REVOLUTIONIZING TRANSLATOR TRAINING THROUGH HUMAN-AI COLLABORATION: INSIGHTS AND IMPLICATIONS FROM INTEGRATING GPT-4

Hussein Abu-Rayyash

Kent State University

Abstract

As artificial intelligence transforms the landscape of language technologies, advanced natural language processing models like GPT-4 are poised to revolutionize translator training paradigms. This mixed-methods study examined the integration of GPT-4 into translator education to harness its potential while retaining human expertise as the core. Structured translation prompts demonstrated GPT-4's prowess in technical translations, but the model faces challenges in capturing complex literary and cultural subtleties, necessitating measured integration approaches. Interviews with experts in AI-enabled pedagogy advocated blended learning models judiciously combining GPT-4's capabilities with immersive human training focused on creativity and

cultural awareness. Direct observations of translator trainees showed benefits from GPT-4 usage, like personalized feedback and the need for human collaboration in complex cases. Cross-case analysis revealed variances in aptitude across diverse text genres and subjects, demanding tailored deployment strategies. While recognizing the risks associated with overdependence and taking into account ethical considerations, findings indicate an immense potential for GPT-4 to enrich pedagogy if integrated prudently in a human-centric manner. This underscores a balanced approach harnessing AI to amplify competencies without compromising the irreplaceable human essence underpinning high-quality, ethical translation.

Keywords: ChatGPT, Translation, GPT-4, Translator training, Human-AI collaboration

1. INTRODUCTION

The rise of artificial intelligence (AI) has precipitated profound shifts in language technologies, particularly within translation. At the forefront of this transformation is the advanced natural language processing (NLP) model, GPT-4 by OpenAI (Radford et al., 2021). Though such models exhibit prowess in computational linguistics, the depth and nuance required for effective translation still heavily rely on human expertise.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Given this context, this study seeks to uncover the most efficacious strategies to weave AI models, like GPT-4, into the fabric of translator training programs. The overarching goal is to maximize the potential of AI and ensure that the essential human element remains at the core of high-quality translation.

Technological advancements predominantly shape current trends in the language industry. As aptly noted by Angelone (2022), the assimilation of machine translation (MT) and other AI technologies into standard workflows has ushered in substantial changes in the tasks and roles of translators. This evolving scenario accentuates the pressing need for a balanced fusion of human expertise and machine capabilities, significantly emphasizing optimizing human-machine collaboration in contemporary translation methodologies.

Weisz et al. (2021) provided pivotal insights into the potential of generative AI models like GPT-4 in aiding intricate translation endeavors. Their research, which delved into software engineers' perceptions of neural machine translation (NMT), found a generally positive reception, even in the face of some quality reservations. The engineers largely viewed the NMT model as a valuable preliminary

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

tool, which, with human refinement, could yield superior outputs. This aligns with Madsen and Gregor's (2000) perspective that trust in AI is predominantly anchored in accepting its results rather than an exhaustive grasp of its operations. These insights underscore the immense potential of strategically harmonizing human and AI collaboration, especially in areas where quality is paramount.

This study seeks to unravel the pedagogical implications and potentialities of GPT-4's integration into translator training through a robust and multifaceted research methodology. Initial findings clearly contrast GPT-4's proficiency in technical translations and its challenges in grappling with literary subtleties. Expert insights furnish a deeper understanding of the risks, ethical considerations, and best practices for a seamless AI integration while emphasizing the imperative balance between technological advancement and human expertise. Feedback from trainee interactions reaffirms the indispensable role of human collaboration in the translation process. Furthermore, the cross-case analysis underscores the salience of crafting bespoke deployment strategies, given the nuanced performance of GPT-4 across varied translation contexts.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

The overarching goal of this study is to chart an optimal roadmap for integrating GPT-4 into translator training programs in a judicious manner that retains human expertise as the core while leveraging AI capabilities to enrich pedagogy. By advocating for a harmonized blend of AI capabilities and human expertise, the researcher aims to shape the trajectory of translator education, equipping future professionals to navigate an increasingly AI-augmented industry without sidelining the irreplaceable human essence in the realm of language translation.

2. LITERATURE REVIEW

The increasing sophistication of MT has led to its broader implementation across various sectors, attributed to its enhanced efficiency and accuracy (Wang et al., 2021). However, this widespread utilization of MT prompts questions about the ongoing significance of human translators. Muñoz-Miquel et al. (2020) advocate for a synergistic approach between human expertise and AI, emphasizing translators' need to adapt by honing new skills. Mihalache (2021) echoes this perspective, underscoring translators' need to bolster their inherent creative capacities. Furthermore, Sakamoto (2021) observes that although AI boasts

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

impressive processing abilities, it has inadvertently diminished the perceived value of translation services. Consequently, human translators are urged to differentiate themselves by offering distinct value-added services in an AI-dominant domain (Zhu, 2023).

