À tradire
ISSN : 2968-3912

4 | 2025
Écrire et traduire à plusieurs : perspectives didactiques

CAT Tools or LLMs? Benefits and Challenges of Translating Collaboratively with Digital Tools: A Case Study at IULM University

Outils de TAO ou LLM ? Avantages et défis de la traduction collaborative à l'aide d'outils numériques : une étude de cas à l'université IULM
CAT Tools o LLM? Benefici e sfide della traduzione collaborativa con strumenti digitali: un caso di studio all'università IULM

Federica Villareale

https://atradire.pergola-publications.fr/index.php?id=605

DOI : 10.56078/atradire.605

Federica Villareale, « CAT Tools or LLMs? Benefits and Challenges of Translating Collaboratively with Digital Tools: A Case Study at IULM University », *À tradire* [], 4 | 2025, 28 avril 2026, 30 avril 2026. URL : https://atradire.pergola-publications.fr/index.php?id=605

Licence Creative Commons – Attribution 4.0 International – CC BY 4.0

CAT Tools or LLMs? Benefits and Challenges of Translating Collaboratively with Digital Tools: A Case Study at IULM University

Outils de TAO ou LLM ? Avantages et défis de la traduction collaborative à l'aide d'outils numériques : une étude de cas à l'université IULM
CAT Tools o LLM? Benefici e sfide della traduzione collaborativa con strumenti digitali: un caso di studio all'università IULM

Federica Villareale

Introduction
Methodology
  Groups and work organisation
  Data collection
Study
  Choosing the best AI-based tool for translation
  Translating with a CAT tool: Smartcat
  Translating with ChatGPT
Conclusions

Introduction

1 In recent decades, remarkable advancements in technology have transformed the way we live, communicate, and work. Consequently, all aspects of translation have been profoundly impacted, resulting in significant changes. For example, the emergence of new professional roles driven by the rise of machine or computer-assisted translation (CAT) tools, the growth of large-scale collaborative non-professional translations, and innovations in translation theory stemming from technological breakthroughs.

2 Furthermore, over the years, the demand for translation has increased as we live in a world where communication is becoming more multimodal and multilingual. Just as the Internet was beginning to spread widely, scholars agreed that the main goal of translation is to convert a text originally written in one language into its equivalent

in another language, in a way that preserves the original's meaning, formal features, and functional roles (cf. Bell, 1991).

3 The advent and development of new technologies such as Neural Machine Translation (NMT), generative AI, and Large Language Models (LLMs) could facilitate the work of linguists. These technologies offer significant support with their ability to simulate natural language—even when transmitting complex messages—and adapt texts to the target culture. After all, as early as 1969, Nida and Taber stated that translation is the reproduction of the closest natural equivalence of the source language in the target language.

4 This led translators to learn how to collaborate with such digital tools, as they can provide a significant advantage in their work, especially regarding time and collaboration. Quoting Valentina Piotto and Giuseppe Sofo (2023: 236):

> L'aspetto particolarmente interessante di queste pratiche è che si basano su una concezione completamente diversa della tecnologia digitale nella traduzione; lungi dall'essere la minaccia "disumanizzante" che viene solitamente percepita, l'apporto tecnologico alla traduzione è visto in questo caso come uno "strumento di convivialità e uno strumento di intervento politico umano", basato su "un allontanamento dal soggetto monadico della traduzione tradizionale nella direzione di una plurisoggettività dell'interazione" [Cronin, 2013: 102].

> The most interesting aspect of these practices is that they are based on a completely different view of digital technology in translation. Instead of being seen as the "dehumanising" threat it is often perceived to be, technology's role in translation is regarded as "a tool of conviviality and an instrument of human political intervention", grounded in "a move away from the monadic subject of traditional translation to a plurisubjectivity of interaction" [Cronin, 2013: 102].

5 As Piquet (2009) already stated, the concepts and tools related to collaborative work are not new. However, over time, they have gained a whole new dimension due to the democratisation of information and communication technologies (ICT) in society and, consequently, in our organisations. When discussing collaborative work, it is essential to consider the substantial technological dimension and the fact

that all available collaborative work tools are evolving rapidly in response to market and user needs. They are maturing from technical, economic, and social perspectives, becoming more user-friendly, easier to install, and more cost-effective in terms of software and performance (Laurenti & Villareale, 2023).

6 There are many useful tools available online for collaborative tasks at present, each with specific technical features. A workgroup planning to use them must consider several factors to ensure the chosen tool is effective and suitable for achieving the goal and does not disrupt the workflow.

7 Selecting the right tool for a specific task is essential in the collaborative process, as not all technologies suit every situation. Making an informed decision can significantly enhance the quality of the work. However, simply choosing the correct tools is not enough to guarantee that the task is performed correctly. It is crucial that all participants genuinely understand what it means to work collaboratively and are fully committed to this approach, rather than trying to stand out. In collaborative work, individual subjectivities merge to form a unified group identity.

8 People working on a collaborative translation task using collaborative tools, like the several machine translation tools available on the market, must remember that all these are imperfect technologies (O'Brien, 2022: 105); therefore, the risk of misunderstandings in written translation remains high, especially because people involved in the communication cannot use their body language to help the translator exactly understand the meaning of their message. Consequently, a thorough revision of the translated text continues to be essential.

9 For this reason, the study, carried out by the International Center for Research on Collaborative Translation [1] (hereinafter the Center) at the IULM University in Milan, investigates how two groups of students collaborated with each other and with different digital tools to understand the benefits that translators can derive from using traditional computer-assisted translation (CAT) tools, machine translation, and LLMs, and the most effective way to organise collaborative translation work. It also examines the most common challenges associated

with using these tools, providing an initial guide for professionals and semi-professionals on how to approach translation tasks with them.

