SPECIALIZED TRANSLATION

THE ROLE AND CHALLENGES OF USING TECHNOLOGY IN INTELLIGENCE

Alias: Evantia DIAMANDOPOLIS

ABSTRACT

As artificial intelligence quickly transforms the landscape of multilingual communication, its role in the high-stakes field of intelligence translation requires careful scrutiny. This study[1] investigates the potential and limitations of AI-driven translation tools in accurately rendering specialized texts where precision and context are crucial. By comparing six leading machine translation systems (Microsoft Translator, Google Translate, Systran, DeepL, Reverso, and ChatGPT) using both qualitative insights and established metrics (BLEU, METEOR, ROUGE, BERTScore), the research presents a detailed view. Although AI improves speed and accessibility, it often struggles to maintain semantic accuracy and domain-specific coherence. The findings support a hybrid approach-combining human expertise with AI capabilities-to produce reliable, secure translations in intelligence work. This paper adds to the growing body of knowledge on how emerging technologies intersect with national security, providing timely insights for linguists, analysts, and policymakers navigating the evolving landscape of digital intelligence.

Keywords: *intelligence translation, machine translation, artificial intelligence, metrics.*

[1] The article is part of the author's graduation paper of the postgraduate programme *Introduction to Intelligence Analysis*, within the Faculty of Intelligence Studies, "Mihai Viteazul" National Intelligence Academy, Bucharest, 2025

Introduction

Artificial intelligence has become a key focus of Europe's digital agenda, influencing both policy and professional practices in translation. Over the past decade, research has increasingly concentrated on integrating AI into translation workflows, with neural machine translation (NMT) emerging as the leading approach. Studies demonstrate improvements in fluency and speed, and in many professional fields, machine translation (MT) is used as a draft tool that is then refined through human post-editing. Language models have a wide range of applications across sectors such as medicine, science, education, cybersecurity, and finance. Both social media platforms and government agencies depend on AI-powered tools to detect and flag content involving hate speech, incitement to violence, or extremism. However, these applications face significant challenges; translation errors are especially common with low-resource languages or emotionally charged texts, often leading to misinterpretation of intent or tone. These limitations underscore the importance of human oversight, especially when automated decisions may have significant legal, ethical, or security implications.

Seen as a central element of society's digital transformation, artificial intelligence has become a priority at both the national (Guvernul României, 2024) and European policy levels (Artificial Intelligence Act (Regulation (EU) 2024/1689)). Within the European Commission, the cutting-edge technology and language tools used by the Directorate-General for Translations (DGT) enhance the efficiency and quality of translation work. The computer-assisted translation (CAT) tool is continually fed with diverse data and high-quality human translations, making it an essential part of DGT operations. Integrated sources include translation memories (Euramis), terminology databases (IATE2), and machine translation (eTranslation). DGT's strategy, Data Strategy@DGT, aims to improve CAT to facilitate real-time collaboration and data exchange among translators. When integrated into secure, ethically governed workflows, AI-assisted translation does not seek to replace human expertise but to strengthen it by providing scalable tools that support professionals in making informed, timely decisions.

It remains unclear, however, how reliable these are for intelligence translation, where accuracy, cultural nuance, and context are essential. Most assessments still rely on general or technical corpora, while systematic research on the automated evaluation of MT engines in intelligence-specific settings remains limited. This leads to a double uncertainty: both the performance of these engines in sensitive areas and the adequacy of the metrics used to evaluate them are not sufficiently studied.

The purpose of this article is to address this knowledge gap through a mixed-methods case study that combines qualitative and quantitative analysis to examine six MT systems against institutional human translations published in the Romanian Intelligence Studies Review (RISR). This study analyzes the quality of translations created by machine translation systems compared to those produced by human translators. By applying BLEU, METEOR, ROUGE, and BERTScore in the quantitative analysis, the research also aims to evaluate how well these metrics reflect translation quality in relation to human qualitative evaluation. The focus is on examining the alignment between automated metric scores and human judgments, to provide insights into the usefulness and limitations of these metrics in specialized translation and to suggest directions for future research.

The research is based on the assumption that AI-based MT can approximate human-level fluency in certain conditions but is limited by terminological accuracy and contextual coherence. The quality of translations produced by artificial intelligence is greatly affected by several factors, including the complexity and nature of the source text, the specificity of the field's terminology, and the context. The research questions focus on two key aspects: the human translator and technological innovations designed to assist the specialist. Therefore, we aim to determine to what extent machine translation can replace the human translator in specialized language for intelligence, and which translation engines perform best.

Theoretical foundations - The use of AI in translation

The development of artificial intelligence in translation stems from linguistic theories and computational advances. Noam Chomsky's theory of generative grammar (Syntactic Structures, 1957) significantly influenced the foundation of later Rule-based machine translation (RBMT) systems in the 1960s and beyond, which focused on formalized linguistic structures and aligned well with computational modeling. Early systems struggled to produce fluent translations, prompting a shift to statistical machine translation (SMT) and later to neural machine translation (NMT) (Jooste, Haque, & Way, 2022). AI-powered translation engines have transitioned from rule-based to data-driven methods, primarily using neural networks and probabilistic models to generate translations. The shift from rule-based systems to NMT and Multimodal machine translation (MMT) systems has improved the fluency and accuracy of machine translations (Sulubacak, et al., 2020). Still, it has also introduced new challenges in managing specialized terminology and understanding contextual nuances.