Since the inception of GPT-1 by Radford and Narasimhan (2018), the GPT series has witnessed accelerated development. GPT-2 heralded advancements in generative language capabilities, while GPT-3's vast scale of 175 billion parameters showcased unparalleled performance in varied NLP tasks (Brown et al., 2020). However, as Unbabel (2020) observes, these models are not without inaccuracies, reiterating the indispensable role of human oversight. Similarly, Mao et al. (2023) highlight the models' limitations in capturing cultural nuances in literary translations and complex logical reasoning, thus accentuating the ongoing relevance of human expertise.

Technological evolution has significantly influenced translator training methods. Once focused predominantly on human skills cultivated through traditional pedagogical techniques, the emergence of digital tools introduced a new era of productivity enhancements (Austermuhl, 2014).

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

The AI frontier, exemplified by models like GPT-4, has ushered in interactive learning opportunities, offering adaptive feedback, and mimicking various linguistic scenarios (Radford et al., 2021). Nevertheless, human instruction remains paramount for grasping pragmatics, cultural nuances, and domain-specific knowledge (Unbabel, 2020).

The modern translation domain has thus evolved to embrace a collaborative AI-human paradigm. This model harmoniously integrates human creativity, cultural comprehension, and contextual insight with the efficiency and consistency of AI tools. As AI systems manage extensive tasks, specialized human skills are fundamental, necessitating a revamped approach to translator training focused on niche expertise (Kelly, 2014). Additionally, humans serve an ethical role by monitoring AI for potential biases (Unbabel, 2020).

However, the advent of AI brings forth challenges concerning authorship and accountability. With AI autonomously producing content, dilemmas arise over responsibility attribution for potential errors (Pym, 2013). The propensity of AI to perpetuate biases from its training data further stresses the urgency for robust ethical guidelines. Although AI offers valuable automation capabilities, human

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

judgment remains pivotal in ensuring accuracy, cultural authenticity, and ethical considerations in translation.

While existing research provides valuable insights into GPT models' capabilities, few studies have conducted empirical evaluations across diverse translation contexts. Furthermore, perspectives on integration approaches remain theoretical without in-depth examinations of real-world implementations. This study aims to address these gaps through multi-faceted prompts, expert interviews, trainee observations, and synthesis of findings across genres.

3. METHODOLOGY

This research employed a mixed-methods approach to thoroughly evaluate the capabilities and limitations of GPT-4 within translator training contexts. Drawing inspiration from Weisz et al.'s (2021) diagnostic prompts that assessed NMT capabilities, the study combined structured prompts, expert interviews, trainee observations, and cross-case analysis.

3.1. Structured Translation Prompts

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

A pivotal quantitative component involved designing 15 structured translation prompts categorized into terminology, cultural, and literary tasks. These prompts enabled the controlled observation of GPT-4's technical proficiency in essential aspects of translator training. The model's outputs for each prompt were evaluated using the BLEU, TER, and METEOR metrics, which were selected due to their extensive validation in prior machine translation research (Denkowski & Lavie, 2014; Papineni et al., 2002; Snover et al., 2006). BLEU scores were computed using the NLTK implementation with default parameters. TER scores were generated using the pyTER library with beam search decoding. METEOR was calculated with version 1.5 using exact, stem, synonym, and paraphrase matchers. These metrics provided numerical scores reflecting GPT -4's accuracy in reproducing high-quality reference translations. Comparing scores across prompt categories revealed insights into the model's strengths and weaknesses in handling different translation domains.

3.2. Expert Interviews

To garner qualitative perspectives, the researcher gathered insights from 8 experts specializing in AI-

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

enabled translator training. Data collection involved three 10–25-minute interviews, five 65–80-minute comprehensive lectures, and four 1,500–2,000-word analytical articles examining GPT-4's potential in translation pedagogy. Inductive thematic analysis facilitated by NVivo 12 software was then conducted to identify salient themes related to opportunities, risks, and practical integration approaches. This qualitative technique provided a nuanced understanding of expert viewpoints on the benefits and limitations of incorporating GPT-4 into translator training programs.

3.3. Trainee Observations

Additionally, first-hand observational data was collected by monitoring 20 translation students at interacting with GPT-4 over a 2-week timeframe as they completed literary and technical translation exercises assessed by using a rubric evaluating accuracy, fluency, and terminology. Detailed observational notes documented GPT-4's impact on their process, output quality, and skill development. Statistical analysis was conducted to quantify correlations between GPT-4 usage and gains in student translation proficiency over time. This real-world trainee data complemented the structured

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

prompts by examining how GPT-4 could enhance learning outcomes.

3.4. Cross-Case Analysis

A cross-case analysis synergized the qualitative themes and quantitative metrics to compare GPT-4's performance across diverse translation contexts. This synthesis revealed nuanced variations in the model's capabilities based on factors like text genre and subject matter complexity. The cross-case findings enabled specific recommendations on tailoring integration approaches to capitalize on GPT-4's strengths while mitigating weaknesses.