Methodology

10 For this research, the Center collaborated with the EMT-DGT (European Master's in Translation – Directorate-General for Translation of the European Commission, Brussels)[2] and the Laboratorio di Redattologia e Traduttologia[3] of the University of Udine. The 20 texts which the 10 students from the IULM University worked on were sourced from the EMT blog, and the Italian translations were then published there. All the texts concerned translation experiences; the style and language were informal, as they are typical of informative blog articles. Two students from the University of Udine helped IULM students during the final revision step (12 students participated in the experiment).

Groups and work organisation

11 This study aimed to evaluate whether using NMT or LLM tools for translation is worthwhile compared to traditional CAT tools, considering that they tend to produce less accurate output requiring closer revision. Consequently, it was essential to assess the quality of the translation relative to the total time spent on the task.

12 Given the numerous NMT and LLM tools available online, it was necessary to select the most suitable one for this particular translation task. Therefore, the work began by first determining how to assess the tools' performance and selecting the appropriate metric to use (see Section "Choosing the best AI-based tool for translation").

13 Once this evaluation was completed, the researchers proceeded to divide the students into two groups of six people each (5 translators/reviewers, 1 editor): one group translating using a free collaborative CAT tool (Smartcat) and the other post-editing the raw output generated by the selected digital tool.

14 Students collaborated to produce a translation ready to be published on the EMT blog, and the workflow was organised as follows (please note that, when talking about "translation", we refer to the entire

workflow a translator follows before delivering their translation to the reviewer, including the self-revision of their own work. In this case, the step of self-revision included more than one translator, as it was a collaborative project):

15 Step 1: **Translation**

Group that worked with CAT-tool (Group A)	**Group that worked on AI output (Group B)**
• Each student first translated and checked a part of the texts on Smartcat individually. • The entire group then collaboratively checked the translations using the Smartcat comment function. • Once the first self-revision was completed, they extracted the bilingual and monolingual Italian files and uploaded them to a Drive folder shared with all students and coordinators. • In the Drive folder, they collaboratively had a final self-revision of all the texts they translated, even offline, using the edit/comment function.	• Each student first translated and checked a part of the texts individually, creating a Word file to be uploaded to a Drive folder shared with all students and coordinators. • The entire group then met on a Microsoft Teams video call and collaboratively checked all the texts in real-time. • They uploaded the final translation to a Drive folder shared with all students and coordinators.

16 Step 2: **Revision/editing**

- In Drive, the two groups collaboratively revised the texts they did not initially translate, creating a copy of the file directly in the Drive folder.
- When the reviewers finished, translators accepted/rejected the changes made.
- When everyone agreed, they uploaded the final revised texts to a Drive folder shared with all students and coordinators.
- The students of the University of Udine revised all the texts one more time, keeping in touch with the IULM work group.
- Texts were ready for publication.

17 Appendix 1 shows precisely how each group proceeded with its work at every step of the workflow.

18 A qualitative analysis of the texts' readability was then carried out, using comparison tables to examine more closely the changes made at each stage of the workflow.

19 Finally, researchers also used students' reports to analyse the time spent on each stage of the translation process.

Data collection

20 As regards the calculation of timing, a proportion was adopted to adjust the results to the sample of characters considered, as the two groups worked on 75,336 and 83,510 characters, respectively. The sum of the total time that each subject spent in each phase of the work to translate and review the total number of characters was then proportioned to the number of characters in the portion of text examined. To collect significant data and fairly compare working times, researchers decided to analyse the translation of a text sample of 5,232 characters for each tool, so that both texts were of the same length (e.g., Total characters = 83,510, Text extract sample characters = 5,232, Total time taken by all participants to translate Total characters = 710 min, total time taken by all participants to translate Text extract characters = 44 min [proportion result 5,232 ÷ 83,510 = x ÷ 710]. That is why the result obtained is not multiplied by the number of participants (five per group), because the number already accounts for the total time.

21 To view all the data and verify the proportions used in each phase, refer to Appendix 1.

Study

Choosing the best AI-based tool for translation

22 The initial step involved assessing which AI-based tool should be used to generate the raw output for post-editing. Researchers analysed the translation of the same text provided by six of the most widely used online tools: Google Translate, Systran, Yandex, DeepL, Microsoft Translator (as NMT), and ChatGPT (as LLM).

23 Researchers evaluated the translation from English into Italian of the article *My Distance Learning*,[4] sourced from the EMT blog, using a manual, quantitative metric called SAE J2450.

24 This metric measures common errors made during translation, regardless of the source or target language, and whether humans or

machines perform the translation. Developed by the Society of Automotive Engineers, the metric is described as "a score sheet that enables evaluators to capture error types and quantities of translation errors" (Woyde, 2001: 38). It provides a translation error scoring system (Translation Quality Score, or TQS) to assess translation quality. It includes seven main error categories (i.e., wrong terms, syntactic errors, omissions/additions, word structure and agreement errors, misspelling, punctuation errors, miscellaneous) and two error severity levels (minor or serious). When assessing translation quality, each error identified by the evaluator is classified into one of the seven categories; after determining its primary category, the evaluator decides whether it is a serious or minor error based on its severity. Both classification levels are subjective judgments by the evaluator. Once the primary and secondary categories are set, each error is weighted and summed. This total is then divided by the number of words in the translated text to produce the translation quality score (Woyde, 2001). A lower score indicates higher translation quality.

