المجلة العربية لعلم الترجمة
Arabic Journal for Translation Studies
Early View
ISSN: 2750-6142
OPEN ACCESS

Accuracy, Risk, and Translator Agency in AI-Assisted Specialised Translation

Eyhab Bader Eddin

Dhofar University, Dhofar, Oman

Corresponding author: linguisteyhab@hotmail.com

ORCID iD : 0000-0003-0096-6334

Received	Accepted	Published online
3/3/2026	22/8/2026	31/8/2026

doi : 10.63939/ajts.9sr3gw56

Cite this article as: Bader-Eddin, E. (2026). Accuracy, Risk, and Translator Agency in AI-Assisted Specialised Translation. *Arabic Journal for Translation Studies, 5*(16). https://doi.org/10.63939/ajts.9sr3gw56

Abstract

This article discusses the role of human–AI collaboration in specialised translation, focusing on accuracy, risk and translator agency in high-stakes legal, medical, technical, financial and institutional contexts. It critically synthesises recent research on machine translation, large language models, post-editing, terminology management, translation quality and professional agency, using a conceptual and analytical approach. The study does not present primary empirical data; rather, it attempts to establish a framework for assessing machine-generated output within specialised translation workflows. It is theoretically innovative in combining accuracy, risk and agency within a single relational model and demonstrating how each translation decision is shaped by system output, professional judgement and potential consequences. Accuracy is therefore understood as a multidimensional requirement encompassing linguistic, terminological and conceptual precision, as well as consistency, register, genre conventions and functional appropriateness. Risk refers to the potential for output to be inaccurate, incomplete or inadequately validated, resulting in errors that could affect legal interpretation, medical safety, technical procedures, institutional responsibility or end-user decisions. These concepts are presented in a six-dimensional analytical model that includes system contribution, accuracy issue, risk level, translator response, decision rationale and impact on final quality. The article argues that computational tools can be used for drafting, terminological support, reformulation, post-editing and quality assessment, but cannot serve as stand-alone authorities in specialised translation. Translator agency is thus redefined rather than eradicated: the translator must assess, edit or reject system output, explain decisions and take responsibility for the final product. The proposed framework offers a basis for future conceptual, empirical, pedagogical and professional studies.

Keywords: AI-Assisted Translation, Human–AI Collaboration, Specialised Translation, Translation Accuracy, Translation Risk, Translator Agency

1. Introduction

The rapid development of artificial intelligence has significantly reshaped the modern practice of translation, especially with neural machine translation, large language models, automatic post-editing, terminology-support tools, and AI-based quality evaluation. Translation is no longer perceived as a human process that is sometimes facilitated by technology but increasingly operates within hybrid workflows where human translators engage with algorithmic systems at various points of text production, evaluation, revision and delivery. The recent studies of machine translation and large language models demonstrate that AI systems can generate more fluent multilingual text and possibly contribute to the productivity of translation, quality evaluation, and integration of terminology (Kenny, 2022; Kocmi & Federmann, 2023; Moslem et al., 2023; Zhu et al., 2024). Nevertheless, the technological advancement also poses some significant questions concerning reliability, accountability and the role of professional translators that is changing.

Such questions are especially pressing when it comes to specialised translation where quality cannot be reduced merely to surface fluency or grammatical correctness. In other domains like medical, legal, technical, financial and institutional translation, precision in terms of terminology, conceptual sufficiency, domain expertise, genre norms, and sensitivity to the communicative implications of inaccuracy are all crucial. A translation can be fluent and yet distort a legal requirement, medical instruction, technical procedure.or institutional policy. This is the reason why specialised translation is more cautious and domain sensitive in the application of AI in comparison to general-purpose translation. Even though AI can help with drafting, finding terminology, paraphrasing and consistency verification, it cannot be considered an independent authority in situations where mistakes can have a professional, legal, financial, or end-user impact.

This article focuses on three dimensions of AI-assisted specialised translation, namely. accuracy, risk, and translator decision-making. It considers accuracy not merely as linguistic correctness, but a multidimensional requirement that entails terminology, meaning, consistency, register and functional adequacy. Risk can be defined as the potential outcomes of erroneous or poorly checked AI output in a high-stakes communicative situation. The human processes that involve evaluating AI suggestions, accepting or rejecting them, revising, or re-evaluating them are known as translator decision-making. In this regard, the role of the translator extends beyond post-editing to incorporate a sense of judgement, responsibility and professional agency.

The main idea of the article is that AI could be used to aid specialised translation, though only under the supervision of a human, and in a risk-conscious practice of a professional translator. The collaboration of humans and AI must thus be seen as a reconfiguration of the expertise of translators, rather than a substitution. The research questions covered in the article are as follows: How to conceptualise accuracy in AI-assisted specialised translation? What are some of the risks involved in the use of AI in

high-stakes translation situations? What are the decision-making processes and agency of translators working with the output of AI? The article begins by describing the collaboration between humans and AI in specialised translation, moves on to the issue of accuracy and risk, and finally describes the decision-making of translators and suggests an analytical framework to be used in future studies and practice.