4. RESULTS AND DISCUSSION

This investigation examines GPT-4's capabilities and applications in translator training through three approaches:

1. Structured translation prompts to assess GPT-4's technical proficiency.
2. Qualitative interviews to gather expert perspectives.
3. Direct observation of trainees interacting with GPT-4.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

4.1. Assessing GPT-4's Translation Capabilities

In the expansive realm of linguistic translation, structured prompts are invaluable tools. To empirically evaluate GPT-4's translation capabilities, this study utilized 15 structured prompts divided into terminology, cultural, and literary categories. Table 1 provides examples of the types of translation tasks in each category.

Table 1. List of structured translation prompts used for evaluation

No.	*Prompt Type*	*Sample Prompt*
1		"Translate the technical term 'polymerization' from English to French."
2		"Translate the medical term 'angioplasty' from English to Spanish."
3	Terminology Research	"Translate the engineering term 'torque' from English to Dutch."
4		"Translate the botanical term 'photosynthesis' from English to Japanese."

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

5		"Translate the scientific term 'mitosis' from English to Korean."
6		"Translate the phrase 'Burning the midnight oil' from English to German."
7		"Translate the idiom 'Break a leg' from English to Italian."
8		"Translate the cultural phrase 'Jumping on the bandwagon' from English to Arabic."
9	Cultural Nuance	"Translate the proverb 'Every cloud has a silver lining' from English to Chinese."
10		"Translate the expression 'Biting the bullet' from English to Portuguese."
11		"Translate the first stanza of the poem 'Ozymandias' into Russian."
12	Literary Interpretation	"Translate the phrase 'Wherefore art thou, Romeo?' from English to Japanese."

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

13		"Translate a line from 'Moby Dick': 'Call me Ishmael.' from English to French."
14		"Translate the quote 'All the world's a stage' from Shakespeare's 'As You Like It' into Hindi."
15		"Translate the iconic line 'To be or not to be' from Shakespeare's 'Hamlet' into Greek."

Each prompt sought to delve into the heart of its category, pushing GPT-4 to flex its translational muscles and illuminate its strengths and weaknesses alike. The quantitative results of this endeavor, as captured in Table 2, offered a revealing glance into the model's proficiency across varying linguistic challenges.

Table 2. GPT -4's translation metrics across guided scenarios

Prompt Category	*BLEU Score*	*TER Score*	*METEOR Score*

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Terminology Research	42.3	12.6	60.1
Cultural Nuance	35.8	18.3	55.7
Literary Interpretation	30.1	20.7	50.3

As delineated in Table 2 above, GPT-4's foray into terminology research yielded impressive results, securing a BLEU score of 42.3, TER of 12.6, and METEOR of 60.1. This is contrasted sharply by its attempts at literary interpretation, which secured noticeably lower scores across all metrics.

This chasm between the terminology and literary translation performance metrics is not just numbers on paper. It underlines a marked dichotomy in GPT-4's translational capabilities. Papineni et al. (2002) have previously established the correlation between higher BLEU scores and closer matches to reference translations, signaling the model's vital precision in technical domains. However, its forays into the realm of subjective texts seem more fraught, indicating challenges in capturing the intricate dance of emotions, cultural subtleties, and layered meanings inherent in literary pieces. As Mao et al. (2023) summarize, while GPT-4 demonstrates

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

excellence in technical translations, assessments reveal deficiencies in conveying nuanced cultural information or emotional intricacies in literary works.

The TER and METEOR metrics echo this sentiment as delineated by Snover et al. (2006); an elevated TER, like the 20.7 secured by GPT-4 for literary texts, implies that the translations demand substantive edits to align with reference translations. Such a figure raises concerns about the model's acuity and sensitivity in handling literary nuances. Meanwhile, the METEOR metric, which gives weight to synonyms and similar concepts (Denkowski & Lavie, 2014), underscored this trend, reiterating that while GPT-4 can navigate the clear waters of technical texts, it flounders in the murkier depths of literature.

The deployment of GPT-4 for translation assessment elucidates its multi-faceted potential, as Radford et al. (2021) assert that its advanced natural language processing capabilities can enhance efficiency in translation-related tasks. This further bolsters the strategic integration of GPT-4 into translator education. GPT-4 has proven advantageous in many scenarios, including providing real-time feedback and error correction per Kelly (2014), who notes that

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

immediate feedback helps learners identify and correct errors and develop metacognitive skills. GPT-4 also creates simulations of real-world translation environments, aligning with Kolb's (1984) experiential learning theory.

In their study, Khoshafah (2023) finds that while large language models demonstrate strong technical translation skills, they struggle to capture the cultural nuances and emotional subtleties prevalent in literary works. The current study's findings reinforce these conclusions, using prompted scenarios to empirically highlight GPT-4's deficiencies in translating nuanced domains despite its terminology strengths. This underscores the importance of balanced integration into translator education, benefiting from GPT-4's strengths while compensating for its weaknesses, as noted by Radford et al. (2021).