According to Andrew Chesterman (2016) translation is governed by more than linguistic rules – it also depends on cultural, political, and strategic norms. All these aspects are challenging for artificial intelligence to understand, especially when applied in sensitive areas, where context plays a crucial role in understanding the text.

Empirical research has shown that while AI-powered translation tools have made significant strides in improving fluency and speed, they often lack the semantic understanding required by specialized fields such as intelligence. Läubli, Rico, and Volk (2018) examined the performance of AI-powered translation engines and found that while AI translations are increasingly fluent, they still struggle with specialized terminology and domain-specific content, such as texts extracted from intelligence or national security articles. This indicates that AI-powered translation engines work better for non-specialized translations, but face challenges when dealing with texts that require domain-specific knowledge (Läubli, et al, 2018). Similarly, Moorkens and O'Brien (2017) have contributed to research on AI-assisted translation, highlighting the importance of human intervention in the post-editing process for translation quality assurance. While AI can speed up translation, human translators remain essential for ensuring accuracy and correcting contextual errors, especially when specialized terms or ambiguities are involved (Moorkens & O'Brien, 2015). Taking into account that social media platforms and institutions with responsibilities in the field are increasingly relying on AI-based tools to monitor hate speech, violent content, and proselytizing. Saadany and Orasan (2021), in the study *BLEU, METEOR, BERTScore: Evaluation of Metrics Performance in Assessing Critical Translation Errors in Sentiment-oriented Text* analyzes the reliability of tools for the automatic evaluation of translations generated by translation engines (machine translation – MT). By comparing the results of the measurement systems, BLEU, METEOR, BERTScore, in the case of translations with severely affected significance versus those that present only distortions of feelings, the authors demonstrated that the measurement systems analyzed need to be improved in order better to capture critical errors in the interpretation of feelings.

The study *Technology Trends in Translation: A Comparative Analysis of Machine and Human Translation*, conducted by Kembaren et al. (2023), provides a detailed analysis of the differences between machine and human translations. The research uses a qualitative methodology based on a literature review to assess the accuracy, consistency, and flexibility of each method. The results show that machine translation excels in speed and consistency, making it ideal for repetitive, standardized translations. However, translations done by human translators offer superior accuracy and flexibility, which is essential for capturing the nuances and linguistic complexity of texts. This research contributes to understanding the role of technology in the translation industry. It highlights that, in many contexts, human translation remains indispensable to ensure the quality and fidelity of the message.

According to the study founded by the European Commission's Directorate-General for Translations on the Status of the Translator Profession in the European Union, (2013) *Studies on translation and multilingualism -The status of the translation profession in the European Union* (Pym et al., 2013), Romania has a system of training specialized translators at the level of university master's studies and a public system of authorization/certification of translators that takes into account specialized training. The translator's authorization is an official document that certifies a translator's professional competence and allows them to carry out certified translations for use in judicial or official proceedings. The Ministry of Justice issues the authorization for judicial proceedings, and the Ministry of Culture issues the authorization for other areas.

The study *The Impact of Artificial Intelligence on the Translation Profession. A Case study of Microsoft Translator* (Mandarić, 2022) explores the growing influence of AI technology on the translation profession, highlighting that, while AI and computer-assisted translation (CAT) tools have significantly transformed the way translators do their work, human translators remain essential for ensuring the quality of translations, especially in complex areas such as literary and technical texts. Research indicates that while AI technologies can accelerate the translation process and provide a valuable foundation for post-editing, they cannot entirely replace the expertise and critical judgment of human translators. In conclusion, the study suggests that technology should be viewed as a supportive tool rather than a threat, helping to streamline the translation process and enabling translators to focus on the creative and nuanced aspects of their work.

The research paper *Implications of Using AI in*

Translation Studies: Trends, Challenges, and Future Direction (Amini, Ravindran, & Lee, 2024) provides an overview of the use of AI in translation studies (TS), covering statistical machine translation, rule-based machine translation, neural machine translation, and hybrid machine translation. The study explores the advantages and limitations of each model and their applications in translation. In addition, various techniques for evaluating the effectiveness of AI models in translation are examined, along with their advantages and limitations, including the management of figurative language (e.g., idiomatic expressions, metaphors) and cultural nuances.

Methodological approach

From a methodological standpoint, the research questions were addressed by analyzing and evaluating translation engines in specialized translation tasks through both qualitative and quantitative comparative analysis. Using a case study, the differences between human and machine-generated translations were examined, focusing on how reliably AI-assisted translation can support human translators. The goal was to identify inaccuracies and gather detailed insights into the progress of AI machine translation technology, along with its current limitations in adapting content and in distinguishing nuances, emotions, interpretations, and language.