2. Literature Review

The latest translation technology literature is more likely to introduce AI-assisted translation as a transition to viewing machine translation as an independent instrument to more human-AI processes. Neural machine translation, large language models, terminology-management systems, automatic post-editing, quality estimation, and AI-assisted revision may be considered AI-assisted translation. Human-AI collaboration, however, cannot be perceived as an equal cooperation between human and machine. Instead, it can be described as a mediated process where AI systems generate, suggest, retrieve, reformulate, or evaluate linguistic material, but human translators are still tasked with the interpretation, validation, revising, and ultimate responsibility. Research on large language models has demonstrated the ability of such systems to assist in translation and post-editing, but also emphasise the importance of caution since AI output can be fluent, convincing, and yet unreliable in meaning or terminology (Kenny, 2022; Raunak et al., 2023).

One of the research strands is the potential of AI in translation in terms of productivity and quality. GPT-based post-editing Work indicates that large language models can enhance output of neural machine translation and minimise some of the large-scale errors, particularly with automatic post-editing (Raunak et al., 2023). Similarly, contextual refinement research shows that LLMs could be useful in sentence- and document-level post-editing, leveraging broader textual context and human corrections made in the past (Koneru et al., 2023). Nevertheless, the main weakness of these studies is also disclosed: performance improvements do not eliminate the need for expert review. Some errors might be fixed by AI systems, and others might be added, such as hallucinated edits or semantically unsafe edits. This complicates their application especially in the workplace where quality is not just a matter of fluency.

A different line of research deals with the translation of terminology and translation of specialised domains. Moslem et al. (2023) demonstrate that large language models can assist in incorporating pre-approved terminology into machine translation processes, especially through the use of terminology-constrained post-editing. Berger et al. (2024) also believe that the indicators of human error can be used to steer the LLMs to more precise corrections in technical fields. These works are significant as they show that AI can be more practical with the limitation of domain-specific resources and human involvement. Simultaneously, they also demonstrate that specialised translation cannot be based on the output of generic AI only. Terminological accuracy is frequently based on

pre-approved terms, translation memories, expert correction, and domain sensitive validation.

Another branch of research is related to the agency of translators, professional control, and human-centred AI. Recent debates on human-centred translation technologies highlight the fact that attitudes to control and autonomy are the key factors in the adoption of AI in the profession by translators (Jiménez-Crespo, 2025). This is particularly pertinent to specialised translation, which contrasts with general translation as it demands conceptual accuracy, knowledge of genre, institutional knowledge and knowledge of consequences. In translation of law, medicine, and technology, errors may affect rights, treatment, safety, compliance, or action of the user. That is why the work of a translator cannot be reduced to routine post-editing. The translator is the evaluator, verifier, risk assessor and decision maker.

Current methods shed light on some key but mostly distinct areas of technology-mediated translation. Research on automatic evaluation and post-editing mainly focuses on system performance and the correction of machine-generated output (Kocmi & Federmann, 2023; Koneru et al., 2023; Raunak et al., 2023). There are terminology-oriented approaches that aim to introduce approved domain-specific terms and enhance terminological control (Moslem et al., 2023), and there are human-centred approaches that emphasise professional autonomy, informed technology use and translators' control over automated processes (Kenny, 2022; Jiménez-Crespo, 2025). These views have merit but tend to treat output quality, terminology, risk and professional agency as separable issues rather than as a set of interdependent issues within the same decision-making process.

The resulting gap is not simply a lack of research on the translation of texts with the help of AI systems, but a lack of an integrated description of the type of translation accuracy problem that may arise from a particular system contribution, the level of risk posed by such a problem in a specific context, and the translator's response to it, which can affect the final quality of the translated text. This article fills this gap by linking these stages through six dimensions of analysis. Its contribution is relational, as it describes not only whether machine output is acceptable but also how professional judgement mediates between technological assistance, its consequences and responsibility for the final translation.

In general, the literature implies that AI can assist specialised translation, but it is only effective with human oversight and guidance, the control of terminology, and risk-sensitive decision-making. The performance of AI, post-editing, integration of terminology, and attitudes of translators have been studied. Nevertheless, the interaction of accuracy, risk, and decision-making of translators as one analytical issue in high-stakes specialised translation has received less attention. This gap motivates the present article which does not regard human-AI collaboration as a technological workflow per se, but as a professionally responsible process that is influenced by domain knowledge, risk assessment, and agency of the translator.

3. Methodological/conceptual approach

The article is based on a conceptual and analytical approach as opposed to empirical research design. It does not present primary data of interviews, surveys, experiments or corpus analysis. Rather, it elaborates a critical review of the recent research on AI-assisted translation, specialised translation, machine translation post-editing, terminology management, translation quality, and translator agency. This method is aimed at explaining the way in which the notions of accuracy, risk, and translator decision-making can be perceived as inseparable aspects of human-AI collaboration in specialised translation.

The study is conceptual and exploratory and does not test or formulate empirical hypotheses. Instead, the research questions posed in the introduction are used as organising inquiries to elucidate the relationships between accuracy, risk and translator decision-making. The central assumption of the study—that specialised translation requires human validation in line with the possible consequences of error—is not stated as a hypothesis to be statistically tested, but rather as a conceptual proposition developed through critical analysis. The proposed relationships can then be formulated as testable hypotheses in empirical research on translators, post-editing tasks, error analysis or specific translation contexts.