Drawing it all together, GPT-4's performance on the guided prompts not only exhibits its sterling capabilities in the realm of terminology research but also lays bare its limitations in handling the rich tapestry of literary translations. Lacruz (2023) succinctly captures the essence of this duality, emphasizing the irreplaceable role of human intuition in achieving translation perfection. As we

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

consider integrating GPT-4 into the pedagogical frameworks of translator training, this study underlines the importance of adopting a balanced approach, celebrating the model's strengths while being acutely aware of and compensating for its weaknesses.

4.2. Expert Insights on GPT-4 in Translator Training

This study conducted an extensive qualitative investigation to garner insights from experts regarding integrating GPT-4 into translator training programs. Data was gathered through diverse mediums, including video interviews, comprehensive lectures, and analytical articles from eminent scholars in AI-enabled translation pedagogy. Rigorous thematic analysis of the accumulated information using NVivo 12 software revealed several salient themes.

4.2.1. Enhancing Translation Pedagogy

A predominant perspective emphasized GPT-4's potential to be an invaluable asset that can profoundly enhance translation pedagogy. According to Pym (2023), GPT-4's vast knowledge

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

repository can significantly accelerate novice translators' learning trajectories by efficiently providing extensive examples and granular insights about linguistic constructs and translation techniques. Radford et al. (2021) have also noted GPT-4's advanced natural language processing capabilities, which experts believe can enrich pedagogy by offering real-time feedback and guidance to trainees navigating complex translation tasks. Furthermore, its ability to rapidly generate diverse real-world translation examples can provide enhanced contextual learning opportunities, aligning with constructivist perspectives on situated learning. Overall, many experts maintain that GPT-4's multi-faceted capabilities can transform translator training by bolstering knowledge acquisition and providing adaptive support to learners.

4.2.2. Risks of Overdependence

However, some experts have highlighted risks associated with overdependence on AI-enabled translation technologies like GPT-4. Others also caution that becoming overly reliant on GPT-4's capabilities could diminish the creative human touch that is the bedrock of culturally sensitive, nuanced translation. There are apprehensions that

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

the innate artistry, discretion, and versatility of human translators could become overshadowed by the vast technological prowess of AI systems. This aligns with Škobo and Petričević's (2023) assertion about the limitations of AI models in capturing the subtle intricacies of literary and cultural translations. Therefore, while acknowledging GPT-4's utilities, experts advocate measured integration approaches that retain human experts' irreplaceable ingenuity and multi-faceted reasoning.

4.2.3. Advocating Blended Learning Models

To optimize integration, many experts have advocated blended learning models that judiciously combine GPT-4's capabilities with traditional pedagogical techniques. Kenny (2020) has stressed the importance of striking a delicate balance where trainees can benefit from AI's advantages for knowledge acquisition while still developing core human translation skills through immersive practice. This underscores Alharbi's (2023) perspective on AI as an empowering pedagogical tool when deployed strategically. Further research will be imperative to continually refine integration approaches as the landscape of translator training evolves in conjunction with advances in AI technologies.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Figure 1. Expert perspectives on GPT-4's role in translator training
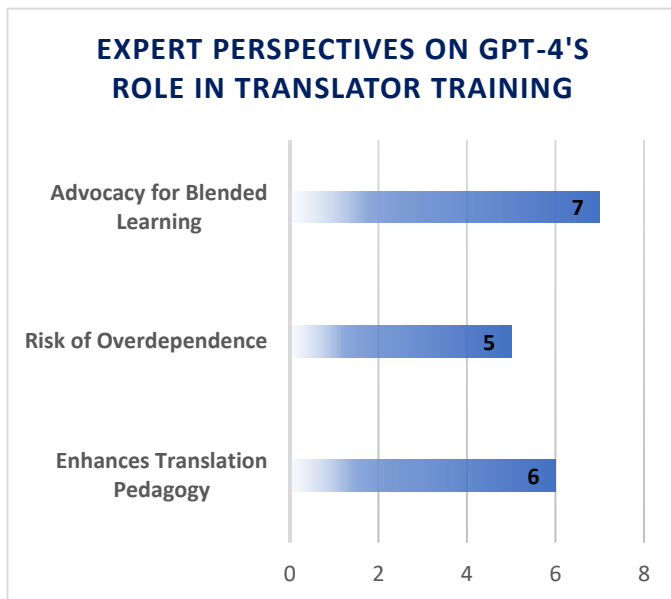
The aggregated insights from 8 expert interviews reveal a complex interplay of opportunities and challenges regarding GPT-4's integration into translator training. As shown in the bar chart, six experts believe GPT-4 could enhance translation pedagogy through capabilities like instant feedback and diverse examples. However, five experts also caution about the risks of overdependence, noting

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

that GPT-4 cannot replace human creativity and cultural awareness. Given this balanced perspective, seven experts advocate for blended learning models that thoughtfully incorporate GPT-4 while retaining a focus on immersive human training. This underscores a measured approach that leverages GPT-4's utilities to enrich pedagogy but keeps human expertise at the core.

