

AI-Assisted Translation and Ethical Awareness in Translator Education: An Exploratory Study

Katty Ferreira 

Department of Languages and Cultures/CLLC, University of Aveiro, Portugal

How to cite: Ferreira, K. (2026). AI-Assisted Translation and Ethical Awareness in Translator Education: An Exploratory Study. In: 12th International Conference on Higher Education Advances (HEAd'26). Valencia, 15-18 June 2026. <https://doi.org/10.4995/HEAd26.2026.21597>

Abstract

The use of digital translation tools, particularly those based on artificial intelligence, has expanded rapidly in higher education, generating expectations of efficiency while also raising concerns about dependency, loss of competences, and professional responsibility. This exploratory study analyses the ethical and professional perceptions of a group of 10 third-year undergraduate students of Translation regarding these tools, based on a questionnaire administered in class in the course "Spanish: Translation Practice". Questionnaire data were analysed descriptively through a combination of quantitative frequencies and qualitative interpretation to explore students' views on risks, accountability, and implications for translator training. The findings are discussed in dialogue with literature on AI ethics, algorithmic opacity, and data protection, offering context-specific insights that may support pedagogical reflection and further research.

Keywords: digital translation tools; professional ethics; translator training; artificial intelligence; empirical study.

1. Introduction

In recent years, the use of automated systems in education and linguistic practice has intensified, affecting both the organisation of teaching and the production of texts and translations. Various authors (Floridi, 2024; Floros, 2021; Bietti, 2019; Borgesius, 2018, among others) have highlighted that these systems not only offer new possibilities for supporting professional work but also raise renewed ethical issues related to algorithmic opacity, data protection, technological dependency, and the redistribution of power between large corporations and users.

Within the specific field of translation, tools such as neural machine translation systems and AI-based writing assistants, such as ChatGPT, have become part of the routine repertoire of students and professionals, reshaping processes of documentation, draft generation, revision,

and post-editing. This scenario requires a rethinking of translator education that integrates technical competence with ethics and responsibility.

This study aims to analyse the perceptions and practices of a group of 10 undergraduate translation students regarding the use of digital translation and artificial intelligence tools, based on a questionnaire administered during a classroom session in the course “Spanish: Translation Practice”. Particular attention is paid to perceived benefits, identified risks, and ethical implications associated with their integration into higher education.

2. Key Theoretical Perspectives on Ethics and Automation in Translator Education

Floridi (2024) argues that the expansion of automated systems exacerbates existing problems, such as discrimination, and generates new ones by combining large volumes of data, opaque decision-making, and the capacity to intervene across multiple domains, including education and work. In this context, it is necessary to consider how the use of such systems in translator training and professional practice may amplify inequalities, consolidate biases, or dilute professional responsibility.

Floros (2021, p. 350) points out that ethical reasoning should be woven throughout translator and interpreter education, such that students are equipped to anticipate and interrogate the moral implications of their decisions in practice and in relation to the use of technologies like AI.

One of the most frequently emphasised issues in the literature is the opaque nature of many data-driven systems. Drawing on Pasquale’s (2015) expression “black boxes,” Borgesius observes that “it is often unclear for somebody why a system makes a certain decision about him or her. Because of the opaqueness of such decisions, it is difficult for people to assess whether they were discriminated against on the basis of, for instance, racial origin.” (2018, p. 10). The author stresses that this lack of transparency complicates the detection of discrimination and the contestation of decisions, which becomes particularly relevant when automated tools are used in evaluative processes or in sensitive tasks such as legal or medical translation, for example.

Alonso-Rodríguez et al. state that “the data used to train machine learning algorithms is affected by biases from certain contexts and people. These systems then internalise partial or discriminatory criteria (about ‘race’, gender, age, etc.) from these sources and can even end up amplifying them.” (2024, p. 84). As these authors show, AI-driven tools may reinforce existing disparities by favouring certain groups while marginalising others. In translation, this can appear in word or style choices that reinforce stereotypes or make minority language varieties disappear. The lack of transparency in AI systems makes it difficult to understand decision-making processes, raising issues of accountability and fairness.

AI may widen inequalities through unequal access to resources and competences. Alonso-Rodríguez et al. (2024) warn that data-based technologies tend to benefit those who already have more resources, widening gaps between institutions, countries, and social groups. As noted by these authors, “the use of AI is also expanding the digital gap, socially, as the countries and groups of people with the most resources are those who benefit from the opportunities it provides. All this compromises educational fairness.” (2024, p. 85). In this context, as Bietti (2019) argues, ethics is often misused in technology policy to justify deregulation and self-regulation.

