



Generative AI and Translation Ethics: A Context-Sensitive Ethical Protocol for Iranian Translators

Maryam Vahedi Velashani 

MA Student, Department of Foreign Languages, Isf.C., Islamic Azad University, Isfahan, Iran; 1142397191@iau.ir

Leila Alinouri* 

Corresponding Author, Assistant Professor, Department of Foreign Languages, Isf.C., Islamic Azad University, Isfahan, Iran; leila1362@iau.ac.ir

Abstract

Generative AI in translation raises ethical questions about transparency, accountability, privacy, fairness, and human oversight, yet context-specific guidance, particularly in Iran, is limited. This study examined Iranian translation students' and professional translators' perceptions of AI ethics and developed a context-sensitive framework and self-assessment checklist. A single-round, Delphi-informed survey of 45 participants (15 undergraduates, 15 MA students, 15 freelance translators) used a 35-item questionnaire covering six ethical dimensions. Descriptive results showed very high agreement on human-AI collaboration and human accountability, and high agreement on disclosure, transparency, privacy, and fairness. Professionals scored lower than students on several dimensions. Drawing on these findings and an existing six-dimensional model, the study proposes a framework of transparency, accountability, human-AI collaboration, privacy, fairness, and ethical governance. Implications for education, practice, and policy include AI literacy, critical oversight, and flexible disclosure guidelines. The checklist translates these principles into practical questions for responsible AI-assisted translation.

Keywords: Accountability, AI-assisted translation, generative artificial intelligence, human-AI collaboration, Iranian translators, translation ethics

Cite this article: Vahedi Velashani, M., & Alinouri, L. (2026). Generative AI and translation ethics: A context-sensitive ethical protocol for Iranian translators. *Translation and Interpreting Research*, 3(11), 67-81.
<https://doi.org/10.22054/tir.2026.96650.1103>



© 2026 – Allameh Tabataba'i University Press

Translation and Interpreting Research is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license.

Print ISSN: 3092-6599

Online ISSN: 3060-7019

Publisher: Allameh Tabataba'i University Press

Published: 30.09.2026

Accepted: 30.08.2026

Received: 01.08.2026

1. Introduction

Artificial intelligence (AI), and generative AI (GenAI) in particular, has become embedded in nearly every professional field, and Translation Studies is at the center of this shift. AI-driven tools allow translators to process volumes of text at speeds that were implausible a decade ago, but this efficiency carries ethical costs. Translators are now routinely required to make judgments that earlier generations rarely faced: when to tell a client that AI was involved, who is answerable when machine-generated output proves wrong, and how to maintain professional standards in a market saturated with AI tools. These are practical decisions made on ordinary working days, and they call for guidance specific to the context in which translators work (Cantos-Gómez, 2025; Floridi & Cowls, 2019).

The nature of translation work is changing. Rather than translating in a purely linear fashion, many translators now spend a substantial share of their time auditing and refining AI output, supervising a process they did not originate (Jiménez-Crespo, 2025; Khany & Barzan, 2026). This shift demands a wider skill set than translator training has traditionally emphasized, combining cultural expertise with working familiarity with the tools. Fairness, accountability, and data privacy, once peripheral concerns for most translators, have moved to the center of professional self-understanding (Aytaş, 2026; Kimera et al., 2024). International statements on AI in translation exist (International Federation of Translators [IFT], 2024), but they rarely reach the day-to-day specificity translators need, for instance when deciding whether to mention AI assistance on delivery. That gap motivates the present study: not another abstract value statement, but a protocol grounded in evidence from the people who will use it.

Iran-specific evidence adds a further layer. Khany and Barzan (2026), interviewing Iranian Translation Studies students, found a mixed emotional response to AI: 60% expressed optimism about efficiency gains, half reported anxiety about deskilling, and 70% voiced concern about job displacement. Their findings suggest that the profession's future depends less on choosing between AI and human labor than on building a collaboration between the two, guided by ethical principles specific enough to be useful.

Ethical guidance for translators has not kept pace with the spread of AI tools. Much of what exists treats AI as a uniform technology and overlooks the linguistic and cultural nuance of professional translation (Ramirez-Polo & Vargas-Sierra, 2023). The gap appears especially pronounced in Iran, where translation-specific ethical guidance on AI use remains scarce (Khany & Barzan, 2026). Lacking localized standards, Iranian translators are often left to resolve ethically complex situations without a professional reference point suited to their own context.

A growing body of literature argues that AI ethics in translation requires context-sensitive reasoning rather than one-size-fits-all rules. Aytaş (2026) argues that transparency in machine translation should be treated as a situated, interpretive process rather than a fixed checkbox, negotiated dialogically between the parties involved. International instruments likewise insist that ethical AI governance respect cultural diversity and linguistic pluralism, principles that bear directly on translation (European Parliament & European Council, 2024).