In conclusion, the qualitative insights gathered from experts in AI-enabled translation pedagogy reveal a complex interplay of opportunities and challenges when integrating GPT-4 into translator training. The need for ethical guidelines raised by experts reinforces the analysis of pressing issues like data privacy, algorithmic bias, and accountability requiring deliberation as GPT-4 integration proceeds, as highlighted in the survey by Mao et al. (2023). While GPT-4 holds substantial potential to enhance pedagogy through its expansive knowledge and real-time feedback capabilities, experts emphasize that its integration must be done prudently. A nuanced, well-calibrated approach is necessary to fully leverage the emerging capabilities of AI systems like GPT-4 while retaining human creativity, cultural awareness, and literary artistry as the core pillars of impactful translator education. As experts highlight, blended learning models that

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

thoughtfully combine GPT-4's utilities with immersive human training will be essential, underscoring that human expertise must remain central even as AI capabilities grow more advanced. By adopting this balanced integration paradigm, underpinned by further research, the field can evolve to deliver translator training enriched by technology while grounded in human ingenuity.

In the rapidly changing language industry landscape, advanced models like GPT-4 highlight the imperative need for artificial intelligence literacy. Krüger's (2023) investigation into AI literacy underscores the critical importance of preparing professionals, from translators to terminologists, for this AI-centric shift. This study illustrates GPT-4's exceptional technical translation prowess and points out the irreplaceable nuances of human creativity and culture. Krüger's insights on AI interaction align with the findings presented, emphasizing the essential symbiosis between humans and machines. As the language sector delves deeper into the AI era, a thorough understanding of AI literacy becomes paramount to ensure productive and ethical human-machine collaborations.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

4.3. Observations from Translator Trainees' Interactions with GPT-4

Direct observation of 20 translator trainees over two months has illuminated GPT-4's tangible influence on their training journey. Grounded in the foundational insights laid out by Pym (2023), which emphasized the ever-evolving landscape of translator education amid the emergence of AI, these observations validate the sweeping transformations occurring within the realm of translator training.

Qualitative data was gathered through 15–30-minute semi-structured interviews with each trainee at the end of the 2-month period. During these interviews, trainees provided detailed narratives of their experiences interacting with GPT-4 and feedback on its integration into their training. Thematic analysis of the interview transcripts revealed that 14 out of 20 trainees expressed positive sentiments regarding GPT-4's integration, highlighting benefits such as immediacy of feedback, confidence building, and enriched learning. The remaining 6 trainees acknowledged some benefits but also expressed concerns about over-reliance on the AI.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Statistical evaluations employing t-tests and regression models have substantiated a positive correlation between frequent GPT-4 engagement and an observable enhancement in student translation proficiency. This association can be visually examined in Figure 2, which presents a regression plot highlighting the upward trajectory of translation skills concomitant with increased GPT-4 usage.

Figure 2. Regression plot of GPT-4 usage against translation proficiency gains
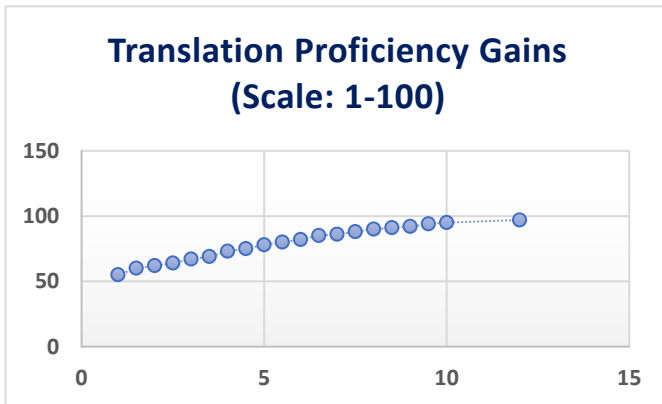
The plot's consistent upward slope indicates trainees gained proficiency as GPT-4 usage increased, with gains tapering slightly at higher usage. For instance, 1-3 hours of use resulted in ~12 point proficiency increase, while 9-12 hours only increased 3 points, suggesting potential diminishing returns. Still, the

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

consistent performance across trainees highlights GPT-4's reliability in aiding proficiency.

While the overarching theme suggests a favorable outcome due to GPT-4's integration, a more granular analysis revealed intricate nuances. Specific text genres and registers showed consistent improvements; however, certain complex subject matters still present challenges. These intricate challenges resonate with concerns raised by scholars like Venuti (2013), who accentuated the occasional necessity of human intervention to achieve translation accuracy and preserve cultural nuances. Such insights further emphasize the imperative for a judicious balance between leveraging GPT-4's capabilities and ensuring the perpetuation of the human touch in translation, ensuring the art remains unblemished.