Qualitatively, focusing on accuracy, terminology, and coherence, machine outputs were contrasted with institutional human translations for accuracy, contextual fit, and stylistic adequacy (Nord, 1991; Chesterman, 2016). Quantitatively, BLEU, METEOR, ROUGE L, and BERTScore were applied to the same material, producing reproducible indicators (Papineni, Roukos, Ward, & Zhu, 2002; Mansuy, 2023; Saadany & Orăşan, 2021). The two dimensions were interpreted sequentially; this dual approach allowed both close textual examination and standardized evaluation. To integrate findings from both qualitative and quantitative approaches into a structured framework, a SWOT analysis was chosen as a complementary, broader perspective that combines internal factors (strengths and weaknesses) with external ones (opportunities and threats). The analysis was conducted by mapping the strengths and weaknesses derived from the characteristics of the translation tools and the recurring patterns identified in the qualitative and quantitative evaluations. At the same time, opportunities and threats reflected the contextual and operational implications. This integration ensured that the SWOT analysis reflected both empirical evidence and practical relevance.

To obtain the numerical results, Python formulas were edited and applied using the AI assistant in Anaconda Jupyter Notebook.

The corpus consisted of three sets of bilingual articles from the Romanian Intelligence Studies Review (RISR). The official Romanian translations generated by human translators served as reference standards in the quantitative analysis. Significantly, these benchmarks date from 2014–2016, before the widespread adoption of NMT; this ensures independence from recent automation but also means some terminology or stylistic conventions may have evolved. Due to space constraints, the article presents Set 3 in detail, as it is selected to represent patterns observed across the corpus.

The six widely used translation tools (Microsoft Translator, Google Translate, Systran, DeepL, Reverso, and ChatGPT) have been selected for their accessibility and widespread use. The translations generated have been quantitatively evaluated using Natural Language Processing (NLP) metrics to compare machine-generated text with a reference text (the published Romanian translation). BLEU measures n-gram precision for translation tasks, while ROUGE measures n-gram recall for summarization. METEOR improves on BLEU by incorporating synonyms and stemming. BERTScore utilizes contextual embeddings from Transformer models to capture semantic similarity, providing a more nuanced assessment that often correlates more closely with human judgment.

By focusing on free, publicly available tools, the study reflects real-world use and verifiable scenarios, while remaining accessible for replication in future research; nonetheless, it acknowledges that free machine translation tools offer lower performance than commercial solutions specialized in translating specific fields, such as intelligence.

Qualitative Results

The third set of translations includes an excerpt from the article Terrorism Serving Geopolitics. The Russian-Ukrainian Conflict as an Example of the Implementation of Aleksandr Dugin's Geopolitical Doctrine and Evgeny Messner's Concept of "Rebel War", by Michał Wojnowski, published in Romanian Intelligence Studies Review (RISR), no. 14/2015, p 44, followed by the Romanian translation version, published in the Romanian edition of this journal, and the translations automatically generated by translation engines.

Set 3 of translations
Source text:
This was reached by removing all the emblems from the soldiers' uniforms, making it unable to identify the country (so-called cryptooperation) and the formation they serve for, as well as by forbidding to communicate, in any form, with the civilians. All this led to many contradictory interpretations of "green people", "soldiers in uniforms" or "gentle people" and inability to define the enemy and taking the right actions. (Wojnowski, 2015)
Target text – the official translation into Romanian:
Acest lucru a putut fi realizat prin îndepărtarea emblemelor de pe uniformele soldaţilor, făcând în acest sens imposibilă identificarea ţării (aşa-numită cripto-operaţiune) şi a formaţiunii din care fac parte, precum şi prin interzicerea relaţionării, sub orice formă, cu civilii. Toate aceste au condus la numeroase interpretări contradictorii ale „omuleţilor verzi", „soldaţilor în uniforme" sau „oamenilor blânzi" şi la incapacitatea de a defini inamicul şi de a lua măsuri oportune. (Wojnowski, 2015)
Comment: The translation published in the Romanian Journal of Intelligence Studies reflects a good understanding of the source text and a correct adaptation to the specific nature of institutional discourse. The wording "Acest lucru a putut fi realizat prin îndepărtarea emblemelor de pe uniformele soldaţilor" is idiomatic and functional. The term "cripto-operaţiune" is used correctly, in the sense that it refers to the concealment of military identity - and not to cryptography, as it erroneously appears in other translations. The expression "relaţionarea cu civilii" may seem slightly abstract or unclear, but it is acceptable in a formal register. The ending of sentence, "incapacitatea de a defini inamicul şi de a lua măsuri oportune," is not only semantically correct, but also terminologically appropriate, with a strategic nuance also present in the original text. Overall, the translation manages to maintain a balance between clarity, terminological accuracy, and formal style, making it suitable for academic publication or institutional use.
Translation with Microsoft translator:
Acest lucru a fost obținut prin îndepărtarea tuturor emblemelor de pe uniformele soldaților, făcându-l imposibil să identifice țara (așa-numita criptooperațiune) și formația pentru care servesc, precum și prin interzicerea comunicării, sub orice formă, cu civilii. Toate acestea au dus la multe interpretări contradictorii ale "oamenilor verzi", "soldaților în uniformă" sau "oamenilor blânzi" și incapacitatea de a defini inamicul și de a lua măsurile corecte. (Microsoft Translator, May 26, 2025)
Comment: The translation provided by Microsoft Translator has several stylistic and semantic limitations that affect the accuracy and clarity of the message. Phrases such as "*făcându-l imposibil să identifice țara*" are grammatically incorrect in Romanian and denote a literal translation from English; the correct construction would be "*făcând imposibilă identificarea țării*." Terminological choices such as "*formația pentru care servesc*" are acceptable in terms of meaning, but inappropriate in a military context, where it would be preferable "*formațiunea din care fac parte*." In addition, machine translation does not capture expressions with a cultural and contextual significance, such as "*green people*," which, in the context of Crimea, is well establish as "*omuleții verzi*," an evocative term of the beginning of the conflict in Ukraine. Although the general meaning is conveyed, the nuances, precision, and contextual adaptation - specific to translation in specialized fields such as intelligence or security studies - are missing. Consequently, human review remains essential to ensure terminological consistency, stylistic fluency and conceptual accuracy.