This study contributes in three related ways. First, it is among the earlier empirical efforts to examine AI ethics in translation from inside an Iranian setting, addressing a geographic imbalance in a literature still weighted toward Western contexts (Bartsch et al., 2025; Ramirez-Polo & Vargas-Sierra, 2023). Second, it operationalizes a six-dimensional ethical framework as a purpose-built questionnaire, for translators. Third, it turns that framework into a practical self-assessment checklist, aiming to close some of the distance between stated ethical principle and what translators do in everyday practice (Jabak, 2025; Ren & Wang, 2025). The study is guided by four research questions:

1. How do Iranian translation students and professional translators believe AI tools should be disclosed to clients and other stakeholders?
2. What approaches to accountability do Iranian translation students and professional translators consider most appropriate when AI contributes to errors or quality issues?
3. What ethical principles do Iranian translation students and professional translators believe should govern collaboration between human translators and AI systems?
4. What practical protocols do Iranian translation students and professional translators recommend for disclosing AI tool use to clients in order to maintain trust?

2. Literature Review

Artificial intelligence, broadly understood as a computer system's capacity to perform tasks that would otherwise require human intelligence, spans machine learning, natural language processing, deep learning, and generative language models. Within Translation Studies specifically, these technologies have reshaped both the mechanics of translation and the professional responsibilities that come with it (Jiménez-Crespo, 2025; Rothwell et al., 2023).

That integration happened gradually, shaped by the broader technological and economic forces that reshaped computational linguistics (Wasserman-Rozen et al., 2024). Early machine translation systems ran on hand-built linguistic rules and produced stiff, context-blind output; statistical methods then learned from bilingual corpora and improved fluency. The most recent wave, neural machine translation, transformer architectures, large language models, and generative AI, has pushed quality further still through better contextual understanding and more natural-sounding output (Kimera et al., 2024; Lai & Gao, 2026).

This history matters here because the ethical questions attached to AI-assisted translation have shifted right alongside the technology: each generation of tools has expanded both what is possible and what needs to be managed responsibly (Shahmerdanova, 2025). Table 1 traces this progression, pairing each technological stage with the ethical concerns that came with it.

Table 1. Evolution of Artificial Intelligence in Translation Studies

Period	Main Technology	Contribution to Translation	Ethical Challenges Emerged
1950s-1980s	Rule-Based MT	First automated translation systems	Lack of transparency, quality issues
1990s-2000s	Statistical MT	Improved fluency and accuracy	Data bias, limited context
2010s	Neural MT	High-quality, natural translation	Opacity, lack of explainability
2017-2019	Transformer Models	Major leap in multilingual translation	Computational cost, embedded bias
2020-2022	Large Language Models	Human-like translation capability	Accountability gaps, misinformation
2023-Present	Generative AI	Full AI-assisted translation support	Authorship concerns, governance gaps

Note. Adapted from Cantos-Gómez (2025), Rothwell et al. (2023), and Shahmerdanova (2025).

A prominent view treats AI not as a substitute for translators but as a collaborator. Where earlier discussions tended to frame AI mainly as labor replacement, more recent work emphasizes a division of labor in which human and machine each contribute what they do best (Bartsch et al., 2025; Schmidt et al., 2025). AI supplies speed, rapid information processing, and consistency; the human translator

supplies creativity, cultural judgment, ethical accountability, and the kind of contextual sense a model cannot fully reproduce (Khany & Barzan, 2026; Ren & Wang, 2025).

Instead of treating AI as a threat to professional standing, the collaborative view treats it as a tool that, handled responsibly, can strengthen that standing. This view rests on the recognition that translation is more than linguistic transfer: fluent output does not substitute for the cultural understanding and creative judgment of a human translator.

Researchers have described several forms this collaboration can take. In what might be called a supervisory arrangement, the translator monitors AI output and steps in only when something goes wrong (Rothwell et al., 2023). A sequential arrangement has AI produce a first draft that the translator then edits (Jiménez-Crespo, 2025). An iterative arrangement has both parties cycling through generation, evaluation, and revision together (Abu-Rayyash, 2023). A more fluid arrangement lets roles shift as the project moves forward (Lai & Gao, 2026). These forms are not mutually exclusive, and a single project may draw on more than one (Khany & Barzan, 2026).

How well the partnership works depends heavily on a translator's AI literacy and willingness to remain critical of machine output, skills that most translation curricula do not yet teach explicitly (Ramirez-Polo & Vargas-Sierra, 2023). These systems are also rarely designed with translators' own input, a power imbalance that can make their use feel imposed rather than chosen (Jiménez-Crespo, 2025). Translators' unease about automation, moreover, is rooted less in fear of being outperformed than in the technology's limitations and market consequences (Vieira, 2020). Any workable ethical framework must therefore take the human and economic side of this partnership as seriously as the technical side.

Bringing AI into professional translation has opened a conversation that goes well beyond how well the software performs (Floridi & Cows, 2019; Kimera et al., 2024). The speed and accessibility gains are real, but using them responsibly means holding onto ethical commitments that protect everyone the translation touches (Aytaş, 2026; European Parliament & European Council, 2024). A growing body of work suggests that human-AI collaboration only holds up over time if it rests on clear governance (Bartsch et al., 2025; Jovari, 2024).

One influential starting point is Floridi and Cows's (2019) framework, which borrows beneficence, non-maleficence, autonomy, and justice from bioethics and adds explicability as a fifth principle. It has been criticized, though, for assumptions that reflect a largely Western frame of reference and for paying too little attention to non-Western professional contexts (Ramirez-Polo & Vargas-Sierra, 2023). This study takes that critique seriously by building a framework meant to adapt general ethical principles to the specific cultural and professional circumstances Iranian translators actually work in.