The article is based on the assumption that specialised translation cannot be assessed by general indicators of linguistic fluency or even surface-level adequacy. In high-stakes fields like legal, medical, technical, financial, and institutional translation, the quality of translation relies on domain-sensitive factors, such as the terminological accuracy, conceptual accuracy, consistency, genre suitability and the possible consequences of error (Kenny, 2022). This is why the article discusses AI-assisted specialised translation as a professional communicative practice where the textual accuracy and real-world risk are interrelated.

The discussion is conducted methodologically by clarifying the concepts and categorising the analysis. To begin with, the article explains the meaning of such concepts as AI-assisted translation, human-AI collaboration, specialised accuracy, risk, translator agency, and decision-making. Second, it also critically analyses the shortcomings of the AI-generated translation output, primarily, the fact that fluency can hide terminological, conceptual, or functional errors. Third, it examines the role of the translator as not just a post-editor, but as an evaluator, verifier, reviser, risk assessor and responsible decision-maker. This enables the article to move beyond a tool-centred view of AI and towards a human-centred description of specialised translation practice.

The article's conceptual contribution lies in bringing together three aspects that are frequently addressed individually are united: accuracy, risk, and translator decision-making. Accuracy is addressed as the quality aspect of the translated text; risk is addressed as the potential implication of inaccurate or unverified output; and decision- making is addressed as the human process in which AI recommendations are accepted, amended,

rejected or re-verified. Based on this, the paper suggests an analytical framework that can be applied to analyse AI-assisted specialised translation workflows.

The literature was selected purposively rather than through a formal systematic protocol. Sources were selected if they: (1) were published between 2022 and 2025; (2) dealt with neural machine translation, large language models, post-editing, terminology management, translation quality or translator agency; (3) explored accuracy, reliability, professional control or human intervention; and (4) provided findings relevant to specialised or high-stakes translation. Peer-reviewed journal articles, conference papers, scholarly books and directly relevant research preprints were included in the review. Research that was not directly related to professional translation decisions or focused primarily on technical model architecture was not given high priority. The resulting body of literature is therefore interpretive but was selected on the basis of explicit thematic, temporal and professional criteria. Special emphasis is placed on studies of reliability, control, post-editing, terminology, and professional responsibility. The purpose is not to provide an exhaustive account of AI-assisted translation, but to identify research that helps explain the changing role and responsibility of the translator in specialised contexts.

The literature used in this article is selected due to its relevance to the key concepts of this article and it is not always selected based on systematic review procedure. The idea is not to exhaust the list of the studies on AI-assisted translation but to fall back on the ones that will assist in clarifying the role of the translator in the specialised context and its influence by AI. Special emphasis is made on the studies of reliability, control, post-editing, terminology and professional responsibility. This is important since AI has not only provided the translator with new tools to add to the toolbox, but it has also changed the way the translators evaluate, revise and take responsibility of the translated text. In this regard, Kenny (2022) is especially useful since she has created her work to perceive machine translation as a technology that needs to be applied in an informed and critical manner, not as a neutral or an absolutely autonomous technology.

The analysis also shows that there are certain common problems with AI-assisted translation and discusses how they impact the practice of specialised translation. These issues include: false fluency, unstable terms, hallucination, omission, excessive confidence in machine output and not knowing when human intervention is necessary. These are the problems that the article discusses through the lens of translator agency since translators are not only polishing the surface of the drafts written by AI. They will also be deciding on the instances when an AI suggestion can be applied, when it requires editing, when it has to be confirmed by a credible source, and when it has to be discarded. It is an anthropocentric viewpoint because Jiménez-Crespo (2025) stresses that the professional control and autonomy are essential when translators are engaged with AI tools. Thus, translator agency is viewed as a significant analytical category to analyse the accuracy and risk management in AI-assisted specialised translation.

This framework is supposed to be an explanatory model that can be improved with the help of the future empirical research. It can support the analysis of the study of

professional translation scenarios, post-editing activities, interviewing of translators, classroom activities and quality-assurance processes. The methodology used in the article is exploratory, critical and framework building in this sense.

4. Discussion: Accuracy, Risk, and Decision-Making of Translators

4.1. Accuracy in AI-Assisted Specialised Translation

Accuracy in AI-assisted specialised translation cannot be perceived as grammatical correctness, surface fluency, or literal correspondence between source and target texts. Accuracy is a multidimensional concept in specialised fields, and entails terminological accuracy, conceptual accuracy, consistency and genre- and register-appropriateness. A translation can be fluent and idiomatic and yet inaccurate, if it uses an imprecise technical term, distorting a legal term, undermining a medical teaching, or not adhering to the communicative norms of the target area. For this reason, specialised accuracy should be measured not only at the linguistic level, but at the conceptual, terminological, functional and professional level as well.

One of the most obvious aspects of specialised accuracy is terminological precision. In legal, medical, technical and financial translation, terms are not interchangeable labels which can be interchanged; they are embedded within disciplinary, institutional and procedural systems. Recent research on terminology integration in machine translation demonstrates that AI systems can enhance the integration of pre-approved terms with the help of terminology resources and post-editing processes (Moslem et al., 2023). But this also shows a weakness: specialised terminology frequently often needs to be externally controlled instead of being left to the probabilistic decisions of AI systems. When an AI system picks a near-synonym or a fluent paraphrase, it might seem reasonable to an ordinary reader but unacceptable to the experts in the field.