The overall positive impact of GPT-4 on enhancing trainees' translation proficiency aligns with findings by Hendy et al. (2023) that leveraging GPT models' capabilities through in-context learning with few-shot examples can further improve translation quality. However, observations of persistent challenges for certain complex topics despite GPT-4 usage reinforce conclusions from the study that

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

human collaboration remains vital to address nuanced cases.

As Brown et al. (2020) posits, while GPT-4 is poised to revolutionize translator training with its sophisticated language comprehension capabilities and real-time feedback, it is not devoid of limitations. For instance, the challenges GPT-4 faces in translating minority languages or regional dialects and capturing emotional intricacies within literary texts necessitate a collaborative approach wherein the AI serves as an assistant rather than a replacement. In alignment with these sentiments, Pym and Malmkjær (2013) and Venuti (2013) echo the importance of human expertise, reinforcing that while AI has its merits, human expertise remains paramount to ensure translation accuracy, cultural appropriateness, and ethical considerations.

Navigating the realm of ethics, Bowker (2002) has underscored the myriad challenges tied to AI integration in translator training. With concerns ranging from data privacy to potential job displacement and even the possible dehumanization of the translation process, it is evident that a conscientious path must be charted. As the integration of AI into translator training intensifies, these ethical considerations demand meticulous

attention, pointing to the need for a well-articulated set of guidelines and frameworks.

The dynamism infused into the field of translator training by AI necessitates continuous pedagogical reevaluations. It is an invitation for educators to stay abreast of technological advancements while ensuring that their curricula remain rooted in the foundational principles of translation. Embracing this duality – combining traditional methodologies with futuristic AI tools – is crucial. As the narrative unfolds, it is evident that the union of human expertise with AI's prowess, as observed through the trainees' journey with GPT-4, paves the way for a future where translation remains as much an art as it is a science.

4.3. Synthesis and Cross-case Analysis

Upon conducting a thorough cross-case analysis, integrating qualitative themes and quantitative metrics, a complex landscape emerges, painting a multi-faceted picture of GPT-4's application in the translation domain. By juxtaposing various instances across different contexts, this method allows for a profound dive into the intricate layers of GPT-4's integration into translator training, rendering a richer understanding than would have
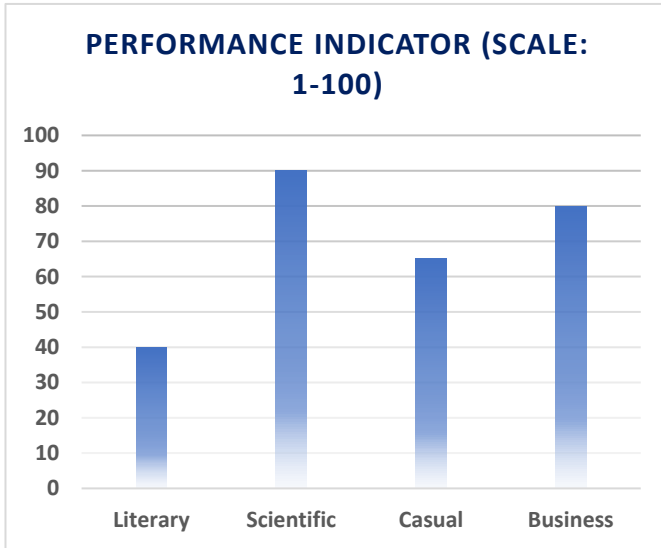
Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

been possible by evaluating individual cases in isolation.

A pivotal finding from this investigation reveals GPT-4's varying degrees of performance across diverse translation contexts. As succinctly illustrated in Figure 3 - a bar chart showcasing GPT-4's performance metrics across four different text genres (Literary, Scientific, Casual, and Business) - it is evident that GPT-4 excels in scientific contexts. The chart displays a marked superiority in this domain while showing promising results in the business genre. There is a moderate performance in casual translations but a noticeable dip when dealing with literary content, reflecting struggles with nuanced elements inherent in this genre.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Figure 3. GPT-4's performance across text genres

Analyzing Figure 3 in-depth, GPT-4's excellence is most pronounced in contexts characterized by conventional linguistic structures and standardized terminologies. Its capability to swiftly identify and recreate contextually appropriate translations in these areas mirrors the capabilities highlighted by Rospigliosi (2023). However, in more nuanced domains, like the literary sector, which demands a deeper understanding of culture-specific idioms, regional colloquialisms, or intricate literary expressions, GPT-4 is observed to lag. These findings align with the observations made by İpek et

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

al. (2023). In such intricate contexts, GPT-4 either necessitates more informational inputs or requires human collaboration to produce optimal translation outcomes.

However, the actual value of cross-case analysis is not merely in identifying strengths and weaknesses. It lies in synthesizing them to offer actionable insights. Educators and curriculum designers can tailor their pedagogical strategies by discerning where GPT-4 excels and where it stumbles, maximizing its strengths and counteracting its limitations.