Translation with Google Translate:

La aceasta s-a ajuns prin îndepărtarea tuturor emblemelor de pe uniformele soldaților, făcându-l în imposibilitatea de a identifica țara (așa-numita criptooperație) și formația pentru care deservesc, precum și interzicerea comunicării, sub orice formă, cu civilii. Toate acestea au dus la multe interpretări contradictorii ale „oamenilor verzi", „soldați în uniforme" sau „oameni blânzi" și incapacitatea de a defini inamicul și de a lua acțiunile corecte. (Google Translate, May 24, 2025)

Comment: In the translation generated by Google Translate, several grammatical and semantic issues stand out. The wording "făcându-l în imposibilitatea de a identifica țara" is incoherent in Romanian: it adopts the English structure "making it impossible," but with no clear subject, it becomes unnatural. Also, "formația pentru care deserves" is not the most semantically appropriate choice for this context. The ending – "și incapacitatea de a lua acțiunile corecte" – it is functional, but stylistically flat and general. Google offers a basic translation, but the lack of contextualization and adaptation, as well as its rigidity of expression, disqualify it for use in specialized or academic contexts.

Translation with Systran:

La aceasta s-a ajuns prin înlăturarea tuturor emblemelor din uniformele soldaților, făcându-l incapabil să identifice țara (așa-numita criptare) și formația pentru care servesc, precum și prin interzicerea de a comunica, sub orice formă, cu civilii. Toate acestea au dus la multe interpretări contradictorii ale „oamenilor verzi", „soldaților în uniforme" sau „oamenilor blânzi" și incapacitatea de a defini inamicul și de a lua măsurile corecte. (https://www.systransoft.com/translate/ Systran, May 26, 2025)

Comment: Systran's translation contains an equivalence error, translating the term "cryptooperation" as "așa-numita criptare," which distorts the original meaning. In the context, "cryptooperation" refers to an operation carried out without any insignia or identifiable elements — not data encryption. This denotes a lack of recognition of specialized terms, contrary to the fundamental principles of terminological translation (Sager, 1990). In addition, constructions such as "în uniformele soldaților" (în loc de "pe uniforme" and "făcându-l incapabil să identifice" are imprecise and misconstrued. The metaphors ("omuleții verzi", "oameni blânzi") are preserved, but integrated into a text marked by confusion. The translation is professionally and stylistically inadequate.

Translation with DeepL:

Acest obiectiv a fost atins prin îndepărtarea tuturor emblemelor de pe uniformele soldaților, făcând imposibilă identificarea țării (așa-numita criptooperațiune) și a formațiunii pentru care aceștia servesc, precum și prin interzicerea comunicării, sub orice formă, cu civilii. Toate acestea au condus la numeroase interpretări contradictorii ale „oamenilor verzi", „soldaților în uniformă" sau „oamenilor blânzi" și la incapacitatea de a defini inamicul și de a lua măsurile corecte. (DeepL, May 26, 2025)

Comment: DeepL provides a coherent, balanced and idiomatic translation. The phrase "Acest obiectiv a fost atins" is a successful stylistic adaptation of "This Was Reached", which avoids literalism. The term "criptooperațiune" is used correctly and in context, and the phrase "formațiunii pentru care aceștia servesc" is grammatical and clear, even if slightly formalized. The phrase "interzicerea comunicării, sub orice formă, cu civilii" is logical and expressive, and the ending – "la incapacitatea de a defini inamicul și de a lua măsurile corecte" – clearly convey the intent of the original. The translation is one of the most balanced of the analyzed set: fluent, stylistically adequate and free of major errors. The main limitation remains the lack of a deeper adaptation of metaphors ("green people", "gentle people") to the target culture.

Translation with Reverso:

Acest lucru a fost atins prin îndepărtarea tuturor emblemelor din uniformele olds', făcându-l în imposibilitatea de a identifica țara (așa-numita cripto-operație) și formarea pentru care servesc, precum și prin interzicerea de a comunica, în orice formă, cu civilii. Toate acestea au dus la multe interpretări contradictorii ale "green people", "soldați în uniformes" sau "people" blând și incapacitatea de a defini inamicul și de a lua acțiunile corecte. (Reverso, May 27, 2025)

Comment: The translation provided by Reverso is affected by obvious processing errors and the lack of a functional adaptation. From the outset, the distortion of the word "soldiers'" as "olds'" compromises the meaning, and the expression "făcându-l în imposibilitatea de a identifica țara" is unidiomatic and confusing. These formulations do not meet the criterion of clarity defined by Chesterman (2016). In addition, "formarea pentru care servesc" is a mistranslation of the term military "formation," replaced by a false friend. This violates the principles of terminological consistency highlighted by Sager (1990). Retaining English words ("green people", "soldați în uniformes") reflects a lack of adaptation, affecting both cohesion and coherence. From Baker's (2018) perspective, the lack of equivalence of the original discourse and the failure of the author's intention, ambiguity, strategic confusion, make translation not functional in any professional context.