Conceptual correctness is also important. Specialised translation involves the precise transfer of the relationships among concepts, procedures, obligations, quantities, classifications and conditions. AI output may preserve the apparent message while distorting the underlying conceptual relations. This is especially concerning since large language models can generate very plausible language even when the underlying content is missing or unsupported. Research on hallucinations in machine translation demonstrates that multilingual systems can produce content that is not related to the source or can omit source information, either of which can be particularly problematic in a professional setting (Dale et al., 2023; Guerreiro et al., 2023). These mistakes are not necessarily apparent, as AI text can be stylistically refined.

Another important factor of accuracy is consistency. Specialised documents may involve the need to repeat the use of a single word, phrase, definition or procedural guidance throughout a document. The lack of consistency in rendering can cause ambiguity particularly in contracts, manuals, clinical texts, and institutional documents. Likewise, the appropriateness of genre and register influence accuracy since the

specialised writing should adhere to the anticipated patterns of professional writing. Even a technically correct sentence can fail because it is either too informal, too general or does not conform to the genre conventions of the target field.

"Common AI-related problems therefore include ambiguous terminology, false equivalents, omissions, hallucinations, over-normalisation and deceptive fluency. The main issue is that fluency may conceal errors. In contrast to the traditional machine translation errors, which can manifest themselves visually, AI-based errors can be rhetorically persuasive. The superficial quality in language should thus not be considered as a sign of specialised accuracy. The accuracy in AI-assisted specialised translation should be checked with the help of domain knowledge, control of terminology, source–target comparison, and human professional judgement.

4.2. Risk in High-stakes Translation Situations

Risk in AI-assisted specialised translation may be defined as the possible impact of inaccurate, incomplete, misleading or insufficiently verified translation output. In this sense, risk is not the same as error. A linguistic, terminological or conceptual issue in the text being translated is an error, and the risk is what that issue might bring about in a specific communicative context. A small stylistic flaw of a general text might not have much practical impact, whereas a misplaced word in a legal text, a mistranslated instruction in a technical manual, or a condition left out in a medical text, can have grave consequences. The interaction of translation error, domain sensitivity, user dependence, and the consequences of action based on the translated text, therefore, constitutes the basis for determining risk. There are a number of risk types, which are particularly applicable to AI-assisted specialised translation. Terminological risk is a situation in which AI chooses ambiguous, inconsistent, obsolete, or untrue equivalents. This can compromise clarity of concepts and professional acceptability. Conceptual risk arises when the translation of the meaning of a legal category, medical diagnosis, technical process, financial condition or institutional procedure is distorted. Procedural risk occurs when the translation distorts procedures, sequences, warnings, duties, or instructions of operation. In technical and medical contexts, this type of risk can directly affect user behaviour. The professional and ethical risk issues are related to accountability, confidentiality, transparency, and the accountability of the translator in using AI tools. End-user risk is a risk that readers, patients, clients, technicians or institutions may use inaccurate translated information when making decisions.

Risk takes different forms and is more serious in different areas. Risk in medical translation can include misinterpretation of symptoms, treatment instructions, informed consent, information on dosage, or safety warnings to patients. The studies of the application of machine translation by medical translators show that the professionals acknowledge the importance of the use of machine translation, but they are still sceptical that the quality of such translation is highly dependent on the text type, knowledge of the field, and human editing (Muñoz-Miquel, 2025). Risk is associated with rights, duties,

liability, procedural validity and institutional interpretation in legal translation. The force of a legal text can be changed by a mistranslated modal verb, condition, legal term or evidential phrase. Risk in technical translation can be related to safety, usability, product compliance, equipment operation or maintenance processes. A translation that is merely fluent may still be dangerous if it misleads the user.

AI introduces an additional layer of risk as its output can be a mixture of surface fluency and unreliability lurking beneath. Research on machine translation hallucination and omission demonstrates that machine translation systems may generate translations with unsupported information or missing source content, and these effects are hard to automatically detect across languages and domains (Benkirane et al., 2024; Dale et al., 2023; Guerreiro et al., 2023). Studies of work on GPT-based post-editing also imply that large language models can enhance translation results and also generate hallucinated edits, implying that automation itself can become a risk factor when viewed as authoritative (Raunak et al., 2023).

Specialised translation with AI should thus be perceived as a risk-managed communication. It is not merely a question of whether AI can deliver acceptable translations, but under what circumstances, what types of texts, with what degree of human supervision and what the consequences would be if errors remained undetected. Risk management involves verification of terminology, comparison between sources and targets, domain sensitive review, documentation of decisions and awareness of AI-generated fluency limits. The human oversight is not a quality improvement, but the prerequisite of professional responsibility in high-stakes situations.