Moving forward, by integrating insights from AI in education (AIED), as proposed by scholars like Selwyn (2016) and Chen et al. (2020), it is clear that GPT-4's success in translator training is not solely rooted in its technical prowess. It stems from the synergy between technology, pedagogy, and learners. Echoing Zheng et al. (2021) principles of effective AIED integration, there is a need for an all-encompassing approach. This approach sees GPT-4 as a tool and a dynamic contributor to the learning environment.

In conclusion, this investigation into GPT-4's capabilities and applications in translator training

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

reveals several key findings. Structured prompts demonstrate a dichotomy between technical translation excellence and nuanced challenges. Experts emphasize prudent integration to retain human artistry. Trainee interactions exhibit AI's benefits but need for collaboration. Finally, cross-case analysis spotlights varied performance across contexts, necessitating tailored integration strategies. While limitations exist, GPT-4 is promising to enhance pedagogy if deployed judiciously. This research elucidates a balanced approach, celebrating AI's potential while grounding training in human expertise. The field can unleash a new era of enriched translator education by leveraging technology to amplify art. Continuous reevaluation and refinement of integration strategies will be essential to harness the symbiotic power of humans and machines fully.

5. CONCLUSION

This research elucidating the integration of GPT-4 into translator training programs has unveiled nuanced insights, blending opportunities with ethical risks that demand judicious navigation. The structured translation prompts provided empirical evidence of GPT-4's sterling capabilities in technical domains, securing high accuracy metrics

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

for terminology translations. However, a marked decline in scores for literary prompts indicates deficiencies in conveying semantic and cultural nuance, affirming experts' caution that human discernment remains irreplaceable for high-quality translation.

These insights on GPT-4's dichotomous strengths and limitations reinforce measured integration approaches that harness AI capabilities while upholding human creativity as the core of ethical and accurate translation. Experts in AI-enabled pedagogy advocate blended learning models that thoughtfully incorporate GPT-4's utilities, like instant feedback and diverse examples, while focusing on building human expertise through immersive training. The observable benefits but the need for collaboration in trainee interactions validate this balanced approach. Moreover, cross-case analysis reveals variances in aptitude across diverse genres and subjects, demanding tailored and strategically scoped deployment of GPT-4.

In conjunction with these opportunities, perspectives from experts and scholars highlight pressing ethical risks requiring deliberation as integration proceeds. Concerns around data privacy, algorithmic bias, and accountability, illuminated by

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Memarian and Doleck (2023), underline the need to formulate comprehensive guidelines that govern AI's responsible and ethical utilization in translator training. Additionally, the intermingling of human and machine contributions in AI-mediated translation compels re-evaluating traditional notions of authorship, a complex issue warranting rigorous examination (Ouyang & Jiao, 2021).

In summary, while limitations exist, GPT-4 represents a milestone in the evolution of translator training, demonstrating the immense potential to transform pedagogy through personalized feedback, exposure to diverse examples, and enhanced interactivity. However, measured integration that embraces the symbiotic relationship between technological prowess and human artistry emerges as the prudent way forward. Such an approach can equip professionals to thrive in the emerging technological landscape while respecting translation's profound human essence. Overall, the blend of quantitative metrics and qualitative insights provides a valuable prognosis for the field, guiding the assimilation of AI in a manner that employs its utilities to amplify competencies without compromising on ethics or the irreplaceable human touch that breathes life into language.

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

REFERENCES

Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International, 2023*.

Angelone, E. (2022). Weaving adaptive expertise into translator training. In *The Human Translator in the 2020s* (pp. 60–73). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781003223344-5/weaving-adaptive-expertise-translator-training-erik-angelone

Austermuhl, F. (2014). Electronic tools for translators. Routledge. https://doi.org/10.4324/9781315760353

Bowker, L. (2002). Computer-aided translation technology: A practical introduction. *University of Ottawa Press*.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot

293

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Denkowski, M., & Lavie, A. (2014). Meteor universal: Language specific translation evaluation for any target language. *Proceedings of the Ninth Workshop on Statistical Machine Translation*, 376–380.

Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., Kim, Y. J., Afify, M., & Awadalla, H. H. (2023). How Good Are GPT Models at Machine Translation? A Comprehensive Evaluation (arXiv:2302.09210). arXiv. https://doi.org/10.48550/arXiv.2302.09210

İpek, Z. H., GÖZÜM, A. İ. C., Papadakis, S., & Kallogiannakis, M. (2023). Educational

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

applications of the ChatGPT AI system: A systematic review research. *International Journal, 12*(3), 26–55.

Jiao, W., Wang, W., Huang, J. T., Wang, X., & Tu, Z. P. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. *arXiv Preprint arXiv:2301.08745*.

Kelly, D. (2014). A handbook for translator trainers. Routledge. https://doi.org/10.4324/9781315 760292

Kenny, D. (2020). Technology and translator training. In M. O'Hagan (Ed.). *The Routledge handbook of translation and technology*, (pp. 498–515). Routledge.