To connect abstract theory with what translators do in practice, this study adopts the six-dimensional model proposed by Al-Kadery and Almotiry (2025), which brings transparency, accountability, human-AI synergy, privacy, fairness, and ethical oversight together into a single approach. This model was chosen over broader AI ethics frameworks because it was built specifically for translation, treating professional expertise and cultural nuance as central rather than incidental (Al-Kadery & Almotiry, 2025; Ramirez-Polo & Vargas-Sierra, 2023). These six areas shaped both the survey instrument and the protocol proposed later in this article. Table 2 summarizes these six principles and the evidence behind each.

Table 2. Core Ethical Principles of AI-Assisted Translation

Principle	Definition	Key Considerations	Research Evidence
Transparency	Clear disclosure and explainability	Disclosure, process transparency	Aytaş (2026); Wasserman-Rozen et al. (2024)
Accountability	Human responsibility for output	Professional responsibility, QA	Bartsch et al. (2025); Schmidt et al. (2025)

Principle	Definition	Key Considerations	Research Evidence
Privacy	Protection of sensitive data	Data storage, ownership	Khan (2025); Kimera et al. (2024)
Fairness	Avoiding discrimination	Representation, bias detection	Ramirez-Polo & Vargas-Sierra (2023)
Human-AI Collaboration	Joint work with clear roles	Competencies, responsibility allocation	Jabak (2025); Khany & Barzan (2026)
Ethical Governance	Institutional frameworks	Professional codes, legal frameworks	European Parliament & European Council (2024); Jovari (2024)

Note. Compiled from the reviewed literature.

Several studies have examined transparency and explainability in AI-assisted systems. Aytaş (2026) carried out a thematic analysis of policy documents and academic literature and found a gap between what companies say about transparency and what they actually deliver in machine translation systems. To address that gap, the study proposed the idea of hermeneutic transparency, a relational framework tying together the technical, ethical, and contextual layers of meaning-making inside machine translation systems. On the legal side, Wasserman-Rozen et al. (2024) argued that the so-called right to explanation in AI often exists more on paper than in practice, a concern echoed by Spitzer et al. (2026), whose experiment on AI-assisted decision-making identified a misinformation effect: incorrect explanations accompanying correct AI advice impaired participants' subsequent reasoning and procedural knowledge.

Also, several studies have examined accountability in AI systems and its implications for professional practice. Using a sequential mixed-methods design, Schmidt et al. (2025) found that framing accountability around process, rather than outcomes alone, does more to build a genuine sense of responsibility among AI developers. Bartsch et al. (2025), reviewing more than 5,000 publications bibliometrically, describe AI accountability as fundamentally a sociotechnical construct rather than a purely technical one. Within translation specifically, Moorkens (2024) introduced the notion of algorithmic norms, whereby translators feel pressured to conform to the demands of algorithmic management, a pressure that sits uneasily with exercising independent professional judgment and responsibility.

Research on human–AI collaboration in translation highlights both the practical benefits of AI and concerns about its limitations and professional implications. Abu-Rayyash (2023) examined how GPT-4 performs in translator training and found that while the model handles technical tasks competently, it frequently misses the cultural and literary subtlety of the source material. Khany and Barzan (2026), interviewing Iranian Translation Studies students, found a similarly layered picture: students valued the efficiency AI offered, yet worried about job security, deskilling, and the erosion of professional agency. In a mixed-methods study combining an analysis of online discussions with a survey of 252 professionals, Rivas Ginel and Moorkens (2025) found that trust in generative AI for core translation tasks is generally low and adoption limited, with prominent concerns about data privacy and job security; some translators nonetheless use tools such as ChatGPT selectively, for example for summarizing or generating ideas.

Privacy and fairness are important ethical considerations in AI-assisted translation, particularly in relation to data protection and professional training. Khan (2025) identified data privacy as a pressing concern in institutional and legal translation, calling for stronger security protocols and more transparent data handling. Kimera et al. (2024) similarly stress that developers bear an ethical obligation to prioritize data protection within neural machine translation systems. Ramirez-Polo and

Vargas-Sierra (2023) point to a related gap in translator education, where ethical competence is often left out of technology-focused curricula altogether.

Research on AI in Translation Studies is expanding, but gaps remain. Much of the existing research concentrates on either how well the software performs or translators' general attitudes toward it, without connecting the two (Cantos-Gómez, 2025; Jiménez-Crespo, 2025). Jiménez-Crespo and Rodríguez (2025) suggest that practitioners' unease is less about any single tool than about the technological ecosystem as a whole, which strengthens the case for a multidimensional ethical framework rather than a narrow one.

A second gap concerns the absence of ethical decision-making models grounded in what translators themselves think, rather than handed down from outside the profession. Ethical principles are discussed at length in the literature but rarely converted into concrete, self-assessable indicators (Bartsch et al., 2025; Schmidt et al., 2025). Within Iran specifically, research on how AI should be integrated ethically into professional translation, while accounting for local cultural, institutional, and professional conditions, remains scarce; Khany and Barzan (2026), who focus on students' attitudes, are a rare exception.

