurate or inappropriate outputs may have significant consequences for individuals and society.

Human oversight remains essential especially wherever AI systems influence decisions or communications that affect rights, safety or public trust. Professional translators and interpreters continue to play a critical role in ensuring linguistic accuracy, cultural appropriateness, ethical decision-making and professional accountability.

The Working Group also recognises the important role of professional translation and interpreting associations, on both national and international levels, in promoting ethical practice, supporting professional development, contributing to public policy discussions and advocating for appropriate governance frameworks. Such organisations have an important role in ensuring that appropriate legal safeguards, transparency requirements, labelling obligations and regulatory measures continue to develop alongside advances in AI technology in order to minimise harm and maintain public confidence.

The ethical principles outlined in this report should therefore be regarded as complementary and mutually reinforcing. Accessibility, accountability, accuracy, confidentiality, fairness and transparency together provide a framework for the responsible use of AI within the translation and interpreting profession.