AI AND MULTILINGUAL COMMUNICATION: ACCURACY, TRUST AND LINGUISTIC EQUITY IN HIGHER EDUCATION AND HEALTHCARE[2]

Laura-Rebeca STIEGELBAUER
Vasile Goldiș Western University of Arad

Email: stiegelbauer.laura@uvvg.ro

Abstract

The ongoing and rapid development of artificial intelligence (AI) is transforming multilingual communication in education and healthcare, two fields where linguistic diversity directly affects access and equity. This article investigates how AI-based translation tools, Google Translate, DeepL and ChatGPT - 4, perform within the multilingual environment of *Vasile Goldis* Western University of Arad in Romania. Using a mixed-methods approach, the study combined quantitative evaluation of translation accuracy with qualitative interviews involving 60 participants, including students, faculty clinicians, and administrative staff. Findings indicate that while ChatGPT – 4 and DeepL achieve high levels of fluency and semantic accuracy, users remain cautious about trust, confidentiality, and cultural sensitivity, especially in high stakes contexts such as clinical documentation. Across both educational and healthcare domains, AI tools were valued for speed and convenience but criticised for their inability to convey nuance, empathy, and intercultural awareness. This article argues that translation accuracy, though necessary, is insufficient on its own. For AI

[2] Article History: Received: 15.08.2025. Revised: 15.09.2025. Accepted:.16.09.2025 Published: 15.11.2025. Distributed under the terms and conditions of the Creative Commons Attribution License CC BY-NC 4.0. Citation: STIEGELBAUER, L.-R. (2025). *AI AND MULTILINGUAL COMMUNICATION: ACCURACY, TRUST AND LINGUISTIC EQUITY IN HIGHER EDUCATION AND HEALTHCARE. Incursiuni în imaginar* 16. *IMAGINARUL ȘI ADAPTĂRILE TEXTULUI LITERAR/ L'IMAGINAIRE ET LES ADAPTATIONS DU TEXTE LITTÉRAIRE/ / LITERARY ADAPTATIONS AND THE IMAGINARY*. Vol. 16. Nr. 1. 339-358. https://doi.org/10.29302/InImag.2025.161.14. No funding was received either for the research presented in the article or for the creation of the article.

to strengthen multilingual communication, institutions must prioritise linguistic equity, critical AI literacy, and ethical oversight, ensuring that technological gains serve support inclusive and context sensitive communication.

Keywords: Artificial Intelligence; Multilingual Communication; Machine Translation; Linguistic Equity; Digital Language Tools.

1. Introduction

In an increasingly interconnected world, multilingual communication plays an important role in shaping social, educational, and healthcare systems. The growing diversity of linguistic communities challenges traditional frameworks of translation and interpretation, pressing institutions to seek scalable, efficient, and context-sensitive solutions. Artificial Intelligence (AI), particularly in the form of neural machine translation (NMT) and language processing systems, has emerged as a powerful force that is transforming how people, professionals, and organizations interact across languages and cultures (Dodigovic, 2005). This article explores how AI is reshaping multilingual communication by enhancing translation accuracy, increasing user satisfaction, and reducing barriers to inclusion, especially in critical sectors such as education and healthcare.

The implications of AI for multilingualism extend beyond surface-level translation tasks. AI tools are no longer limited to dictionary-based output; instead, they engage with semantic nuances, contextual understanding, and adaptive learning. For example, NMT systems like DeepL and Google Translate have evolved to recognize idiomatic expressions and register-specific language, enabling more fluent and culturally appropriate translations (Freeth & Treviño, 2024). These technologies, when combined with speech recognition and real-time subtitling, are also proving effective in educational contexts by supporting learners from linguistically diverse backgrounds (de Oliveira, 2023). However, the advancement of such tools also introduces ethical, pedagogical, and epistemological concerns. Critics argue that the *invisibility* of AI translators mirrors historical debates over the erasure of human translator agency (Venuti, 1995; Freeth

& Treviño, 2024), encouraging renewed attention to the politics of language technology.

