



Incorporating Collaborative and Active Learning Strategies in the Design and Deployment of a Master Course on Computer-Assisted Scientific Translation

Marco Zappatore¹ 

Accepted: 19 July 2023 / Published online: 7 August 2023
© The Author(s) 2023

Abstract

This research aims to address the current gaps in computer-assisted translation (CAT) courses offered in bachelor's and master's programmes in scientific and technical translation (STT). A multi-framework course design methodology is proposed to support CAT teachers from the computer engineering field, improve student engagement, and promote computer-supported education, together with a balanced coverage of the most relevant topics in the CAT domain. STT is currently in high demand in many fields, requiring translators with sector-specific language skills and considerable computer literacy in order to manage translation projects with complex structures, and format heterogeneity. However, many STT curricula often lag behind current market demands, focusing primarily on language and translation theory, with less emphasis on CAT technologies and tools. Moreover, the lack of shared course design guidelines hinders the introduction of innovative teaching approaches based on collaborative learning. A novel multi-framework CAT course design methodology, named CATDeM, is proposed, based on the integration between an official European translation competence framework, real-life-mimicking laboratorial activities, and computer-supported collaborative learning, enriched with discussion case studies and role-playing experiences. A real-life case study is examined to illustrate and evaluate the implementation of CATDeM in two consecutive editions (2020/2021 and 2021/2022) of a one-semester compulsory CAT course in a M.A. degree in STT at the University of Salento (Italy). Students' perceptions of translation technology and role-plays, as well as their attitudes towards the proposed CAT course are evaluated through a post-grading self-assessment questionnaire. Achieved results indicated successful student engagement and self-assessed improvement in translation, technical, and interpersonal skills. The importance given by students to role-playing experiences mimicking professional scenarios was also highlighted, paving the way for CATDeM to be adopted in similar contexts.

Keywords Computer-supported collaborative learning · Computer-assisted translation · Higher education · Scientific translation · Roleplaying-based education

✉ Marco Zappatore
marcosalvatore.zappatore@unisalento.it

¹ Department of Engineering for Innovation, University of Salento, Via Monteroni sn, 73100 Lecce, Italy

Abbreviations

AI/DL	Artificial intelligence/deep learning
BPM	Business process modelling
CAT	Computer-assisted translation
CLT	Corpus linguistics tool
CSCL	Computer-supported collaborative learning
DMEMO	Define-model-execute-monitor-optimise
EMT	European master's in translation
GAI	Generative artificial intelligence
GAI-MT	Generative artificial intelligence for machine translation
GLAID	Group learning activities instructional design
ICT	Information and communication technology
LSP	Language service provider
MAHT	Machine-aided human translation
MT	Machine translation
MTE	Machine translation engine
MTPE	Machine translation post-editing
PTW	Professional translation workflow
QA	Quality assurance
QC	Quality control
SLMS	Software localisation management system
STB	Simulated translation bureau
STT	Scientific and technical translation
SW-L10N	Software localisation
TBMS	Translation business management system
TM	Translation memory
TMS	Translation management system
TQA	Translation quality assurance
TQMS	Translation quality management system
UI/UX	User interface/user experience

1 Introduction

Scientific and technical translation, or STT as defined by Olohan (2015), typically entails translating extensive volumes of data and documents, encompassing hundreds or even thousands of files in different formats, into multiple target languages (Garcia, 2015) in order to make them accessible to scientists, researchers, clinicians, students, technical specialists, and even ordinary users of electronic devices. Representative examples of technical documents include patents, user manuals, scientific articles and reports, engineering specifications, safety data sheets, medical guidelines, clinical procedures, product catalogues, specialist glossaries, as well as text strings from application software and content from Web pages. These documents typically contain technical jargon, many abbreviations, significant amounts of quantitative data along with units of measurements, visual plots and tables, pictures and graphics (Olohan, 2015, 2021).

In such a peculiar scenario, not only the correctness, quality, and consistency of the translation are crucial but also the role of technology, as it is simultaneously a driving factor and an essential enabler (Rothwell & Svoboda, 2019) that has shaped the scientific

translator's workflow (Bowker, 2015). As a direct consequence, over the last decade, STT teaching curricula and translator training programmes have included courses on technologies that support the translation workflow, with the aim of complementing translation competences with technological competences (PACTE group et al., 2018). Technologies are now perceived as fundamental components along with linguistic, translation, and intercultural skills for the education of students (Gaspari et al., 2015; Rico, 2017), thanks to the definition of specific, multi-skill competence frameworks (Toudic & Krause, 2017).

In addition to the fundamental and ever-present need for scientific translation skills, modern translators must acquire at least a profitable operational knowledge of computer-assisted translation (CAT) tools (Krüger, 2016). Essentially, these software applications support the translator's daily activities: the human translates digital documents, while the software manages a database of all the translations, preserves the graphical layout of the document, and automates repetitive tasks. These solutions are also known as Translation Management Systems (TMS)¹ (Shuttleworth, 2015).

Realistic problem-based educational contexts (which are more widely used in STEM pedagogy) are also gaining importance in translation education (Mellinger, 2018). This is facilitated by the adoption of new technologies that allow the implementation of the long-theorised approach of *situated translation* (Bowker & Marshman, 2010; Risku, 2002), according to which the translator's actions and problem-solving skills are intertwined with the overall situation in which these actions take place and these skills are required.

One might therefore assume that all the necessary conditions are in place for university translation courses to finally catch up with the surrounding technological landscape, but this is not yet the case. The heterogeneity of course programmes, content structures, and learning objectives is evidenced by the significant flow of scientific literature focusing on the teaching of specific topics, ranging from traditional CAT systems (Shuttleworth, 2017) and terminology management (Montero Martínez & Benítez, 2009) to machine translation post-editing (Arenas & Moorkens, 2019) and even to programming resources for machine translation (Krüger, 2021). Similarly, educational strategies and the assessment of pedagogical effectiveness vary considerably, from context-based approaches (Killman, 2018) to autonomous learning (Shuttleworth, 2017) and virtual learning (Wang & Wang, 2021). Moreover, CAT courses do not take advantage sufficiently of the fruitful collaboration between the computer science sector and the applied translation fields, with the former perceiving the latter as secondary and the latter perceiving the former as too computer-dependent. Furthermore, several studies have pointed out that the gap between the current state of STT teaching and the market demand for skilled translators requires appropriate pedagogical interventions (Rothwell & Svoboda, 2019; Zhang & Vieira, 2021).

Based on such premises, this paper proposes a broader and more systematic perspective on how to design Master CAT courses in which translation technology and STT are intertwined, without focusing only on a specific topic or sub-topic. In addition, special attention is given to how to provide students with more engaging opportunities to acquire the technical and professional skills required by the market and to bring them closer to the real working environment of STT.

¹ Since the alternate use of CAT and TMS acronyms is very often a source of ambiguity, for the rest of this paper CAT will identify the domain of computer-assisted translation (e.g., CAT course, CAT teaching, etc.), while TMS will indicate the platforms supporting CAT activities. Less recently, TMS platforms were also known as Machine-Aided or Machine-Assisted Human Translation (MAHT) tools (Feder, 2003; Isabelle, 1993).

More specifically, this work presents a multi-framework methodology, called CATDeM, to design CAT courses, enriched with discussion case studies and role-playing elements, with the aim of easing the workload of educators and instructors as well as improving student engagement and learning outcomes. CATDeM is aimed specifically at instructors in the field of computer-engineering and has three objectives: (1) *modelling*: to leverage frameworks this typology of instructors is already familiar with, in order to speed up the course design process; (2) *technology*: to guide the instructor rigorously through the technical peculiarities of STT and CAT fields; (3) *education*: to provide the instructor with a holistic perspective on STT and CAT teaching. To the author's knowledge, the proposed methodology is the first of its kind at the time of writing this paper, thanks to the composition of several frameworks.

Another important aim is to propose a balance between the different topics and sub-topics related to CAT in a M.A. degree in STT. In some cases, the translation technology component is reduced to a single course, while in others it is possible to benefit from several courses dedicated to vertical topics, where the introductory CAT course is then followed by courses on MT, terminology management, software localisation, and so on (Malenova, 2019; Motiejūnienė & Kasperavičienė, 2019; Zhang & Vieira, 2021). This diversity depends on the specific academic context and, therefore, the modular and self-consistent design methodology proposed in this paper is intended to be adaptable to the specific needs of the teacher designing a single CAT course, in order to take into account all possible scenarios.

Before introducing CATDeM, an analysis of the tools that populate the translation technology landscape will be proposed according to functional categories that represent relevant topics for any CAT course. It will be explained how a pivotal component of CATDeM is borrowed from business process modelling (BPM Resource Center, 2014), to take advantage from a rigorous ICT perspective so that the CAT course can not only be modelled but also operated, monitored, and optimised to make it replicable in different academic contexts and compatible with professional STT workflows. CATDeM is further enriched with computer-supported collaborative learning and active learning elements (De Hei et al., 2016), such as discussion case studies and role-playing experiences, whose mapping to the course structure and learning goals will be thoroughly discussed, together with the proposed teaching approach and learning platforms.

A real-life case study will then be examined to illustrate the implementation of CATDeM in two consecutive editions of a one-semester compulsory course on computer-assisted scientific translation in a M.A. degree in STT and Interpreting² at the University of Salento (Italy), covering the academic years 2020/2021 and 2021/2022. Students' perceptions of translation technology and their attitudes towards the proposed CAT course will be evaluated through a post-grading self-assessment questionnaire.

The paper is structured as follows: extended motivations and research questions are detailed in the remaining part of Sect. 1. The theoretical background and the related works are presented in Sect. 2, while the ecosystem of translation technologies is sketched in Sect. 3. Course design strategies and their implementation in a CAT course acting as a pilot case study are thoroughly described in Sect. 4. The achieved results are discussed in Sect. 5. Section 6 provides the answers to the research questions. Conclusions are drawn

² For the sake of brevity, the case study of the “*computer-assisted scientific translation course of the M.A. degree in scientific and technical translation and interpreting*” will be referred to as “*CAT course*” in the rest of the paper.

in Sect. 7, together with limitations and further developments of the study. Finally, a list of acronyms is reported to support the reader, along with the list of software tools mentioned throughout the paper.

1.1 Motivations

The language services market, mainly driven by the business globalisation, the need for customised digital content, and the ever-expanding medical sector, grew by 4.77% year-on-year in 2022 despite global economic uncertainties (Slator, 2023) and is expected to grow at a compound annual growth rate (CAGR) of 5.3% over the period 2023–2028 (Mordor Intelligence, 2023). In terms of market size, translation and localisation together represent the largest segment of language services (i.e., 43%), followed by interpreting (15.1%), dubbing (10.7%), and machine translation (5.9%) (Hickey, 2023a). The translation and localisation segment, in turn, is driven by several industry verticals, in varying proportions, from government (18.1%) to healthcare (10.2%), from life sciences (9.5%) to retail and e-commerce (2%) (Hickey, 2023b). Some of these industry verticals are more closely related to scientific and technical translation than others, as certain sectors require accurate translation of specialised and technical content to ensure effective communication, compliance with standards, and global dissemination of knowledge within their respective fields. More specifically, technology, IT, software, automotive, aviation, manufacturing, healthcare, life sciences, financial and legal, when aggregated, account for 47.1% of the verticals listed in (Hickey, 2023b), thus representing one of the largest shares of the translation services market. In addition, several verticals (e.g., utilities, healthcare and pharmaceuticals, electronics and semiconductors, IT services and gaming, etc.) have shown steadily increasing demand in the recent years (DePalma, 2021).

These insights clarify how required a scientific and technical translator can be nowadays. This role, as anticipated in the introductory section, requires professionals skilled not only in the CAT field but also on Machine Translation (MT) engines (Alcina, 2008), which represent a powerful asset (if used properly by the translator) and typically require human post-editing intervention before automatically translated texts can be delivered to the client. Similarly, translation project managers have to deal with complex software systems that allow them to manage the entire service provisioning pipeline effectively, from the initial negotiation with the client up to the delivery of the final product (Shuttleworth, 2015). From a deployment-oriented perspective, a significant subset of these tools has shifted over the last decade to cloud-based software-as-a-service (SaaS) solutions that allow: (1) translators to benefit from collaborative functionalities and remote access to resources, and (2) language service providers (LSPs) to avoid the need for on-premises equipment, thus increasing accessibility, productivity, and market opportunities. Moreover, TMS and MT engines (or MTEs, for brevity) are only the best known elements in a broader collection of technological enablers that translators, as well as trainers and educators, has to confront with: at least ten different categories, encompassing more than 800 IT tools, can be identified, according to a recent market report by Nimdzi (Akhulkova et al., 2022).

Designing and implementing a CAT course is, therefore, a multifaceted challenge, made more difficult by the lack of widely available curriculum design guidelines. It is necessary to identify carefully not only the enabling information and communication technology (ICT) tools but also the most appropriate topics and the most effective teaching strategies. Only in this way will students be able to meet the ever-increasing market demand for professionally qualified scientific translators, a trend that was already consolidating in the early 2000s

(Bacon, 2002), but has now become even more strategic, as the flow of technical publications has doubled in the last decade (Olohan, 2015). Companies wishing to offer localised products on a global scale or requiring the secure translation of patents in foreign markets, businesses demanding complex translations compliant with safety instructions detailed in technical standards, IT vendors requiring adequate global localisation of the graphical user interface of their software products, all need professional translators capable of transposing technical-scientific knowledge from one language into another (Olohan, 2021; Scheibengraf, 2023).

Another challenge for STT students is to cope adequately with technology advances by learning how to use them effectively (Doherty, 2016) and without being locked to a specific commercial software solution or approach that could lead to high licensing costs, excessive dependency on proprietary formats as well as on vendor-provided updates and customer support. While in more ICT-driven areas such as engineering education, students achieve progressively the necessary skill set to adapt to different and rapidly evolving software environments or to use alternative free and open source tools (Álvarez Ariza & Nomesqui Galvis, 2023; Pinto et al., 2019; Srinivasa et al., 2021), the same effectiveness could be more difficult to reach for STT students. To this end, computer-supported collaborative learning (CSCL) (Lipponen et al., 2004) and active learning activities, largely exploited in scientific curricula at higher education level (Hmelo-Silver & Jeong, 2021; Ma et al., 2020) but still overlooked in translation teaching, can represent an effective educational asset to exploit.

1.2 Research Questions

As anticipated in the introductory section, this paper proposes a CAT course design methodology called CATDeM to systematise the STT educational context, where ICT and enabling software platforms are too often relegated to a secondary role or focused solely on vertical sub-topics (e.g., machine translation post-editing or terminology management). To achieve this goal, the following research questions (RQs) will be answered:

- (RQ1) What is the technological landscape from which the educators should draw, depending on pedagogical requirements and market demands?
- (RQ2) How to support computer scientists and researchers involved as instructors in CAT courses (in B.A./M.A. in translation studies)?
- (RQ3) How to bring computer-supported active learning and computer-supported collaborative learning strategies from the computer science area (where they are usually adopted) to CAT courses?
- (RQ4) Are these activities effective and how they are perceived by students?

To this end, the rationale of a CAT course designed and implemented at the University of Salento (Italy) and incorporating collaborative and active learning activities will be proposed, together with a detailed data analysis of its effectiveness.

2 Theoretical Background and Related Works

2.1 Computer-Assisted Translation and Machine Translation Teaching

CAT courses are nowadays available in the majority of translation teaching curricula, either at bachelor or master level (Rothwell & Svoboda, 2019). A common format relies on

laboratory-oriented courses, which are mainly dedicated to making students operationally competent in using a specific software tool (Malenova, 2019), with few courses focusing on machine translation and even fewer on software/videogame localisation. Collaborative learning practices are even rarer and a full technology-driven training for modern translators is far from being achieved systematically on a large scale.

The inherent interdisciplinary nature of the CAT approach would require tight cooperation and interaction between educators from the computer science sector and educators from the applied-translation area, so to provide students with a real competitive advantage, but CAT courses do not benefit from such a combined approach and these two categories of educators often operate separately.

Consequently, even though the wide availability of ICT tools has made their application in translation teaching inevitable (Ivanova, 2016) and even though TMSs have long been considered essential in translator education (Bowker, 2002), the real scenario has been mostly static for about two decades in terms of novel educational strategies for technology-assisted language teaching. Moreover, two typical obstacles can be identified.

The first hindrance is related to students' digital competence (also referred to as digital literacy), intended as the ability to apply five basic digital skills (i.e., information and data literacy, online communication and collaboration, digital content creation, safety, problem-solving) in a confident, critical, and responsible way to a defined context (Brolpito, 2018). These basic skills vary significantly depending on geographical and social contexts, thus creating potentially relevant digital divides from country to country. According to Eurostat, digital skills of young Europeans (i.e., those aged 16–29) exhibit considerable differences, ranging from 93% of Finnish youngsters having at least one basic digital skill to 49% and 46% in Bulgaria and Romania, respectively (Eurostat, 2022). Moreover, digital skills of B.A./M.A. students are generally lower than those of computer science students and diversified among different humanities faculties (Vodă et al., 2022; Yoleri & Nur Anadolu, 2022), thus limiting significantly the possibility to cover additional topics as complex as software and video game localisation or technical foundations of MT. In order to incorporate those topics effectively in translation education curricula, more advanced digital skills are required, such as prior knowledge of programming languages, data structures, UI/UX (User Interface/User eXperience) design principles as detailed in Microsoft's localisation guides (Microsoft, 2022), and AI/DL (Artificial Intelligence/Deep Learning) fundamentals (Maxwell Chandler & O'Malley Deming, 2011; O'Hagan, 2020).