This study sets out to close these gaps in four ways: by using a structured questionnaire to describe how Iranian translators perceive the core dimensions of AI ethics; by converting ethical principles into concrete, self-assessable indicators; by offering a practical checklist that bridges theory and everyday practice; and by proposing a context-sensitive ethical framework built for the Iranian translation profession specifically.

3. Methodology

This study used a quantitative descriptive design organized around a structured questionnaire. Selected principles of the Delphi method informed that design, most notably its reliance on knowledgeable participants able to offer informed judgments on an emerging professional issue. Consistent with the study's aims, a single-round, Delphi-informed descriptive survey was used to capture participants' perceptions at one point in time rather than across iterative rounds.

A descriptive approach fit the subject matter because the ethical questions raised by generative AI in translation remain underexplored, especially inside Iran. Because the goal was to describe perceptions rather than establish cause and effect, a descriptive survey was more defensible than an experimental or correlational design.

The study participants were drawn from three groups representing different stages of professional development and levels of practical experience in translation: undergraduate translation students, MA students in Translation Studies, and professional freelance translators. Each group comprised 15 participants, yielding a total sample of 45. This structure enabled comparisons across groups with varying levels of academic training and professional experience.

Participants were recruited through purposive convenience sampling, based on their translation background, familiarity with AI tools, accessibility, and willingness to participate. Undergraduate students were required to be in their final year and to have completed at least four semesters of translation coursework, while MA students needed to have completed at least two semesters of graduate study. Professional translators were required to have a minimum of three years of experience and to be actively involved in AI-assisted translation projects.

Although purposive convenience sampling limits the statistical generalizability of the findings, it is appropriate for exploratory descriptive research aimed at identifying patterns rather than estimating population parameters (Creswell & Creswell, 2018).

Data were collected using a structured questionnaire developed specifically for this study. Its items were informed by key themes in the literature on AI governance and translation ethics (Aytaş, 2026; Kimera et al., 2024; Ramirez-Polo & Vargas-Sierra, 2023). The final instrument contained 35 questions, including three demographic questions and 32 Likert-type items examining participants' ethical stances on generative AI. The latter were organized around the study's six core dimensions and aligned with its research questions. Responses were recorded on a five-point Likert scale ranging from strongly disagree to strongly agree. Table 3 presents the six dimensions, the number of items assigned to each, and their corresponding sources.

To ensure linguistic and cultural equivalence, the questionnaire was initially drafted in English and then translated into Persian using a forward–backward translation procedure. Two translators with expertise in Translation Studies independently produced the Persian versions. A third translator, who was not given access to the original English version, then back-translated the questionnaire into English. Discrepancies between the original and back-translated versions were subsequently discussed and resolved by the research team.

Table 3. Questionnaire Dimensions

Ethical Dimension	Number of Items	Source
Transparency and Disclosure	6	Aytaş (2026); Wasserman-Rozen et al. (2024)
Accountability and Responsibility	6	Bartsch et al. (2025); Schmidt et al. (2025)
Human-AI Collaboration	6	Jabak (2025); Khany & Barzan (2026)
Privacy and Confidentiality	5	Khan (2025); Kimera et al. (2024)
Fairness and Bias Mitigation	5	Ramirez-Polo & Vargas-Sierra (2023)
Ethical Governance and Professional Practice	4	European Parliament & European Council (2024); Jovari (2024)

The four research questions were mapped onto specific item sets: Q1-Q6 addressed RQ1 (Transparency and Disclosure); Q7-Q12 addressed RQ2 (Accountability and Responsibility); Q13-Q18 addressed RQ3 (Human-AI Collaboration); and Q19-Q32 addressed RQ4 (Practical Protocols and Ethical Governance). Privacy and Fairness were not treated as separate research questions but as cross-cutting dimensions folded into the item sets for RQ2 and RQ4.

Content validity was established through a close review of the relevant literature, from which items were derived in alignment with the conceptual framework and research questions. Face validity was examined by a three-member expert panel comprising a Translation Studies specialist with 12 years of experience, an AI ethics in translation specialist with 8 years of experience, and an educational research specialist with 15 years of experience. Based on their feedback, minor revisions were made to improve the clarity and precision of individual items.

Data were collected electronically via Google Forms over six weeks, distributed through email, translation-related social media groups, and professional networks. Every response was checked for completeness, duplication, missing data, and response consistency, and any questionnaire with substantial gaps or inconsistencies was excluded.

Before completing the questionnaire, participants were told the study's purpose, the voluntary nature of participation, the confidentiality procedures in place, and their right to withdraw at any point. Informed consent was obtained beforehand, and responses were collected anonymously throughout. All information was handled confidentially, no identifying details were collected, and findings are reported only in aggregate form; the data were stored on a password-protected computer accessible solely to the researcher.

Data were analyzed using Microsoft Excel. Given the single-round, Delphi-informed descriptive design, rather than an experimental or comparative one, analysis relied on descriptive statistics only: frequencies, percentages, means, and standard deviations to surface the prevailing patterns in participants' ethical attitudes. No inferential tests such as ANOVA, correlation, or factor analysis were run, and the study makes no claim of statistical significance between groups. Mean scores were interpreted using the following bands: 1.00-1.80 (very low agreement), 1.81-2.60 (low agreement), 2.61-3.40 (neutral), 3.41-4.20 (high agreement), and 4.21-5.00 (very high agreement).