As Jackson (2020) emphasizes, intercultural communication requires not only linguistic competence but also critical awareness of power dynamics and cultural embeddedness. AI, while technically adept, often lacks this human reflexivity. Thus, this article also considers the limitations of AI-mediated communication, particularly in situations where trust, nuance, and emotional resonance are essential, such as doctor-patient interactions or classroom dialogue. In these contexts, AI becomes a tool of augmentation rather than substitution.

This article adopts a mixed-methods approach, combining quantitative evaluation of AI translation accuracy with qualitative interviews exploring user experiences in higher education and healthcare at Vasile Goldis Western University of Arad. By integrating these two strands, the study provides an empirically grounded assessment of how AI operates in real multilingual contexts, where clear communication is critical for both academic success and clinical safety.

The central argument is that while translation accuracy remains an important benchmark, it is insufficient on its own. AI can only strengthen multilingual communication if its integration is guided by ethical responsibility, user training, and a commitment to linguistic equity. In other words, the study positions AI not only as a tool for more fluent translation, but also as a means of supporting inclusive, context-sensitive communication across languages and cultures.

2. Literature Review & Theoretical Grounding

The convergence of artificial intelligence (AI) and multilingual communication demands an interdisciplinary foundation that blends technological innovation with sociolinguistic awareness and intercultural theory. This section provides a combination of the most relevant scholarly contributions that frame AI's role in multilingual communication.

2.1. AI and Translation Accuracy

AI-driven translation tools have advanced rapidly, evolving from rule-based systems and statistical machine translation (SMT) to neural machine translation (NMT), which processes language at the level of entire sequences rather than distinct phrases. As Dodigovic (2005) anticipated in early applications of AI in second language acquisition, language-learning environments enriched by AI not only provide corrective feedback but also model higher-order processing. Contemporary systems like DeepL, Google Translate, and GPT-4 or even GPT-5[3] draw upon massive language corpora to generate more accurate, context-aware outputs, particularly in high-resource languages.

Despite these improvements, translation remains imperfect because accuracy remains uneven. Freeth and Treviño (2024) highlight persistent errors in idiomatic expressions, specialized terminology, and low-resource languages. Moreover, automated systems are often transactional rather than relational, privileging speed and fluency over fidelity to cultural nuances (Venuti, 1995). These limitations underscore the need to assess translation technologies not only by technical benchmarks but also by their capacity to function in real multicultural contexts such as classrooms and healthcare environment.

2.2. Intercultural Competence and Human Dimensions

To assess the role of AI in multilingual contexts, it is essential to understand the human dimensions of intercultural communication. Jackson (2020) builds upon Byram's model of intercultural communicative competence (ICC), which encompasses not only linguistic proficiency but also attitudes, knowledge, skills of interpreting and relating, and critical cultural awareness. AI's ability to perform syntactic translation does not automatically translate to intercultural competence. The nuance, ambiguity, and culturally embedded knowledge required for effective communication across cultures remain challenges for machine learning systems. House (2020) similarly

[3] At the time of this research, GPT-5 had not yet been released and was therefore not part of the analysis

emphasizes that translation and interpreting are not purely linguistic processes but intercultural acts that rely on negotiation of meaning. This observation highlights a key limitation of AI: even when it achieves lexical accuracy, it frequently fails to capture the pragmatic and cultural subtleties necessary for effective cross-cultural communication.

This theoretical gap has led scholars to propose frameworks that incorporate critical digital literacy into AI tool design. In education, de Oliveira (2023) emphasizes the necessity of aligning AI-assisted tools with students' language development trajectories, arguing that technologies should scaffold, not supplant, authentic communication in multilingual classrooms. Such models require educators to function as mediators between AI outputs and pedagogical intentions.