Table 1. TQS of different tools

	Google Translate		Systran		Yandex		Microsoft		DeepL		C
	Serious	**Minor**	**Serious**	**Minor**	**Serious**	**Minor**	**Serious**	**Minor**	**Serious**	**Minor**	**Serio**
Wrong terms	2	1	4		4		2	1	3		
Syntactic errors	4		3	1	7	1	3		1	1	
Omissions			1								
Word structure and agreement errors		1	1	1		1	1	1			
Misspellings											
Punctuation errors											
Miscellaneous										1	
TOTAL	14		20		24		14		10		
Number of words	799		787		805		808		794		
TQS	1.75		2.54		2.98		1.73		1.26		

25 As shown in the table, ChatGPT scored higher, while DeepL performed better compared to the other NMTs. However, as these two types of technologies (NMT and LLMs) operate in very different ways, their translation outputs can differ considerably, and it is essential to consider the benefits and limitations that might arise from using each tool.

26 Machine translation, which utilises various algorithms, patterns, and large databases of existing translations, takes a source text, divides it into words and phrases—segments—and substitutes these with corresponding words and phrases in another language (the target) (Smartcat, 2022). In other words, MT is explicitly made for translating written texts.

27 Conversely, Large Language Models use deep learning techniques to understand, summarise, generate, and predict new content. Once trained, an LLM provides a basis for AI to be used for various practical purposes (Stryker, 2023). One such purpose is to generate and translate texts by taking an input prompt and leveraging the learned knowledge to predict the next word or phrase. The model iteratively produces the output, considering the context of the input and its previous predictions. Therefore, LLMs were not created solely for translation, but they can perform this task.

28 Considering all this, after evaluating the pros and cons of each tool, it was decided to utilise ChatGPT as the foundation for the translation since it has no typing restrictions, can be asked to edit sentences endlessly, and exhibits superior rephrasing abilities. Nonetheless, one must remain vigilant with the source text because, when translating from English into Italian, rephrasing can differ significantly from the original, altering the meaning of the translation; therefore, thorough post-editing is always necessary.

Translating with a CAT tool: Smartcat

29 To compare the benefits of collaborative translation work performed using a CAT tool versus an LLM, a decision was made to evaluate the texts produced based on changes made at each translation step and to analyse the time required by the translators.

30 To collect significant data and fairly compare working times, researchers decided to analyse the translation of a text sample of 5,232 characters (length of the source text translated with Smartcat: "Innovating for accessibility: Sign language at the University of Geneva's FTI")[5] for each tool, so that both texts were of the same length. Researchers then compiled a table highlighting the steps of the working process, the changes made to the text during each of these steps, and, finally, the type of changes made, to understand the problems encountered that led to a variation in the following step.

Table 2. Workflow, time, and changes – Smartcat

	Step 1: Original text > first translation	**Step 2:** First translation > internal revision within the Group A (A1 on A1, A2 on A2, A1 on A2, A2 on A1)	**Step 3:** Internal revision > revision outside Group A (B on A, A on REV(B))	**Step 4:** Editing of texts translated by Group A	**Total time spent**
Time spent	86 min.	54 min.	57 min.	46 min.	**243 min. 4 h**
Most significant changes			✓	✓	
			Rendering in Italian		

31 See the appendix for the detailed timing by translator in total and proportioned to the number of characters in the portion of text examined.

32 As expected, the first translation produced on Smartcat took the participants about 1.5 hours, as they had to create a new text in another language from scratch. Although the first translation was already of medium-high quality, the revision steps took 2.6 hours, bringing the total effort to 4 hours. Most changes occurred during the external revision phase, which was carried out by other translators, and in the final editing process before publication. These two tasks required a total of 1.7 hours and mainly involved the Italian rendering.

33 The following table displays some of the final changes made to the first translation, focusing on improving fluency in Italian and selecting terms more appropriate for the context and target language.

Table 3. Examples of changes

Error type	Original text	Step 1	Step 2	Step 3	Step 4
Rendering in Italian	Accessibility is about making sure that people with disabilities and/or with special needs have access to society on an equal basis.	Accessibilità significa assicurare alle subjecte con disabilità e/o con speciali esigenze di avere accesso alla collettività in maniera equa.	-	Accessibilità significa assicurare alle subjecte con disabilità e/o con **bisogni speciali** di avere accesso alla **società** in maniera equa.	-
	We had two main goals in setting up this programme: improving the inclusion of Deaf people by making the workplace more accessible, and making information more accessible to a wider public by training communication	Avevamo due obiettivi principali quando abbiamo messo a punto questo programma: da una parte l'inclusione delle persone sorde nei luoghi di lavoro rendendoli più accessibili, dall'altra **parte** far sì che le informazioni raggiungano un pubblico più ampio	-	Avevamo due obiettivi principali quando abbiamo messo a punto questo programma: da una parte l'inclusione delle persone sorde nei luoghi di lavoro rendendoli più accessibili, dall'altra far sì che le informazioni raggiungano un pubblico più ampio	**Due erano gli obiettivi principali quando è stato messo a punto il programma**: da una parte **includere** le persone sorde nei luoghi di lavoro rendendoli più accessibili, dall'altra **convogliare le informazioni** verso un pubblico più ampio

Translating with ChatGPT

34 The same type of analysis was conducted on the edited text initially translated by ChatGPT. In this case, translators were asked to work on another text, always sourced from the EMT blog ('*Can you see me? Can you hear me?' New teaching and learning environments and the new 'normal'*) [6]; researchers then took into consideration for their analysis the first part of the text (5,232 characters), to correctly compare working times and data.