Translation with ChatGPT-4o:

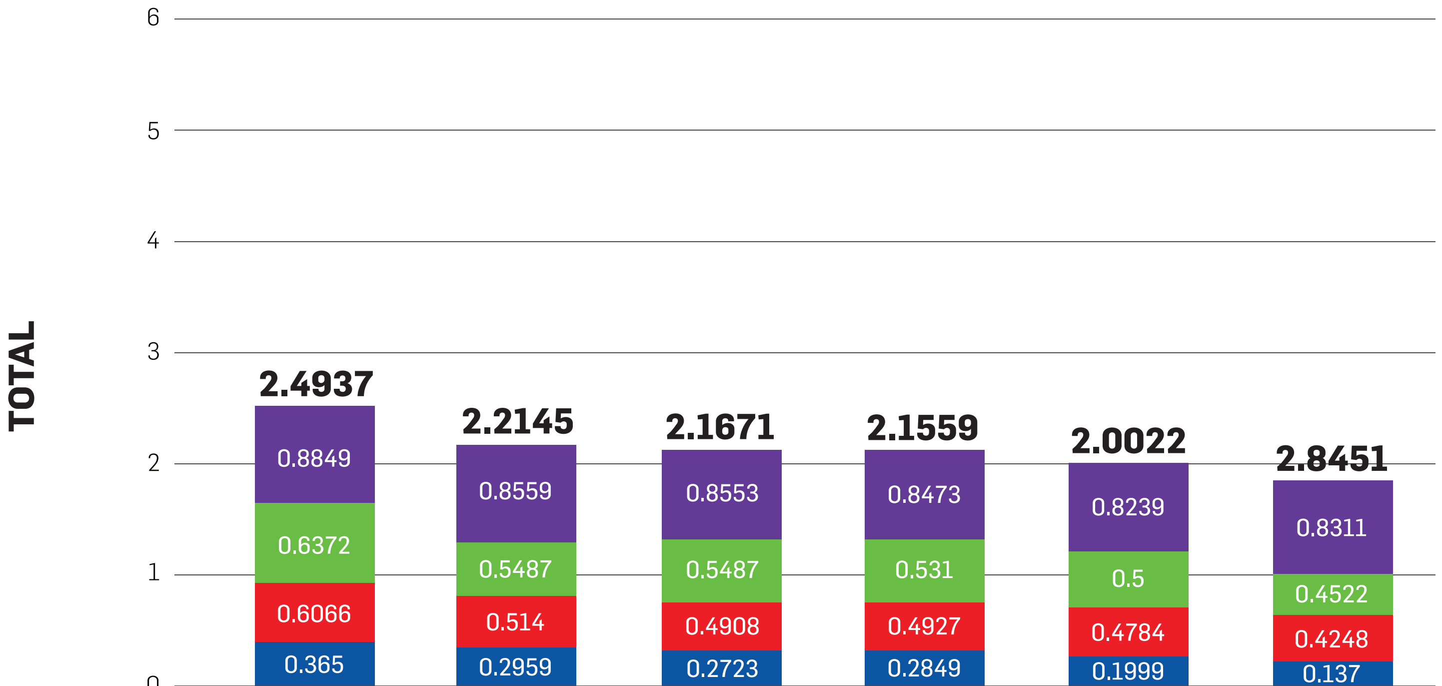
Această situație a fost posibilă prin eliminarea tuturor însemnelor de pe uniformele soldaților, ceea ce a făcut imposibilă identificarea țării (așa-numita criptooperațiune) și a unității în care aceștia activează, precum și prin interzicerea oricărei forme de comunicare cu populația civilă. Toate aceste măsuri au condus la interpretări contradictorii referitoare la „omuleții verzi", „soldații în uniforme" sau „oamenii blânzi" și la incapacitatea de a identifica inamicul și de a adopta reacțiile adecvate.

ChatGPT- (GPT-4o, released in June 2024) (The translation is produced by ChatGPT-turbo model (version 09.04.2024) using the official ChatGPT interface available on May 26, 2025).

Comment: This translation stands out for its stylistic balance, semantic fidelity and professional register. The phrase "eliminarea tuturor însemnelor de pe uniformele soldaților" is accurately adapted, and "criptooperațiune" is logically integrated into the sentence. "Unitatea în care aceștia activează" is a good choice adapted to the military context. Phrases such as "omuleții verzi" and "adoptarea reacțiilor adecvate" are expressive and consistent with the register of academic discourse. Although it does not introduce creative equivalences, the translation maintains clarity, a sober tone, and functional terminology, offering one of the most robust automated variants, comparable to the human translation in terms of style.