The second obstacle is related to CAT educators, who usually come from computer science/engineering backgrounds. As a direct consequence, it is highly plausible that they master scientific and technical genres in written/spoken English as both text producers and text consumers, since they are specialists who need to communicate frequently with other specialists in the same field. However, it is definitely less likely, that they know the translation strategies needed to deal with other languages and cultural contexts (Olohan, 2021), how to interact with LSPs, or the professional translation pipeline regulated by international standards (ISO, 2015). Therefore, the main challenge in this case is to avoid that students are trained more from a pure software tool-oriented operational perspective and less from an applied-translation perspective.

Limitations can also be observed with regard to MT topics in STT teaching (Wittkowsky, 2014). MT is usually referred to in terms of pre- and post-editing activities (Toledo Báez, 2018), with some introductory concepts supporting the basic distinction between statistical, neural, and hybrid MT (Zhang & Vieira, 2021), but without delving into the latest approaches, such as graph convolutional networks (GCN) (Suman Banerjee & Khapra, 2019; Zhang et al., 2023) or federated learning (FL) (Lu et al., 2022).

GCN allows achieving sentence-structure-aware NMT in densely connected deep architectures without suffering from performance degradation as the architecture scales, and is more suitable for hardware acceleration, thus promising to outperform traditional NMT approaches (Guo et al., 2019).

FL has proven to be beneficial for multilingual MT, as it allows machine-learning models to be trained across multiple decentralised devices (or edge computing nodes) without sharing raw data, thus preserving data privacy and security while enabling collaborative model training. As multilingual MT is often applied to large volumes of sensitive data that require complex training of mixed-domain translation models, FL can be used to overcome the need for costly data preparation techniques (Passban et al., 2022) and communication overhead between neural nodes (Roosta et al., 2021), achieving better performance than traditional MT approaches based on centralised learning (Weller et al., 2022). Although the working principles of FL are beyond the scope of a CAT course, STT students would benefit from discussing FL-related topics, such as centralised vs decentralised learning, data privacy, data distribution, and resource constraints.

Another crucial challenge in MT is how MT models need to be trained differently for low-resource languages (i.e., languages for which limited corpora are available) and for high-resource languages. STT students should be taught about the technological approaches to cope with low-resource languages, such as model compression and knowledge distillation (Gou et al., 2023; F. Wang et al., 2021), which can increase MT accuracy and reduce the operational costs of MT service providers in the post-deployment phase (Jooste et al., 2022).

Finally, the emergence of Generative AI (GAI) technologies as alternative enablers to multilingual translation (GAI-MT) should be also considered as a no longer avoidable topic to be discussed in STT courses. GAI-MT promises accuracy, speed, and scalability (Hariri, 2023; Nguyen et al., 2022) but students should be adequately trained on how to interact properly with GAI-enabled systems through appropriate prompt engineering techniques (Roumeliotis & Tselikas, 2023) in order to improve the quality of the translation output.

2.2 Competence Frameworks: An Underestimated Opportunity

Competence models started to appear in the late 2000s as a schematisation to combine students' competences together. In (Göpferich & Jääskeläinen, 2009), the translation competence is associated with skills on communication, domain, strategy, psychomotricity, and translation-specific tools. Similarly, the European Master's in Translation (EMT) (Toudic & Krause, 2017) is a framework for selecting European universities that offer translation courses satisfying a set of evaluation criteria in five competence areas: language and culture, translation, technology, personal and interpersonal, and service provision. If we focus specifically on the EMT technology competence, it requires to train students in:

1. Using relevant IT applications (e.g., office software) and IT resources (e.g., cloud-based applications).
2. Using effectively search engines, corpus-based tools, text analysis tools and TMSs.
3. Pre-processing, processing and managing files, media/multimedia sources.
4. Understanding the basics of MT and its impact on the translation process.
5. Assessing the relevance of MT in a translation workflow.
6. Using the appropriate MT system where relevant.

7. Applying adequate ICT tools to support language and translation (e.g., workflow management software).

European universities can apply for a renewable 5-year *EMT quality label*, which ascertains their effectiveness in incorporating all five competence areas in their curricula. Unfortunately, the evaluation process is extremely selective and only 3% of EU universities have been successful so far, thus making the framework limitedly effective in shaping the overall European translation teaching sector on the one hand, and splitting universities into EMT and non-EMT ones, on the other hand. Moreover, only a minority of international LSPs are currently aware of the EMT (ELIS Research, 2022). Since the potentially beneficial effects of EMT on CAT course design strategies can be directly observed in a very limited number of academic institutions, it does not yet represent a large-scale game-changer, but rather an interesting methodological reference on which to base new STT/CAT teaching approaches.

That is why, although EMT concept seems to many as not much more than a missed opportunity and the path to its achievement a relevant obstacle to overcome, it can represent the common grounding element on which to build rigorous and technology-driven translation teaching curricula, even without necessarily trying to apply for it. The validity of the EMT framework is also demonstrated by an interesting initiative promoted by nearly a dozen EMT-universities that introduced in their CAT courses some laboratorial activities that replicate the processes of real translation agencies, under the definition of Simulated Translation Bureaus (STBs) (Buyschaert et al., 2017). The STB concept may pave the way for the simulation of a cooperative learning context in translation teaching but, again, several challenges can be identified. Firstly, the number of institutions involved is very small and their geographical coverage is limited: no more than one or two institutions per European country are involved, with many countries not represented at all, and there are no similar initiatives in non-EU countries. Secondly, the STB design strategies adopted vary so much from one university to another that it is practically impossible for a non-participating CAT instructor to design and launch a STB as well as to integrate it into a standard STT curriculum without a considerable preparatory work and without appropriate guidelines to support the adaptation of the STB to the specificities of the STT curriculum.

Moreover, in-depth knowledge of LSP workflows is needed. It is also noteworthy that computer scientists involved in CAT courses are usually not aware of competence frameworks or LSP/STB workflows. Similarly, a large proportion of translation teachers still perceive technology competences as being primarily associated with purely supportive computer tools needed to speed up translation.

In order to fill this gap, Sect. 4 will also propose a set of design principles for embedding a STB (as a role-playing experience) in STT curricula.

2.3 Computer-Supported Collaborative Learning (CSCL) and Active Learning

The CSCL approach has long been accepted as a useful educational strategy in universities, especially in STEM disciplines (Flores et al., 2015), as it supports group-learning activities with the aim of improving students' individual knowledge as well as group and social skills (Ludvigsen & Mørch, 2010). Since the early 2000s, a considerable stream of scientific literature has been devoted to CSCL (Barron, 2000; Lipponen et al., 2004), thanks to technological advances favouring the collaboration among students who form a team (*social constellation*), physically or virtually, with the aim of achieving a specific set of learning objectives (Strijbos, 2011).

Effective CSCL requires a clear definition of individual student responsibility and a clear determination of whether a positive interdependence exists between individual outcomes and group performance (Wang, 2009). Both of these elements are supported by (at least) a teacher who plays a coordinating role, through scripting (Dillenbourg, 2002) and scaffolding (Rienties et al., 2012; Shin et al., 2020), which could progressively fade out in order to let students acquire greater autonomy. Teacher-led coordination is pivotal, especially at the beginning, as students must receive indications on how to carry out group tasks, activate problem-based learning dynamics, and exploit available software platforms properly (Hernández-Sellés et al., 2020).

When it comes to the design of CSCL experiences, several approaches are available, depending on whether educators focus more on knowledge, communication, collaboration, motivation, or pragmatism as learning objectives (Meijer et al., 2020). Very often, however, these elements are addressed separately, without fruitfully combining the different components of the learning environment. The Group Learning Activities Instructional Design (GLAID) framework was proposed in 2016 (Strijbos, 2016) in order to support educators in designing CSCL experiences with eight design components (C_i) and five stages (S_i):

- **S₁. Analysis:** *profiling* of students, teachers, and curriculum.
- **S₂. Design:**
 - C_1 : *interaction* patterns among students to achieve pre-defined targets,
 - C_2 : individual and group *learning objectives* and outcomes,
 - C_3 : formative and/or summative *assessment* strategies focusing on both individual students and groups.
- **S_{3A}. Development (didactics):**
 - C_4 : description of student *task features* (e.g., type, duration, frequency, etc.),
 - C_5 : *structuring* and *timing* of the CSCL experience,
 - C_6 : student *guidance* (i.e., scaffolding) policies and dynamics, along with the expected role of the educators.
- **S_{3B}. Development (logistics):**
 - C_7 : organisation of student *group constellation* (e.g., composition, type, number of participants, level of heterogeneity, etc.),
 - C_8 : *facilities* enabling teaching and learning in the CSCL experience,
- **S₄. Implementation:** performing and monitoring the instructional process.
- **S₅. Evaluation:** analysis of learning processes and learning outcomes.

Group-learning and active-learning, some of the core elements in CSCL, need careful planning in order to avoid disappointing results when transferred to CAT courses in STT teaching. Group-learning requires an orchestration between supporting technologies and pedagogy (Pierre Dillenbourg, 2013), but students must have acquaintance with them and teachers from the computer area should be trained in the best pedagogy strategies when translation studies are involved. Similarly, active learning is much more exploited in computer science courses to promote student-to-student and student-to-teacher interactions to

improve learning outcomes (Chilukuri, 2020; Sobral, 2021), while it does not represent a primary typology of learning activities in CAT courses, except for the STBs implemented in a few European universities (Loock et al., 2017), whose diversity, however, shows that no common design methodologies have been adopted so far. Therefore, the GLAID framework will be adopted to design one of the CSCL-based activities in the proposed CAT course, as explained in Sect. 4.

Table 1 summarises the various related studies, frameworks and the educational strategies and approaches discussed in Sect. 2, together with their positive aspects, limitations, and their prospected use in CATDeM.

3 Technological Landscape

According to Nimdzi's 2022 annual technology report (Akhulkova et al., 2022), there are more than 800 software applications supporting scientific translators. These tools are grouped into nine categories, covering computer-assisted translation, machine translation, and audio-visual translation. Such a wide range of alternatives is somewhat intimidating for students and even for teachers, who very often decide, as a consequence, to focus primarily on the tool that is currently favoured by the market, since it is plausible to assume that when students start their professional careers they will be required to use that specific tool by most translation agencies and bureaus. Consequently, many CAT courses today cover Trados Studio,³ the platform that holds 80% of the global translation market share (Verified Market Research, 2022). Moreover, it is worth to note that the tools for corpus linguistics are not listed in Nimdzi's report, while they should be included instead because of their role in training MT engines. This category alone consists of more than 200 additional applications, according to various lists and repositories directly maintained by communities of computational linguists (Corpus-Analysis.com, 2022).

The following subsections will now deal with the most relevant categories of tools for the CAT domain.

3.1 Translation Management Systems (TMS) and Translation Business Management Systems (TMBS)

These tools, which are fundamental to any modern translator (Shuttleworth, 2015), support the translation of language assets in multiple digital formats according to a rigorous processing pipeline (RWS Group, 2021). Once the documents to be translated (i.e., *source documents*) are provided as input, an automated workflow guides the translator through project configuration settings to a translation editor where source documents can be actually translated. The translator focuses solely on the textual content. The TMS controls all non-translation tasks and provides the final documents (i.e., *target documents*) as output, applying the graphical layout and structure of the original documents to the translated text.

The TMS also allows managing and reusing linguistic resources (RWS Group, 2021) via dictionaries, glossaries and, most importantly, translation memories (TMs), which are bilingual databases of aligned sentences (i.e., every source sentence is stored along with its

³ All software applications (and their URLs) and mentioned in this paper are listed in a dedicated list placed before the bibliography.

Table 1 Summary of the related studies discussed in Sect. 2

Related work	Short description	Positive aspects	Limitations	Proposed use in CATDeM
European Master's in Translation (EMT) (Toudic & Krause, 2017)	A reference framework for translation education programmes. It defines student competences in five areas (i.e., language and culture, translation, technology, personal and interpersonal skills, service provision). To be EMT accredited, universities in the EU must include all the competences in their programmes	Competences are equally important and interdependent Competences are further broken down into sub-competencies Technological and service-oriented skills are not overlooked Competences and sub-competencies represent a methodological reference in CAT education	Evaluation process is highly selective Only 3% of EU universities are qualified as EMT accredited Very limited large-scale impact on CAT course design strategies Little known to the majority of international LSPs Not adopted outside Europe	To be used as the common grounding element on which to design rigorous and technology-driven CAT courses in B.A./M.A. degrees All EMT competences must be at least partially covered in the designed CAT course programme
Simulated Translation Bureau (STB) (Buysschaert et al., 2017)	A set of laboratorial activities replicating the operational workflow of real translation agencies/bureaus, typically integrated into CAT courses	Participating students are provided with more engaging hands-on experiences The STB offers a cooperative learning context for students The STB particularly matches with personal and interpersonal skills and service provision skills from EMT	Only a dozen universities in the EU offer STB experiences No similar activities in non-EU countries STB design strategies are highly variegated, depending on the CAT courses they belong to Computer scientists involved as CAT teacher may not know LSP workflows in details	To be used as a role-playing educational experience offered during the instructional part of the designed CAT course
Computer-Supported Collaborative Learning (CSCL) (Flores et al., 2015)	An educational strategy, especially indicated in STEM teaching, which leverages IT solutions for enhancing group-learning activities among students	Fosters group and social skills Enables flexible learning, social interaction, and critical thinking Improves the importance of distributed expertise among participants	Difficult to be applied effectively without a proper design methodology for the CSCL experience Individual accountability difficult to be estimated Need for instructor guidance (e.g., through scaffolding)	To be used during the designed CAT course (and especially during the role-playing experience) to enhance learning outcomes and favour collaborative knowledge construction among students

Table 1 (continued)

Related work	Short description	Positive aspects	Limitations	Proposed use in CATDeM
Group Learning Activities Instructional Design (GLAID) (Strijbos, 2016)	A framework to support educators in designing CSCL experiences more rigorously	Based on 8 components and 5 stages that provide the educator with a clear course design reference (all those referred to CSCL)	Components and course design stages can be perceived as complex The adequate orchestration of supporting technology and pedagogy is not simple to achieve	To be used to model rigorously the CSCL component included in CATDeM

translation in a given language, the first time it is delivered by the human translator). Each time a new sentence needs to be translated, the TMS searches the TM: if a match is found, the corresponding translations can be retrieved and used/modified, thus saving a considerable time and effort.

The creation and management of translation projects is also an important feature of TMSs. Moreover, translation companies can rely on special TMSs specifically tailored to business aspects, known as Translation Business Management Systems (TBMSs), which offer quoting, invoicing, client history, personnel management and payment, and which can be integrated with TMSs. As can be easily guessed, translation students and STT students in particular need to be familiar with TMSs (Alkan, 2016; Tarasenko et al., 2019), while TBMSs are not covered usually.

3.2 Software Localisation Management Systems (SLMS)

Software localisation⁴ is the process of adapting a software product or video game for a specific market, so that all its components (e.g., UIs, graphics, navigational elements, dialogues, measurement units, character sets and fonts, local regulations, copyright claims, etc.) are not only translated but also perceived by end users as specifically designed for their linguistic and cultural context. Localisation is much more than just the translation of text documents: software localisers face many more challenges and pitfalls than translators. First, the timing is different: usually, digital documents are translated once the source text has been finalised, whereas software products are very often localised during the development process (or in the early stages of testing) so that they can be shipped to multiple markets simultaneously. Second, multiple adaptations may be required to the UI (e.g., form resizing, dialogue adaptation, etc.) as well as to any graphics containing visible and readable text. Third, the translated elements must fit back into the original software product, so it is necessary to check that all functionalities work properly. Software localisers are typically involved in close collaboration with software developers and have to adapt to different working schemes (e.g., Agile, DevOps) (Lenker et al., 2000).

SLMSs allow translators to integrate the translation workflow seamlessly into the software development pipeline, eliminating unnecessary code snippets that would only slow down the translation process and would pose significant risks of accidental code misplacement. Consequently, these platforms can be used only after appropriate training, which requires explaining in advance: (1) the file formats that a software localiser will encounter in daily activities (e.g., JSON, XML, YAML files, etc.), (2) the software development pipeline, (3) the structure of a software product and the placement of translatable assets. For these reasons, SLMSs are rarely addressed in CAT courses (Zhang & Vieira, 2021) and related hands-on activities are very difficult to implement, even though the market demand for software localisers is growing rapidly, especially in the mobile app sector (Omar et al., 2020; Schroeder, 2018).

⁴ Software localisation is frequently known in the field as SW-L10N, where 10 is the number of letters between “L” and “N” (W3C, 2007).

3.3 Translation Quality Management Systems (TQMS)

Their pivotal role is to support translators in identifying, tracking, and resolving issues hampering the quality of translated documents (Vela-Valido, 2021), according to specific standards such as ISO 17100:2015 (ISO, 2015) on regulatory requirements and client guidelines. TQMSs apply different types of evaluation criteria, such as spelling and grammar correctness, terminology consistency, text meaning and style, readability level, adaptation of culture-specific aspects, adequacy of formats (e.g., dates, units of measurements, and currencies). They can focus on *quality control* (QC, basic quality controls to check whether requirements are met) or *quality assurance* (QA, planned procedures to manage quality controls). TQMSs are covered rarely in CAT courses (Zhang & Vieira, 2021), although they represent a must for any professional LSP (especially those working on the international market). As an educational replacement, special TMSs with TQA functionalities, such as MemsourceQA (Memsource, 2021), can be used.