Table 4. Workflow, time, and changes – ChatGPT

	Step 1: Original text > raw output	**Step 2:** Raw output > First individual MTPE of the translation	**Step 3:** Collaborative revision (video call on Microsoft Teams)	**Step 4:** Collaborative revision > revision outside Group B (A on B, B on REV(A))	**Step 5:** Editing of texts translated by Group B	**Total time spent**
Time spent	3 min.	81 min.	55 min.	52 min.	34 min.	**224 min. 3.75 hours**
Most significant changes		✓ Calques Grammar errors Comprehension			✓ Rendering in Italian Gender neutrality	

35 See the appendix for the detailed timing by translator in total and proportioned to the number of characters in the portion of text examined.

36 Again, as expected, although the translation time is very short – only 3 minutes, the revision time has increased significantly to 3.75 hours. However, in such cases, only the post-editing and revision time should be considered for the workload, as the translation was generated very quickly by the tool.

37 Most of the editing was done during the initial post-editing of the raw text to address calques, grammatical errors, and comprehension issues. It is the most important part of the workflow: the translator works on a translation that, even though it is already good and understandable, must be carefully examined, thinking about every word and sentence, trying to render the text as suitable as possible for their target market. The changes made during this step not only correct calques and grammatical or comprehension errors, but are also intended to make the Italian text as fluid and comprehensible as possible. However, an external revision is still needed, since a different subject, reading the translated text for the first time and comparing it with the original text, may notice errors or inaccuracies that escaped the translator's eye, who is accustomed to their own translation and understanding. That is why we had Steps 4 and 5.

38 In fact, the editor working in the final phase aimed to improve the Italian translation and, particularly, to address gender-neutrality issues—a common problem in Italian when a text is directed towards an unspecified audience. The machine tends to translate all adjectives and nouns referring to people of all genders with the overused masculine form in Italian. Still, for inclusivity reasons, it is always better to use gender-neutral words in Italian.

Table 5. Examples of changes

Error type	Original text	Step 1	Step 2	Step 3	Step 4	Step 5
Gender neutrality	but the key is to be adaptive, flexible and patient.	ma la chiave è essere **adattabili, flessibili e pazienti**.	-	-	-	ma la chiave è **sapersi adattare con flessibilità e pazienza**.
Casts/Comprehension	Teaching distance learning works in the scenarios already mentioned and it is a solution whilst we minimise the risk of infection on campus and wait for the pandemic to abate.	L'insegnamento a distanza funziona negli scenari già menzionati ed è una soluzione mentre **cerchiamo di minimizzare** il rischio di infezione in **campus** e attendiamo che la pandemia **diminuisca**.	L'insegnamento a distanza funziona negli scenari già menzionati ed è una soluzione mentre **riduciamo** al minimo il rischio di infezione **in presenza** e attendiamo che la pandemia **si attenui**.	-	-	L'insegnamento a distanza funziona negli scenari già menzionati **ed è una soluzione per ridurre al minimo il rischio di contagio in presenza mentre attendiamo che la pandemia si attenui**.
	exercise lessons	lezioni di **esercizio**	lezioni di **ginnastica**	-	-	-
Rendering in Italian	Can you see me? Can you hear me?	'Riesci a vedermi? Riesci a sentirmi?'	-	-	-	**Mi vedete? Mi sentite?**

Conclusions

39 It must be noted that this is a pilot study involving only two workflows; future research could involve larger groups of translators and analyse workflows using different digital tools (e.g., NMT, CAT tools integrated with AI, etc.).

40 As demonstrated, although the time needed to generate the first translation from the original text in the LLM workflow is significantly reduced (3 vs 86 min.), the overall time for translators to deliver the final version has not decreased substantially (3.75 vs 4 h). This is because AI does not always understand the context and linguistic subtleties, sometimes producing sentences that are hard to understand or even inaccurate – which are easier to identify, but also sentences that, despite appearing grammatically and syntactically correct, do not truly reflect the meaning of the original text ("transparent errors"). The translator then must review the raw output very carefully, correcting errors and adapting the language to the specific context, as only an expert linguist can make the necessary adjustments. Clearly, this step takes time.

41 On the other hand, the use of CAT tools allows translators to immediately produce a more accurate version of the text in the target language, requiring less revision. However, this forces them to invest much more time in this step, especially if, as in this case, they do not have a Translation Memory available. In fact, students reported that "working from a shared Word file would have been much faster".

42 Taking everything into account, it is impossible to say with certainty whether one workflow is more convenient for translators than the other. What is certain, however, is that these two types of work are different and require different skills, as the post-editors of AI output must have a very deep understanding of how machines work and what errors they most frequently make.

43 In this context, it would be beneficial for translators to receive specialised training on the most common machine errors in translation. This would enable them to identify these errors more swiftly and thus reduce the time spent on post-editing the translated text. Researchers from the Centre conducted a study on the advantages for students of engaging with AI-based NMT and LLM systems, as well as receiving targeted training on how these tools operate and the types of errors they tend to generate. The findings of the research were presented at the PRIN *International Conference*, held in Bergamo on July 16-17, 2025 [7].

44 Ultimately, it is essential for translators, particularly when working with AI-translated texts, to retain their creativity, as some linguistic

problems cannot be solved by machines alone. Only a human translator's creativity, intuition, and skill can overcome these challenges. For example, as previously mentioned, when working with language pairs in which one language faces issues with gender neutrality, the translator's resourcefulness can produce an accurate version that faithfully reflects the source text's meaning while respecting the subtleties of the target language.

45 Another aspect we observed is that using a real-time communication tool (Microsoft Teams or WhatsApp, in this case) accelerates certain steps of the workflow and communication (Laurenti & Villareale, 2023), as reported by students involved. In fact, the possibility of asynchronous communication via the comment feature of CAT tools undoubtedly gives translators more flexibility in managing their work and time, but it also extends the delivery time for the final text.