Qualitative Results

Graph 1. Graphical representation of the quantifiable parameters resulted from the computerized evaluation of the machine translations in Set 3



	DeepL	Microsoft	Google	Systran	Reverso	ChatGPT-4o
Total	2.4937	2.2145	2.1671	2.1559	2.0022	2.8451
BERT F1	0.8849	0.8559	0.8553	0.8473	0.8239	0.8311
ROUGE-1 F1	0.6372	0.5487	0.5487	0.531	0.5	0.4522
METEOR	0.6066	0.514	0.4908	0.4927	0.4784	0.4248
BLEU	0.365	0.2959	0.2723	0.2849	0.1999	0.137

Discussion

The analysis of the seven translations confirms that, in the translation of intelligence texts, terminological fidelity, stylistic clarity, and contextual adequacy are essential for the accuracy and effectiveness of communication. A clear balance distinguishes official human translation between semantic consistency, correct terminological adaptation, and expression in a professional register. Terms such as "cripto-operațiune" are correctly adapted, and the sentence structure respects both the original meaning and the conventions of specialized translation. Although the wording is slightly stiff in some passages, the translation overall is superior. Among the machine translations, ChatGPT's provides the clearest, most coherent, and most adapted rendering of the source text. Phrases such as "adoptarea reacțiilor adecvate" or "unitatea în care aceștia activează" denote a good understanding of the military register and an increased capacity for idiomatic adaptation. However, this variant also remains limited terminologically, as it does not provide cultural context or reinterpretations of metaphors. DeepL stands out for its clarity and fidelity, but remains close to a literal translation. Google Translate, Systran, and Reverso have serious shortcomings: grammatical mistakes, terminological confusion ("deservesc," "criptare"), and calques that undermine the text's coherence. These limitations make them unsuitable for translations in sensitive areas, such as intelligence. In conclusion, this study confirms that, although machine translation has advanced significantly, specialized translations in the field of intelligence and security remain dependent on human expertise, especially for interpreting metaphors with significant connotations, adapting acronyms, and selecting institutionally accepted terminology. AI-based tools can serve as practical support, but they cannot replace the contextual competence and interpretive responsibility of the human translator in areas of high geopolitical and terminological significance.

Findings of the quantitative analysis – graph interpretation

Graph 1 (Graphical representation of the quantifiable parameters resulting from the computerized evaluation of the machine translations in Set 3) reflects the comparative performance of the six translation tools evaluated, based on the scores obtained in the four metrics: BLEU, METEOR, ROUGE-1 F1, and BERTScore F1.

DeepL stands out as the leader, with a total score of 2.4937, supported by a very high BERT F1 score (0.8849), which indicates efficient retention of semantic meaning, as well as solid performance on METEOR (0.6066) and ROUGE (0.6372). However, the BLUE score (0.365) remains relatively modest, suggesting that the formulations are not necessarily formally close to the reference but may be freer or paraphrased.

On the following positions are Microsoft translator (2.2145), Google Translate (2.1671), and Systran (2.1559), with similar scores between them, all having balanced values at the level of general significance (BERT) and lexical content (ROUGE), but lower performance in terms of formal fluency (BLEU). Reverso, although with a more modest overall score (2.0022), comes closer to the others with a decent BERTScore (0.8239), confirming the trend observed in previous sets: it retains the general meaning but with a simpler, less formal expression. ChatGPT-4o, on the other hand, is once again in last place (1.8451), with low scores, especially for BLUE (0.137) and METEOR (0.4248), suggesting a freer, less faithful translation in terms of structure and terminology.

It is important to emphasize that, even in this case, the scores from automatic assessments do not always reflect the true quality of the translation. Some systems can score high because they reproduce text almost exactly, preserving the words' form but ignoring natural flow, logical clarity, or stylistic adaptation. We also highlight that, in specialized translation—especially in the field of intelligence—human evaluation is crucial for terminological accuracy, contextual relevance, and discursive coherence, beyond what the algorithms' figures suggest.

SWOT Analysis: AI in Intelligence Translation

Table 2. SWOT Analysis

STRENGHTS	WEAKNESSES	OPPORTUNITIES	THREATS
rapid processing of large volumes of text in various languages;	inability to interpret ambiguities or cryptic language;	acceleration of human analysis in crisis situations;	risks regarding the confidentiality of translated data;
access to extensive and up-to-date terminology databases;	reduced sensitivity to the geostrategic and operational context;	development of secure internal AI models;	potential for bias influencing the decision-making process;
efficiency in translating repetitive content;	tendency toward literal translations;	support for rare languages or dialects;	excessive trust without human validation;
integration with automatic analysis and detection systems.	errors in rendering novel or specialized terms.	evaluation and traceability in machine translations.	cyber vulnerabilities associated with the use of AI-based translation platforms.

To rigorously evaluate the potential and limitations of using artificial intelligence in specialized translation within the intelligence field, a structured SWOT analysis was conducted, aligned with the European Union's AI and cybersecurity regulations and strategic directions. This approach enabled the identification of significant advantages of applying AI in intelligence translation, including the rapid processing of large volumes of multilingual texts and access to updated terminology databases (Strengths). However, machine translation engines still face significant challenges in interpreting complex language and often produce overly literal or incomplete translations, thereby compromising the terminological precision essential to this domain (Weaknesses).