3.4 Machine Translation Engines (MTE)

These are software solutions capable of automatically translating a source text into another (set of) language(s) without any human intervention (Garg & Agarwal, 2018; Xu & Li, 2021). Statistical methods of MT (SMT) have been studied for years but today the deep learning MT (namely, neural MT or NMT) offers better performance when properly trained, especially when translating languages with different word order. NMT relies on encoder-decoder architectures, where each source sentence is encoded into a numerical representation and then is used by the decoder to output the target sentence.

In language courses, students leverage MTEs with profitable learning outcomes (Raad, 2020), while students in translation courses should focus on other aspects, such as: (1) evaluating or post-editing (i.e., correcting) the MTE output; (2) pre-editing (i.e., simplifying) source texts to help the MTE achieve better translation quality; (3) understanding how MT algorithms actually work; (4) appropriately using MT also in CAT contexts (Toledo Báez, 2018); (5) exploring MT functionalities when embedded in automatic subtitling and dubbing applications (Koponen et al., 2020). However, courses/modules dedicated to MT usually focus on manual post-editing techniques, overlooking a whole branch of translation quality assurance (TQA) and text annotation tools that would help students significantly.

3.5 Corpus Linguistics Tools (CLT)

Considerable quantities of texts (known as *text corpora*) are needed to train both NMT and SMT machine translation algorithms, especially for scientific texts (Tehseen et al., 2018). Usually, the larger the number of words/sentences in the corpus, the higher the resulting quality of the MT algorithm (Rikters, 2018). In order to build these corpora quickly, texts are automatically extracted from multiple sources (e.g., Web repositories, digitised documents, etc.). The problem with automatic text extraction is that gathered data are noisy, because of incomplete or wrong sentences that are very often present alongside the correct ones. Since the quality of the training dataset directly affects the resulting quality of the MT algorithm (especially when NMT is involved) (Rikters, 2018), it is crucial to select

Table 2 Software tools for translation students and professional translators (Akhulkova et al., 2022)

Category ¹	Applications
TMS	Trados Studio, Memsource, memoQ, SmartCat, OmegaT, ...
TBMS	Plunet, XTRF, TranslationProjex, ...
SLMS	Passolo, Phrase, Lokalise, Poedit, Pootle, ...
TQMS	LexiQA, Okapi CheckMate, MemsourceTQA, TQAuditor, ...
MTE	AWS Translate, Google Translate, DeepL, Bing MS Translator, ...
CLT	AntConc, CATMA, TextSTAT, Sketch Engine, CorpKit, ...

¹*TMS*: translation management system, *TBMS*: translation business management system, *SLMS*: software localisation management System, *TQMS*: translation quality management system, *MTE*: machine translation engine, *CLT*: corpus linguistics tool

carefully the sources, adopt correct corpus-building procedures, pre-process the resulting dataset whenever possible, and assess its overall quality.

Even if translators are not involved directly in MTE configuration and training process, they can exploit CLTs, which essentially are speech and language processing software applications (Anthony, 2013), to create *parallel text corpora* usable as a source of translation suggestions for both MT engines and human translators. A parallel text corpus is a collection of monolingual source texts, stored together with their translation into one or more languages (*parallel*): each text is chunked into sentence-like segments linked in a 1:1 relationship (*alignment*) to the segments corresponding to their translation in another language stored in the corpus. In addition, CLTs allow the user to analyse collections of texts and to identify typical/unusual language use patterns, concordances, co-occurrences.

Very rarely dealt with in CAT courses (Zhang & Vieira, 2021), CLTs represent an important addition to the so-called *translator's workbench* (Delpech, 2014), as they can be used in several other language and translation-related activities. CLTs should also be considered when designing the educational offerings in this sector as they provide a convenient way to perform *language data management* (LDM) (Li & Rafiei, 2018), which is becoming increasingly important not only for MT (Eetemadi et al., 2015; Rikters, 2018), but also for automatic text labelling and annotation (Gries & Berez, 2017), text summarisation (Tawmo et al., 2022), sentiment analysis (Stuart et al., 2016), and to create natural language interfaces to data repositories (Quamar et al., 2022). Unfortunately, LDM is largely underestimated in translation studies (X. Zhang & Vieira, 2021).

Table 2 lists the most significant software tools for the CAT tools domain by the categories considered so far.

4 CATDeM

The second research question focuses on proposing a set of course design guidelines for computer scientists and researchers involved in teaching CAT courses (B.A./M.A. in translation studies). Normally, educators focus more on technical aspects and there is a risk of overlooking highly significant elements of the STT education as well as in-domain competence frameworks. In order to introduce a fundamental ICT-oriented perspective in translation studies, which is usually underestimated despite the growing use of cloud-based platforms by professional translators and language service providers (LSPs), this section

presents course design guidelines and their implementation in a real course (a 1st-year, one-semester compulsory course in the M.A. in STT and interpreting at the University of Salento, Italy).

4.1 Design Methodology

In order to make the CAT course structure outputted by CATDeM replicable in other universities and compatible with the professional translation workflows (PTWs) adopted by LSPs, a multi-stage design methodology was borrowed from business process modelling (BPM), called the *Define-Model-Execute-Monitor-Optimise* (DMEMO) cycle (BPM Resource Center, 2014). The core concept in DMEMO is to identify goals, requirements, and constraints so that a process can be modelled and subsequently executed. The process execution must be monitored against pre-defined evaluation metrics and then improvement actions can be introduced depending on the monitoring outcomes. After that, the DMEMO cycle starts again.

To transfer the DMEMO cycle to the CAT context, it has been assumed that the course represents the process and that each cycle coincides with an academic year. Each stage (Fig. 1) is defined as follows.

- *Define*: this is a teacher-only activity that needs to involve educators from both computer science and translation studies. It entails the definition of targeted skills (i.e., all EMT competences), purposes (i.e., learning objectives) and technological enablers (e.g., tools, needed educational software licences, etc.). The original DMEMO cycle requires that business modelling strategies are set during this stage: from the CAT-oriented perspective, such strategies coincide with the CSCL component needed to simulate a translation bureau (Sect. 4.5.2).
- *Model*: this stage defines (1) involved students and teachers, (2) required resources and technologies, (3) evaluation approaches, (4) collaborative/active learning strategies (i.e., discussion case studies, role-plays).
- *Execution*: it starts with a preparation period, to set up the necessary tools and platforms, then the actual work period starts (i.e., classroom instructional hours and laboratory activities).
- *Monitor*: it covers evaluation and assessment. Formative pre-grading assessments are provided by CAT and translation teachers for scaffolding purposes, during case studies and role-playing activities. Summative regular grading covers course topics (2/3 overall score) and role-playing (1/3 overall score). Summative post-grading is carried out via an online survey and only involves students that have been assessed. Monitoring results are crucial to guide the final stage.
- *Optimise*: monitoring outcomes are examined to identify potential improvements for the next academic year. Collaborative thinking sessions among involved educators are typically exploited to that end.

4.2 Learning Goals

The proposed learning outcomes are grounded on the EMT framework (Toudic & Krause, 2017), so to create an appropriate balance between the expectations of the educators from the field of translation studies and the typical teaching approach adopted in computer

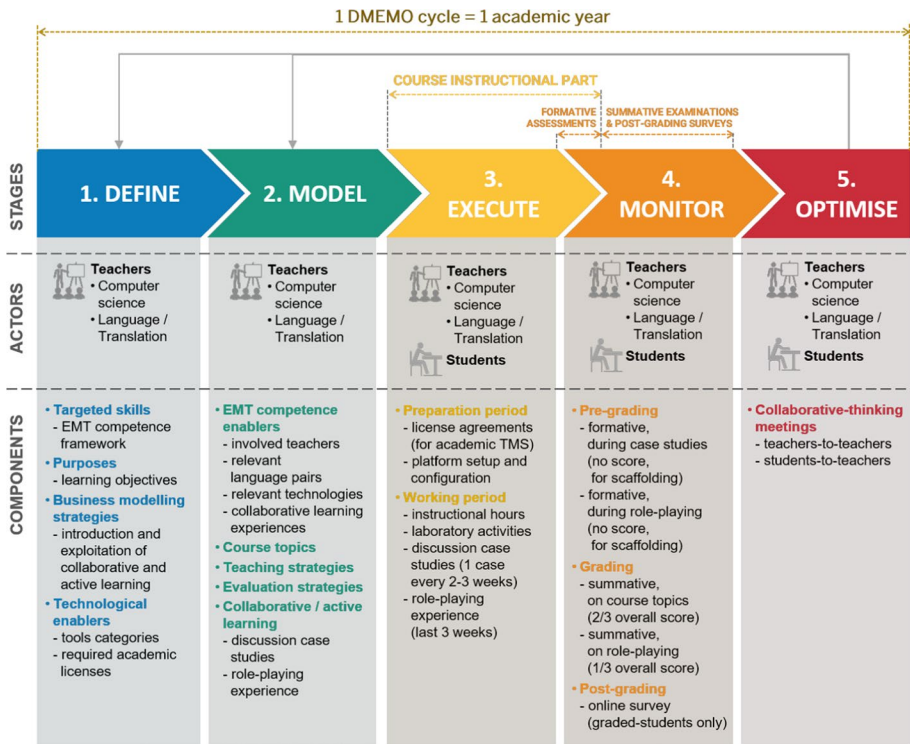


Fig. 1 Proposed CAT course design workflow based on the D MEMO cycle (BPM Resource Center, 2014)

science, which is based on skills-lab-like experiences and on a strong presence of computer-supported activities. Therefore, the EMT is leveraged as the common knowledge ground where shared learning goals are defined and CSCL dynamics are activated (Clark et al., 2002). In this way, students who complete the course will be able to achieve the following learning goals (LGs):

- (LG1) mastering current technology-driven trends in CAT and MT, in order to be capable of working profitably and independently in business-like scenarios;
- (LG2) learning the advantages of TMS platforms for STT, along with their limitations with other translation typologies;
- (LG3) knowing and directly experiencing PTW, as well as contextualising TMS platforms in PTW;
- (LG4) understanding translation quality principles, industry reference standards, and resource sharing (e.g., translation memories, termbases, etc.) to boost cooperative work;
- (LG5) acquiring MT basics, along with pre-editing, post-editing, and MT-TMS coupling dynamics;
- (LG6) exploring advanced MT topics, such as technologies involved (statistical vs neural), training principles (e.g., corpus-based training), and translation quality assurance procedures;

Table 3 Mapping of course topics, learning goals, and collaborative learning activities

Course topic*	Learning goal**								Collaborative and active learning activity	
	LG1	LG2	LG3	LG4	LG5	LG6	LG7	LG8	Role-playing	Discussion case study
T1	✓								✓	
T2		✓	✓	✓					✓	
T3			✓	✓	✓	✓				✓
T4				✓					✓	
T5			✓	✓					✓	
T6			✓	✓			✓			✓
T7			✓	✓		✓	✓	✓		✓

*Course topics T1: computer-Assisted Translation basics and terminology. T2: translation Management Systems. T3: Machine Translation: from training to post-editing. T4: translation quality assurance and standards. T5: professional translation workflow. T6: software and video game localisation. T7: fundamentals on language data management

**Learning Goals LG1: mastering technology-driven trends in CAT and MT. LG2: learning features, advantages and limitations of TMSs. LG3: knowing the professional translation workflow. LG4: understanding reference standards. LG5: acquiring MT basics. LG6: exploring MT advanced topics. LG7: experiencing software localisation features. LG8: approaching the language data perspective

(LG7) experiencing software localisation tasks and examining how they differ from traditional translation;

(LG8) approaching the cross-disciplinary importance of a data-oriented perspective when dealing with language data.

4.3 Content Structure

Students are introduced to a series of EMT-compliant topics (T) of increasing complexity from a computer-oriented perspective. As a second objective, those topics are designed to cover the broadest possible range of enabling tools, while at the same time stimulating a constant awareness of the targeted professional dimension. Current and future trends in the sector are constantly brought into focus, to stimulate student interest and motivation. Table 3 relates course topics to learning objectives and anticipates their relationship to collaborative learning activities (thoroughly explained in Sect. 4.5).

Topics are organised as in the following.

- T1. *CAT basics and terminology.* The course begins with an introduction to the basic concepts and definitions of the CAT domain. The significance of adopting computer-assisted solutions to improve translation quality, efficiency, and consistency is discussed with real-world examples. All the necessary elements common to most TMSs are explained to students, so that they can understand the role and the importance of translation projects, translation memories, glossaries, and translation quality checkers. Contextual awareness is provided in the form of typical application scenarios, with the aim of fostering students' critical thinking and situated learning.

- T2. *TMS*. These platforms represent the core content of any CAT course: here, several TMSs are presented, highlighting their common features and their ability to perform the same tasks (e.g., creating a translation project, populating a translation memory, exporting a terminology database, etc.), even if with different approaches. Indeed, one of the most underestimated difficulties for students in this field is the risk of becoming “locked in” to a particular software product, so that they continue to use the same tool they were introduced to, even long after they have finished their studies. This issue is mainly due to the fact that students start CAT courses with little computer acquaintance and, therefore, they do not possess enough operational autonomy in advance to be able to switch from one TMS to another with confidence depending on the necessities (e.g., document availability, backward compatibility of file formats, translation delivery constraints, compliance with standards, etc.). For this reason, the TMS-specific course module addresses the market-leading, desktop-based TMS (i.e., Trados Studio), a cloud-based TMS used to perform collaborative learning activities (i.e., Memsource,⁵ see Sect. 4.5.2), and a freeware desktop tool (i.e., OmegaT) that is a valid alternative to the other two (since Trados Studio is a commercial product and Memsource is free for academic purposes only).
- T3. *MT: from training to post-editing*. When it comes to MT, many students perceive it as a handy way of achieving quick and ready-to-use target documents. The first part of this topic aims to debunk such misconception in favour of a rigorous and systematic approach to how MTEs work and how they are trained, according to their type (e.g., neural vs statistical). Subsequently, the basics of pre-editing (Miyata & Fujita, 2016) and post-editing (Arenas & Moorkens, 2019) are introduced, involving students in practical discussion case studies (see Sect. 4.5.1). Specific attention is paid to comparing translators and post-editors, who differ in their methodological approaches and the software tools they need to accomplish their tasks.
- T4. *TQA and standards*. This topic is closely related to T2 and T3, being TQA crucial for distributing machine-translated or human-translated texts. Manual and automatic TQA procedures must be considered. To address the first type, students must be instructed on error taxonomies (Lommel, 2018; Mariana et al., 2015; Vardaro et al., 2019), which help them detecting and classifying errors by type and severity (Shi et al., 2019). As for automatic procedures, reference metrics such as BLEU (Papineni et al., 2002), METEOR (Satanjeev Banerjee & Lavie, 2005), or TER (Snover et al., 2006) must be examined. Students may also benefit from the use of text annotation tools.
- T5. *PTW*. A detailed analysis of typical LSP activities is proposed to the students, with a constant focus on tools and to real scenarios. Reference standards (ISO, 2015) are discussed and different professional roles (e.g., project manager, translator, reviewer, post-editor, etc.) are compared. This topic will be directly and thoroughly addressed during the role-playing activity (Sect. 4.5.2).

⁵ The cloud-based CAT tool Memsource was used to implement and evaluate the pilot case study for this research. In October 2022, Memsource unified its brand identity with Phrase (under the new name of Phrase TMS), a software localisation platform acquired by Memsource the previous year (Phrase, 2022). Currently, the module named Phrase TMS offers document translation features while the module called Phrase Strings provides software localisation functionalities. In order to be consistent throughout the paper with the activities performed during the pilot case study, Memsource will be referenced as the tool actually used. It is also important to point out that the majority of functionalities offered at that time by Memsource are now also available in Phrase without any significant modification.

- T6. *SW-LION*. This type of translation is proposed to the students by means of a 6-criteria comparison with document translation.
- i. *Process lifecycle*: the two translation processes are compared and their key features are highlighted.
 - ii. *Subject domain*: students learn that a software application contains at the same time texts about different domains, such as IT (e.g., technical settings), in-domain content, laws and regulations (e.g., disclaimers), dialogues and narratives (e.g., in video games), and so on.
 - iii. *Source text structure*: in software localisation, the source text may change several times, depending on the development approach adopted (e.g., DevOps, Agile methodology, etc.) (Kasakliev et al., 2019). Therefore, students are involved in examining typical software development workflows in order to understand when and how frequently software localisation tasks are required.
 - iv. *Text and file formats*: these are examined as they normally depend either on the specific programming language or on the specific string file format (e.g., .xls/.xlsx, .json, strings, etc.).
 - v. *Context-awareness*: students are instructed on how to manage decontextualised strings coming from a software app, whose logical order may be apparently missing and which usually contains a considerable amount of variables and source-code comments. These aspects are conveyed and examined by means of translation kits⁶ prepared by the teacher.
 - vi. *Localisation testing*: students learn how to evaluate the quality of translated text assets (e.g., language testing checklists, software testing approaches and cycles, etc.) (Mangiron, 2018; Sharma & Bhatia, 2014). This topic is supported by a dedicated discussion case study (Sect. 4.5.1).
- T7. *LDM fundamentals*. The last topic introduces students to a concept with which they are not familiar. Considering texts as *language-related data*, instead of simple elements to translate, is something they are not able to perceive at the beginning of the course, although this aspect is becoming more and more crucial, especially in the EU (ELRC Consortium, 2019). However, previous topics (especially, T1-T3), help to pave the way towards LDM from two perspectives. First, the *language data lifecycle* (Mattern, 2022), where data are collected from existing sources (e.g., text corpora) or created (e.g., by extracting them from interviews), catalogued and organised, annotated (to enrich with additional information content) or described, processed (e.g., simply translated, or used to train MTEs), preserved/archived, shared or even exchanged in a dedicated marketplace (TAUS, 2021), and eventually deleted. Second, thanks to a dedicated case study (Sect. 4.5.1.), it will be explained how to handle language data under specific conditions (e.g., translation of sensitive documents, use of cloud-based TMSs, localisation of text assets from unreleased video games), so that data ownership, copyright, and security requirements are considered (Moorkens & Lewis, 2019; Seinen & Van Der Meer, 2020).