46 We can then reflect on the practice of collaborative translation. In their reports, students recognised the practical usefulness of collaborative translation, which helps translators improve both their translation skills and soft skills. Collaborative work promotes discussion and the exchange of ideas, enables translators to understand better the source text, and enhances their ability to identify and correct errors in a translated text. It also boosts their confidence in their own abilities and skills.

47 In light of what has been said so far, it can be concluded that collaborative translation, carried out with the aid of digital tools, not only improves the final quality of the translated text but also enhances translators' working conditions (Laurenti and Villareale, 2023). It allows them to save time on mechanical tasks and invest it in creative tasks that require human input. Therefore, technology proves to be a valuable aid to translators, provided they know how to utilise it effectively.

Bell Roger Thomas, 1991, *Translation and Translating: Theory and Practice*, London and New York, Longman.

Cronin Michael, 2013, *Translation in the Digital Age*, London, Routledge, 2013.

Farrell Michael, 2018, "Machine translation markers in post-edited

machine translation output", *Proceedings of the 40th Conference Translating and the Computer*, AsLing, p. 50-59.

Farrell Michael, 2023, "Current evidence of post-editese: differences between post-edited neural machine translation output and human translation revealed through human evaluation", in *Proceedings of the International Conference Hit-IT 2023*, p. 52-63, [https://doi.org/10.26615/issn.2683- 0078.2023_005 (https://doi.org/10.26615/issn.2683-0078.2023_005)], viewed on 27 August 2025.

Kamaluddin Mohamad Ihsan, Rasyid Moch. Wildan Khoerul, Abqoriyyah Fourus Huznatul, SAEHU Andang, 2024, "Accuracy Analysis of DeepL: Breakthroughs in Machine Translation Technology" *JEEF: Journal of English Education Forum*, Vol. 4, No. 2, p. 122-126.

Laurenti Francesco, Villareale Federica, 2023, "Traduzione, ambienti virtuali e nuove tecnologie: strumenti informatici per la collaborazione", *Testo a Fronte. Teoria e pratica della traduzione*, Vol. 69, No. 2, p. 201-220.

Metha Alifa S, Hidayah Annisa Ghaisani Dzatil, Aditya Rahmadsyah, Wihadi Marwito, 2024, "Artificial Intelligence Meet Language as Technology Advances in Translation Tools", *INJUCHUM: International Journal of Computer in Humanities*, Vol. 1, No. 1, p. 51-58.

Nida Eugene, Taber Charles, 1969, *The theory and practice of translation*, Leiden, E.J. Brill.

O'brien Sharon, 2022, "How to deal with errors in machine translation: Post-editing", in Dorothy KENNY (ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*, Berlin, Language Science Press, p. 105-120, [http://doi.org/10.5281/zenodo.6759982], viewed on 27 August 2025.

Piotto Valentina, Sofo Giuseppe, 2023, "Collaborazione e strumenti digitali nel campo della traduzione", in Francesco Laurenti (Ed.), *La traduzione collaborativa. Tra didattica e mercato globale delle lingue*, Rome, Aracne, p. 223-241.

Piquet Alexandre, 2009, *Guide pratique du travail collaboratif : Théories, méthodes et outils au service de la collaboration*, Brest, Télécom Bretagne.

Smartcat, 2022, *What is machine translation and how does it work?*, 26 May, [https://www.smartcat.com/blog/what-is-machine-translation/], viewed on 27 August 2025.

Steigerwald Emma, Ramírez-Castañeda Valeria, Brandt Débora, Báldi András, SHAPIRO Julie Teresa, Bowker Lynne, Tarvin Rebecca, 2022, "Overcoming language barriers in academia: Machine translation tools and a vision for a multilingual future", *BioScience*, Vol. 72, No. 10, p. 988-998, [https://doi.org/10.1093/biosci/biac062], viewed on 27 August 2025.

Stryker Cole, 2023, "What are large language models (LLMs)?", *IBM*, 2 November, [https://www.ibm.com/think/topics/large-language-models], viewed on 27 August 2025.

Woyde Rick, 2001, "Introduction to the SAE J2450 Translation Quality Metric", *Language International*, p. 37-39.

Yulianto Ahmad, Supriatnaningsih Rina, 2021, "Google Translate vs. DeepL: A quantitative evaluation of close-language pair translation (French to English)", *AJELP: The Asian Journal of English Language and Pedagogy*, Vol. 9, No. 2), p. 109-127, [https://doi.org/10.37134/ajelp.vol9.2.9.2021], viewed on 27 August 2025.

Appendix

Group A: Translation with Smartcat

- 6 people involved (5 translators/reviewers, 1 editor)
- 75,336 total translated characters
- Sample of characters in the portion of text examined: 5,232

Group A divided into two Subgroups as follows:

- Subgroup A1: subject 1, subject 2
- Subgroup A2: subject 3, subject 4, subject 5

Group B: Translation with ChatGPT

- 6 people involved (5 translators/reviewers, 1 editor)
- 83,510 total translated characters
- Sample of characters in the portion of text examined: 5,232

Group B divided into two Subgroups as follows:

- Subgroup B1: subject 1, subject 2
- Subgroup B2: subject 3, subject 4, subject 5

Organisation and timing of the workflow with Smartcat

Please note that throughout the text-editing process, students kept in touch via a WhatsApp group and updated each other on their progress almost daily.

Step 1: Division of the parts to be translated based on the number of characters and translation.

In each subgroup (A1 and A2), the members divided up the assigned texts, trying to keep a fair number of characters per subject. Then, each subject translated their own part individually. Afterwards, each member of the subgroup revised the translations of the other members of their subgroup using Smartcat's comment function (A1 on A1, A2 on A2).

Step 2: Collaborative review using Smartcat's comment function.