4. Results

4.1. Demographic Characteristics

The final sample comprised 45 participants, 15 in each of the three groups (33.3% each). Female participants made up 60.0% of the sample and male participants 40.0%. Just under half (48.9%) fell within the 18-25 age range, unsurprising given the inclusion of undergraduate and graduate students, while professional translators spanned a wider age range. Table 4 presents the full breakdown.

Table 4. Demographic Characteristics of Participants

Characteristic	Category	Frequency	Percentage
Gender	Male	18	40.0%
	Female	27	60.0%
Age Group	18-25 years	22	48.9%
	26 years and above	23	51.1%
Group	BA Translation Students	15	33.3%
	MA Translation Studies Students	15	33.3%
	Professional Freelance Translators	15	33.3%

4.2. Results for Research Questions 1-4

Table 5 presents the mean scores for each research question by group; the sections below interpret each result in turn.

Table 5. Summary of Mean Scores by Research Question and Group

Research Question	BA Students	MA Students	Freelancers	Total	SD
RQ1: Transparency and Disclosure	4.09	4.11	4.00	4.07	0.76
RQ2: Accountability and Responsibility	4.65	4.51	4.43	4.53	0.64
RQ3: Human-AI Collaboration	4.75	4.53	4.59	4.62	0.62
RQ4: Practical Protocols	4.20	4.14	4.03	4.12	0.76
Grand Mean	4.42	4.32	4.26	4.33	0.68

4.2.1. Transparency and Disclosure

Responses across all three groups were favorable toward transparency and disclosure in AI-assisted workflows (total mean = 4.07). Undergraduate ($M = 4.09$) and MA students ($M = 4.11$) showed a slightly stronger preference for openness about AI use than the professional group ($M = 4.00$). Asked whether translators should discuss the extent of AI use with clients before a project (Item Q2), more than three-quarters of the sample agreed, with only 6.7% opposed. A similar pattern held for Item Q3, where most participants agreed that clients have a right to know when AI is involved.

4.2.2. Accountability and Responsibility

A strong consensus emerged that ultimate responsibility for translation quality has to rest with the human professional ($M = 4.53$, $SD = 0.64$). The item on the need for human oversight (Q9) drew very high endorsement ($M = 4.71$), with over 70% strongly agreeing and no one disagreeing. Accountability thus appears to be treated as a professional given rather than a matter of preference.

4.2.3. Human-AI Collaboration

Participants expressed strong belief in the value of human-AI collaboration ($M = 4.62$, $SD = 0.62$). The highest-rated of the items reported here was the statement that combining human and AI effort improves translation quality (Q17, $M = 4.74$). This pattern suggests that participants regard human judgment and ethical training as central to collaboration with AI.

4.2.4. Practical Protocols

Findings point to clear support for standardized protocols governing AI disclosure ($M = 4.12$, $SD = 0.76$). The call for formal professional guidelines (Q30) resonated strongly ($M = 4.29$). Students showed a somewhat greater inclination toward proactive disclosure than freelancers did, a difference that may reflect training and career stage, although the design cannot separate these from age.

4.3. Cross-Question Comparative Summary

Participants placed the most weight on collaboration ($M = 4.62$) and accountability ($M = 4.53$), suggesting that human oversight is regarded as the most critical element. Transparency and practical protocols also drew high, though somewhat lower, agreement. Group differences were small and not uniform: professional translators gave the lowest mean on RQ1, RQ2, and RQ4, MA students on RQ3, and undergraduates the highest mean on RQ2 to RQ4. In keeping with the exploratory nature of the design, these should be read as observed patterns rather than statistically confirmed differences.

4.3.1. Cross-Cutting Dimensions: Privacy and Fairness

Privacy and fairness were not set up as standalone research questions but were woven into the item sets for accountability (RQ2) and practical protocols/governance (RQ4). Items on data protection and consent (Q19-Q23) and on bias auditing and cultural sensitivity (Q24-Q28) sat within these sets. Mean agreement was high for privacy and confidentiality ($M = 4.07$) and for fairness and bias mitigation ($M = 4.11$), alongside accountability ($M = 4.53$) and human-AI collaboration ($M = 4.62$), which suggests that participants treat data protection and cultural fairness as built into responsible AI-assisted translation rather than as an afterthought (Khan, 2025; Kimera et al., 2024; Ramirez-Polo & Vargas-Sierra, 2023).

4.3.2. Transparency: A Context-Sensitive Principle

The broad endorsement of transparency (overall $M = 4.07$) confirms its place as a cornerstone of professional integrity for Iranian translators, but the difference between students and professionals merits attention. Students scored slightly higher, which may reflect the more idealized ethics taught at university, whereas practitioners may take a more pragmatic line shaped by market competition and the demands of client contracts (Khany & Barzan, 2026); these small differences were not tested inferentially.

This pattern fits transparency as a context-sensitive principle that has to be balanced against professional realities rather than applied as a blanket rule, a reading consistent with Aytaş's (2026) call to treat transparency as situated and dialogic rather than uniform. It also echoes Wasserman-Rozen et al. (2024), who argue that meaningful transparency depends on professional judgment and context as much as on formal disclosure.