⁶ The *translation kit* is a collection of files (translatable source and supporting references) supplied to translators by software developers.

4.4 Teaching Approach and Learning Platforms

The instructional part of the course spans over 14 consecutive weeks and involves a combination of complementary teaching strategies to keep students engaged throughout the entire period. To this end, the very first frontal lessons about CAT principles and definitions are soon followed by a series of practical, hands-on laboratory sessions enriched with computer-supported collaborative learning (CSCL) and active learning activities. These two approaches are triggered by group-works on selected discussion case studies (Sect. 4.5.1) and a role-playing experience (Sect. 4.5.2). While the case studies are 1-week-long activities distributed regularly throughout the course, role-playing is carried out in the last three weeks of the course.

It is noteworthy that the collaborative and active learning dimensions are fundamental to allow students perceiving the proposed topics as an interesting landscape to explore in the light of their prospective professional career rather than just a stack of overly complex ICT-related notions that they will have to struggle against due to their initial scarce expertise on informatics.

Moreover, since CAT educator may come from the computer-engineering sector (which is the norm in this scenario and which is also true for the proposed case study), there may exist a provenance divide with students and other educators from traditional courses on translation. In order to bridge such a gap and not to reduce the effectiveness of the course, the educator should assume the role of facilitator and leverage appropriately the advanced features offered by any online learning platforms already available. Another relevant aspect is that the expected improvements in learning outcomes should be shared in advance and explicitly with students, otherwise learning platforms will be perceived as an additional operational burden to keep during the course. Instead, online learning platforms have long been regarded as valuable teaching aids, especially when they enable cooperative work, under the definition of groupware platforms (Ellis & Wainer, 2004). To this end, Microsoft Teams is used, since it was adopted at this university to enable remote teaching and distance learning during the SARS-CoV-2 pandemic and, therefore, students are already familiar with it. The platform offers typical features such as document sharing, instant chat messaging, and audio-visual conferencing. Microsoft Teams also allows supporting students via a course-specific Wiki and as a full groupware for the role-playing experience (Sect. 4.5.2).

4.5 Collaborative and Active Learning Activities

4.5.1 Discussion Case Studies

Three course topics are linked to typical case studies from the translation industry sector: MT, SW-L10N and LDM. The case studies are presented according to a pre-defined analysis pattern (Table 4), which specifies:

- (1) Involved stakeholders,
- (2) Initial motivations and requirements,
- (3) Enabling tools
- (4) Needed tasks/interventions,
- (5) Required standards, process workflows, and guidelines,
- (6) Expected challenges,

(7) Possible mitigation strategies.

Students learn how to deal with specific situations in real-life scenarios. Similarly, students are also encouraged to approach work contexts in a systematic and methodical way; they are provided with tools, supporting files, and teaching materials for each case study. The active learning component is initiated by the teacher/facilitator, who starts the discussion by explaining the scenario, then provides adequate context awareness, and finally guides the students through the different steps of the analysis pattern listed above. A demo session is also performed by the teacher for each case study and then the students are requested to try the same procedure on their own (possibly in groups). It is crucial for the teacher to encourage the students to develop critical-thinking and decision-making approaches, so that expected the challenges and mitigation actions elicited in the analysis pattern are properly considered. Finally, in order to boost student involvement, the case studies are incorporated into the course as non-evaluative activities, so that group discussions are stimulated without any anxiety due to evaluation expectations.

The first discussion case study refers to a translation agency that decides to use a MT pipeline to translate a set of source documents supplied by an engineering company, from English into Italian language. The documents describe 5G wireless communication systems and potential risks to human health. The MT is required because the agency is supposed not to have any in-house or freelance translators with the needed expertise on that specific topic and language pair. A freely available MTE (i.e., DeepL) is selected and it must be used in conjunction with a typical TMS platform (i.e., Trados Studio), assuming that a valid DeepL API key is available (so that students have to check autonomously how the integration procedure has to be done). Since the adopted MT pipeline requires a full human post-editing intervention, students are required to refer to the corresponding TQA and MTPE guidelines (Lommel, 2018; Mariana et al., 2015; Vardaro et al., 2019), as well as to the ISO 17100:2015 standard (ISO, 2015). Students are also encouraged to reflect on the usefulness of a text annotation tool (i.e., CATMA) to post-edit the MT output. In the real-life scenario discussed, technical challenges and operational difficulties are expected: the former can be tackled by testing in advance how to integrate the APIs of the selected MTE with the agency's TMS platform, while the latter can be mitigated by ensuring a freelance human post-editor with adequate in-domain competence for (at least) the target language is available.

In the second discussion case study, the localisation from English into Italian language of a Python-based video game is discussed. Students are provided with a translation kit prepared by the teacher starting from an already published free open-source visual novel named "Under What?" (Gartman, 2019). The choice of a Python video game allows students to interact with code snippets with relatively low syntax complexity. The interaction with code snippets is facilitated by the teacher, who shows the differences between the original location of translatable strings in the source code and their typical translator-friendly appearance in a SLMS platform, where those strings are displayed as key-string pairs. Similarly, the selection a visual novel means that the majority of the translatable text assets come from dialogues and narrative sections, so that the typical issues related to decontextualised and non-sequential strings can easily emerge during the discussion. In addition to the SLMS (i.e., Phrase), another useful tool allowing students to examine the software developer's perspective is a game development environment: in this case, an IDE for Python visual novels is used (i.e., Ren'py). Since this case study focuses on a client

Table 4 Analysis pattern of the discussion case studies

Analysis pattern		Case study	
		Machine translation (MT)	
		Software localisation (SW-L10N)	
		3. Language data management (LDM)	
1. Stakeholders	Client: engineering company Service provider: translation agency	Client: independent software house Service provider: translation agency	Client: researcher, Service provider: translation agency
2. Initial motivations and requirements	MT needed to translate documents about 5G technology and health risks due to lack of in-house/freelance technical translators in that language pair	SW-L10N needed to translate a Python-based visual novel. Compliance with the software house's Agile software development approach is required	A researcher has written a scientific article in English and now needs it translated into Italian. The agency uses cloud-based TMS platforms only
3. Enabling tools	MTE (DeepL) with API key TMS (Trados Studio) Text annotation tool for TQA (CATMA)	SLMS (Phrase) Python game development and testing IDE (Ren'Py)	Cloud-based TMS (Memsourse)
4. Tasks	Full MTPE (Eng → Ita)	Agile SW-L10N (Eng → Ita) during two Agile sprints	Translation of a scientific manuscript (Eng → Ita)
5. Standards, guidelines, process work-flows	TQA guidelines Full MTPE guidelines ISO 17100:2015 standard	SW-L10N guidelines Agile software development lifecycle Agile translation and localisation workflow Software testing procedures	GDPR Copyright rules about language data and individual segments Technical documentation of the TMS
6. Challenges	MTE + TMS integration Lack of in-domain expertise for human post-editors	De-contextualised, non-sequential strings In-game texts (string length, special fonts, etc.) Tight deadlines (Agile sprints) Accurate software testing Studying the Agile localisation workflow accurately Adopting a testing procedure and a proper checklist	Identification of personal data Copyright handling on source/target text, segments Retrieval of the conditions of use from the TMS documentation Studying the GDPR Examining "Terms and Conditions of Use" clauses in the TMS end-user contract
7. Mitigation strategies	Testing the MTE + TMS integration in advance Consulting freelance in-domain experts with adequate competence (target language)		

Table 4 (continued)

Analysis pattern	Case study		
Machine translation (MT)	Software localisation (SW-L10N)	3. Language data management (LDM)	
Supplied files	Source documents about 5G technologies (Ofcom, 2022; UK Government, 2021) and guidelines about radiofrequency safety (ICNIRP, 2022)	Translation kit created by the teacher from a free, open-source video game (Gartman, 2019)	Scientific paper on a novel data mining algorithm (open access) (Yiwen, 2017)

company that works according to the Agile software development lifecycle (Martin, 2014), the main challenge for students is to understand how Agile software localisation works, thus considering how in this case developers typically push new code to the project codebase at the end of every Agile sprint and how translators/localisers are required to work in a continuous delivery mode. Students have also to learn how the SLMS platform synchronises the codebase with the translation project. In order to manage the non-sequential order of pushed strings, students are encouraged to understand the typical challenges encountered by software localisers and how to address them (Xia et al., 2013). Typical game localisation issues such as string length boundaries, special fonts, and variables are also widely discussed (Mangiron, 2018). Suggested mitigation strategies include testing procedures and checklists. This case study can also be acted out with students involved as localisers and the teacher acting as both facilitator and software developer. An interesting variation (whose feasibility is under examination for the next academic year) would be to involve also students from a M.Sc. degree in computer science and have them play the role of a real software house.

The third discussion case study covers the LDM field and focuses on how to handle legal rights in a translation supply chain, where the client, the agency, and the translators hold different intellectual property rights. This scenario starts with the assumption that the agency receives a request for an English-to-Italian translation of a scientific manuscript on a novel data mining algorithm. The adopted research article (Yiwen, 2017) has indeed already been published and is openly accessible. The agency is supposed to rely solely on a cloud-based TMS (i.e., Memsourse), whose servers are obviously not located on the agency's premises. Students are first introduced to the basics of data protection requirements, represented in this case by the European General Data Protection Regulation (GDPR) (European Union, 2022), as well as to the core concept of *personal data*, which entails data to a broader extent than its US counterpart PII (*personally identifiable information*) (Seinen & Van Der Meer, 2020). Students learn how to identify and manage personal data within language data and, then, they are challenged by the teacher to distinguish between copyright on the whole source text, and copyright on individual source segments.⁷ The discussion continues by introducing, in an inquiry-based approach, the differences between copyright on source text and copyright on target text. Since the cloud-based TMS is the only tool needed in this case study, students are asked to think critically about the physical location of the source text, once it has been uploaded to the cloud repository, and of the target text, before it is downloaded from the same repository. Consequently, it is explained how to gather the necessary information from the technical documentation of the TMS.

4.5.2 Role-Playing

The STB was incorporated as the last part of the instructional period, weighted at one third of the overall examination grade and designed according to the GLAID framework (Table 5). Students were required to self-group into translation teams, each one representing a fictional translation bureau. Every team had one project manager (PM), (at least) two translators, and one reviewer. Team members were free to choose their PM. They also

⁷ When a TMS processes a source document, the text is partitioned into (portion of) sentences, according to automatic segmentation rules based on strong punctuation marks. Every (portion of) sentence is called a segment. During the translation activity, source segments are then associated with translations (target segments) in a dedicated Translation Memory.

defined a fictional company logo and a fictional company name to be used throughout the role-playing experience. The academic-free edition of Memsourse was the enabling cloud-based collaborative platform for every team, as it provided a significant subset of functionalities typical of real LSPs.

The role-playing was initiated by translation/language teachers, acting as fictional clients and asking each team the provisioning of a translation service for a given set of source documents, within a given deadline, for a given language pair. Then, each PM negotiated a fictional quote (according to the typical per-word payment rate for the selected language pair), received the source documents from the fictional client, created a translation project in the available TMS platform and assigned tasks to translators and reviewers. Upon translation and revision, a translation quality check was performed and the PM returned targets document to the client with a fictional invoice.

5 Evaluation

5.1 Participants and Sampling Strategy

As anticipated in the previous sections, CATDeM has been evaluated in a real-life case study spanning two academic years (a.y.): it was used to design and implement two consecutive editions of a compulsory CAT course, which were delivered in 2020/2021 a.y. and 2021/2022 a.y., respectively. Students from 2021/2022 a.y. will be considered as the study group in this section, while students from 2020/2021 a.y. will be considered as the control group. The course was held in the M.A. in STT and Interpreting at the University of Salento (Italy). Students of other M.A. degrees in language studies at the same university could also choose it as an elective. However, none of these degrees offers any other course in translation technology to students, who only benefit from a compulsory module on basic computer science during the B.A. degree in order to build up their computer literacy. The instructional part (56 h, 9 ECTS) was hybrid: students mainly attended face-to-face, but they could also participate remotely via Microsoft Teams. These settings were the same in the two academic years considered. In both the study group and control group, only students who actively participated to the instructional part of the course (i.e., attended more than 70% of the lessons) were included.

As detailed in Table 6, in 2021/2022 a.y., 24 out of 26 enrolled students (92.3%) actively participated, while 2 students did not show up after the enrolling. The sample consisted of 19 female (79.2%) and 5 male (20.8%) participants. Only two of them (8.3%) attended the course outside prescribed periods. The control group is represented by the 31 actively participating students out of 40 enrolled in 2020/2021 a.y., and even if absolute numbers are slightly different, the demographic breakdown of the control group is very similar in both years. As for the contents, LDM was not covered in 2020/2021, nor were the case studies. The role-playing experience was present, albeit with a different setup. In terms of scoring, the percentage of students who passed the exam after 2 out of 7 attempts was 91.6% in 2021/2022 and 64.5% in the previous year, while the average score was quite similar.

The role-playing experience, summarised in Table 7, was a compulsory, in-course, 3-week activity in 2021/2022 a.y., whereas in the previous a.y. it was an elective, post-course, 6-month activity. The grading mechanism was also different: it was part of the course grade in 2021/2022 and it was linked to additional ECTS in 2020/2021. As an

Table 5 GLAID-based role-playing activities, schematised as in (De Hei et al., 2016)

Stage	Component
(S ₁) Analysis	<p>Characterisation of:</p> <p>Students: MA students in STT, MA students in foreign languages</p> <p>Teachers: CAT teacher from the computer science area, translation/language teachers</p> <p>Curriculum: compulsory module in a compulsory CAT course</p> <p>Collaborative goals: foster student personal and interpersonal skills defined in EMT framework</p> <p>Global goals: introduce students to the technology-driven PTW</p>
(S ₂) Design	<p>C₁. Interaction:</p> <p>Mode: online and in person</p> <p>Support: role-based, by peers in each translation team</p> <p>Collaboration: translation and revision tasks self-distributed in each translation team to achieve the best results</p> <p>C₂. Learning objectives:</p> <p>Individual: completing responsibly assigned tasks; complying with deadlines; avoiding complaints from peers/supervisors; avoid unauthorized data access and intentional damage to the TMS</p> <p>Group: provide the requested translation service as agreed with the fictional client</p> <p>C₃. Assessment:</p> <p>Summative (self-assessment): every student evaluates starting knowledge and achieved learning outcomes via two (pre/post) questionnaires</p> <p>Formative (technical): the CAT teacher periodically controls TMS usage and support requests</p> <p>Formative (translational): translation teachers assess the quality of target documents</p>
S _{3A} . Development (didactics)	<p>C₄. Task features:</p> <p>Task typology: (1) interacting with fictional clients; (2) using TMS; (3) managing translation projects; (4) translating documents; (5) reviewing documents; (6) complying with deadlines; (7) interacting with other team members; (8) assessing peers; (9) self-assessing</p> <p>Task meaningfulness: tasks must be EMT-compliant, engaging for students, grounded on real-life use cases, suitable for collaborative tasks</p> <p>Task frequency: each translation team must receive at least 1 task</p> <p>Task duration: each task should not last more than 1 week</p> <p>C₅. Structuring:</p> <p>Students: self-grouped into translation teams at the beginning of the experience</p> <p>Team: each team works as a fictional translation bureau</p> <p>Teachers (language/translation): play the role of fictional clients</p> <p>Teachers (CAT): play the role of platform administrator and provide technical support</p> <p>C₆. Guidance:</p> <p>Executor: teachers; students (only when instructing supervised peers—e.g., PM supervising translators)</p> <p>Teacher's role: promotes inquiry-based learning (<i>pedagogical</i>); motivates students and other teacher (<i>social</i>); supports teams, manages the TMS; resolves issues (<i>organisational</i>)</p> <p>Mode: classroom lectures and online meetings, via instant messaging and emails</p> <p>Timing: start before the beginning of the experience</p> <p>Duration: supervision of teachers via scaffolding with progressive fade out until the first half of the experience; supervising students should be active until the end of experience</p>

Table 5 (continued)

Stage	Component
S _{3B} . Development (logistics)	<p>C₇. Group constellation: Group type: each group of students is a translation team Group size: no less than 1 PM, 2 translators, 1 reviewer Number of groups: sufficient to cover all the students attending the course</p> <p>C₈. Facilities: Teaching resources: scripts about each role Learning resources: handouts and scripts about roles, TMS use and task management Time duration: the experience covers the last 3 weeks of the instructional part of the course: Space location: activities are performed both in classroom and online (via a cloud-based TMS) IT platforms: Mensource: cloud-based TMS (as online CAT for students; as back-end with control elements for teachers) Microsoft Teams: to offer file sharing, communication, time and task management, activity workflow, and WIKIs (where to store resources)</p>
S ₄ . Implement	<p>Monitoring: Instructional process: participating students are monitored both during classroom activities by the supervising teacher and online, via TMS/Teams usage logs and reports</p>
S ₅ . Evaluation	<p>Evaluation: Summative (overall, educator-provided): one third of the course examination grade Summative (student self-assessment): via online survey administered to all graded students Formative (in-domain, educator-provided): quality tests by translation teachers on target documents</p>

elective, the role-playing in the previous a.y. was attended by 51.6% of active students, while the female-to-male gender ratio was 3.8 in 2021/2022 and 4.3 in 2020/2021.