2.3. Ethics, Equity and Translator Visibility

The rise of AI in translation raises ethical questions about authorship, responsibility, and fairness. Venuti's (1995) concept of "translator invisibility" resonates strongly in the age of machine translation, where the labour and intervention of human mediators are often erased. Freeth and Treviño (2024) caution that AI can exacerbate this invisibility, presenting polished texts without transparency about underlying decisions or biases.

Equity is another concern. Grozdanoff, Popov, and Serafimova (2023) argue for ethical AI systems trained on diverse datasets to avoid privileging dominant languages and marginalizing minority ones. In high-stakes contexts like healthcare, Martin and Crichton (2020) show how overreliance on literal translations can endanger patient safety. These perspectives converge on the need for responsible AI governance, where translation tools are integrated not only to improve efficiency but also to safeguard inclusivity, transparency, and linguistic justice.

3. Methodology

This study employs, as stated before, a mixed-methods approach to evaluate the role of artificial intelligence (AI) tools in multilingual communication, particularly within higher

education and healthcare contexts in Romania. The rationale for this design stems from the need to both quantify translation accuracy and capture experiential insights in contexts where multilingual communication is not just a pedagogical tool but a necessity for academic success and clinical safety. This study combines quantitative analysis of translation accuracy using standard evaluation metrics and qualitative analysis of user perceptions through interviews. By integrating these two approaches, the study captures both the measurable performance of AI tools and the lived experiences of users who rely on them in multilingual settings.

3.1. Research Context and Design

The research was conducted within the institutional framework of Vasile Goldiș Western University of Arad, a comprehensive university in Western Romania offering multilingual undergraduate programs. These include Medicine (taught in Romanian, English, and French) as well as Applied Modern Languages, where students are trained in professional translation and interpretation. These programs provided a rich and diverse setting for examining AI's role in both academic and clinical multilingual scenarios.

A convergent parallel design was chosen, combining quantitative methods (to evaluate translation performance) and qualitative methods (to capture attitudes, perceptions, and contextual reflections). This combination ensures that findings are grounded in both objective language metrics and real-world user experiences.

3.2. Participants

The participant group consisted of 60 individuals affiliated with the university and its healthcare partnerships:

- 33 medical students (from Romanian, English, and French language programs)
- 10 Applied Modern Languages students (future translators)
- 8 professors of English/French and AML alumni working as professional translators
- 5 clinicians from local hospitals involved in student placements

- 4 administrative staff who regularly use multilingual documentation

Inclusion criteria required participants to have direct experience with AI translation tools within academic, administrative, or clinical interactions over the past 12 months.

The participants from the Medicine programs were students in years 4 to 6, while the 10 students from the Applied Modern Languages program were in their 2nd and 3rd years; all study documents were collected between January and March 2025.

3.3. Data Collection Tools and Procedures

Three mainstream AI translation tools were selected for analysis: Google Translate, DeepL, and OpenAI's GPT-4 (*free version*) for generative tasks. The tools were assessed in typical university-related scenarios such as:

- Translating medical terminology in clinical case studies
- Rendering multilingual emails, consent forms, and course materials
- Assisting student translations in real-time or for assignment preparation.

Regarding the quantitative strand, 10 short source texts of 150–200 words were submitted to each platform. Reference translations were produced by professional translators associated with the university's professors of English and French belonging to the department. Translation accuracy was evaluated using BLEU and TER scoring systems (Dodigovic, 2005). BLEU measures the degree of overlap between the machine translation and the reference translation by comparing n-grams (unigram (one word), bigram (two words), trigram (three words), and so on), producing a score between 0 and 100%, where higher values indicate closer similarity. TER, in contrast, calculates how many edits (insertions, deletions, substitutions, or shifts) are required to transform the machine translation into the reference translation, with lower percentages indicating greater accuracy. In this study, BLEU was used to capture overall fluency and lexical similarity, while TER highlighted the effort needed for post-editing. Together they provided a balanced view of translation quality, combining surface-level accuracy with

practical usability. Errors were further categorized thematically to identify patterns such as false friends, register mismatches, and cultural misinterpretations.