4.3.3. Accountability: The Non-Negotiable Foundation

Accountability produced one of the study's clearest findings ($M = 4.53$). Very high agreement on the necessity of human review ($M = 4.71$) may indicate a reconfiguration of professional identity in the AI era: rather than being displaced, translators appear to be moving into a role centered on critical evaluation and ethical oversight, effectively the last safeguard in an AI-assisted workflow (Jiménez-Crespo, 2025).

These descriptive results align with the view that responsible AI adoption depends on real human oversight and continuous critical evaluation (Schmidt et al., 2025; Bartsch et al., 2025). Strong agreement that translators need skills to evaluate AI output critically (Q15, $M = 4.53$) and training in AI ethics (Q16, $M = 4.49$) suggests that capable AI tools raise rather than lower the competencies expected of translators, including AI literacy, ethical awareness, and the ability to detect inaccuracies and bias (Spitzer et al., 2026).

4.3.4. Human-AI Collaboration: An Ethical Process

The high agreement on human-AI collaboration (overall $M = 4.62$) suggests participants treat AI integration not as a purely technical matter but as an ethical process that demands sustained critical engagement. The strongest agreement among the items reported, on the statement that combining human and AI effort increases quality ($M = 4.74$), suggests participants see AI as genuinely supportive technology, provided it works under human supervision and inside clear ethical limits (Abu-Rayyash, 2023).

4.4. Practical Protocols and the Integrated Framework

Support for clear professional guidelines (overall $M = 4.12$) reflects a recognition that ethical practice cannot rest on individual judgment alone. Professional translators' lower scores on three of the four research questions point to a recurring tension: ethical ideals are widely embraced in principle, but putting them into practice runs up against the complexities of real working conditions (European Parliament & European Council, 2024; Jovari, 2024).

4.4.1. Theoretical and Practical Implications

These findings add a human-centered strand to the literature on Translation Studies and AI ethics, showing that generative AI does not remove the need for human translators so much as reshape their role and widen the competencies the job demands (Jabak, 2025; Jiménez-Crespo, 2025). Where earlier discussions stayed at the level of universal principle, these findings suggest that how principles such as transparency and accountability are interpreted in practice depends heavily on cultural context, educational background, professional tradition, and working conditions. Ethical AI governance in translation, in other words, needs to combine universal values with real contextual flexibility (Bartsch et al., 2025; Schmidt et al., 2025).

The main contribution lies in operationalization. Earlier work has laid out essential ethical principles (Al-Kadery & Almotiry, 2025; Floridi & Cowls, 2019), but relatively few studies have translated them into practice-oriented tools for translators. The self-assessment checklist offered here gives Iranian translators a tool they can use in practice.

Practically, translators stand to benefit from building AI literacy and practicing straightforward transparency: disclosing AI use to clients clearly and early, exercising sound judgment when reviewing AI output, and pursuing ongoing professional development in AI ethics. Professional associations should draft clear, context-sensitive guidelines covering transparency, accountability, privacy, and fairness, and policymakers should build regulatory frameworks that support responsible AI use without undercutting translators' professional rights or client interests (European Parliament & European Council, 2024).

From an educational standpoint, the somewhat higher agreement among undergraduates on RQ3 suggests the next generation of translators will need competencies beyond linguistic and cultural knowledge, and university programs should build in coursework on AI literacy, critical evaluation, and ethical reasoning (Abu-Rayyash, 2023; Ramirez-Polo & Vargas-Sierra, 2023; Ren & Wang, 2025). Rather than pushing students toward full dependence on AI tools or reflexive resistance, translation education should cultivate a balanced view in which AI is a supportive technology that still requires continuous human judgment (Khany & Barzan, 2026).

Institutionally, the weight participants gave to practical protocols (RQ4) and human accountability (RQ2) suggests that professional trust calls for protocols that are clear yet flexible enough to cover disclosure, confidentiality, accountability, data protection, and quality assurance together. Associations, educational institutions, and developers should keep ethical standards current as a shared undertaking (Bartsch et al., 2025).

4.4.2. Proposed Context-Sensitive Ethical Framework

Building on these descriptive findings and grounded in the six-dimensional framework proposed by Al-Kadery and Almotiry (2025), this study proposes a context-sensitive ethical framework for Iranian translators, organized around six interconnected pillars:

1. Dimension 1: Transparency and Responsible Disclosure. AI use should be disclosed to clients before a project begins; the extent of AI involvement should be communicated clearly; and translators should maintain transparency throughout the translation process rather than only at the outset.
2. Dimension 2: Human Accountability and Final Authority. Final authority over translation quality remains with the human translator; AI-generated output must be critically reviewed

before it is accepted; and translators are accountable for errors that originate in AI-assisted work.

3. Dimension 3: Human-AI Collaboration. AI functions as a supportive tool rather than a replacement for human judgment; collaboration needs clear role definition and a clear allocation of responsibility; and translators should maintain professional autonomy over decision-making throughout.
4. Dimension 4: Privacy and Confidentiality. Client data must be protected at every stage of AI-assisted translation; AI tools used should comply with recognized data protection standards; and clients should give informed consent before their data is processed through AI systems.
5. Dimension 5: Fairness and Bias Mitigation. AI output should respect cultural diversity and linguistic variation; translators should actively monitor for bias in AI-generated translations; and inclusive practices should guide which AI tools get selected and how they are used.
6. Dimension 6: Ethical Governance and Professional Standards. Professional associations should establish clear AI ethics guidelines; educational institutions should build AI ethics into their curricula; and ethical standards should be reviewed and updated on an ongoing basis rather than fixed once and left alone.