5.2 Data Collection and Data Visualisation

A post-grading online survey was administered to the study group, after passing the exam, to determine:

1. Student self-assessment of competence improvement (question Q1),
2. Student perception of the course topic's usefulness with respect to future professional career (Q2),
3. Student opinion about the usefulness of collaborative/active learning activities in improving theoretical knowledge (Q3) and practical/interpersonal skills (Q4),
4. Student-perceived level of difficulty in learning how to use the proposed tools (Q5) and in using them (Q6),
5. Student feedback on: most interesting topic and collaborative/active learning activity (Q7),
6. STUDENT feedback on topics and collaborative/active learning activities to improve and/or reduce (again Q7).

Table 6 Comparison between study group and control group (over the entire CAT course)

	Study group (academic year 2021/2022)	Control group (academic year 2020/2021)
<i>Demographics</i>		
Enrolled students	26	40
Active participants (attendance > 70%) to instructional part*	24 (92.3% of enrolled)	31 (72.5% of enrolled)
Female participants	19 (79.2% of active)	24 (74.2% of active)
Male participants	5 (20.8% of active)	7 (22.6% of active)
Outside-prescribed-term participants	2 (8.3% of active)	4 (12.9% of active)
<i>Contents</i>	From T1 to T7 (see Table 3)	From T1 to T6 (see Table 3)
<i>Collaborative and active learning activities</i>		
Discussion case studies	✓	
Role-playing experience	✓	✓
<i>Scoring after 2 (out of 7) exam attempts</i>		
Participants who passed the exam	22 (91.6% of active)	20 (64.5% of active)
Average score	28.9 out of 30 pts	29.2 out of 30 pts

*In both groups, only the active participants have been included in the study sample

Table 7 Comparison between study group and control group (over the role-playing experience only)

	Study group (academic year 2021/2022)	Control group (academic year 2020/2021)
<i>Demographics</i>		
Participants	24 (100% of active)	16 (51.6% of active)
Female participants	19 (79.2% of active)	13 (81.2% of active)
Male participants	5 (20.8% of active)	3 (18.8% of active)
Male-to-female gender ratio	0.26	0.23
Female-to-male gender ratio	3.8	4.3
<i>Role-playing experience (details)*</i>		
Curriculum	Compulsory, in-course	Elective, post-course
Time duration	3 weeks	6 months
Task frequency	1 per week	1 per month
Task average duration	1 week	1 month
Scoring mechanism	1/3 of the course grade	Additional ECTS
Average student constellation size	3.25 members	3.2 members

*Only the differing aspects are shown. GLAID framework was used in both the academic years to design the role-play

All questions were mandatory. Table 8 lists each question, the elements covered, and the corresponding Likert scale or request. Questions Q1, Q3–Q6 adopted a 5-point Likert scale with a neutral central item but different assessment scales depending on the question (Vagias, 2006). The answers to these questions will be reported as diverging

simple horizontal stacked bar charts, centred on the neutrality line, so to make the distribution of respondents' opinions immediately visible (Evergreen, 2019).

Question Q2, which asked respondents to quantify how useful the topics covered in the course could be in terms of a translator's future professional career, used a progressive intensity 5-point Likert scale with the addition of a sixth item corresponding to the "unsure" option (score: 0). The option was deemed necessary in order not to force respondents to make such an assessment, as some of them may choose a different career at the end of their studies or may not yet have a clear idea of their professional expectations. For the "0" item in this particular scale, a decision had to be made whether to include these responses in the statistical analysis or to exclude them (i.e., complete case analysis), as they are missing valid responses. It was decided to exclude the 0-scores from the analysis as they correspond to respondents who do not actively express their opinion on the usefulness of a given topic, thus making these responses fundamentally different from the others and, therefore, not part of the overall distribution. From a data visualisation perspective, it was decided to use a chart with 100% horizontal stacked bars (Evergreen, 2019), since the assessment scale adopted does not have a neutrality line in the middle.

Question Q7 asked students to indicate the most interesting topic as well as the most interesting collaborative and active learning activity. Similarly, they were asked to indicate what they would like to emphasise (in a forthcoming edition of the course) and what they would like to shorten. The visualisation chosen in this case is a three-column heatmap (Evergreen, 2019), where the elements of interest are placed side by side and the adopted colour gradient scale immediately shows those with the highest number of responses.

The response rate in the study group was 75%, with 18 respondents for each question. Although 18 respondents may seem relatively few, it is important to consider the context of the case study. Firstly, the study group had a limited number of components, as it consisted of 24 actively participating students out of a total of 26 enrolled students, as shown in Table 6. The questionnaire was only administered to actively participating students, so the maximum number of students invited to participate was actually 24. In general, such a relatively small sample size is not uncommon for specialised courses, particularly in niche areas such as STT and especially considering the typical annual enrolment rates for the M.A. degree to which the CAT course covered in this case study belongs, which has an average of 30 students enrolled per year. Therefore, achieving a response rate of 75% indicates a high level of student engagement and represents a significant proportion of the total eligible student population, thus suggesting a willingness to provide feedback on the course. Secondly, while a larger sample of respondents would be desirable in order to make generalisations, the goal of our case study was not to make broader assumptions but rather to gain insights into the self-perceived skills and attitudes of the students who had taken the course designed according to CATDeM. Therefore, by focusing on a specific cohort (i.e., the 2021/2022 cohort) we aimed at exploring the effectiveness of the course structure and at identifying areas for further improvement.

5.3 Results and Findings

The first six questions in the survey were all associated with Likert scales and each covered different aspects, one per sub-question, as shown in Table 8. More specifically, all

Table 8 Post-grading online survey for students

ID	Question	Sub-question	Evaluation scale/request
Q1	To what extent do you agree with the statement: "The course has improved my starting skills"?	All five EMT competence types: Language, Translation, ICT tools, Team work, Interpreting)	5. Strongly agree 4. Agree 3. Neutral 2. Disagree 1. Strongly disagree
Q2	Do you feel this topic will be useful in your future professional career?	All course topics: T1 (CAT), T2 (TMS), T3 (MT + MTPE), T4 (TQA), T5 (PTW), T6 (SW-L10N), T7 (LDM)	5. Extremely 4. Very 3. Moderately 2. Slightly 1. Not at all 0. Unsure
Q3	How useful were role-playing and case studies in improving your theoretical knowledge?	The role-playing experience (STB) and the three case studies (MT, SW-L10N, LDM)	5. Definitely useful 4. Useful 3. Neutral 2. Useless 1. Definitely useless
Q4	How useful were role-playing and case studies in improving your practical and operational skills?		5. Very easy 4. Easy 3. Neutral 2. Difficult 1. Very difficult
Q5	Please rate how you perceived learning the tools proposed in this course	4 tool categories: TMS with QA functions, SLMS, MTE, CLT	
Q6	Please rate how you perceived using the tools proposed in this course		
Q7	Please indicate the topic and the collaborative learning activity you found most interesting. Also indicate what you would like to improve and what would you like to reduce	All course topics, case studies and the roleplaying experience	1 preference per topic, 1 preference per collaborative learning activity 3 requests: most interesting element, element to improve, element to reduce

five EMT competencies were covered by Q1, all seven course topics by Q2, role-playing and all three discussion case studies by Q3 and Q4, and the four main tool categories by Q5 and Q6, for a total of 28 sub-questions. Question Q7 asked for two responses about the perceived level of interest, need for improvement, and need for reduction in course topics and collaborative learning activities, for a total of 6 additional responses per respondent.

Before examining the corresponding Likert charts, a statistical analysis is now provided. Descriptive statistics are reported in Table 9. For each question and corresponding topic, the following elements are considered: number of valid/excluded responses, mean, median (MED), standard deviation (STD), interquartile range (IQR), range, skewness and kurtosis, minimum (MIN), and maximum (MAX) values.

All sub-questions received 18 valid responses with the exception of the sub-questions on the perceived usefulness of software localisation (Q2.T6(SW-L10N)) and language data management (Q2.T7(LDM)) for future professional career, where 3 and 4 respondents respectively chose the 0-score option “*unsure*”. Those responses were then excluded from the analysis as not representative of a clear opinion of the respondent. Questions Q1, Q3, and Q4 have the highest means, with the highest value (4.667) corresponding to the expected usefulness of TMS in a translator’s professional career (Q2..T2(TMS)). The positive attitude of the students towards TMS is confirmed by the fact that the same question shows the lowest STD (0.485), indicating the greatest agreement among the students’ perceptions. The lowest means are found in questions Q5 and Q6, with the lowest value (2.444) associated with the perceived usage difficulty of software localisation management systems (Q5.SLMS), although this question has a STD of 1.149, indicating greater variability and dispersion of responses. Similarly, the difficulty of using TMS has a large STD, suggesting that students reacted differently to the hands-on activities with that type of applications. As for the dispersion of the middle 50% of the data (IQR), the lowest dispersion is in Q1.Interp_Skills, suggesting that the students perceived the course as very effective in improving interpersonal skills (this is confirmed by $Mean=4$ and $STD=0.594$), and in Q3.CS1(MT), indicating that the students perceived the discussion case study on MT very useful ($Mean=3.944$, $STD=0.639$). Skewness is negative in 20 out of 28 sub-questions, thus indicating a longer left tail in the distribution of responses, which together with the other descriptive statistics confirms that the majority of respondents have positive perceptions and attitudes. Kurtosis was negative for 23 of the 28 sub-questions, meaning that their distributions are relatively flat compared to a normal distribution. Therefore, the combined analysis of skewness and kurtosis for those questions indicates that students’ negative perceptions and attitudes are fewer than their positive ones, which is also visually confirmed by the ridgeline plot (Holtz, 2023) showing the combined distribution of responses in Fig. 2. From that figure is also noticeable how students perceived LDM: it was considered as less useful for professional career (Q2.T7(LDM)) but certainly relevant for improving theoretical knowledge (Q3.CS3(LDM)) and operational skills (Q4.CS3(LDM)). As for SLMS, it was perceived as difficult both from a learning perspective (Q5.SLMS) and from a usage perspective (Q6.SLMS). Figure 2 also shows that the range is larger especially in Q5 and Q6 (except for Q6.MTE).

Correlation analysis was then carried out between Q3 and Q4, and between Q5 and Q6, as they covered the same aspects from different perspectives (Table 8). Spearman’s rank correlation (ρ) was preferred to Pearson correlation, because respondents were provided with Likert ordinal scales. Statistical significance was calculated using the p -value parameter to check whether the observed correlation occurred by chance.

Table 10 reports the correlation analysis for Q3 and Q4, whose heatmap is shown in Fig. 3 on the left. The highest and most significant positive correlation ($\rho=0.67$, $p=0.02$)

Table 9 Descriptive statistics for the post-grading online survey

Question.Sub-question	Valid/Excl	Mean	MED	STD	IQR	Skewness	Kurtosis	Range	Min	Max
Q1.Language	18/0	3.611	4	0.608	1	0.408	-0.513	2	3	5
Q1.Translation	18/0	4.167	4	0.707	1	-0.25	-0.776	2	3	5
Q1.ICT_tools	18/0	4.444	5	0.705	1	-0.915	-0.252	2	3	5
Q1.TeamWork	18/0	4.556	5	0.616	1	-1.085	0.387	2	3	5
Q1.Interp_Skills	18/0	4.000	4	0.594	0	0.000	0.425	2	3	5
Q2.T1(CAT)	18/0	4.611	5	0.502	1	-0.498	-1.987	1	4	5
Q2.T2(TMS)	18/0	4.667	5	0.485	1	-0.773	-1.594	1	4	5
Q2.T3(MT + MTPE)	18/0	4.222	4	0.647	1	-0.230	-0.411	2	3	5
Q2.T4(TQA)	18/0	4.222	4	0.732	1	-0.383	-0.906	2	3	5
Q2.T5(PTW)	18/0	3.833	4	0.618	0.75	0.093	-0.101	2	3	5
Q 2.T6(SW-L10N)	15/3	3.000	3	0.845	0.5	-0.819	0.969	3	1	4
Q2.T7(LDM)	14/4	2.857	3	1.027	1.5	-0.669	-0.355	3	1	4
Q3.Roleplaying	18/0	4.222	4	0.548	0.75	0.159	0.181	2	3	5
Q3.CS1(MT)	18/0	3.944	4	0.639	0	0.041	-0.143	2	3	5
Q3.CS2(SW-L10N)	18/0	4.167	4	0.618	0.75	-0.093	-0.101	2	3	5
Q3.CS3(LDM)	18/0	4.278	4	0.752	1	-0.529	-0.933	2	3	5
Q4.Roleplaying	18/0	4.278	4	0.575	1	-0.022	-0.255	2	3	5
Q4.CS1(MT)	18/0	4.222	4	0.647	1	-0.230	-0.411	2	3	5
Q4.CS2(SW-L10N)	18/0	4.556	5	0.616	1	-1.085	0.387	2	3	5
Q4.CS3(LDM)	18/0	3.778	4	0.732	1	0.383	-0.906	2	3	5
Q5.TMS + QA	18/0	3.222	3	1.166	1.75	0.016	-0.675	4	1	5
Q5.SLMS	18/0	2.444	2	1.149	1	0.545	-0.127	4	1	5
Q5.MTE	18/0	4.167	4	0.707	1	-0.250	-0.776	2	3	5
Q5.CLT	18/0	2.944	3	0.873	1	-0.482	-0.188	3	1	4
Q6.TMS + QA	18/0	3.667	4	1.029	1	-0.324	-0.871	3	2	5
Q6.SLMS	18/0	3.167	3	1.295	2	-0.163	-0.994	4	1	5
Q6.MTE	18/0	4.333	4	0.594	1	-0.210	-0.472	2	3	5
Q6.CLT	18/0	3.167	3	0.707	1	-0.250	-0.776	2	2	4

is between Q3.Role-Playing and Q4.CS1(MT), namely between the perceived usefulness of role-playing for improving theoretical knowledge and discussion case studies for operational skills in MT. This is also confirmed by the fact that during the role-play activities and the discussion case studies, students frequently compared their opinions on the benefits and limitations of MT. The usefulness of the discussion case study on SW-L10N for improving theoretical knowledge (Q3.CS2(SW-L10N)) and the usefulness of the discussion case study on language data management for improving operational skills (Q4.CS3(LDM)) showed the lowest correlation ($\rho = -0.446$, $p = 0.063$).

Table 11 reports the correlation analysis for Q5 and Q6. As it can be seen in the corresponding heatmap (Fig. 3, on the right), there are more positive correlations between these two questions. The highest and most significant ones are between Q5.SLMS and Q6.SLMS ($\rho = 0.879$, $p < 0.001$), and between Q5.CLT and Q6.SLMS ($\rho = 0.886$, $p < 0.001$), indicating that learning and using those two classes of applications was perceived in the same way by the students.

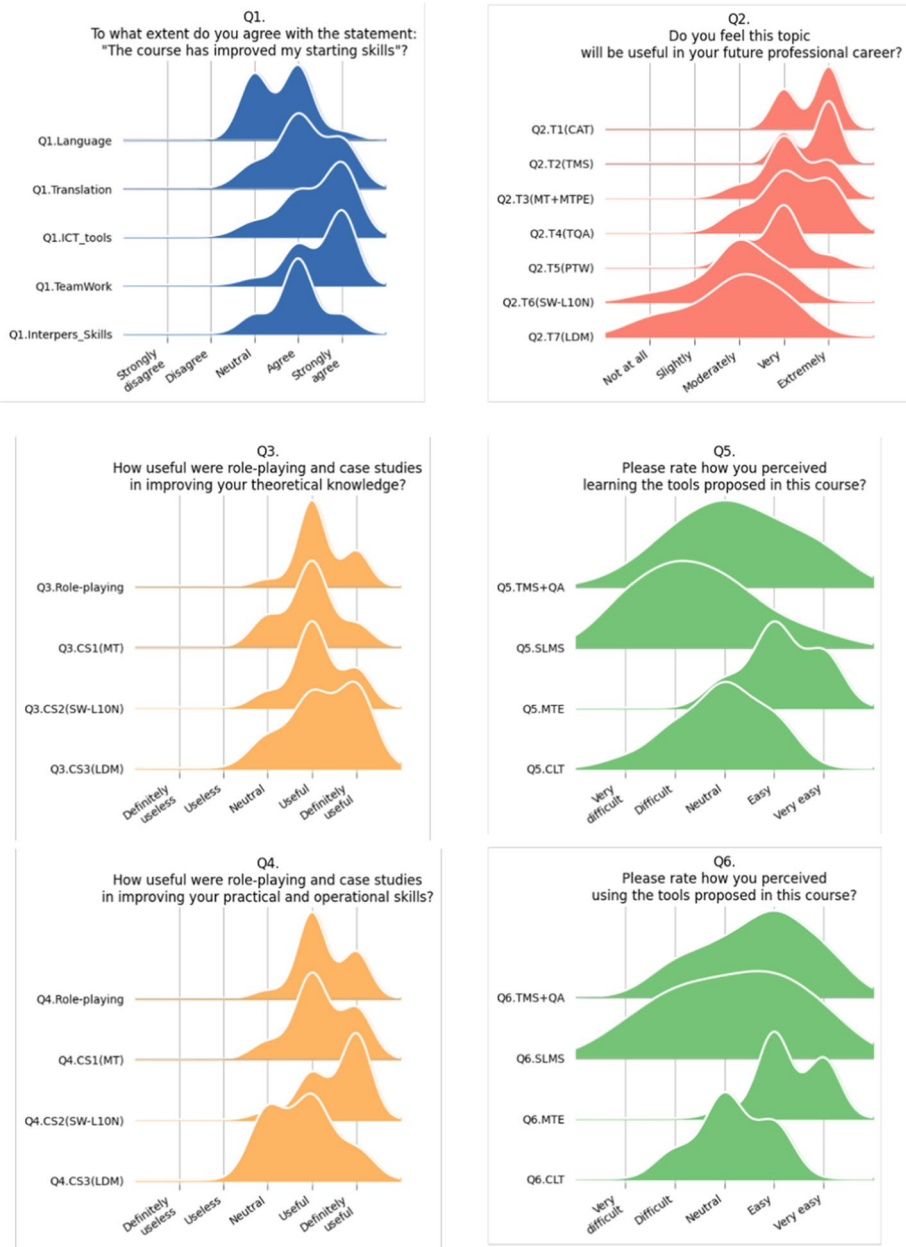


Fig. 2 Ridgeline plots showing distribution of responses by question and by sub-question. Questions Q3 and Q4 shares the same colour as they are related. The same applies to questions Q5 and Q6

In the remaining part of this subsection, Likert charts (i.e., horizontal bar charts) are presented for each question, to provide an analysis of their findings that complements the ridgeline plots in Fig. 2. In question Q1 (Fig. 4), students were asked to self-assess their improvement in relation to the EMT set of competences (Toudic & Krause, 2017).