Regarding the qualitative strand, semi-structured interviews were then conducted (in-person or online) with the participants to gain insights into perceptions of trust, cultural sensitivity, and practical usability. All participants consented to the use of their contributions and information for the purposes of this study, with the assurance that their identities would remain anonymous both during the research process and in the published results.

3.4. Analytical Framework

Quantitative data were processed in SPSS (Statistical Package for the Social Sciences)[4]. Errors were coded into categories: lexical, syntactic, semantic, pragmatic, and cultural. Qualitative data were transcribed and analysed using thematic coding in NVivo[5]. Thematic coding in this study, involved systematically tagging text segments with codes that represent recurring ideas, allowing patterns to emerge across participants' accounts. NVivo facilitated the process by enabling the organization of interview excerpts into nodes such as *trust in AI tools*, *ethical concerns* and *pedagogical use,* which were then grouped into broader themes. This process was aligned with frameworks drawn from translation ethics and critical pedagogy. Freeth and Treviño (2024) emphasized the importance of

[4] It is a widely used software for **quantitative data analysis**, particularly in fields like education, psychology, linguistics, sociology, healthcare, and communication studies. (*see*: Karamurugan, S., & Govindarajan, B. (2022). Statistical package for the social science. *International Journal of Business and Economics Research*, 8(4), 616–618. https://www.researchgate.net/publication/367127751

[5] Thematic coding is a method used in qualitative research to identify, analyse, and interpret patterns (or "themes") wi**th**in data such as interview transcripts, open-ended survey responses, focus group discussions, or even text from digital sources. (*see*: Hollywood, E., & Meehan, B. (2020, September). *How using NVivo enhanced the thematic analysis of various data sources in research involving children, parents and grandparents*. Paper presented at Trinity College Dublin, School of Nursing & Midwifery. https://www.researchgate.net/publication/344783360)

translator visibility and ethical responsibility, which informed the coding of data related to authorship, transparency, and the perceived invisibility of AI. de Oliveira (2023) advances critical language pedagogy, calling for AI to be positioned as a scaffolding tool rather than a substitute for linguistic development. This framework guided the interpretation of themes related to student learning, their ability to make informed and independent choices and their dependence on AI tools.

A specific emphasis was placed on identifying where AI-supported communication either enhanced or hindered multilingual equity. This localized, real-world methodological structure enabled a meaningful assessment of how AI tools operate in a transdisciplinary, multilingual academic environment, supporting both language learning and healthcare communication. By combining thematic coding with these frameworks, the analysis moved beyond surface-level descriptions to address how participants negotiated ethical and educational concerns in their engagement with AI translation tools. Moreover, findings in both strands, namely, the quantitative data and the qualitative data were compared to provide understanding of AI's role in multilingual communication.

4. Results & Analysis

4.1. Quantitative Findings: Translation Accuracy and Tool Performance

The results are presented in two strands, following the mixed-methods design: (1) quantitative evaluation of translation accuracy and (2) qualitative analysis of user perceptions. Together, these findings highlight both the technical performance of AI tools and their perceived value and limitations in real multilingual contexts. Thus, the quantitative analysis focused on evaluating the linguistic performance of three AI translation tools: Google Translate, DeepL, and GPT-4 (free version) across three core domains: lexical accuracy, grammatical correctness, and cultural or pragmatic appropriateness. The same 10 short texts (150–200 words each) were translated from Romanian and French into English, and vice versa. These texts reflected authentic materials

drawn from university syllabi, medical discharge summaries, and administrative forms.

Translations were scored using **BLEU (Bilingual Evaluation Understudy)** and **TER (Translation Edit Rate)** metrics, as explained in subchapter 3.3. Human-translated versions drafted by professional translators (AML Alumni) and professors of English and French within the foreign languages department of the university were used as standards.