Regarding opportunities, the integration of artificial intelligence, supported by continuous human monitoring and revision, represents both a development opportunity and a strategic necessity, especially in crisis situations where speed and accuracy are crucial (Opportunities). Tactically, developing in-house models that comply with cybersecurity standards and are continuously updated terminologically can ensure successful implementation. In the context of rapid technological expansion, threats should not be underestimated: unwarranted reliance on AI may generate errors that can be exploited for disinformation or diminish human judgment in final evaluations (Threats).

These conclusions are directly correlated with the EU legislative framework. Under the EU Artificial Intelligence Act — the first regulation on AI — AI used in the intelligence field is classified as a high-risk system, requiring strict compliance with accuracy and cybersecurity standards from design to operation. Under the European regulatory framework, AI systems must prevent cyberattacks, errors, and data breaches and include robust procedures for backup, detection, response, and recovery (European Commission, 2025). The European Union Agency for Cybersecurity (ENISA) emphasizes that the continuous development of AI security standards is necessary, not only to protect infrastructures but also to maintain public trust and professional interoperability. In the field of intelligence, these legislative provisions and best practices translate into a responsibility to ensure the confidentiality of sensitive information, data protection, and traceability in automated translation (ENISA, 2022).

Conclusions

To obtain a comprehensive view of machine translation system performance, the quantitative analysis was complemented by an in-depth qualitative assessment that included comparisons with reference human translations. These excerpts, taken from sources published between 2014 and 2016, were chosen precisely to avoid contamination by recent machine translation technologies, thereby ensuring an authentic human standard before the widespread adoption of neural machine translation tools. In the automated assessment, these human versions were introduced as a benchmark, and the translations generated by the AI systems were treated with the candidate code. The evaluation assessed how closely a machine translation aligned with the coherence, accuracy, and contextual adaptation of the human version provided.

This correlation between the two methods supports the core conclusion of the research: machine translation cannot operate without human judgment. The quality of a translation, especially in sensitive areas like intelligence, depends not only on algorithms and databases but also on the human ability to understand context, interpret nuances, and make responsible stylistic and terminological choices. Humans should not compete with automated tools but should master and critically use them, including during the evaluation stage. In this way, the mixed research method, which combines quantitative analysis with qualitative interpretation, offers a broader and more balanced perspective, validated by both objective data and reasoned human judgments. This integrative approach enables both measuring performance and gaining a deep understanding of the meaning and limitations of AI-assisted translation.

An important finding from this comparison is that, although technology has advanced significantly and modern translation engines have real-time access to large linguistic and terminological databases, translation errors still occur, especially in the rendering of specialized terms in the field of intelligence.

While this study offers valid arguments for assessing AI-assisted translation in specialized areas, certain limitations must be recognized. First, the research was confined to three sets of texts and a small number of translation engines. Although this selection reflects the main translation tools available to the general public, it does not cover the entire range of AI translation technologies, especially those used by government or defense agencies, which often remain inaccessible due to security restrictions. Additionally, the reference translations used as benchmarks were drawn from human translations published several years ago (2014-2016). While this choice ensures independence from recent AI influence, it may not fully represent the current development of intelligence terminology or language use. Moreover, human translations are subject to interpretation and stylistic differences, introducing some subjectivity into the evaluation. Future research should aim to expand the dataset to include more diverse source texts and less common language pairs, particularly those relevant to security settings. In contrast, the automated evaluation

metrics used (BLEU, METEOR, ROUGE, and BERTScore) offer valuable insights but are not perfect. These metrics can penalize contextually appropriate reformulations and often overlook semantic nuance or accuracy, which are crucial in intelligence contexts translation.

The theoretical and methodological new aspects of this study arise from using quantifiable parameters to evaluate and analyze machine translations, leading to new research questions in this field. For example, research studies can be conducted to evaluate the performance of machine translation evaluation systems. All these steps are necessary to assess, implement, and continuously train and adjust AI-based systems. Ideally, machine translation (MT) systems are not static entities. The technology behind MT is constantly evolving, enhancing its performance and adaptability over time. Therefore, it is reasonable to expect that TM systems will also improve as research advances, and this is where research plays a crucial role. Researchers need a clear framework to identify areas for evaluation and improvement. Quantifiable parameters enable an objective comparison between different approaches, allowing researchers to determine which method is more effective and adjust the system accordingly.

By exploring the potential of implementing AI to improve the speed and accuracy of translations in the fields of intelligence and national defense, future research could have several important implications for both translation practice and the broader field of translation technology. The findings from this research can serve as a relevant starting point for academic and professional discussions on the use of artificial intelligence in specialized translation, particularly in intelligence. Additionally, these results can be shared at conferences on applied linguistics, translation technology, and ethical AI, thereby enhancing dialogue among researchers, practitioners, and stakeholders within institutions.

REFERENCES:

Amini, M., Ravindran, L., & Lee, K.-F. (2024). Implications of using AI in translation studies: Trends, challenges and future direction. Asian Journal of Research in Education and Social Sciences, 6(1), 740-754. doi:https://doi.org/10.55057/ajress.2024.6.1.67

Baker, M. (2018). Account, Translation and Conflict: A Narrative (Classics Edition ed.). London & New York: Routledge.