The protection of personal data emerges as a central axis in current discussions. Floridi (2024) emphasises that AI raises questions about privacy, data protection, and individual rights. In his words, “as AI systems amass, analyse, and manipulate immense amounts of personal information, questions emerge about individual rights and safeguarding sensitive data, especially in contexts like education, work, and public health.” (2024, p.4).

Another recurring concern is the potential dehumanisation of processes (O'Brien, 2012) that have traditionally required professional judgement, contextual sensitivity, and ethical responsibility. Selwyn (2019), warns that the systematic replacement of human activities by automated processes may weaken relational and ethical dimensions that are not replaceable. Applied to translation and translator education, this reflection reinforces the need to preserve the translator's role as a responsible agent capable of interpreting context, making informed decisions, and responding to clients and readers, and to conceive translator training as a space where not only technical competences but also ethical and relational criteria are developed.

3. Methodology

This study adopts a descriptive-exploratory, cross-sectional design based on an online questionnaire administered to undergraduate translation students. The participants were 10 ($n = 10$) third-year undergraduate students of Translation at University level enrolled in the course Spanish: Translation Practice at a Portuguese university. The questionnaire was completed during a single face-to-face class session (120 minutes), with 15-30 minutes allocated for completion. Participation was voluntary and anonymous, and students were informed in advance about the academic and research purposes of the study, as well as the exclusive use of data for pedagogical and research analysis. The research instrument consisted of an ad hoc questionnaire created using Google Forms. Items were designed based on recurring ethical issues highlighted in the literature on AI and translation practice (e.g., responsibility, opacity, bias, and data protection). It was structured to collect data on students' familiarity with and frequency of use of automated translation tools; the stages of the translation process in which these tools are employed; perceived benefits and risks; levels of trust in the outputs; responsibility for errors; handling of personal data; perceptions of the added value of the

professional translator; awareness of ethical debates; and recommendations for responsible use in training and professional contexts. The questionnaire contained 23 items, including 15 closed-ended questions (single-choice, multiple-choice, and Likert-type scales) and 8 open-ended questions aimed at eliciting concrete experiences and reflective opinions. Closed-ended items allowed for the identification of general trends regarding use, perceived usefulness, trust, and risks, while open-ended responses provided qualitative insights into students' experiences, ethical concerns, and proposed strategies for responsible integration. Responses were exported in CSV format for analysis. Quantitative data were analysed descriptively through frequencies and percentages, given the small sample size, without inferential statistical testing. Qualitative responses were analysed through an inductive coding process aimed at identifying recurring themes and categories. Given the limited dataset, this qualitative component is intended as a structured interpretation of students' written comments rather than a full thematic analysis in the sense of large-scale qualitative research.

4. Results

This section presents a selection of results from the questionnaire, focusing on those items that are directly aligned with the aim of the present article, namely students' perceptions of ethical risks, responsibility, trust, data protection, and the role of the professional translator. While the questionnaire also included items on familiarity with AI tools and general use patterns, these are only briefly reported as they provide contextual information for the findings presented below. As the questionnaire was anonymous, participants are identified in this section using numerical codes (Student 1, Student 2, etc.). Percentages are included only to summarise response distribution within this small group and should not be interpreted as statistically representative.

4.1. Use of AI Tools in Translation Practice

All 10 students (100%) reported having used AI-based translation tools at some point. In terms of familiarity, 8 students (80%) indicated a medium level and 2 students (20%) a high level of familiarity. Regarding frequency of use, 5 students (50%) reported frequent use, while the remaining students indicated occasional or sporadic use. Students reported using AI tools at different stages of the translation process. Most indicated using them during drafting or proposal stages (70%, $n=7$) and during revision or reformulation (60%, $n=6$). A smaller number reported using AI before translation, during documentation or text analysis (30%, $n=3$). No student reported using AI as the sole means of translation. The most frequently reported tasks included searching for terminological equivalents (70%, $n=7$), synonym search and reformulation (60%, $n=6$), and stylistic correction (50%, $n=5$). Fewer students reported using AI to translate full segments or drafts (30%, $n=3$) or for post-editing their own translations (30%, $n=3$). Perceived benefits were primarily practical. The most frequently selected benefits were time-saving (70%,

n=7), support for cohesion and syntax (60%, n=6) and facilitation of work (50%, n=5). One student reported that AI contributed to improving translation quality.