4.1.1. Overall Performance Scores

Tool	BLEU score (avg)	TER Score (avg)
DeepL	78.5	17.3
Google Translate	72.4	20.8
Chat-Gpt-4	82.9	14.5

GPT-4 outperformed all other systems in terms of fluency and semantic cohesion, especially with complex or culturally nuanced texts. DeepL scored second highest, offering robust syntactic accuracy, especially in medical translations but struggling slightly with idiomatic expressions in French-Romanian-English conversions. Google Translate produced the weakest results, with frequent register mismatches and pragmatic errors.

4.1.2. Error Types

Errors were categorized into:

- **Lexical mismatches** (e.g., false friends, synonyms with inappropriate connotation)
- **Syntactic errors** (e.g., subject-verb agreement, clause structure)
- **Cultural misrendering** (e.g., idioms or culturally loaded terms)
- **Terminological inconsistency** (particularly in medical documents).

DeepL and GPT-4 showed better consistency with domain-specific terminology (e.g., *comorbidități*, *bilet de externare*) when compared to Google Translate, which often failed to adapt register appropriately. Errors in cultural nuance were most notable in

Google Translate, which occasionally produced literal translations of idioms like *"mi s-a umflat măseaua pentru că m-a tras curentul"* as *"my face is swollen because I got electrocuted"* rather than the intended "my tooth is swollen because it was drafty" - a clear example of pragmatic failure. DeepL had a better translation, one that could convey the source language message *"My tooth is swollen because I got a draft."*

4.2. Qualitative Findings: User Perceptions and Experiences

In parallel with quantitative analysis, the 60 participants (medical students, translation students, healthcare professionals, and university staff) also provided insights through semi-structured interviews. Thematic analysis revealed five key areas: **trust, convenience, anxiety, perceived accuracy,** and **cultural adequacy.**

4.2.1. Trust and Credibility

Medical students in the French-taught program expressed caution when using AI tools for clinical documentation. One respondent noted:

"In French, if a word has even a small error in a medical sentence, it can confuse the meaning and with AI, I sometimes don't trust it without checking in a dictionary." *(Student, 3rd Year, Medicine in French)*

Educators echoed this concern. Colleagues teaching at the Applied Modern Languages Program reported that students were increasingly submitting assignments with unedited AI-generated translations, which lacked "voice and cultural sensitivity" (Freeth & Treviño, 2024, p. 66).

4.2.2. Convenience and Time-Saving

Across disciplines, convenience was viewed as the biggest advantage. Healthcare professionals stated that AI tools helped them prepare multilingual consent forms or appointment reminders more quickly. Administrative staff used Google Translate and ChatGPT-4 (*free version*) for routine cross-linguistic communication with international students. Yet they emphasized the need to **verify output manually**, especially in legal documents.

4.2.3. Anxiety and Over-Reliance

Applied Modern Languages (AML) students revealed a paradox: while AI helped reduce anxiety in translation tasks, it also generated new anxiety about losing language skills. As one student put it:

"I'm afraid we'll stop learning how to *really* translate if we only rely on AI. It's helpful, but addictive." *(Student, 2nd Year, AML program)*

This concern aligns with the pedagogical warnings in de Oliveira (2023), who urges that AI should scaffold, not replace, language learning.

4.2.4. Perceived Accuracy and Use Cases

GPT-4 was praised for producing fluent and readable text, but many participants observed that **its translations lacked terminological consistency** over multiple documents if the prompt was not accurately and explicitly written. In contrast, DeepL was seen as more accurate in medical phrasing, though less versatile for casual dialogue or student tasks.