4.3 Agency and decision-making of a translator

In AI-assisted specialised translation, translators should remain central to the analysis rather than being treated as secondary correctors of machine output. Although post-editing is usually introduced as the prevailing type of human intervention in machine translation workflows, this framing can understate the complexity of the task of the translator. “Rather than merely repairing linguistic defects after AI has produced a draft, translators evaluate the appropriateness of AI application, check the quality of the results, amend or reject recommendations, refer to external sources, explain their decisions, and take responsibility for the final text. This is especially true in specialised translation, where fluency is only one of the factors of acceptability of a translation, and terminological accuracy, conceptual accuracy, domain conventions, and the impact of error are also relevant.

The recent discussions on human-centred AI in translation focus on the importance of control and autonomy of professional translators as the key to the ethical and efficient use of AI technologies (Jiménez-Crespo, 2025). This is particularly applicable due to the fact that AI systems might seem authoritative and generate output that needs to be reviewed by experts. Research on large language models and post-editing demonstrates that AI can enhance the quality of translation and eliminate some kinds of errors, but it can also cause hallucinated or semantically unsafe edits (Raunak et al., 2023). This implies that the translator does not only have the responsibility to correct the blatant errors. It entails

making decisions regarding when AI assistance is appropriate, when its results are reliable and when human reformulation is required.

There are a number of decision points that are thus pivotal to AI-assisted specialised translation. The first concerns whether to use AI at all. This choice is determined by the type of text, confidentiality, client requests, sensitivity of the domain, the terminology resources at hand and the potential consequences of error. An internal document with routine information might permit some AI support, but a legal agreement, clinical teaching, patent document or technical safety guide might enforce a more restrictive control or avoidance of unjustified AI output. The second decision point is on whether to accept or reject AI suggestions. Acceptance is not a passive process; it entails comparing the source-text, domain analysis and making a decision as to whether the proposed rendering has retained the intended meaning and purpose.

The third point of decision is related to terminology verification. In specialised translation, terminology is frequently associated with institutional standards, legal classes, medical taxonomies, technical specifications or company-approved glossaries. The AI-generated terminology can be fluent and inaccurate, inconsistent or unsuitable to the context. Terminology integration studies indicate that AI can be used to facilitate the use of terms with approved terminology resources, although human correction and human correction and validation remain necessary (Moslem et al., 2023). Translators are thus forced to make decisions on when a term can be accepted, when it has to be verified with some authority and when a seemingly correct counterpart cannot be adopted due to its inability to fit into the field.

A fourth decision point is the extent of intervention that is needed. Certain AI-generated segments might be subject to minor revision, whereas some have to be reformulated or retranslated. Whether to post-edit or rewrite hinges on the severity of the mistake, the amount of specialised information, the trustworthiness of the AI output, and the time required to fix the portion. In other instances, a lot of post-editing can prove to be inefficient and unreliable compared to translating the segment from scratch. This is in line with the argument that translators practice strategic judgement as opposed to merely using mechanical corrections.

There are various factors that influence these decisions. Translator expertise is also important since only translators with strong linguistic and domain competence can detect the nuanced issues of conceptual or terminological nature. Decision-making is also influenced by time pressure, which can be tempting to rely more on AI results or decrease the levels of verification. Another significant aspect is trust in AI: too much trust can be a motivation to over-use AI output, whereas too little trust can diminish the usefulness of AI help. The type of text also is a factor, as promotional, institutional, legal, medical, and technical texts have varying levels of accuracy needs. Lastly, the magnitude of potential error influences the degree of care that is needed. The greater the potential impact, the more the verification, documentation and human review is required.

Agency of translators is thus a key human element of specialised translation. It does not just mean the ability of the translator to interfere with AI output, but also their professional right to determine how AI is to be used, restricted, criticised or denied. In high-stakes specialised translation, the human agency is not a hindrance to the efficiency of technology; it is the requirement that renders AI-assisted translation professionally viable.

5. Proposed Analytical Framework

The increased application of AI to specialised translation requires an analytical framework that is both specifically designed for the interaction between machine output, specialised accuracy, risk, and translator decision-making. The available discourses tend to concentrate on either the performance of the AI, post-editing work, the integration of terminologies, or the attitudes of the translators. These areas are critical but do not give a complete picture of how translators decide in the case of high stakes specialised applications of AI output. A dedicated framework is thus required since the core problem is not whether the output of AI is right or wrong but how different types of AI contribution generate various accuracy problems, risk, translator reactions, and impact on the overall quality.

The first dimension of the suggested framework is the kind of AI contribution. AI can be used as a complete draft generator, segment translator, terminology suggester, reformulator, quality-checker, or even automatic post-editor. These functions cannot be considered the same since they have various degrees of reliance on AI. A terminology suggestion, for example, can be simpler to check than a complete translated paragraph which has several embedded conceptual choices in it. The type of AI contribution can be determined to make the human judgement contribution to the workflow clearer.

The second dimension is the kind of accuracy problem. The issues of accuracy in specialised translation can be terminological inaccuracy, conceptual distortion, inconsistency, omission, addition, hallucination, register mismatch or inappropriateness to the genre. This dimension is required as not every error is of the same kind or severity. A minor stylistic flaw is not a false legal counterpart or a mistranslated medical teaching. Categorising the accuracy issue will enable the researcher and practitioner to go beyond the overall quality labels and pinpoint the problem that needs to be addressed.