4.2. Perceived Risks and Trust in AI Output

Students identified several risks associated with the use of AI tools. The most frequently reported concerns were dependency (80%, n=8), loss of skills and ethical issues (70%, n=7), contextual errors (60%, n=6), and lack of reliability (50%, n=5). One student summarized the risks as follows: “Dependency, contextual errors, unreliability, loss of skills, ethical issues, and reduction in the quality of texts produced, without creativity” (Student 8). When asked about trust in AI-generated translations, 9 students (90%) reported that they do not fully trust the output produced by these tools.

4.3. Perceived Added Value of the Professional Translator

When asked about the added value of professional translators in the era of AI, 40% of students (n=4) reported “much added value,” 30% (n=3) “moderate added value,” and 20% (n=2) “indispensable added value.” One student indicated that professional translators offer no added value. When specifying concrete aspects of this added value, students most frequently mentioned quality, cultural adaptation and localization, responsibility and accountability, and specialized terminological knowledge. Ethical judgement and sensitivity to communicative or emotional aspects were also mentioned by a smaller number of students. Student 9 identified “quality, cultural adaptation, specialization” as concrete benefits of human translation.

4.4. Responsibility, Data Protection, and Ethical Awareness

90% of students (n=9) indicated that the human translator is primarily responsible to the client if an AI-assisted translation contains a serious error. No student attributed responsibility to the AI tool, the provider company, or the client. One student commented: “It depends on the situation, but I believe that responsibility essentially lies with the person who decided to use it, which is sometimes the translator, but other times it is imposed by the client who is more interested in saving time.” (Student 8). Regarding data protection, 80% of students (n=8) reported that they would never input personal data into AI platforms, while 20% (n=2) would do so only with explicit written client authorization. Reported confidentiality measures included avoiding the insertion of full texts, anonymizing identifying details, and awareness that information entered into AI platforms may not remain under user control. Finally, 60% of students (n=6) reported having heard of concepts such as “AI ethics” or “responsible AI.” When asked to identify the most relevant ethical implications of AI use in translation, the most frequently selected issues were confidentiality and data protection (80%, n=8), devaluation of work and job precariousness (80%, n=8), authorship and attribution of responsibility (50%, n=5), and cultural or ideological bias (50%, n=5). Student recommendations for responsible AI

use emphasized caution against over-reliance, with 4 students explicitly warning against using AI for complete translations rather than targeted tasks like terminology or reformulation. Key themes included prioritizing confidentiality (Student 4), maintaining careful oversight of machine translation outputs (Student 3), and recognizing AI's role as support rather than substitute. Student 10 provided the most comprehensive guidance: "Be very aware that AI can be support, but not a substitute for translation work... use must be punctual and restricted... best is to reformulate the request and never copy/paste." Regarding AI integration in translator training, 5 of 10 students (50%) responded "Yes", 3 (30%) said "No", and 2 (20%) said "Maybe". Among students favouring AI integration, proposed approaches emphasized ethical training, practical application, and critical evaluation. Recommendations included teaching ethical boundaries and usage limits (Student 8), critical evaluation exercises (Student 10), practical assignment support (Student 4), optimal tool usage for efficiency (Student 2), and explicit ethics training combined with case studies: "Ethics - for new translators to understand that AI cannot be used for everything. Practical cases - since AI is here to help, at least we learn how to use it." (Student 5).

5. Discussion

The results presented in this study suggest a pattern of use of AI-based translation tools that is both widespread and cautious, reflecting a tension already identified in the literature between technological affordances and ethical awareness. Students integrate these tools pragmatically into their translation practice, mainly as support for drafting, reformulation, and terminological search, while explicitly rejecting their use as autonomous translators. This selective and instrumental use suggests that, at least at this stage of training, AI is not perceived as a replacement for human expertise but as an auxiliary resource whose value depends on the translator's judgement. Although students acknowledge concrete benefits such as time-saving and support for cohesion and syntax, trust in AI output remains limited. A clear majority of respondents stated that they do not fully trust AI-generated translations, and responsibility for errors is consistently attributed to the human translator rather than to the tool or its provider. This small-scale result illustrates issues discussed in the literature about the difficulty of delegating professional accountability to opaque systems (Borgesius, 2018; Pasquale, 2015) in relation to algorithmic decision-making. Even under external orders for AI adoption (e.g., from clients), students defend the translator's final responsibility.