This framework is operationalized as a practical checklist (see Table 6).

Table 6. Ethical Protocol Checklist for Translators

Dimension	Ethical Indicator	Self-Assessment Question	Yes	No	N/A
Transparency	Disclosure of AI use	Did I inform the client about AI use before starting the project?	☐	☐	☐
Transparency	Process transparency	Did I document the AI tools and processes used in this translation?	☐	☐	☐
Accountability	Human review	Did I review all AI-generated content before finalizing the translation?	☐	☐	☐
Accountability	Quality assurance	Did I document the quality assurance procedures followed?	☐	☐	☐
Human-AI Collaboration	Role definition	Did I clearly define my role versus the AI's role in this project?	☐	☐	☐
Human-AI Collaboration	Professional judgment	Did I exercise professional judgment in accepting or rejecting AI outputs?	☐	☐	☐
Privacy	Data protection	Did I ensure secure handling of client data throughout the process?	☐	☐	☐
Privacy	Client consent	Did I obtain client consent for AI data processing?	☐	☐	☐
Fairness	Bias monitoring	Did I check AI outputs for cultural or linguistic bias?	☐	☐	☐
Fairness	Cultural sensitivity	Did I ensure cultural appropriateness of the final translation?	☐	☐	☐
Ethical Governance	Professional standards	Did I comply with professional ethics guidelines?	☐	☐	☐
Ethical Governance	Continuous learning	Have I sought professional development opportunities to improve my understanding of AI ethics in translation?	☐	☐	☐

5. Conclusion

This study examined the ethical landscape surrounding generative AI in translation and used what it found to build a context-specific protocol for Iranian practitioners. The results show a consistent commitment to transparency, human accountability, and collaborative practice, along with a modest but recurring difference between student and professional perspectives that may reflect the practical realities of working in the field.

Generative AI changes what translation work looks like, but it does not replace the human translator; it expands the expertise the job requires (Jiménez-Crespo, 2025). By turning these findings into a practical checklist, the study offers professionals a living resource they can use directly and update as the technology and the profession change together.

Several limits deserve acknowledgment. First, the findings are rooted in Iran's specific socio-cultural setting, which limits how far they generalize elsewhere. Second, the exploratory, single-round design captured initial perspectives rather than a settled consensus, and reliance on self-reported data may not fully match actual translation behavior. Third, the convenience sample of 45 calls for caution in any broader interpretation, and group differences were not tested inferentially. Fourth, the questionnaire was validated through literature-based content review and expert judgment only, and the proposed framework and checklist have not yet been tested in practice.

The framework opens several directions for further work. Cross-cultural and comparative studies are needed to see how cultural, institutional, and economic context shapes translators' ethical perceptions of AI, including comparisons between Iranian translators and their counterparts elsewhere, and between differently regulated environments such as the EU and Iran (Bartsch et al., 2025). Multi-round Delphi studies involving wider stakeholders, such as agency managers, technology developers, and end users, could refine and validate the framework (Skulmoski et al., 2007), and larger probability-based samples analyzed with inferential statistics would allow the formal hypothesis testing this exploratory design could not support.

Observational and experimental research into how translators actually interact with AI tools, through eye-tracking, keystroke logging, or think-aloud protocols, would add depth beyond self-reported data (Schmidt et al., 2025). Longitudinal studies could clarify whether the caution shown by professional translators persists over time or gives way to greater trust as familiarity grows (Khany & Barzan, 2026). Finally, intervention studies applying the proposed protocol in real translation settings and tracking its effect on decision-making, client trust, and translation quality would offer the clearest test of these principles.

AI Declaration

Two AI tools were used in preparing this manuscript. Grammarly was used for grammar and readability editing. Claude (Sonnet 5.5, Anthropic) was used in October 2026 to align the manuscript with the journal's formatting guidelines, to edit language and shorten text to meet the word limit, to suggest corrections to wording and source descriptions, and to check reference details against online sources. All suggestions were reviewed and verified by the author against the study data and the cited sources. AI was not used to collect or analyze data, to generate results, or as an author. The author takes full responsibility for the accuracy, originality, and integrity of this work.