Table 10 Correlation analysis (Spearman's rank correlation and p-value) for questions Q3 and Q4

Sub-question	Param	Q3. Role- playing	Q3. CS1 (MT)	Q3. CS2 (SW-L10N)	Q3. CS3 (LDM)	Q4. Role- playing	Q4. CS1 (MT)	Q4. CS2 (SW-L10N)	Q4. CS2 (SW-L10N)
Q3.Role-playing	ρ	-							
	<i>p-value</i>	-							
Q3.CS1 (MT)	ρ	-0.091	-						
	<i>p-value</i>	0.720	-						
Q3.CS2 (SW-L10N)	ρ	0.040	-0.291	-					
	<i>p-value</i>	0.874	0.241	-					
Q3.CS3 (LDM)	ρ	0.142	0.279	-0.142	-				
	<i>p-value</i>	0.574	0.262	0.574	-				
Q4.Role-playing	ρ	-0.002	0.515	-0.132	0.537	-			
	<i>p-value</i>	0.993	0.029	0.602	0.021	-			
Q4.CS1 (MT)	ρ	0.670	-0.075	0.059	-0.138	-0.014	-		
	<i>p-value</i>	0.002	0.769	0.817	0.584	0.958	-		
Q4.CS2 (SW-L10N)	ρ	0.107	-0.075	-0.115	0.189	-0.166	0.122	-	
	<i>p-value</i>	0.673	0.767	0.649	0.453	0.510	0.631	-	
Q4.CS3 (LDM)	ρ	0.022	-0.008	-0.446	0.327	0.307	-0.158	0.238	-
	<i>p-value</i>	0.930	0.976	0.063	0.185	0.215	0.530	0.341	-

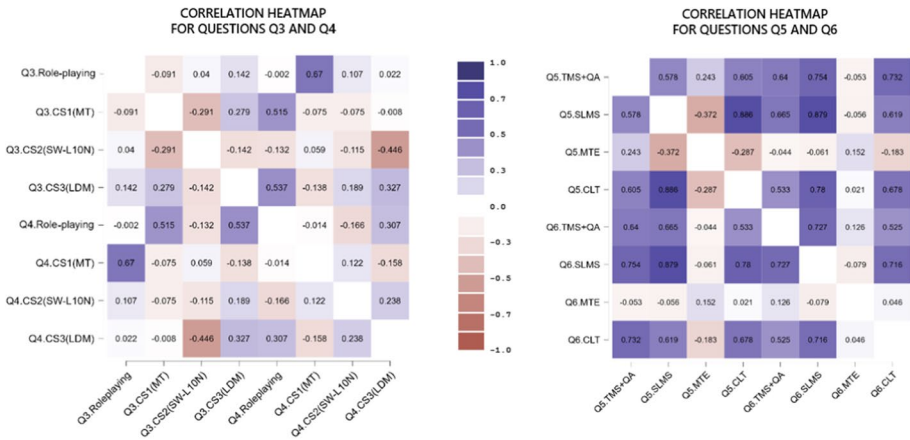


Fig. 3 Correlation heatmaps for questions Q3 and Q4 (on the right), Q5 and Q6 (on the left)

No negative responses were collected. As expected, the language competence scored lower than the others, since the course was not aimed at improving language skills already acquired in other courses. The competences related to ICT tools and teamwork scored best, showing that students perceived a significant increase in their skills. The role-playing experience contributed most to this result, as it will be clarified by the analysis of the other questions. It is also worth to mention that there were very few neutral scores, thus indicating that students clearly perceived the effectiveness of the course.

When asked about their future career and how useful they perceived the course topics, students answered remarkably positively (Fig. 5). Also in this case, five out of seven topics received only positive ratings (i.e., either moderately, very, or extremely useful). TMS was the best-performing topic, with more than two thirds of respondents stating that they expected these platforms to be crucial for their work. Similarly, CAT competences and TQA guidelines/standards were given considerable importance. The last two course topics received more mixed responses. Some respondents chose the “unsure” option: this can be explained because the proposed course is a first-year one, so students may need more time before deciding on their career. A small subset of respondents were definitely convinced that they would not exploit SW-L10N and LDM in the future, thus suggesting that they were not considering from the beginning to apply for jobs requiring these specific competences and skills.

The survey then asked students (Fig. 6) to what extent collaborative and active learning was considered useful in improving theoretical knowledge (question Q3) and practical/operational skills (question Q4). As there were no negative responses to either question, this provides a first immediate feedback on how effective the proposed course structure was perceived by the students.

Moreover, several further insights emerge by breaking down answers by type of activity and skill. Firstly, role-playing was considered the most effective experience, with operational skills even slightly more important. Secondly, the two case studies on MT and SW-L10N performed equally well but the second one received the best evaluation in terms of operational skills (*Extremely useful* for 61% of respondents). This result is a clear indication that the majority of students perceived the topic as a valuable educational source, even if some of them decided not to follow that specific path (see question Q2). Moreover, it can

Table 11 Correlation analysis (Spearman's rank correlation and p-value) for questions Q5 and Q6

Sub-question	Param	Q5. TMS+QA	Q5.SLMS	Q5.MTE	Q5.CLT	Q6. TMS+QA	Q6.SLMS	Q6.MTE	Q6.CLT
Q5.TMS+QA	ρ	-							
	<i>p-value</i>	-							
Q5.SLMS	ρ	0.578	-						
	<i>p-value</i>	0.012	-						
Q5.MTE	ρ	0.243	-0.372	-					
	<i>p-value</i>	0.332	0.129	-					
Q5.CLT	ρ	0.605	0.886	-0.287	-				
	<i>p-value</i>	0.008	<.001	0.249	-				
Q6.TMS+QA	ρ	0.640	0.665	-0.044	0.533	-			
	<i>p-value</i>	0.004	0.003	0.862	0.023	-			
Q6.SLMS	ρ	0.754	0.879	-0.061	0.780	0.727	-		
	<i>p-value</i>	<.001	<.001	0.811	<.001	<.001	-		
Q6.MTE	ρ	-0.053	-0.056	0.152	0.021	0.126	-0.079	-	
	<i>p-value</i>	0.834	0.825	0.546	0.933	0.619	0.756	-	
Q6.CLT	ρ	0.732	0.619	-0.183	0.678	0.525	0.716	0.046	-
	<i>p-value</i>	<.001	0.006	0.467	0.002	0.025	<.001	0.855	-

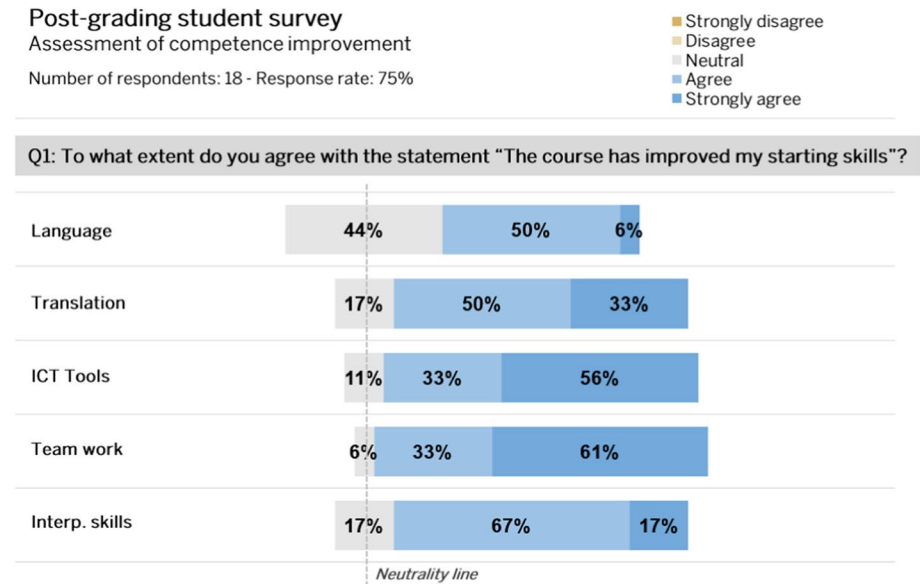


Fig. 4 Student self-assessment of competence improvement at the end of the course (question Q1)

be assumed that only when students are given a clear understanding of a specific methodology, technology, or platform they are supplied with the necessary elements to make a motivated choice about their professional career. The theoretical effectiveness of MT received a slightly higher percentage of neutral responses (22%), thus indicating that more students were already informed to some extent about typical MT challenges. Finally, the LDM case study received significantly positive results for theoretical knowledge and the highest percentage of neutral responses (38%) for operational skills. This can be explained in two ways: firstly, it was the case study with less operational content compared to the other two; secondly, as it also involved legal aspects (with which the students are not familiar), some respondents did not understand really the practical implications of LDM. Consequently, this perception could be mitigated by modifying the case study structure: for instance, experts in data privacy/protection and legal regulations can be invited to hold special seminars during the course to provide students with more contextual awareness.

Questions Q5 and Q6 focused on the tool categories covered in the course: TMS with QA functionalities, SLMS, MTE, and CLT. In these situations, it is always necessary to distinguish between the difficulty of learning to use a tool and the difficulty of actually using it. Overall, the learning phase was perceived as more difficult: Fig. 7 shows that the brown-shaded responses (i.e., *Very difficult* and *Difficult*), are greater for question Q5 compared to Q6, per tool category.

For instance, TMS + QA was *Very difficult* to learn for 6% of respondents and *Difficult* for 22%, while it was difficult to use for only 17% of respondents. SLSM achieved the highest perceived difficulty in both learning (22% + 33%) and using (11 + 22%), but also in that case the perceived difficulty decreased by 22%. SLSMs, due to their inherent complexity and proximity to the field of computer engineering, were the ones representing the most challenging category for the students of the course, thus indicating the need for more preparatory educational steps.

Post-grading student survey

Assessment of topic usefulness

Number of respondents: 18 - Response rate: 75%

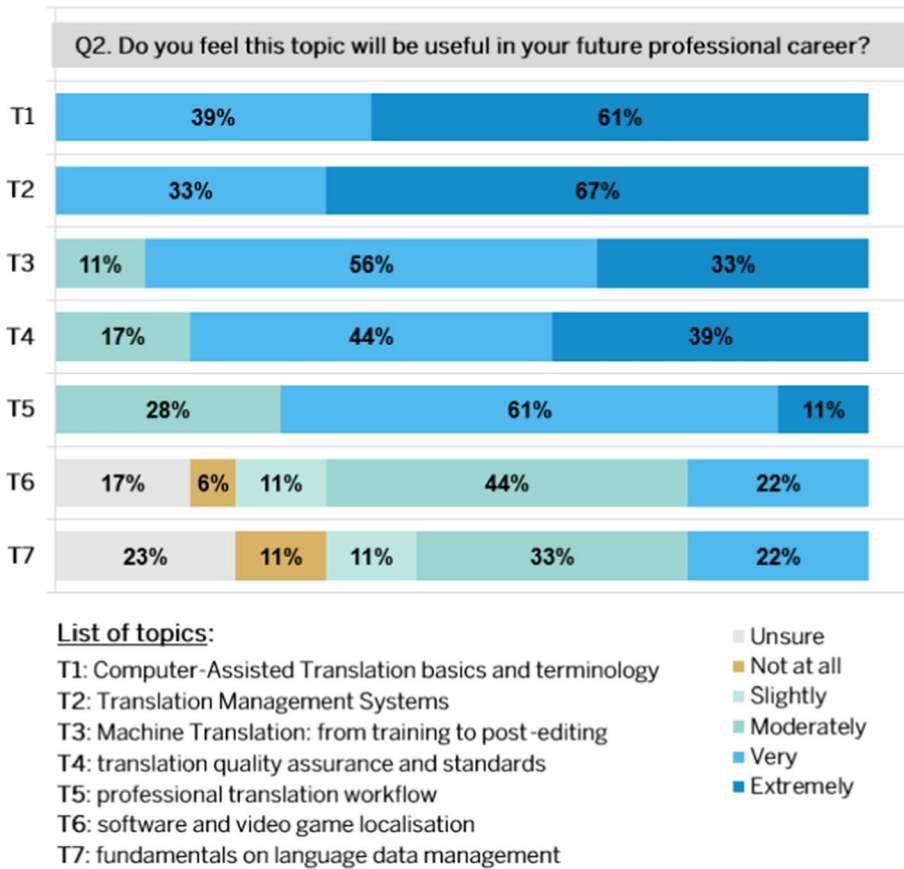


Fig. 5 Student self-assessment of topic usefulness for future professional career (question Q2)

Several other findings emerge from the breakdown of responses. First, MTE was rated as the least difficult (both to learn and to use). Second, the first three categories of tools received few neutral responses about their use: therefore, at the learning stage, students split between the relatively few continued to struggle with the tools and those who used the tools effectively. Third, CLT had the highest number of neutral answers about use: this indicates the relative simplicity of these tools and may suggest that students were less engaged or interested in the activities requiring a CLT. Therefore, question Q7 focused on students' level of interest, as now explained.

When asked what they found more interesting in the course (only one preference among topics and only one preference among collaborative and active learning activities), students answered as depicted in Fig. 8. TMS was considered the most interesting topic, followed at some distance by MT and MTPE, and PTW. As a confirmation of what emerged from previous questions, LDM received the lowest score. It is also important to note that the answer "none" was never selected. As for the case studies, MT and SW L10N received large

Post-grading student survey

Perceptions on collaborative/active learning activities

Number of respondents: 18 - Response rate: 75%

- Definitely useless
- Useless
- Neutral
- Useful
- Definitely useful

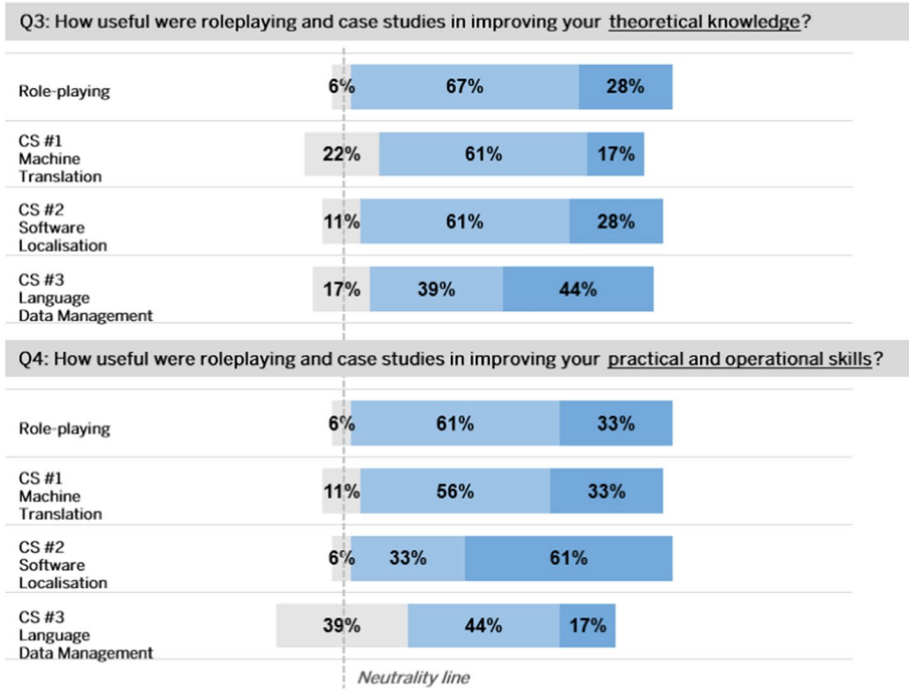


Fig. 6 Student perceptions about collaborative and active learning activities (questions Q3, Q4)

preferences, while LDM was the one with the lowest score. Also in this case, no students opted for the “none” answer.

In terms of what should be improved, the heatmap in Fig. 8 denotes that more than half of the respondents ($n=10$ out of 18) would not improve any topic further and that would not reduce any topic as well ($n=15$), thus indicating that they believe the course is already well balanced. Some students ($n=3$ per topic) proposed to emphasise TQA and SW L10N. Similarly, LDM was identified as the only topic that should be reduced ($n=3$).

On the contrary, the collaborative and active learning activities received a significant number of requests to emphasise even more the role-playing experience ($n=8$) and the case study dedicated to software localisation ($n=5$): this clearly indicates how students were involved in these experiences and how much they enjoyed them. At the same time, a large majority of respondents ($n=14$) suggested that none of the collaborative experiences should be reduced.