Once the translations were ready, subgroup A1 revised subgroup A2's translations (A1 on A2), and subgroup A2 revised subgroup A1's translations (A2 on A1). Again, each subject worked individually, proofreading all the texts translated by the other sub-group (e.g. subjects of A1 proofread all the texts translated by subgroup A2).

Step 3: Revision of texts from Group A by components of Group B.

Two subgroups were created within Group B (B1, with two members, and B2, with three members), and they revised the texts both individually and collaboratively, providing comments and suggesting changes.

A member of the group created a table to better visualise the titles of the articles translated by Group A and the number of characters in each text. Students then decided how many characters each subject would review (approximately 21,000) and selected the texts accordingly.

Finally, Group A had to approve/refuse the changes proposed by Group B.

Step 4: Editing.

The editing of all the texts translated by Group A was performed by a single student of the University of Udine.

Time spent by participants to translate, review, and edit all 75,332 characters			
Step 1: Original text > first translation	**Step 2:** First translation > internal revision within Group A (A1 on A1, A2 on A2, A1 on A2, A2 on A1)	**Step 3:** Internal revision > revision outside Group A (A on B, A on REV(B))	**Step 4:** Editing of texts translated by Group A

Subject 1	360 min.	180 min.	90 min.	
Subject 2	240 min.	210 min.	145 min.	
Subject 3	180 min.	150 min.	290 min.	
Subject 4	210 min.	120 min.	145 min.	
Subject 5	240 min.	120 min.	145 min.	
Time total	1230 min.	780 min.	815 min.	660 min.

Time spent by participants proportioned to the number of characters in the portion of text examined		
	Total time	**Time for the sample**
Step 1	1230 min.	86 min.
Step 2	780 min.	54 min.
Step 3	815 min.	57 min.
Step 4	660 min.	46 min.
Total	3485 min. 58 h	243 min. 4 h

Organisation and timing of the workflow with ChatGPT

Please note that throughout the text-editing process, students kept in touch via a WhatsApp group and updated each other on their progress almost daily.

Step 1: Division of the parts to be translated based on the number of characters and creation of output with LLM.

The group was divided into two subgroups (B1, with two members, and B2, with three members). Each member of the group uploaded the original English texts to ChatGPT and extracted the raw output.

Step 2: Self-post-editing of the LLM's raw output.

Each member of the group post-edited their own texts using the Word review tool or the comments section. Participants then exchanged texts and revised them: B1 revised B2, and vice versa.

Step 3: Collaborative real-time review using Microsoft Teams.

First video call: 135 minutes

Second video call: 41 minutes

Step 4: Revision of texts from Group B by components of Group A.

Group A revised the translations from Group B. Again, the work was divided between subgroups A1 and A2.

The last step for Group B was to approve/refuse the changes proposed by Group A.

Step 5: Editing.

The editing of all the texts translated by Group B was performed by a single student of the University of Udine.

Time spent by participants to translate, review, and edit all 83,510 characters					
	Step 1: Original text > raw output	**Step 2:** Raw output > First individual self-revision of the translation	**Step 3:** Collaborative revision	**Step 4:** Collaborative revision > revision outside Group B (B on A, B on REV(A))	**Step 5:** Editing texts translated by Group B
Subject 1	6 min.	200 min.	176 min.	155 min.	
Subject 2	8 min.	275 min.	176 min.	120 min.	
Subject 3	10 min.	360 min.	176 min.	110 min.	
Subject 4	20 min.	297 min.	176 min.	260 min.	
Subject 5	5 min.	155 min.	176 min.	180 min.	
Time total	49 min.	1287 min.	880 min.	825 min.	540 min.

Time spent by participants proportioned to the number of characters in the portion of text examined		
	Total time	**Time for the sample**
Step 1	49 min.	3 min.
Step 2	1287 min.	81 min.
Step 3	880 min.	55 min.
Step 4	825 min.	52 min.
Step 5	540 min.	34 min.
Total	3581 min. 60 h	225 min. 3.75 h

1 International Center for Research on Collaborative Translation, IULM University, Milan. [https://www.iulm.it/wps/wcm/connect/iulm/minisiti-en/international-center-for-research-on-collaborative-translation], viewed on 20 August 2025.

2 [https://commission.europa.eu/education/skills-and-qualifications/develop-your-language-skills/european-masters-translation-emt_en], viewed on 20 August 2025.

3 [https://redattologia.uniud.it/], viewed on 20 August 2025.

4 [https://european-masters-translation-blog.ec.europa.eu/articles-emt-blog/my-distance-learning-2021-01-24_en?prefLang=cs], viewed on 8 September 2025.

5 [https://european-masters-translation-blog.ec.europa.eu/articles-emt-blog/innovating-accessibility-sign-language-university-genevas-fti-2021-11-24_en], viewed on 25 August 2025.

6 [https://european-masters-translation-blog.ec.europa.eu/articles-emt-blog/can-you-see-me-can-you-hear-me-new-teaching-and-learning-environments-and-new-normal-2020-10-15_en], viewed on 25 August 2025

7 [https://www.esp-xr.eu/events/conference], viewed on 25 August 2025.

English

In today's increasingly globalised world, translators cannot avoid collaborating with technological tools that offer a wealth of possibilities. The advent and proliferation of generative AI and Large Language Models have brought several advantages to translators, particularly in terms of saving time. However, it is important not to overestimate these systems' capabilities, bearing in mind that human intervention is still necessary to ensure the highest quality. This study investigates how two groups of students collaborated with each other and with different digital tools to understand the benefits that translators can derive from using traditional computer-assisted translation (CAT) tools, machine translation, and LLMs, and the most effective way to organise collaborative translation work. It also examines the most common challenges associated with using these tools, providing an initial guide for professionals and semi-professionals on how to approach translation tasks with them.