Students' identification of dependency and loss of skills as major risks further reflects concerns raised in the literature regarding the long-term effects of automation on professional competence and autonomy. The fear of becoming overly reliant on AI tools echoes broader concerns about dehumanisation and the erosion of judgement and contextual sensitivity discussed by O'Brien, (2012), Selwyn (2019), among others. In the context of translator education, these concerns are particularly relevant, as training is not limited to acquiring

technical efficiency but involves developing interpretative, cultural, and ethical capacities that cannot be automated.

Students place data protection and confidentiality at the core of their ethical awareness. The fact that most respondents report avoiding the insertion of personal or sensitive data into AI platforms, or doing so only with explicit client authorization, suggests a practical awareness of the risks associated with data misuse and loss of control. This awareness reflects the concerns articulated by Floridi (2024) regarding the accumulation and manipulation of personal information by AI systems. Anonymization and partial input strategies show students aligning AI tools with professional norms amid opaque data processing.

Students also show awareness of broader professional concerns such as bias, devaluation of work, and precariousness. These responses point to a wider awareness of how AI may affect the translation profession as a whole. Such concerns echo with the arguments of Alonso-Rodríguez et al. (2024), who warn that data-driven systems tend to reproduce and reinforce existing inequalities and partial perspectives. Although students do not frame these issues in theoretical terms, their answers indicate an implicit understanding that the use of AI tools is linked to broader economic and professional dynamics, including the visibility, value, and stability of translation work.

Finally, attitudes towards the integration of AI into translator training suggest neither rejection nor uncritical acceptance. Students who support its inclusion consistently emphasise the need for ethical guidance, critical evaluation, and restricted use, rather than technical instruction alone. This position implicitly challenges approaches that frame AI merely as a productivity tool and reinforces the view that ethics should be an integral component of translator education, rather than an afterthought or a form of self-regulation, as alerted by Bietti (2019).

6. Conclusion

This study set out to explore how undergraduate translation students perceive and use AI-based translation tools, with particular attention to ethical risks, responsibility, data protection, and the role of the professional translator. Based on an exploratory questionnaire, the findings do not aim at statistical generalisation but offer situated insights into student perceptions that may support pedagogical reflection and future research. Since data were collected in a classroom context, responses may have been influenced by contextual factors and social desirability, particularly regarding responsibility and confidentiality. Nevertheless, the results suggest that AI tools are already embedded in translation training, although their use is marked by caution, selectivity, and a strong sense of human responsibility.

Rather than embracing AI as an autonomous solution, students consistently position it as a support tool whose outputs must be evaluated, corrected, and ethically managed by the

translator. Concerns about dependency, loss of competences, confidentiality, and professional accountability echo issues widely discussed in contemporary debates on AI and highlight their relevance within translator education. The added value of the human translator is primarily located in areas that resist automation, such as cultural adaptation, contextual understanding, ethical judgement, and responsibility towards clients and readers.

These findings suggest that the challenge may lie not in rejecting AI, but in integrating it while preserving professional agency, ethical awareness, and critical judgement. Translator education should therefore be conceived as a space where technological competence and ethical reflection develop together, in response to ongoing transformations in the field, without losing sight of its human and social dimensions.

References

- Alonso-Rodríguez, A. M^a. (2024). Towards an Ethical Framework for Artificial Intelligence in Education [Hacia un marco ético de la inteligencia artificial en la educación]. *Teoría de la Educación. Revista Interuniversitaria*, 36(2), 79-98. <https://doi.org/10.14201/teri.31821>
- Bietti, E. (2019). From Ethics Washing to Ethics Bashing. A View on Tech Ethics from Within Moral Philosophy. In: *Proceedings of ACM FAT* Conference (FAT* 2019)*. ACM, New York, NY, USA, 1-10. <https://doi.org/10.1145/3351095.3372860>
- Borgesius, F. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy. <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decisionmaking/1680925d73>
- Floridi, L. (2024). The Ethics of Artificial Intelligence: exacerbated problems, renewed problems, unprecedented problems - Introduction to the Special Issue of the American Philosophical Quarterly dedicated to The Ethics of AI. <http://dx.doi.org/10.2139/ssrn.4801799>
- Floros, G. (2021). *Ethics of translator and interpreter education*. In A. Shlesinger & S. B. Kim (Eds.), *The Routledge Handbook of Translation and Ethics* (pp. 338–350). Routledge.
- O'Brien, S. (2012). Translation as human–computer interaction. *Translation Spaces*. 1. 10.1075/ts.1.05obr.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.