Post-grading student survey

Perceptions on tool categories

Number of respondents: 18 - Response rate: 75%

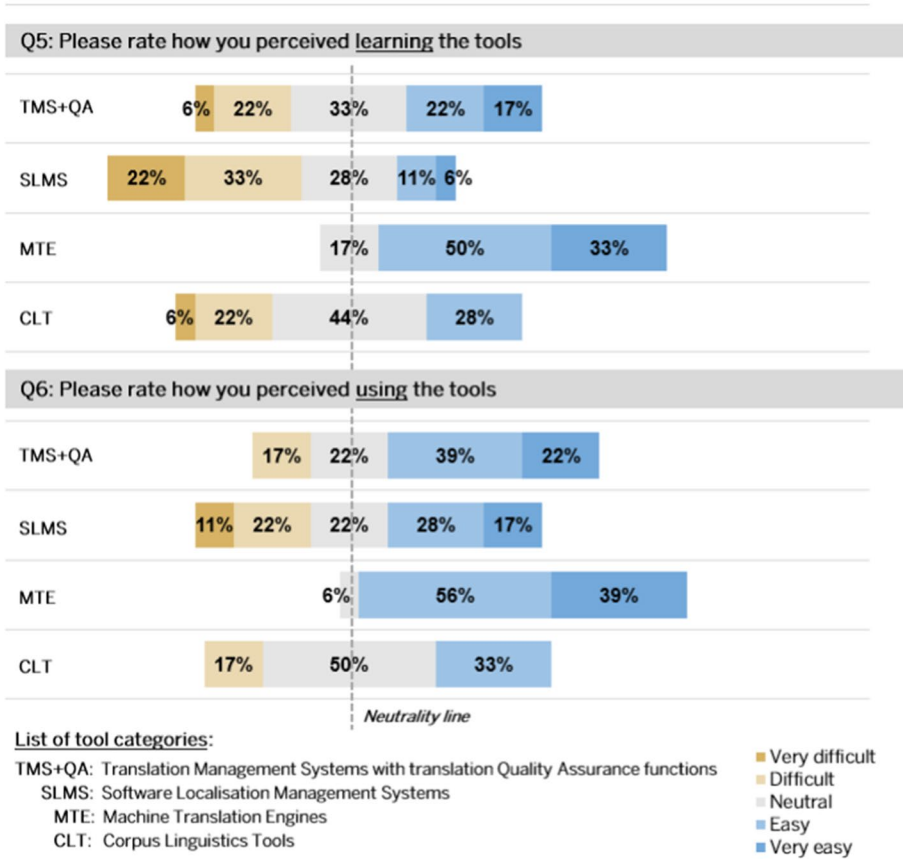


Fig. 7 Student perceptions on the proposed tools (questions Q5, Q6)

6 Discussion

Before examining in the following subsections how RQs (Sect. 1.2) have been answered, it is important to consider CATDeM, the CAT course design methodology proposed in this study, in terms of the different frameworks it integrates. A direct comparison against each of the composing elements (i.e., EMT; STB, CSCL, GLAID) is not feasible, as they have different purposes and CATDeM is not aimed at replacing them. Taken individually, those frameworks offer complementary features that have been combined in CATDeM in order to provide STT students with a holistic learning experience. More specifically, the original aim of EMT is to act as a reference for the evaluation of translation teaching curricula, while in CATDeM it is used as a comprehensive reference to align all the other components to the skills categories EMT refers to, starting from the STB. Since different STB

Post-grading student survey

Final assessment on topics and collaborative/active learning activities

Number of respondents: 18 - Response rate: 75%

Q7: Please indicate the topic and the collaborative learning activity you found most interesting. Also indicate what you would like to improve and what you would like to reduce.*

		Most interesting	Should be improved	Should be reduced
Topics	T1 (CAT)	0	0	0
	T2 (TMS)	7	0	0
	T3 (MT+MTPE)	3	1	0
	T4 (TQA)	2	3	0
	T5 (PTW)	3	0	0
	T6 (SW LION)	2	3	0
	T7 (LDM)	1	1	3
	none	0	10	15
Collaborative / Active Learning Activities	Roleplaying (STB)	7	8	0
	CS #1 (MT)	5	3	0
	CS #2 (SW LION)	4	5	2
	CS #3 (LDM)	2	2	2
	none	0	0	14

* For each of the three requests (i.e. interest, improvement, reduction), respondents were only allowed to indicate one preference among the topics and one preference among the collaborative activities.

Fig. 8 Student assessment on topics and collaborative/active learning activities (question Q7)

implementations have been described so far in the literature, it has been shown how CAT-DeM can support multiple STB formats but the shorter format experimented with the study group (see Table 6) was preferred by the students and was more manageable by the CAT teacher involved in the case study. As for CSCL activities, their role in CATDeM is to be incorporated in the CAT course design so that they can be aligned with the specific learning objectives outlined by the EMT. Finally, the GLAID framework was the gluing element used to guide the instructional design of CATDeM so that any CAT course designed according to is pedagogically sound. The role of GLAID is particularly relevant as it also allows to shape rigorously the role-playing experiences and the discussion case studies.

6.1 Answering RQ1: An Ecosystem of Tools

Section 3 analysed the suitability of current technological enablers for CAT courses and supported the subsequent definition of the course design guidelines. Therefore, it is now possible to answer to RQ1. The six categories of tools discussed in Sect. 3 represent a set of competences that every CAT educator should foster in students, according to a progressive level of learning/usage difficulty that depends on the following criteria:

- i. Relevance to student training,
- ii. Estimated IT expertise required to students (both to learn and then to use the system),
- iii. Importance for the professional career of a scientific translator.

As summarised in Table 12, systems belonging to the TMS group should definitely be placed first, because of their importance for translators' training and future career. Immediately afterwards, students should learn why and how to use TQMS, in order to understand that simply delivering a translated product is not enough and that proper quality checks need to be carried out according to the specified requirements. The relatively low ICT expertise required to use TQMS would not be a significant burden. Subsequently, MT and MTEs can be introduced, as they also benefit from quality assurance tools. SLMS and CLT platforms should be introduced later, towards the end of the CAT course, due to the high level of ICT expertise required and in order to capitalise on the technical/technological knowledge that students have already achieved in dealing with the previous tool categories. In this way, such tools will just represent a further level of complexity rather than a height too difficult to climb. Finally, the feasibility of including also TBMS should be considered carefully, as they are primarily focused on business management and human resource management rather than translation. Therefore, they should be included in CAT courses only on a case-by-case basis. Table 12 summarises these considerations and also lists some applications for each tool category.

In addition, it is important to make students aware of the functional relationships that exist between these tool categories, as the different applications should not be used separately but jointly, in order to work as an ecosystem truly capable of supporting every stage of the translation workflow. Figure 9 depicts a complete usage pattern including the main application categories, the stakeholders involved, and their typical actions.

6.2 Answering RQ2

RQ2 referred to course design guidelines for computer scientists and researchers involved as CAT instructors, who tend to focus more on technical aspects and risk overlooking important aspects on STT education and competence frameworks. Nevertheless, the ICT-oriented perspective must be fostered in translation studies, since it is very often underestimated, despite a growing use of cloud-based platforms by professional translators and language service providers (LSPs).

Moreover, many European universities offer translation curricula that consist mainly of courses on translation theory, language and linguistics, intercultural communication, interpreting, and literature, while CAT and MT courses usually accounts for no more than 10 out of 120 ECTS and only introduce students to the basics of TMS, leaving little space for other activities such as role-playing and language data management.

Table 12 Categories of software tools supporting translation students and professional translators

Category ¹	Applications ²	ICT expertise required to students ³	Importance for student training	Importance for professional career ⁴	Suggested priority in CAT course ⁵
TMS	<i>Trados Studio, Memsource, memoQ, SmartCat, OmegaT, ...</i>	Medium	High	Very High	I
TBMS	Plunet, XTRF, TranslationProjex, ...	High	Low	Medium	VI (optional)
SLMS	Passolo, <i>Phrasex</i> , Lokalise, Poedit, Pootele, ...	Very high	Medium	Medium	IV
TQMS	LexiQA, Okapi CheckMate, <i>MemsourceTQA</i> , TQAuditor, ...	Low	High	High	II
MTE	AWS Translate, <i>Google Translate, DeepL, Bing MS Translator, ...</i>	Low	High	High	III
CLT	AntConc, CATMA, TextSTAT, Sketch Engine, CorpKit, ...	Medium	Medium	Medium	V

¹Category: TMS: Translation Management System, TBMS: translation Business Management System, SLMS: Software Localisation Management System, TQMS: Translation Quality Management System, MTE: Machine Translation Engine, CLT: Corpus Linguistics Tool

²Applications: See the Appendix for direct URLs. The tools indicated in bold were used in the proposed course (see Sect. 4)

³ICT expertise: Low: students can start using the tool without any specific training

Medium: students require specific training before using the tool proficiently

High/Very High: students require an intensive training on specific ICT topics before becoming capable of using the tool

⁴Career: Low: tools required only for very specific translation-related jobs or to senior professional profiles

Medium: tools required not in the short term or to a subset of translators only

High: tools that are daily useful for any translator's professional career

Very High: tools that are indispensable for any translator's professional career

⁵Priority: The order according to which the various tool categories should be addressed in a CAT course, from I (first) to VI (last)

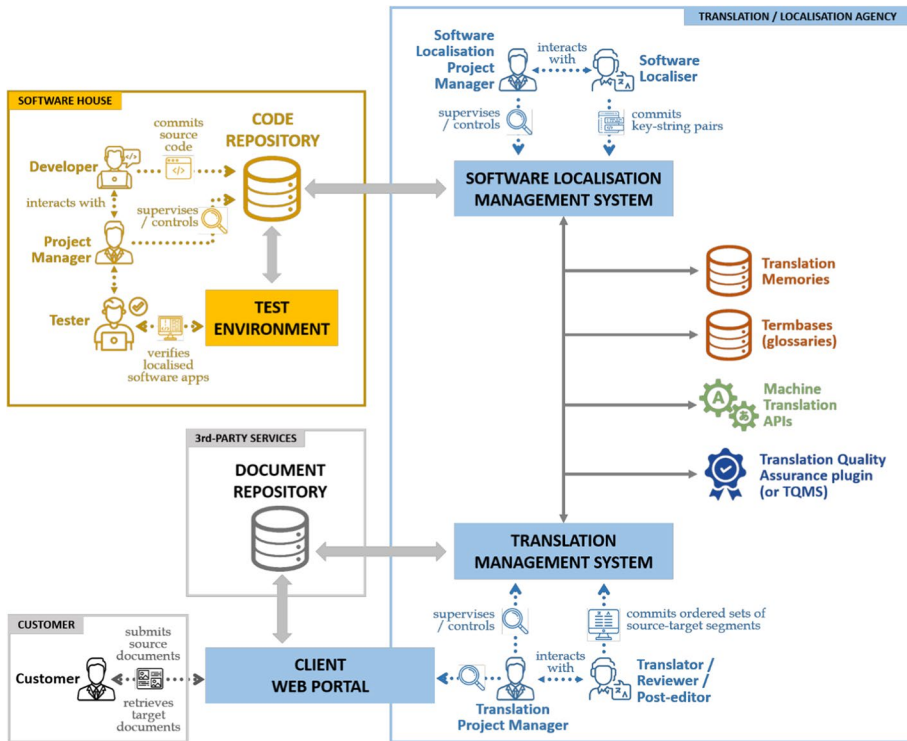


Fig. 9 Typical usage pattern of a computer-assisted translation ecosystem exploiting TMS, SLMS, TQA/TQMS, in combination with MTE/MT APIs (icons are from flaticon.com)

The proposed CATDeM, based on the DMEMO cycle (BPM Resource Center, 2014), sets out a rigorous and replicable course structure, inspired by the continuous improvement approach and encouraging active cooperation between computer and language educators. The learning goals are based on the EMT framework (Toudic & Krause, 2017), thus making the course perfectly suited for this non-STEM educational scenario, while the adopted content structure includes the most recent topics on computer-assisted translation, machine translation, software localisation, and language data management, thus also incorporating a broad range of technological enablers. The teaching approach relies on collaborative and active learning activities, enabled by a groupware platform and based, respectively, on discussion case studies and a role-playing experience. Finally, the role-playing is designed according to the GLAID framework (De Hei et al., 2016), which helps to create a context in which student can practice their future professional career with a high degree of simulation fidelity.

6.3 Answering RQ3 and RQ4

Incorporating collaborative and active learning activities into CAT courses (RQ3) is challenging due to the heterogeneous backgrounds of the educators the abundance of technological resources, and the low level of computer literacy of students. The proposed GLAID

framework (Sect. 4) systematised the design of the role-playing experience, which should represent the core collaborative activity in any CAT course.

The control group and the study group, associated with two consecutive academic years, had comparable course content while the role-playing experience was different, and the discussion in Sect. 5.1 provided interesting insights. In the control group, role-playing was an elective post-course activity and was perceived differently depending on whether students participated or not. Overall, the level of student involvement and satisfaction expressed by participants in that academic year demonstrated the inherent validity of the role-playing approach but its operational window (i.e., one semester) was considered too long, as it overlapped with other curricular courses. The redesign effort carried out with the study group in 2021/2022, thanks to the continual improvement opportunity offered by the DMEMO cycle and always systematised through the GLAID framework, allowed to make the role-playing experience even more attractive for the students. As soon as it became a compulsory activity of the CAT course, the role-play was no longer affected by the students' biased perceptions. Moreover, all the participants experienced a high degree of realism in such important skills as project management and organisational dynamics, which are essential to perform their tasks with an adequate perception of functional fidelity (Davidovitch et al., 2009). The proposed CATDeM, allows integrating collaborative and active learning in the CAT scenario, thus answering to RQ3.

The effectiveness of role-playing (RQ4) can be ascertained with the results reported in the last two rows of Table 6: even though the average score showed a difference of 0.3 out of 30 points (i.e., 28.9 in 2021/2022 and 29.2 in 2020/2021), the number of students who passed the exam after the first two calls was 27.1% higher in the study group. This demonstrates a greater effectiveness when role-playing is compulsory and conducted over a relatively limited period of time (i.e., 3 weeks instead of 6 months). In this way, students were highly engaged by the opportunity to experience a professional work simulation and they did not perceive it as an additional educational burden. Interestingly, there was another indication that duration of the role-play was too long for the control group: in the elective version, participants started to drop out their engagement after the first two months. In the study group, instead, there were no requests to reduce the duration of the experience and some students ($n=5$, 20.8%) asked to extend it by another week.

The other element to quantify reported in RQ4 related to students' perceptions. The outcomes of the questionnaire examined in Sect. 5.3 disclosed several insights in this regard. Firstly, students perceived considerable improvements in their ICT-related competences, especially with the help of the role-playing activity. As for the topics, TMS was obviously perceived as fundamental, but SW-L10N and LDM revealed an interesting result: students perceived them as valuable even if they did not plan to apply for jobs in these subfields. This apparent mismatch suggests that more work is needed to engage students further. A suitable approach would be to design cross-course role-playing activities specifically devoted to software localisation, where translation students play the role of software localisers and computer engineering students recruited from a M.Sc. course could play the role of software developers. From an educational perspective, this would achieve two complementary aims. On the one hand, the computer engineering students would be introduced to localisation-oriented programming practices (which are very often excluded from software engineering and application development courses). On the other hand, the translation students would benefit from a real software development context and, above all, they could interact with their peers. From an organisational perspective, the activity could present a certain degree of complexity and, therefore, its practical feasibility is currently being investigated for the next academic year. An even easier improvement could be

offered by slightly modifying the case study related to LDM, by inviting experts and company representatives to involve students in seminars dealing with implementation examples of language data management.

The role-playing experience was perceived as the most effective activity, so successful that some students even asked for it to be extended, and its usefulness for both educational and professional paths was clearly (and rightly) recognised by the students. The perception of how different it can be to learn and use a software application in this field was also interestingly documented by the results of the questionnaire. In many cases, even the software applications with the most difficult learning phase were perceived as not so difficult to use, which also indicates that the educational approach adopted was effective in making the students capable of using these tools.

Overall, the students were very satisfied with the structure of the course: the role-playing experience and the discussion case study on SW-L10N also received many requests for improvement, suggesting that the inclusion of collaborative and learning activities incorporation can be even more beneficial in future iterations of the course.

6.4 Limitations

The CATDeM methodology proposed in this paper was discussed in terms of the case study deployed at the University of Salento and its effectiveness in improving students' perceptions and attitudes towards CAT technologies in STT teaching. However, there are some limitations that will be addressed in a forthcoming study. Firstly, the generalisation of the results obtained can be achieved by replicating the case study in different universities or by including also bachelor students. In this way, it would be possible to diversify and increase the sample size of the post-grading questionnaire presented in this study. With regard to the evaluation and validation procedure, not only student self-assessment post-grading questionnaires could be used, but also pre-/post-course competence surveys and questionnaires for the teachers involved. Finally, a long-term impact assessment could be carried out by conducting a follow-up study to track the performance of the students involved in this research also in the subsequent part of their translation courses or in their first steps in the professional translation sector.