Français

Au cours des dernières décennies, les progrès technologiques considérables ont transformé notre façon de vivre, de communiquer et de travailler. En conséquence, tous les aspects de la traduction ont été profondément affectés, entraînant des changements significatifs. Par exemple, l'émergence de nouveaux rôles professionnels a été favorisée par l'essor des outils de traduction automatique ou assistée par ordinateur (TAO), la croissance des traductions collaboratives à grande échelle réalisées par des non-professionnels, ainsi que les innovations dans la théorie de la traduction découlant des avancées technologiques.

De plus, au fil des années, la demande de traduction s'est accrue, dans un monde où la communication devient de plus en plus multimodale et multilingue. Au moment même où Internet commençait à se généraliser, les chercheurs s'accordaient à dire que l'objectif principal de la traduction était de convertir un texte initialement rédigé dans une langue en son équivalent dans une autre, de manière à en préserver le sens, les caractéristiques formelles et les rôles fonctionnels (cf. Bell, 1991).

L'avènement et le développement de nouvelles technologies, telles que la traduction automatique neuronale (TAN), l'IA générative et les grands modèles de langage (LLM), peuvent faciliter le travail des linguistes. Ces technologies offrent un soutien significatif grâce à leur capacité à simuler le langage naturel — même lors de la transmission de messages complexes — et à adapter les textes à la culture cible. Après tout, dès 1969, Nida et Taber affirmaient que la traduction consiste à reproduire l'équivalence naturelle la plus proche de la langue source dans la langue cible.

Comme l'a déjà souligné Piquet (2009), les concepts et les outils liés au travail collaboratif ne sont pas une nouveauté. Cependant, au fil du temps, ils ont pris une dimension totalement nouvelle grâce à la démocratisation des technologies de l'information et de la communication dans la société et, par conséquent, au sein de nos organisations. Lorsqu'on parle de travail collaboratif, il est essentiel de prendre en compte la dimension technologique, ainsi que le fait que tous les outils de travail collaboratif disponibles évoluent rapidement pour répondre aux besoins du marché et des utilisateurs. Ils gagnent en maturité sur les plans technique, économique et social, devenant plus intuitifs, plus faciles à installer et plus avantageux du point de vue des logiciels comme des performances (Laurenti & Villareale, 2023).

De nombreux outils destinés aux activités collaboratives sont aujourd'hui disponibles en ligne, chacun présentant des caractéristiques techniques spécifiques. Un groupe de travail qui envisage de les utiliser doit prendre en compte différents facteurs afin de s'assurer que l'outil choisi est efficace et adapté à la réalisation de l'objectif, sans perturber le flux de travail.

Le choix de l'outil approprié pour une tâche spécifique est essentiel dans le processus collaboratif, car toutes les technologies ne conviennent pas à toutes les situations. Prendre une décision éclairée peut considérablement améliorer la qualité du travail. Cependant, le simple fait de choisir les bons outils ne suffit pas à garantir que la tâche sera correctement exécutée. Il est

essentiel que tous les participants comprennent bien ce que signifie travailler de manière collaborative et s'engagent pleinement dans cette approche, plutôt que de chercher à se démarquer. Dans le travail collaboratif, les subjectivités individuelles se fondent pour former une identité de groupe unifiée.

Quiconque travaille sur une tâche de traduction collaborative à l'aide d'outils collaboratifs, tels que les divers outils de traduction automatique disponibles sur le marché, doit garder à l'esprit qu'il s'agit de technologies imparfaites (O'Brien, 2022 : 105). Par conséquent, le risque de malentendus dans la traduction écrite reste élevé, d'autant plus que les personnes impliquées dans la communication ne peuvent pas utiliser le langage corporel pour aider le traducteur à comprendre exactement le sens de leur message. Une révision approfondie du texte traduit reste donc essentielle.

C'est pourquoi cette étude, menée par l'International Center for Research on Collaborative Translation de l'université IULM de Milan, analyse la manière dont deux groupes d'étudiants ont collaboré entre eux et avec différents outils numériques, afin de comprendre les avantages que les traducteurs peuvent tirer de l'utilisation des outils traditionnels de traduction assistée par ordinateur (TAO), de la traduction automatique et des LLM, ainsi que la manière la plus efficace d'organiser le travail de traduction collaborative.

L'objectif de cette étude était d'évaluer si l'utilisation d'outils de traduction automatique neuronale (TAN) ou de LLM présentait un avantage par rapport aux outils de TAO traditionnels, étant donné que ces outils ont tendance à produire des résultats moins précis qui nécessitent une révision plus approfondie. Par conséquent, il fallait évaluer la qualité de la traduction au regard du temps total nécessaire pour accomplir la tâche.

Compte tenu du grand nombre d'outils de TAN et de LLM disponibles en ligne, il était nécessaire de sélectionner celui le mieux adapté à cette tâche de traduction particulière. Les chercheurs ont donc évalué la traduction de l'anglais vers l'italien de l'article « My Distance Learning », disponible sur le blog de l'EMT, à l'aide d'une mesure quantitative manuelle appelée SAEJ2450.

À l'issue de l'évaluation, il a été décidé d'utiliser ChatGPT comme base pour la traduction, car cet outil ne présente aucune limite de saisie, permet de modifier les phrases à l'infini et fait preuve d'excellentes capacités de reformulation.

Une fois cette étape terminée, les chercheurs ont divisé les étudiants en deux groupes de six personnes chacun (cinq traducteurs/réviseurs, un éditeur) : un groupe traduisait à l'aide d'un outil de TAO collaboratif gratuit (Smartcat) et l'autre post-éditait le résultat brut généré par l'outil d'IA sélectionné ; les étudiants ont collaboré pour produire une traduction prête à être publiée sur le blog de l'EMT.