The third dimension is the level of risk. The level of risk can be low, medium, or high depending on the domain, text function, audience and consequences of possible error. A term that is mistranslated in a draft may be of minor risk, but the same type of mistranslation in a contract, patient-information leaflet, safety manual, or regulatory document could have serious repercussions. Risk level thus connects textual accuracy with the impact in the real world. This is necessary owing to the fact that specialised translation is not only a linguistic activity, but also a type of professional communication with possible legal, clinical, technical or institutional implications.

The fourth is translator response. The responses can be acceptance of the AI suggestion, slight revision, term checking, outside source checking, rejection of the suggestion,

complete re-formulation of the segment, domain expert escalation or uncertainty documentation. This dimension places the agency of the translator at the centre of the framework. It acknowledges that quality is not produced by AI output as such, but as a result of the informed reaction of the translator to the output.

The fifth dimension concerns the rationale behind the translator's decision. The choices made by translators can be informed by domain knowledge, client requirements, terminology, translation memory, institutional guidelines, ethical concerns, time constraints, or perceived risk. The decision rationale should be included as it can be viewed as a manifestation of professional judgement and be subjected to systematic analysis. It can also enable future empirical research to investigate the reasons behind the acceptance, adaptation, or rejection of AI recommendations in various specialised settings by translators.

The sixth dimension is the influence on the final quality. This dimension is used to evaluate the effect of the response of the translator on the terminological accuracy, conceptual accuracy, consistency, readability, genre suitability, and risk management. It relates the decision-making process to the final translated product.

The framework can be used in future conceptual and empirical research to offer categories to analyse AI-assisted specialised translation workflows. It can be applied to case studies, translator interviews, think-aloud protocols, post-editing experiments, or error analysis, or classroom-based translator-training research. In a wider sense, it assists in changing the discourse on a small-scale issue of AI performance to a more comprehensive description of specialised translation as a risk-conscious, human-monitored, professionally responsible communication.

Table 1 summarises the six key dimensions of the proposed framework to make it more operational and clear. The table shows the relationships among AI contribution, accuracy problems, the level of risk, the response of the translator, the reasoning behind the decision and the overall quality of the translation. It is meant to be a tool for researchers, trainers, and professional translators. It would help them concentrate on AI-assisted specialised translation not only as a technological process, but as a risk-conscious and human-monitored decision-making process.

Dimension	Focus of analysis	Examples in AI-assisted specialised translation	Relevance to translator agency
Type of AI contribution	Identifies the function performed by AI in the translation workflow.	Full draft generation; segment-level translation; terminology suggestion; reformulation; quality checking; automatic post-editing.	Clarifies the extent to which the translator depends on AI output and where human judgement is required.
Type of accuracy issue	Classifies the nature of the problem found in AI-generated or AI-assisted output.	Terminological inaccuracy; conceptual distortion; inconsistency; omission; addition; hallucination; register mismatch; genre inappropriateness.	Helps translators determine whether the problem requires minor revision, verification, rejection, or full reformulation.
Risk level	Assesses the seriousness of the	Low risk in internal drafts; moderate risk in institutional communication; high	Encourages translators to adapt the level of verification

	possible consequences of an error.	risk in legal, medical, technical, financial, or regulatory texts.	and intervention to the communicative stakes of the text.
Translator response	Describes the action taken by the translator in response to AI output.	Accepting, revising, verifying terminology, consulting external sources, rejecting the suggestion, reformulating, escalating to a domain expert, or documenting uncertainty.	Places the translator's decision-making role at the centre of the workflow.
Decision rationale	Explains why a particular translator response is chosen.	Domain expertise; client requirements; terminology databases; translation memories; institutional guidelines; ethical concerns; time constraints; perceived risk.	Makes professional judgement visible, accountable, and available for reflection or empirical study.
Impact on final quality	Evaluates how the translator's intervention affects the final translation.	Improved terminological precision; conceptual accuracy; consistency; readability; genre appropriateness; functional adequacy; risk control.	Connects translator agency to the quality, reliability, and professional acceptability of the final text.

Table 1. Six-dimensional analytical framework for AI-assisted specialised translation

5.1. Illustrative Application of the Framework

The practical operation of the framework can be demonstrated through two brief translation cases. These cases represent recurrent problems in specialised translation and are included to illustrate the framework rather than to provide empirical evidence or evaluate a particular AI system.

Case 1: Contractual obligation

Consider the English contractual sentence, "The supplier shall notify the client within five working days." A machine-generated Arabic version might render it as "يجوز للمورّد إخطار العميل خلال خمسة أيام عمل" ("The supplier may notify the client within five working days"). The AI contribution is a complete segment translation, while the accuracy problem is a conceptual error involving legal modality: shall expresses an obligation, whereas يجوز expresses permission. Because this difference may affect the parties' contractual duties, the risk is high. The appropriate translator response is to reject the modal choice and revise the sentence to "يجب على المورّد إخطار العميل خلال خمسة أيام عمل" The decision is justified by the legal function of the source sentence and the need to preserve its obligatory force. The intervention therefore improves conceptual accuracy and reduces the possibility of legal misinterpretation.