Hence, we could say that AI was most trusted for low-risk, repetitive translation tasks (emails, appointment letters) and **least trusted for high-stakes contexts such as academic assessment or critical medical decisions.**

4.2.5. Cultural Sensitivity and Adaptability

Participants from all groups noted that AI tools still struggled with **register and tone**. For instance, GPT-4 translated formal Romanian phrases too informally into English, occasionally undermining the seriousness of medical content. Cultural connotations were not always well preserved, especially in emotionally sensitive messages (e.g., condolence notices or treatment refusal statements).

As Jackson (2020) argues, true intercultural communication requires more than language substitution, it requires *localised meaning*, empathy, and awareness of the communicative context.

Taken together, the results suggest that AI tools can deliver high levels of technical accuracy, with GPT-4 and DeepL

performing particularly well. Yet, users' judgments about their usefulness extended well beyond metrics. Trust, convenience, and cultural appropriateness were decisive factors shaping perceptions, and these were often absent in otherwise fluent translations. The findings therefore reinforce the central argument of this article, namely that translation accuracy, while necessary, is insufficient on its own. Effective multilingual communication requires systems that also inspire trust, respect cultural nuance, and operate within ethical frameworks that safeguard all parties involved.

5. Discussion & Implications

The findings of this study confirm that artificial intelligence (AI) has the potential to support multilingual communication in meaningful ways, but their usefulness depends on far more than technical accuracy. While quantitative results demonstrated that GPT-4 and DeepL achieved strong scores on BLEU and TER metrics, participants consistently evaluated these tools in relation to broader communicative needs, such as trust, cultural appropriateness, and ethical responsibility. This dual perspective highlights a key lesson: translation accuracy is a necessary benchmark, but it is insufficient on its own for contexts where human well-being, academic integrity, and intercultural relations are at stake. This section explores the broader implications of using AI in multilingual education and healthcare, especially within the academic setting of Vasile Goldiș Western University of Arad, where linguistic diversity is not only valued but structurally embedded in its programmes.

5.1. Accuracy Is Not Always Enough

From a purely technical perspective, GPT-4 and DeepL produced the most accurate translations when compared to Google Translate. These results were measured using two standard tools: BLEU and TER.

The BLEU score (Bilingual Evaluation Understudy) is a method that compares the AI's translation to a human-translated version by checking how many word patterns they have in common. The higher the score, the closer the AI's result is to that

of a human. In our study, GPT-4 scored above 80 out of 100, indicating very high similarity.

The TER score (Translation Edit Rate), on the other hand, looks at how many corrections a human would need to make in the AI's translation. A lower TER means fewer edits are needed. GPT-4 and DeepL showed low TER scores, which means their translations were not only fluent, but also required minimal revision.

However, despite these strong scores, interview data revealed that many users still hesitate to fully trust AI translations, especially in fields like medicine, where even small errors can lead to serious misunderstandings. For instance, medical students expressed concern about using AI-generated translations in clinical settings, fearing they might miss nuanced or critical meanings in Romanian[6]. This shows that technical performance is not the only factor that matters, users also evaluate AI tools based on trust, emotional tone, and cultural sensitivity.

5.2. Language Learning and Professional Identity

In the Applied Modern Languages program, students shared that while AI helps them complete assignments more efficiently, it can also create a risk of over-reliance. Some students admitted to using AI without fully understanding the translation choices it made. This raised concerns about losing essential skills such as paraphrasing, cultural adaptation, and critical thinking, all of which are vital for professional translators.

This tension mirrors, as stated before, what de Oliveira (2023) describes as the need to *scaffold* learning with digital tools, rather than replace it. From a pedagogical point of view, AI should be integrated into the curriculum as a support tool, helping students develop editing, prompting and evaluation skills. Simply allowing AI to take over the translation process may weaken students' confidence in their own abilities and dilute their linguistic creativity.

At a university like ours, where translation is taught as a professional and ethical practice, students must learn how to

[6] In Romania, by their fourth year, medical students must have conversational proficiency in Romanian to begin clinical training and patient interaction.

critically engage with AI output, post-edit when necessary, and understand when not to use AI at all. This form of AI literacy is becoming just as important as knowing a foreign language.