Chesterman, A. (2016). Memes of translation: The spread of ideas in translation theory. John Benjamins. Preluat de pe https://vdoc.pub/download/memes-of-translation-the-spread-of-ideas-in-translation-theory-3jn4728go5b0

Chomsky, N. (1957). Syntactic structures. Mouton. Preluat pe 02. 06, 2025, de pe https://stanford.edu/class/psych205/papers/Chomsky-1957.pdf

ENISA. (2022). enisa.europa.eu. (ENISA, Ed.) Retrieved 05. 06, 2025, from https://www.enisa.europa.eu/: https://www.enisa.europa.eu/

European Commision. (2025). EU AI Act: first regulation on artificial intelligence. Retrieved 05 06, 2025, from Topics European Parliament: https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

European Parliament and Council of European Union. (2024, June 13). REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Official Journal of EU. Retrieved 09 03, 2025, from EU Artificial Intelligence Act: http://data.europa.eu/eli/reg/2024/1689/oj

Guvernul României. (2024). Secretariatul general al guvernului. Preluat pe 05. 03, 2025, de pe gov.ro: https://sgg.gov.ro/1/wp-content/uploads/2024/07/ANEXA-1-10.pdf

Jooste, W., Haque, R., & Way, A. (2022). Knowledge Distillation: A Method for Making Neural Machine Translation More Efficient. MDPI Information, 13, 88. doi:10.3390/info13020088

Kembaren, W. F., Hasibuan, A. K., & Natasya, A. (2023). Technology Trends in Translation: A Comparative Analysis of Machine and Human Translation. Absorbent Mind: Psychology and Child Development, 3(2), 169-183. doi:https://doi.org/10.37680/absorbent_mind.v3i2.4486

Läubli, S., Sennrich, R., & Volk, M. (2018). Has machine translation achieved human parity? A case study on evaluating human and machine translation quality. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 4791-4796). Association for Computational Linguistics. doi:10.18653/v1/D18-1512

Mandarić, K. (2022). The Impact of Artificial Intelligence on the Translation. Osijek. Retrieved 06 05, 2025, from https://urn.nsk.hr/urn:nbn:hr:142:845158

Mansuy, R. (2023). https://plainenglish.io/blog/evaluating-nlp-models-a-comprehensive-guide-to-rouge-bleu-meteor-and-bertscore-metrics-d0f1b1. Retrieved 06. 05, 2025, from https://plainenglish.io: https://plainenglish.io/blog/evaluating-nlp-models-a-comprehensive-guide-to-rouge-bleu-meteor-and-bertscore-metrics-d0f1b1

Moorkens, J., & O'Brien, S. (2015). Post-Editing Evaluations: Trade-offs between Novice and Professional Participants. The 18th Annual Conference of the European Association for Machine Translation (EAMT 2015), 1. Antalya, Turkey. Retrieved 5 6, 2025, from https://www.researchgate.net/publication/275031846_Post-Editing_Evaluations_Trade-offs_between_Novice_and_Professional_Participants/link/5551f19008ae980ca6069e54/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

Moorkens, J., & O'Brien, S. (2017). Assessing User Interface Needs of Post-Editors of Machine Translation. În Human Issues in Translation Technology (pg. 127-148). Routledge. Retrieved on June 5, 2025, de pe https://www.taylorfrancis.com/chapters/edit/10.4324/9781315648934-17/assessing-user-interface-needs-post-editors-machine-translation-joss-moorkens-sharon-brien

Nord, C. (1991). Text analysis in translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis. Amsterdam: Rodopi.

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. The 40th Annual Meeting of the Association for Computational Linguistics (ACL) , (pp. pp. 311-318.). Philadelphia. doi:10.3115/1073083.1073135

Pym, A., Sfreddo, C., Chan, A. L., & Grin, F. (2013). The Status of the Translation Profession in the European Union. Anthem Press. doi:10.2782/64819

Saadany, H., & Orăşan, C. (2021). BLEU, METEOR, BERTScore: Evaluation of Metrics Performance in Assessing Critical Translation Errors in Sentiment-oriented Text. În I. Ltd. (Ed.), Proceedings of the Translation and Interpreting Technology Online Conference, (pg. 48–56). doi:10.26615/978-954-452-071-7_006

Sager, J. C. (1990). A Practical Course in Terminology Processing. John Benjamins Publishing Company.

Sulubacak, U., Caglayan, O., Grönroos, S. A., Rouhe, A., Elliott, D., Specia, L., & Tiedemann, J. (2020). Multimodal machine translation through visuals. Machine Translation, 34, pp. 97–147. Retrieved 09 03, 2025, from https://doi.org/10.1007/s10590-020-09250-0

Wojnowski. (2015). Terorismul servind scopuri geopolitice. Conflictul ruso-ucrainean ca exemplu al doctrinei lui Aleksandr Dugin și conceptul lui Evgeny Messner de „Război rebel". Revista Română de Studii de Intelligence(14), p. 41. Retrieved 02. 02, 2025, from https://www.ceeol.com/search/article-detail?id=562985

Wojnowski. (2015, 06 14). Terrorism Serving Geopolitics. The Russian-Ukranian Conflict as An Example of The Implementation of Aleksandr Dugin's Geopolitical Doctrine and Evgeny Messner's Concept of "REBEL WAR". Romanian Intelligence Studies Review, p. 44. Retrieved on February 02, 2025, de pe https://www.ceeol.com/search/article-detail?id=562985