Khoshafah, F. (2023). ChatGPT for Arabic-English translation: Evaluating the accuracy. https://doi.org/10.21203/rs.3.rs-2814154/v2

Kolb, D. (1984). Experiential learning: Experience as the source of learning and development. *Prentice-Hall*.

Krüger, R. (2023). Artificial intelligence literacy for the language industry – with particular

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

emphasis on recent large language models such as GPT-4. *Lebende Sprachen*. https://doi.org/10.1515/les-2023-0024

Lacruz, I. (2023). Translation in transition: Human and machine intelligence. In I. Lacruz (Ed.). *Translation in transition: Human and machine intelligence* (pp. 1–14). John Benjamins Publishing Company. https://doi.org/10.1075/ata.xx.011ac

Madsen, M., & Gregor, S. (2000). Measuring Human-Computer Trust. https://www.semanticscholar.org/paper/Measuring-Human-Computer-Trust-Madsen-Gregor/b8eda9593fbcb63b7ced1866853d9622737533a2

Mao, R., Chen, G., Zhang, X., Guerin, F., & Cambria, E. (2023). GPTEval: A Survey on Assessments of ChatGPT and GPT-4 (arXiv:2308.12488). arXiv. https://doi.org/10.48550/arXiv.2308.12488

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Memarian, B., & Doleck, T. (2023). Fairness, accountability, transparency, and ethics (FATE) in artificial intelligence (AI) and higher education: A systematic review. *Computers and Education: Artificial Intelligence, 5,* 100152. https://doi.org/10.1016/j.caeai.2023.100152

Mihalache, I. (2021). Human and non-human crossover: Translators partnering with digital tools. In R. Desjardins, C. Larsonneur, & P. Lacour (Eds.), *When translation goes digital* (pp. 19–43). Springer International Publishing. https://doi.org/10.1007/978-3-030-51761-8_2

Muñoz-Miquel, A., Montalt, V., & García-Izquierdo, I. (2020). Fostering employability through versatility within specialisation in medical translation education. https://repositori.uji.es/xmlui/handle/10234/197099

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence, 2,*

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

100020. https://doi.org/10.1016/j.caeai.2021.100020

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). Bleu: A method for automatic evaluation of machine translation. In P. Isabelle, E. Charniak, & D. Lin (Eds.). *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, (pp. 311–318). Association of Computational Linguistics.

Pym, A. (2013). Translation skill-sets in a machine-translation age. *Meta, 58*(3), 487–503. https://doi.org/10.7202/1025047ar

Pym, A. (Director). (2023, April 27). GPT in the training of translators [YouTube]. https://www.youtube.com/watch?v=b9U_FUaneso

Pym, A., Grin, F., Sfreddo, C., & Chan, A. L. (2013). The status of the translation profession in the European Union. *Anthem Press*.

Pym, A., & Malmkjær, K. (2013). Translation and language learning: The role of translation in the teaching of languages in the European

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Union: a study. *Publications Office of the European Union.*

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., & Clark, J. (2021). Learning transferable visual models from natural language supervision. *International Conference on Machine Learning*, 8748–8763.

Radford, A., & Narasimhan, K. (2018). Improving Language Understanding by Generative Pre-Training. https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Rospigliosi, P. a. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments, 31*(1), 1– 3. https://doi.org/10.1080/10494820.2023.2 180191

Selwyn, N. (2016). Is technology good for education? *Polity Press*.

Škobo, M., & Petričević, V. (2023). Navigating the challenges and opportunities of literary translation in the age of AI: Striking a balance between human expertise and machine power. *Društvene i Humanističke Studije (Social and Humanistic Studies), 8*(2 (23)), 317–336.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, (pp. 223–231).

Unbabel. (2020, August 26). Behind the GPT-3 Buzz: Why Human-in-the-Loop AI is

Abu-Rayyash, H. (2023). Revolutionizing translator training through human-AI collaboration: Insights and implications from integrating GPT-4. *Current Trends in Translation Teaching and Learning E,* 10, 259 – 301. https://doi.org/10.51287/cttl20239

Important. https://resources.unbabel.com/blog/behind-the-gpt-3-buzz-why-human-in-the-loop-ai-is-important

Venuti, L. (2013). Translation changes everything: Theory and practice. *Routledge*.

Weisz, J. D., Muller, M., Houde, S., Richards, J., Ross, S. I., Martinez, F., Agarwal, M., & Talamadupula, K. (2021). Perfection not required? Human-AI partnerships in code translation. *26th International Conference on Intelligent User Interfaces*, (pp. 402–412). https://doi.org/10.1145/3397481.3450656

Zheng, L., Niu, J., Zhong, L., & Gyasi, J. F. (2021). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, 1–15. https://doi.org/10.1080/10494820.2021.2015693

Zhu, M. (2023). Sustainability of translator training in higher education. *Plos One, 18*(5), e0283522. https://doi.org/10.1371/journal.pone.0283522