5.3. Beyond Words: The Importance of Cultural Awareness

Translation is never just about replacing words. As Jackson (2020) explains, true communication between cultures requires an understanding of tone, context, emotion, and relationships. This is something AI, no matter how advanced, still struggles to grasp.

Participants in our study, who were coming from various cultural backgrounds, frequently mentioned that AI-produced texts were sometimes "too cold" or "too mechanic," especially when dealing with emotionally sensitive topics such as patient care, formal complaints, or condolences. In such cases, even a grammatically correct sentence may come across as disrespectful or inappropriate.

This reinforces the idea that AI can assist communication but not replace the human capacity for empathy and intercultural judgment. In healthcare, in classrooms, and in professional translation, AI must be seen as a collaborative tool, one that supports but does not override human communication strategies. It should be seen as our compliant assistant as long as we are a skilful leader.

5.4. Responsible Integration and Linguistic Equity

There is also a growing conversation about the ethical side of using AI in multilingual environments. Grozdanoff, Popov, and Serafimova (2023) argue that AI systems must be developed in ways that respect linguistic diversity and promote inclusion, not uniformity. Our findings support this view.

At Vasile Goldiș Western University, where we teach medicine in three languages and train future translators, AI offers many opportunities to support student learning and patient communication. For example, it can help generate multilingual glossaries, translate course materials, or simplify administrative communication with international students (our International Relations Department, Erasmus Department and Foreign Students Department). However, these tools should be introduced with care, ensuring both staff and students receive

proper training on how to evaluate the output and make ethically sound decisions about when and how to use them.

Without adequate training and critical use, AI tools risk reinforcing dominant language norms, which may marginalize speakers of minority languages. If used as a substitute rather than a support, such technologies could deepen existing inequalities instead of fostering inclusion.

5.5. Reflecting on the Bigger Picture

Taken together, these findings suggest that AI offers powerful tools for multilingual communication, but it cannot replace the human side of language. Metrics like BLEU and TER show AI's speed and fluency, yet users stress that trust, and cultural fit matter just as much.

At our university, where students are trained to work across cultures and languages, this means developing AI literacy that respects both the power and the limits of technology. AI can help democratize access to information and support linguistic inclusion but only when it is used with awareness, responsibility, and critical thinking. This article therefore advocates for a balanced approach, where the integration of AI is always guided by ethical principles and a commitment to linguistic justice, hence, it should be seen as a complement to, rather than a replacement for, human multilingual competence.

6. Privacy and Confidentiality

In fields like medicine and healthcare, confidentiality is a non-negotiable ethical principle. AI translation tools that rely on cloud-based processing (e.g., DeepL or Google Translate) may store or transmit sensitive information during translation. This poses a significant risk, especially when translating patient records or clinical notes. For this reason, students and professionals must be trained to avoid inputting confidential data into public AI

systems unless data protection can be guaranteed. ChatGPT has the *temporary button*[7] which could be used.

Furthermore, as AI tools become more integrated into university infrastructure, institutional policies must address **data privacy**, including how and where AI systems store the content generated or translated by users. Without proper regulation, AI can easily become a silent recorder of sensitive institutional or personal information. Hence, the responsible integration of AI into multilingual communication depends not only on accuracy and usability but also on safeguarding confidentiality.

7. Limitations of This Study

While this study aimed to provide a comprehensive analysis of AI in multilingual communication, focusing on functional texts in academic and medical domains, several limitations must be acknowledged. First, the sample size, particularly for the qualitative interviews, was relatively small and limited to a single institution. As such, the findings may not be generalizable to all Romanian or European universities. Second, the study focused primarily on written translation tasks, leaving out important modes of multilingual interaction such as **oral interpretation, real-time communication,** and **sign language translation**, which are increasingly relevant in multilingual academic and medical environments.