Case 2: Medical dosage instructions

Consider the instruction, "Take one tablet twice daily." An inaccurate Arabic rendering such as "تناول قرصين مرة واحدة يومياً" ("Take two tablets once daily") preserves the total number of tablets but changes the required dosing schedule. The AI contribution is again a segment-level translation, but the accuracy problem concerns both meaning and functional appropriateness. The risk is high because the change may affect treatment effectiveness or patient safety. The translator should reject the proposed version, verify the instruction

against the approved medical information, and render it as "تناول قرصاً واحداً مرتين يومياً" The rationale for intervention is based on medical accuracy, patient safety, and the communicative function of dosage instructions. The final version preserves both the quantity and frequency specified in the source text.

These cases demonstrate that an apparently fluent translation cannot be evaluated independently of its context and possible consequences. They also show how the framework connects the type of technological contribution to the accuracy problem, risk level, translator response, decision rationale, and effect on final quality.

6. Implications for practice and training

The above discussion has significant implications for the specialised translation practice. When AI-assisted translation is viewed as a risk-sensitive process, not as a tool which is based on productivity alone, then professional practice should not be limited to uncritically accepting the output of AI-assisted translation. Translators should treat AI as a tentative source of suggestions, rather than a final producer of meaning, in areas of specialisation. This is especially significant since the large language models and machine translation systems can generate fluent text that might seem credible, but still has terminological, conceptual, or contextual mistakes (Guerreiro et al., 2023; Raunak et al., 2023). The professional value of AI is thus not as much based on its independent output but the extent to which it is critically incorporated into human controlled translation processes.

The first implication is that there is a necessity of critical AI usage. The translators ought to be in a position to know when AI assistance is needed, when it should be restricted and when it is not needed. This decision is based on sensitivity of the text, confidentiality, expectation of the client and the potential repercussions of the mistake. AI can be applied to draft, reformulate, compare or support terminology, but the output should be checked against the original text, domain norms and communicative purpose. The second implication is related to terminology checking. As specialised translation is very much dependent on approved and context-specific terminology, translators should not rely solely on AI-generated equivalents. Terms should be verified with glossaries, standards, translation memories, institutional databases, expert sources, and client-specific materials (Moslem et al., 2023).

The third implication is that there is a necessity of risk awareness. Professional translators need to be trained to differentiate between low- and high-risk translation scenarios and to vary the application of AI in these scenarios. Legal, medical, technical, financial, and institutional texts are those that may need more stringent review since they can have an impact on rights, safety, treatment, compliance, or decision-making. This results in a fourth implication: human review which is domain sensitive. It is not just that grammar or style should be reviewed but that terminological accuracy, conceptual accuracy, consistency, genre conventions, omissions, additions and potential hallucinations should be reviewed.

All these implications apply to translator training as well. AI literacy must be created through training programmes, although AI literacy does not imply learning to prompt or use tools. It is possible to do this by means of such tasks as prompt evaluation, terminology verification, detection of hallucinations, risk classification and reflective justification of decisions. It must involve awareness of AI limitations, being aware of false fluency, checking terminology, risk assessment, and recording decisions. Students should be trained to consider the reasons why they accept, revise, reject or reformulate AI output. In this regard, reflective decision-making is a fundamental professional competence. The education of future translators must thus not only equip them with the skills to be effective AI users, but also to be critical, ethical, and responsible in the application of AI in the specialised context.

7. Conclusion

This article has argued that AI has the potential to assist specialised translation. However, it has to be regulated by human expertise, domain knowledge, and risk-conscious professional judgement. The increased use of neural machine translators, large language models, terminology systems, and post-editing systems that use AI has certainly transformed the processes of translation. Nonetheless, AI use does not eliminate the need for human translators. Instead, it makes human expertise more visible and more necessary, especially in areas of specialisation where accuracy cannot be reduced to fluency, speed, or formal correctness.

The main issue of the article has been the correlation between accuracy, risk and decision-making by translators. The accuracy of AI-assisted specialised translation should be seen as a multidimensional notion of terminological accuracy, conceptual accuracy, consistency, linguistic appropriateness to genre, and linguistic functional appropriateness. Risk is the potential outcomes of errors that are not identified or not thoroughly examined, particularly those that are high-stakes like legal, medical, technical, financial and institutional translation. The connection between these two dimensions is through translator decision-making, which demonstrates how professional translators evaluate the output of AI, determine the accuracy of terminology, amend or discard proposals, and explain their final decisions. The role of the translator in this system is not a passive post-editor but a dynamic evaluator, verifier, risk-assessor and responsible decision-maker.

The article has also highlighted that apparent linguistic quality does not ensure specialised accuracy. Even AI-generated translations may be fluent, coherent and persuasive and include false equivalents, vague terms, omissions, hallucinations or conceptual errors. This is of particular concern to specialised translation since the mistakes might not be immediately apparent to non-experts and they can have actual professional or end-user impacts. This is why AI-assisted specialised translation should be considered a risk-managed communication, but not as a mere automated text generation.