References

- Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E*, 10, 259-301. <https://doi.org/10.51287/cttl20239>
- Al-Kadery, S. A. K., & Almotiry, M. M. G. (2025). Artificial intelligence and ethical practices: A study in the realm of translation. *International Journal of Linguistics and Translation Studies*, 6(2), 1-15. <https://doi.org/10.36892/ijlts.v6i2.563>
- Aytaş, G. (2026). Transparency and accountability in machine translation: Are new ethical norms emerging in the age of AI? *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 26(1). <https://doi.org/10.18037/ausbd.1811507>
- Bartsch, S. C., Nguyen, L. H., Schmidt, J.-H., Du, G., Adam, M., Benlian, A., & Sunyaev, A. (2025). The present and future of accountability for AI systems: A bibliometric analysis. *Information Systems Frontiers*, 27, 2463-2484. <https://doi.org/10.1007/s10796-025-10636-9>
- Cantos-Gómez, P. (2025). AI in specialised translation: Terminology, quality and the evolving role of human translators. *Ibérica*, 50, 19-44. <https://doi.org/10.17398/2340-2784.50.19>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- European Parliament & European Council. (2024). *The EU Artificial Intelligence Act* (Regulation (EU) 2024/1689). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- International Federation of Translators. (2024). *Machine translation in the age of artificial intelligence* [Position paper]. Babel. <https://doi.org/10.1075/babel.00411.ift>
- Jabak, O. (2025). AI-Human collaboration for faithful translation (AHCFT): A theoretical framework. *Dragoman*, 23, 178-191. <https://doi.org/10.63132/ati.2025.aihuma.5125>
- Jiménez-Crespo, M. A. (2025). Human-centered AI and the future of translation technologies: What professionals think about control and autonomy in the AI era. *Information*, 16(5), Article 387. <https://doi.org/10.3390/info16050387>
- Jiménez-Crespo, M. A., & Rodríguez, S. A. (2025). Is it AI or PE that worry translation professionals: Results from a human-centered AI survey. In P. Bouillon et al. (Eds.), *Proceedings of Machine Translation Summit XX: Volume 1* (pp. 407–419). European Association for Machine Translation. <https://aclanthology.org/2025.mtsummit-1.32/>
- Jovari, B. (2024). Artificial intelligence ethics in organizational human resources management. *International Journal of Management, Accounting and Economics*, 11(7), 930-950. <https://doi.org/10.5281/zenodo.12752414>
- Khan, M. I. (2025). Impact of AI on legal and institutional translation: Issues and strategies. *International Journal of Translation and Interpretation Studies*, 5(6), 1–15. <https://doi.org/10.32996/ijtis.2025.5.6.1>
- Khany, R., & Barzan, P. (2026). Optimism, anxiety, and ethics: Translation students' attitudes and stances on AI-assisted workflows and professional identity. *Media and Intercultural Communication: A Multidisciplinary Journal*, 4(1), 1-17. <https://doi.org/10.22034/mic.2025.535661.1051>

- Kimera, R., Kim, Y.-S., & Choi, H. (2024). Advancing AI with integrity: Ethical challenges and solutions in neural machine translation. *arXiv*. <https://doi.org/10.48550/arXiv.2404.01070>
- Lai, J., & Gao, L. (2026). A study on human-AI collaboration modes in lyric translation from a translator-centered perspective. *International Journal of Linguistics, Literature and Translation*, 9(2), 166–171. <https://doi.org/10.32996/ijllt.2026.9.2.17>
- Moorkens, J. (2024). “I am not a number”: On quantification and algorithmic norms in translation. *Perspectives: Studies in Translation Theory and Practice*, 32(3), 477-492. <https://doi.org/10.1080/0907676X.2023.2278536>
- Ramirez-Polo, L., & Vargas-Sierra, C. (2023). Translation technology and ethical competence: An analysis and proposal for translators’ training. *Languages*, 8(2), Article 93. <https://doi.org/10.3390/languages8020093>
- Ren, X., & Wang, R. (2025). Integrating human-AI collaboration into translation education: A comprehensive protocol for assessment, diagnosis, and strategy development. *PLoS ONE*, 20(12), Article e0338089. <https://doi.org/10.1371/journal.pone.0338089>
- Rivas Ginel, M. I., & Moorkens, J. (2025). Translators’ trust and distrust in the times of GenAI. *Translation Studies*, 18(2), 283-299. <https://doi.org/10.1080/14781700.2025.2507594>
- Rothwell, A., Moorkens, J., Fernández-Parra, M., Drugan, J., & Austermuehl, F. (2023). *Translation tools and technologies*. Routledge. <https://doi.org/10.4324/9781003160793>
- Schmidt, J.-H., Bartsch, S. C., Adam, M., & Benlian, A. (2025). Elevating developers’ accountability awareness in AI systems development: The role of process and outcome accountability arguments. *Business & Information Systems Engineering*, 67(1), 109-135. <https://doi.org/10.1007/s12599-024-00914-2>
- Shahmerdanova, R. (2025). Artificial intelligence in translation: Challenges and opportunities. *Acta Globalis Humanitatis et Linguarum*, 2(1), Article 02500108. <https://doi.org/10.69760/aghel.02500108>
- Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research*, 6, 1–21. <https://doi.org/10.28945/199>
- Spitzer, P., Holstein, J., Morrison, K., Holstein, K., Satzger, G., & Kühl, N. (2026). Don’t be fooled: The misinformation effect of explanations in human-AI collaboration. *International Journal of Human–Computer Interaction*, 42(12), 8996-9024. <https://doi.org/10.1080/10447318.2025.2574511>
- Vieira, L. N. (2020). Automation anxiety and translators. *Translation Studies*, 13(1), 1-21. <https://doi.org/10.1080/14781700.2019.1694567>
- Wasserman-Rozen, H., Gilad-Bachrach, R., & Elkin-Koren, N. (2024). Lost in translation: The limits of explainability in AI. *Cardozo Arts & Entertainment Law Journal*, 42(2), 391–438. <https://larc.cardozo.yu.edu/cardozoaelj/vol42/iss2/5>