En ce qui concerne le calcul des temps, les résultats ont été réduits au prorata de l'échantillon de caractères examiné, les deux groupes ayant travaillé respectivement sur 75 336 et 83 510 caractères. La somme du temps passé par chaque sujet à chaque étape du travail pour traduire et

réviser le nombre total de caractères a ensuite été rapportée au nombre de caractères de la portion de texte examinée. Afin de recueillir des données significatives et de comparer équitablement les temps de travail, les chercheurs ont décidé d'analyser la traduction d'un échantillon de texte de 5 232 caractères pour chaque outil, de sorte que les deux textes aient la même longueur.

Comme prévu, la première traduction réalisée sur Smartcat a pris environ une heure et demie aux participants, car ils ont dû créer un nouveau texte dans une autre langue en partant de zéro. Bien que cette première traduction fût déjà d'une qualité moyenne à élevée, les étapes de révision ont nécessité 2,6 heures, portant le temps total consacré à 4 heures. La plupart des modifications ont été apportées lors de la phase de révision externe, effectuée par d'autres traducteurs, et lors du processus d'édition finale avant la publication. Ces deux activités ont pris au total 1,7 heure et ont principalement concerné le rendu en italien.

En revanche, en travaillant avec ChatGPT, bien que le temps de traduction soit très court – seulement 3 minutes –, le temps de révision a considérablement augmenté, atteignant 3,75 heures. La majeure partie du travail d'édition a été effectuée lors de la post-édition initiale du texte brut afin de corriger les calques, les erreurs grammaticales et les problèmes de compréhension.

Il convient de souligner qu'il s'agit d'une étude pilote portant uniquement sur deux flux de travail ; des recherches futures pourraient inclure des groupes plus importants de traducteurs et analyser des flux de travail utilisant différents outils numériques (par exemple, la TAN, les outils de TAO intégrant l'IA, etc.).

Comme démontré, bien que le temps nécessaire pour générer la première traduction à partir du texte original dans le flux de travail LLM soit considérablement réduit (3 contre 86 min.), le temps total passé par les traducteurs pour livrer la version finale n'a pas diminué de manière substantielle (3,75 contre 4 heures). Cela s'explique par le fait que l'IA ne saisit pas toujours le contexte et les subtilités linguistiques, produisant parfois des phrases difficiles à comprendre, voire inexactes – qui sont plus faciles à repérer –, mais aussi des phrases qui, bien qu'apparemment correctes sur le plan grammatical et syntaxique, ne reflètent pas fidèlement le sens du texte original (« erreurs transparentes »). Le traducteur doit donc relire très attentivement le résultat brut, en corrigeant les erreurs et en adaptant le langage au contexte spécifique, car seul un linguiste expérimenté peut apporter les modifications nécessaires. Il est évident que cette étape prend du temps.

D'autre part, l'utilisation d'outils de TAO permet aux traducteurs de produire immédiatement une version plus précise du texte dans la langue cible, qui nécessite moins de révisions. Cependant, cela les oblige à consacrer beaucoup plus de temps à cette étape, surtout si, comme dans le cas présent, ils ne disposent pas d'une mémoire de traduction.

Compte tenu de tout cela, il est impossible d'affirmer avec certitude qu'un flux de travail est plus avantageux pour les traducteurs qu'un autre. Ce qui est certain, cependant, c'est que ces deux types de travail sont différents et

requièrent des compétences différentes, car les post-éditeurs de textes générés par l'IA doivent avoir une compréhension approfondie du fonctionnement des machines et des erreurs qu'elles commettent le plus souvent.

Italiano

Nel mondo di oggi, i traduttori non possono prescindere dal collaborare con strumenti tecnologici: l'avvento e la diffusione dell'IA generativa e dei modelli linguistici di grandi dimensioni (LLM) hanno comportato diversi vantaggi, in particolare in termini di risparmio di tempo. Tuttavia, è importante non sopravvalutarne le capacità, tenendo presente che solo l'intervento umano può garantire la massima qualità della traduzione. Obiettivo del presente studio è comprendere i vantaggi che i traduttori possono trarre dall'uso dei tradizionali strumenti di traduzione assistita (CAT), dalla traduzione automatica e dagli LLM, nonché il modo più efficace per organizzare il lavoro di traduzione collaborativa. Pertanto, è stato indagato il modo in cui due gruppi di studenti hanno collaborato tra loro e con diversi strumenti digitali al fine di produrre una traduzione di alta qualità. Lo studio esamina inoltre le sfide più comuni associate all'uso degli strumenti digitali basati su IA.

Mots-clés

traduction collaborative, outils basés sur l'IA, collaboration à l'aide d'outils numériques

Keywords

collaborative translation, AI-based tools, collaborating with technologies

Parole chiave

traduzione collaborativa, strumenti basati su IA, collaborazione con strumenti digitali

---

Federica Villareale

IULM University of Milan

Dr. Federica Villareale is a Doctoral Researcher at IULM University of Milan and collaborates with the International Center for Research on Collaborative Translation. Her research concentrates on human-machine collaboration and interaction, with particular focus on digital and AI-based tools for translation. In 2024, she received the 1st International MA Thesis/PhD Dissertation on Collaborative Translation Award for her thesis on collaborative translation and revision. She has participated in numerous international conferences, such as the *PRIN International Conference*, the VII and VIII *International Congress on Science and Translation*, the II International Congress *Translation and Cultural Sustainability*, or the International Congress *Scenari Multimediali e Didattica della Traduzione*. She

has also authored scientific articles presenting research on emerging technologies for translation, including those involving NMTs and LLMs.
federica.villareale.l@gmail.com