The proposed analytical framework offers a more systematic way of analysing this process as it connects AI results, specialised accuracy, risk assessment, and translator

decision-making to a model. Future conceptual analyses, empirical research, self-reflection as a professional, and training of translators can be guided by the framework through taking into account the nature of the contribution made by AI, the nature of the problem of accuracy, the degree of risk, the reaction of the translator, the reason behind the decision, and the implications on the final product. Its importance lies in the fact that it sheds light on the instances where human judgement is applied in the workflows of AI-assisted translations, and it also shows how quality is not only generated by the technology, but through informed judgement of the translator, verification and intervention. The framework also aims to shift the discussion on the passive question of whether AI will replace translators in the specialised fields, to the more proactive question of how can human expertise be used to guide, constrain, evaluate and improve the application of AI in specialised translation? To conclude Responsible human–AI collaboration therefore depends on preserving translator agency at the centre of specialised translation practice.

Acknowledgements

The author has no acknowledgements to declare.

Ethical approval

Ethical approval was not required for this study because it adopted a conceptual and analytical approach based exclusively on the critical review of published literature. The study did not involve human participants, animals, interviews, surveys, experiments, personal data, or corpus-based empirical data.

Author contributions

Eyhab Bader Eddin was solely responsible for the conceptualisation of the study, literature selection and review, methodological design, development of the analytical framework, critical analysis, writing of the original manuscript, and review and editing of the final version.

Disclosure statement

The author declares that there are no competing interests or conflicts of interest associated with this research.

Funding

This research received no external funding.

About the authors

Eyhab Bader Eddin, BA, MA, PhD, CL, MCIL, MITI is an applied linguist, poet, interpreter, translator and currently an Assistant Professor of Translation at Dhofar University, Oman. He formerly taught in Syria, Oman, Kuwait and Saudi Arabia on both

MA and BA programmes. With a PhD titled 'Semantic Problems in A. J. Arberry's Translation of the Suspended Odes (Mu'allaqat), he is passionately interested in Classical Arabic and how it can be functionally translated into English. His MA dissertation's title, 'A Comparative Stylistic Study of the Poetry of Gibran Khalil Gibran and James Elroy Flecker', submitted at the University of reading, the UK was an interesting dissertation bringing together literature, and linguistics from a critical standpoint. His research areas of interest are Arabic and English stylistics, translation theory, syntax, translation theory and its application along with Quranic discourse and stylistics. He is interested in Classical Arabic poetry. He published in the fields of linguistics and translation in such refereed journals as Translation Journal, the British Journal of Middle Eastern Studies, ZDMG, etc.

ORCID

Eyhab BADER EDDIN https://orcid.org/0000-0003-0096-6334

Data Availability Statement

No new data were created, collected, or analysed in this study. The article is conceptual and analytical in nature and is based entirely on the critical examination of previously published literature cited in the reference list.

References

Benkirane, K., Gongas, L., Pelles, S., Fuchs, N., Darmon, J., Stenetorp, P., Adelani, D. I., & Sánchez, E. (2024). Machine translation hallucination detection for low and high resource languages using large language models. In *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 9647-9665). Association for Computational Linguistics.

Berger, N., Riezler, S., Exel, M., & Huck, M. (2024). Prompting large language models with human error markings for self-correcting machine translation. In *Proceedings of the 25th Annual Conference of the European Association for Machine Translation* (pp. 638-648). European Association for Machine Translation.

Dale, D., Voita, E., Lam, J., Hansanti, P., Ropers, C., Kalbassi, E., Gao, C., Barrault, L., & Costa-jussà, M. R. (2023). HalOmi: A manually annotated benchmark for multilingual hallucination and omission detection in machine translation. *In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (EMNLP) (pp. 638-653). Association for Computational Linguistics.

Guerreiro, N. M., Alves, D., Waldendorf, J., Haddow, B., Birch, A., Colombo, P., & Martins, A. F. T. (2023). Hallucinations in large multilingual translation models. *Transactions of the Association for Computational Linguistics, 11*, 1500-1517. https://doi.org/10.1162/tacl_a_00615

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

Kenny, D. (Ed.). (2022). *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Language Science Press. https://doi.org/10.5281/zenodo.6653406

Kocmi, T., & Federmann, C. (2023). Large language models are state-of-the-art evaluators of translation quality. In *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (pp. 193-203). European Association for Machine Translation.

Koneru, S., Exel, M., Huck, M., & Niehues, J. (2023). *Contextual refinement of translations: Large language models for sentence and document-level post-editing*. arXiv. https://doi.org/10.48550/arXiv.2310.14855

Moslem, Y., Romani, G., Molaei, M., Haque, R., Kelleher, J. D., & Way, A. (2023). Domain terminology integration into machine translation: Leveraging large language models. In *Proceedings of the Eighth Conference on Machine Translation* (pp. 902-911). Association for Computational Linguistics.

Muñoz-Miquel, A. (2025). Approaching machine translation in the medical and health field: An exploratory study. *The Journal of Specialised Translation, 44*, 22-43. https://doi.org/10.26034/cm.jostrans.2025.8453

Raunak, V., Sharaf, A., Wang, Y., Awadallah, H. H., & Menezes, A. (2023). Leveraging GPT-4 for automatic translation post-editing. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 12009-12024). Association for Computational Linguistics.

Zhu, W., Liu, H., Dong, Q., Xu, J., Huang, S., Kong, L., Chen, J., & Li, L. (2024). Multilingual machine translation with large language models: Empirical results and analysis. In *Findings of the Association for Computational Linguistics: NAACL 2024* (pp. 2765-2781). Association for Computational Linguistics.