Additionally, the evaluation of AI output was limited to select language pairs (e.g., Romanian-English, Romanian-French, French-English) and a just three specific tools (ChatGPT-4, DeepL, Google Translate). Future studies could expand this scope by including lesser-known AI tools, multimodal translation systems, or voice-based AI interpreters. Despite these limitations, the study offers meaningful insight into how AI is currently reshaping, and challenging, multilingual communication in academic and professional contexts. By recognizing these limitations, the study underscores that its conclusions should be understood as

[7] The "**Temporary" Chat button** in ChatGPT lets you start a conversation that is not saved to your chat history, does not use memory, and is excluded from training, working like a private or incognito mode.

exploratory rather than definitive. They nevertheless provide a valuable basis for further research into how AI can be responsibly integrated into multilingual education and healthcare.

8. Conclusion

This article examined how AI translation tools are being used in the multilingual context of higher education and healthcare at Vasile Goldis Western University of Arad. By combining both translation performance metrics with user perspectives, the analysis highlighted both the promise and the limitations of Ai in bridging linguistic gaps. GPT-4 and DeepL achieved high technical accuracy, yet participants consistently emphasised that effective multilingual communication depends not only on lexical correctness but also on trust, tone and cultural sensitivity.

The findings point to three central implications. First, AI cannot replace human judgement in sensitive environments such as medical communication or academic assessment, where intercultural awareness and empathy are essential. Second, universities should embed AI literacy into curricula and staff training, enabling users to critically evaluate automated output and understand when and how AI should be used. Third, institutions must commit to linguistic equity by supporting tools and policies that serve less-resourced languages alongside dominant ones.

In conclusion, while translation accuracy remains a valuable benchmark, the overarching contribution of this study is to argue for the responsible integration of AI in ways that promote inclusion, ethical practice, and sustainable multilingual communication. Only by aligning technological capacity with linguistic justice can AI contribute meaningfully to equitable access in education and healthcare.

References:

Books and Book Chapters

Dodigovic, M. (2005). *Artificial intelligence in second language learning: Raising error awareness*. Multilingual Matters.

de Oliveira, L. C. (2023). *Supporting multilingual learners' academic language development: A language-based approach to content instruction.* Routledge.

Freeth, P. J., & Treviño, R. (Eds.). (2024). *Beyond the translator's invisibility: Critical reflections and new perspectives.* Leuven University Press.

Grozdanoff, B. D., Popov, Z., & Serafimova, S. (Eds.). (2023). *Rationality and ethics in artificial intelligence.* Cambridge Scholars Publishing.

Hollywood, E., & Meehan, B. (2020, September). *How using NVivo enhanced the thematic analysis of various data sources in research involving children, parents and grandparents.* Paper presented at Trinity College Dublin, School of Nursing & Midwifery.

House, J. (2020). Translation, interpreting, and intercultural communication. In J. Jackson (Ed.), *The Routledge handbook of language and intercultural communication* (2nd ed., pp. 173–190). Routledge.

Jackson, J. (Ed.). (2020). *The Routledge handbook of language and intercultural communication* (2nd ed.). Routledge.

Karamurugan, S., & Govindarajan, B. (2022). Statistical package for the social science. *International Journal of Business and Economics Research, 8*(4), 616–618.

Martin, G. S., & Crichton, J. (2020). Intercultural communication in health care settings. In J. Jackson (Ed.), *The Routledge Handbook of Language and Intercultural Communication* (2nd ed., pp. 503–520). Routledge.

Tipton, R. (2024). *The Routledge guide to teaching ethics in translation and interpreting education.* Routledge.

Venuti, L. (1995). *The translator's invisibility: A history of translation.* Routledge.

Legal and Government Documents

U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations.* https://tech.ed.gov/files/2023/05/AI-Future-of-Teaching-and-Learning.pdf