7 Conclusions and Future Work

Compared to the professional STT market, the corresponding educational sector seems to be moving at a slower pace, as technologies and enabling tools are very often considered only from an operational perspective, and overshadowed by traditional teaching focused on language/translation theory and mainly requiring a basic knowledge of translator-supporting tools. Consequently, CAT courses represent a distinct area in higher education curricula, which is significantly undervalued and where basic computer-supported teaching strategies play a secondary role. There is an important domain mismatch between CAT educators, who very often come from the computer science/engineering sector, and the B.A./M.A. programmes to which their courses belong. This gap is widened by the lack of shared course design guidelines addressing for this very specific sector. In order to overcome these limitations, this paper presented a multi-framework CAT course modelling methodology called CATDeM by relying on several design strategies and by incorporating computer-supported collaborative and active learning activities in the form of three

discussion case studies and a role-playing experience. CATDeM was then implemented in a compulsory course belonging to a M.A. degree in scientific-technical translation and interpreting at the University of Salento (Italy). The participating students were asked, after passing the examination, to evaluate in an online survey their improvement on multiple competences, the usefulness of the course topics, the usefulness of the collaborative and active learning activities, the difficulty in learning and using the proposed tools, and their overall perception of the course. The results of the survey (response rate: 75%) showed that the students definitely appreciated the structure of the course and its educational approach, thus confirming that the modelling strategy adopted is appropriate for the considered teaching scenario as it improves students' attitudes towards the different technologies supporting the STT domain. In addition, the proposed role-playing experience achieved greater effectiveness compared to its previous implementation, achieving greater effectiveness. The promising results of this study will be used to further refine the proposed methodology, by investigating additional aspects such as more direct involvement of real LSPs and the cooperation with computer engineering courses that could provide a real software development environment where students could learn how to integrate the translation workflow with the typical software development lifecycle.

Appendix

List of technological enablers mentioned in the text and corresponding URLs (accessed: 18-June-2023).

Technological enabler	URL
AntConc	https://www.laurenceanthony.net/software/antconc/
AWS Translate	https://aws.amazon.com/translate/
Bing MS Translator	https://www.bing.com/translator
CATMA	https://catma.de/
CorpKit	https://interrogator.github.io/corpkit/
DeepL	https://www.deepl.com/translator
EZTitles	https://www.eztitles.com/
Google Translate	https://translate.google.com/
LexiQA	https://lexiqa.net/
Lokalise	https://lokalise.com/
MateSub	https://matesub.com/
memoQ	https://www.memoq.com/
Memsource	https://www.memsource.com/
Microsoft Teams	https://www.microsoft.com/en-us/microsoft-teams/education
Okapi CheckMate	https://okapiframework.org/wiki/index.php/CheckMate
OmegaT	https://omegat.org/
Passolo	https://www.trados.com/products/passolo/
Phrase TMS (former Memsource)	https://phrase.com/
Plunet	https://www.plunet.com/
Poedi	https://poedit.net/
Pootle	https://pootle.translatehouse.org/
Ren'py	https://renpy.org/

Technological enabler	URL
Sketch Engine	https://www.sketchengine.eu/
SmartCat	https://www.smartcat.com/
Subtitle Edit	https://nikse.dk/SubtitleEdit/
TextSTAT	https://neon.niederlandistik.fu-berlin.de/en/textstat/
TQAuditor	https://tqauditor.com/
Trados	https://www.trados.com/products/trados-studio
TranslationProjex	https://translationprojex.com/
VisualSubSync	https://www.visualsubsync.org/home
WinCaps	https://broadstream.com/products/wincaps/
XTRF	https://xtrf.eu/

Author Contributions Conceptualisation, Methodology, Formal analysis and investigation, Writing—original draft preparation, Writing—review and editing, Resources, Data curation, Data visualisation and graphics, Supervision: MZ.

Funding Open access funding provided by Università del Salento within the CRUI-CARE Agreement. No funding was used.

Data Availability This manuscript has no associated data or the data will not be deposited

Declarations

Conflict of interest No competing interests are reported.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akhulkova, J., Hickey, S., & Hynes, R. (2022). *The Nimdzi Language Technology Atlas*. <https://www.nimdzi.com/language-technology-atlas/>
- Alcina, A. (2008). Translation technologies: Scope tools and resources. *Target*, 20(6), 79–102.
- Alkan, S. C. (2016). Use of cloud-based translation management systems in translation education. *Participatory Educational Research*, *spi16*(1), 43–46. <https://doi.org/10.17275/per.16.spi.1.5>
- Álvarez Ariza, J., & Nomesqui Galvis, C. (2023). RaspyControl Lab: A fully open-source and real-time remote laboratory for education in automatic control systems using Raspberry Pi and Python. *HardwareX*, 13, e00396. <https://doi.org/10.1016/j.ohx.2023.e00396>
- Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2), 141–161.
- Arenas, A. G., & Moorkens, J. (2019). Machine translation and post-editing training as part of a master's programme. *Journal of Specialised Translation*, 31, 217–238.
- Bacon, J. (2002). Scientific translation and interpreting: Bright career prospects. *Science*. <https://www.science.org/content/article/scientific-translation-and-interpreting-bright-career-prospects>

- Banerjee, S., & Lavie, A. (2005). METEOR: An automatic metric for mt evaluation with improved correlation with human judgments. In *Intrinsic and extrinsic evaluation measures for machine translation and/or summarization, proceedings of the workshop ACL 2005*, 65–72
- Banerjee, S., & Khapra, M. M. (2019). Graph convolutional network with sequential attention for goal-oriented dialogue systems. *Transactions of the Association for Computational Linguistics*, 7, 485–500. https://doi.org/10.1162/tacl_a_00284
- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences*, 9(4), 403–436.
- Bowker, L. (2002). *Computer-aided translation technology: A practical introduction*. University of Ottawa Press.
- Bowker, L. (2015). Computer-aided translation: Translator training. In S.-W. Chan (Ed.), *Routledge encyclopedia of translation technology* (pp. 88–104). Routledge.
- Bowker, L., & Marshman, E. (2010). Toward a model of active and situated learning in the teaching of computer-aided translation: Introducing the CERTT project. *Journal of Translation Studies*, 13(1–2), 199–226.
- BPM Resource Center. (2014). *Understanding BPM and related improvement methodologies*. http://www.what-is-bpm.com/get_started/bpm_methodology.html
- Brolpito, A. (2018). *Digital skills and competence, and digital and online learning*. https://www.etf.europa.eu/sites/default/files/2018-10/DSC_and_DOL_0.pdf
- Buyschaert, J., Fernandez-Parra, M., & van Egdome, G. W. (2017). Professionalising the curriculum and increasing employability through experiential learning: The cases of INSTB. *Current Trends in Translation Teaching and Learning E (CTTLE)*, 4, 78–111.
- Chilukuri, K. C. (2020). A novel framework for active learning in engineering education mapped to course outcomes. *Procedia Computer Science*, 172, 28–33. <https://doi.org/10.1016/j.procs.2020.05.004>
- Clark, R., Rothwell, A., & Shuttleworth, M. (2002). *Integrating language technology into a postgraduate translation programme*.
- Corpus-Analysis.com. (2022). *Tools for corpus linguistics*. <https://corpus-analysis.com/>
- Davidovitch, L., Parush, A., & Shtub, A. (2009). The impact of functional fidelity in simulator-based learning of project management. *International Journal of Engineering Education*, 25(2), 333–340.
- De Hei, M. S., Sjoer, E., Admiraal, W., & Strijbos, J. W. (2016). Teacher educators' design and implementation of group learning activities. *Educational Studies*, 42(4), 394–409.
- Delpech, E. M. (2014). *Comparable corpora and computer-assisted translation* (1st ed.). Wiley-ISTE.
- DePalma, D. A. (2021). *The language sector in eight charts*. CSA Research. <https://csa-research.com/Blogs-Events/Blog/ArticleID/785/The-Language-Sector-in-Eight-Charts>
- Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL* (pp. 61–91). Open Universiteit Nederland.
- Dillenbourg, P. (2013). Design for classroom orchestration. *Computers and Education*, 69, 485–492. <https://doi.org/10.1016/j.compedu.2013.04.013>
- Doherty, S. (2016). The impact of translation technologies on the process and product of translation. *International Journal of Communication*, 10, 947–969.
- Eetemadi, S., Lewis, W., Toutanova, K., & Radha, H. (2015). Survey of data-selection methods in statistical machine translation. *Machine Translation*, 29(3–4), 189–223. <https://doi.org/10.1007/s10590-015-9176-1>
- ELIS Research. (2022). *ELIS 2022. European language industry survey 2022. Trends, expectations and concerns of the European language industry*. <https://commission.europa.eu/system/files/2022-03/elis2022-report.pdf>
- Ellis, C., and Wainer, J. (2004). Groupware and Computer Supported Cooperative Work. In *Multigent systems: A modern approach to distributed artificial intelligence* (pp. 425–455). The MIT Press.
- ELRC Consortium. (2019). *Sustainable language data sharing to support language equality in multilingual Europe: Why language data matters* (ELRC White Paper).
- European Union. (2022). *Complete guide to GDPR compliance*. <https://gdpr.eu/>
- Eurostat. (2022). *Being young in Europe today: digital world*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Being_young_in_Europe_today_-_digital_world&oldid=566905
- Evergreen, S. D. H. (2019). *Effective data visualization: The right chart for the right data* (2nd ed.). Sage Publications Inc.
- Feder, M. (2003). Machine-assisted human translation: Its position. *Perspectives: Studies in Translation Theory and Practice*, 11(2), 135–143. <https://doi.org/10.1080/0907676X.2003.9961469>

- Flores, M. A., Veiga Simao, A. M., Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: A study in higher education. *Studies in Higher Education*, 40(9), 1523–1534.
- Garcia, I. (2015). Computer-aided translation systems. In S.-W. Chan (Ed.), *Routledge encyclopedia of translation technology* (1st ed., pp. 68–87). Routledge.
- Garg, A., & Agarwal, M. (2018). Machine translation: A literature review. *ArXiv*, 17. <https://arxiv.org/pdf/1901.01122.pdf>
- Gartman, D. (2019). *Under What? A downloadable game for Windows*. <https://gartman.itch.io/underwhat>
- Gaspari, F., Almaghout, H., & Doherty, S. (2015). A survey of machine translation competences: Insights for translation technology educators and practitioners. *Perspectives: Studies in Translatology*, 23(3), 333–358.
- Göpferich, S., & Jääskeläinen, R. (2009). Process research into the development of translation competence: Where are we, and where do we need to go? *Across Languages and Cultures*, 10(2), 169–191.
- Gou, J., Sun, L., Yu, B., Wan, S., Ou, W., & Yi, Z. (2023). Multilevel attention-based sample correlations for knowledge distillation. *IEEE Transactions on Industrial Informatics*, 19(5), 7099–7109. <https://doi.org/10.1109/TII.2022.3209672>
- Gries, S. T., and Berez, A. L. (2017). Linguistic annotation in/for corpus linguistics. In *Handbook of Linguistic Annotation* (pp. 379–409). Springer Netherlands. https://doi.org/10.1007/978-94-024-0881-2_15
- Guo, Z., Zhang, Y., Teng, Z., & Lu, W. (2019). Densely connected graph convolutional networks for graph-to-sequence learning. *Transactions of the Association for Computational Linguistics*, 7, 297–312. https://doi.org/10.1162/tacl_a_00269
- Hariri, W. (2023). Unlocking the potential of ChatGPT: A comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing. *Technology*, 15(2). <http://arxiv.org/abs/2304.02017>
- Hernández-Sellés, N., Muñoz-Carril, P. C., & González-Sanmamed, M. (2020). Interaction in computer supported collaborative learning: An analysis of the implementation phase. *International Journal of Educational Technology in Higher Education*, 17(1), 13. <https://doi.org/10.1186/s41239-020-00202-5>
- Hickey, S. (2023a). *Language industry services: Market split by service and segment leaders in 2023a*. <https://www.nimdzi.com/language-industry-services-market-split-by-service-and-segment-leaders-in-2023a/>
- Hickey, S. (2023b). *Language industry verticals: Market split in percentage and segment leaders in 2023b*. <https://www.nimdzi.com/language-industry-verticals-market-size-by-segment-leaders/>
- Hmelo-Silver, C. E., & Jeong, H. (2021). Benefits and challenges of interdisciplinarity in CSCL research: A view from the literature. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2020.579986>
- Holtz, Y. (2023). *Ridgeline plot*. From Data to Viz. <https://www.data-to-viz.com/graph/ridgeline.html>
- ICNIRP. (2022). *5G Radiofrequency—RF EMF*. <https://www.icnirp.org/en/applications/5g/index.html>
- Isabelle, P. (1993). Machine-aided human translation and the paradigm shift. In *Proceedings of the fourth machine translation summit, Kobe, Japan*, 177–179. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Machine-Aided+Human+Translation+and+the+Paradigm+Shift#0>
- ISO. (2015). *ISO 17100:2015. Translation Services: Service Requirements*.
- Ivanova, O. (2016). Translation and ICT competence in the globalized world. *Procedia: Social and Behavioral Sciences*, 231, 129–134.
- Jooste, W., Haque, R., & Way, A. (2022). Knowledge distillation: A method for making neural machine translation more efficient. *Information*, 13(2), 88. <https://doi.org/10.3390/info13020088>
- Kasaklijev, N., Somova, E., & Gocheva, M. (2019). Green mobile application development through software localization. *International Journal on Information Technologies and Security*, 11(4), 3–16.
- Killman, J. (2018). A context-based approach to introducing translation memory in translator training. In G. Conception (Ed.), *Translation, globalization and translocation* (pp. 137–159). Palgrave Macmillan.
- Koponen, M., Sulubacak, U., Vitikainen, K., & Tiedemann, J. (2020). MT for subtitling: Investigating professional translators' user experience and feedback. In *Proceedings of the 14th conference of the association for machine translation in the Americas, 2014*, 79–92. <https://aclanthology.org/2020.amta-research.pdf>
- Krüger, R. (2016). Contextualising computer-assisted translation tools and modelling their usability. *Trans-Kom*, 9(1), 114–148.
- Krüger, R. (2021). An online repository of python resources for teaching machine translation to translation students E. *Current Trends in Translation Teaching and Learning*, 8, 4–30. <https://doi.org/10.51287/cttle20212>
- Lenker, M., Anastasiou, D., & Buckley, J. (2000). Workflow specification for enterprise localisation. *The International Journal of Localisation*, 9(1), 10.

- Li, Y., & Rafiei, D. (2018). Natural Language Data Management. In *Natural language data management and interfaces* (pp. 15–46). Springer International Publishing. https://doi.org/10.1007/978-3-031-01862-6_3
- Lipponen, L., Hakkarainen, K., & Paavola, S. (2004). Practices and orientations of CSCL. In J. W. Strijbos, P. A. Kirschner, & R. L. Martens (Eds.), *What we know about CSCL* (pp. 31–50). Kluwer Academic Publishers.
- Lommel, A. (2018). *Metrics for translation quality assessment: A case for standardising error typologies* (pp. 109–127). https://doi.org/10.1007/978-3-319-91241-7_6
- Loock, R., Kerremans, K., Vandepitte, S., & Van Egdom, G.-W. (2017). Teaming up in translator education: simulated translation bureaux as a means of increasing professionalism and employability. *BAAHE 2017: Let's Inter-Act! innovative teaching practices in english studies*.
- Lu, J., Liu, H., Jia, R., Wang, J., Sun, L., & Wan, S. (2022). Towards personalized federated learning via group collaboration in IIoT. *IEEE Transactions on Industrial Informatics*. <https://doi.org/10.1109/TII.2022.3223234>
- Ludvigsen, S. R., and Mørch, A. I. (2010). Computer-supported collaborative learning: Basic concepts, multiple perspectives, and emerging trends. In *International encyclopedia of education* (pp. 290–296). Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.00493-0>
- Ma, X., Liu, J., Liang, J., & Fan, C. (2020). An empirical study on the effect of group awareness in CSCL environments. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2020.1758730>
- Malenova, E. (2019). Cloud technologies in a translation classroom. *Trans-Kom*, 12(1), 76–89.
- Mangiron, C. (2018). Game on! burning issues in game localisation. *Journal of Audiovisual Translation*, 1(1), 122–138. <https://doi.org/10.47476/jat.v1i1.48>
- Mariana, V., Cox, T., & Melby, A. (2015). The multidimensional quality metrics (MQM) framework: A new framework for translation quality assessment. *Journal of Specialised Translation*, 23, 137–161.
- Martin, R. C. (2014). *Agile software development, principles, patterns, and practices* (1st ed.). Pearson Education Limited.
- Mattern, E. (2022). The linguistic data life cycle, sustainability of data, and principles of solid data management. In *The open handbook of linguistic data management* (pp. 61–72). The MIT Press. <https://doi.org/10.7551/mitpress/12200.003.0009>
- Maxwell Chandler, H., & O'Malley Deming, S. (2011). *The game localization handbook*. <http://books.google.com/books?hl=fr&lr=&id=e2y1dmqbWgUC&pgis=1>
- Meijer, H., Hoekstra, R., Brouwer, J., & Strijbos, J. W. (2020). Unfolding collaborative learning assessment literacy: A reflection on current assessment methods in higher education. *Assessment and Evaluation in Higher Education*, 45(8), 1222–1240.
- Mellinger, C. D. (2018). Problem-based learning in computer-assisted translation pedagogy. *HERMES: Journal of Language and Communication in Business*, 57, 195–208. <https://doi.org/10.7146/hjclb.v0i57.106205>
- Memsources. (2021). *Quality Assurance (QA) Overview*. Memsources help center. <https://help.memsources.com/hc/en-us/articles/360013675800-Quality-Assurance-QA-Overview>
- Microsoft. (2022). *Quick Start localization style guides*. <https://docs.microsoft.com/en-us/globalization/localization/styleguides>
- Miyata, R., and Fujita, A. (2016). Understanding pre-editing for black-box neural machine translation. In *Proceedings of the 16th conference of the european chapter of the association for computational linguistics*, 1539–1550.
- Montero Martínez, S., & Benítez, P. F. (2009). Terminological competence in translation. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 15(1), 88–104. <https://doi.org/10.1075/term.15.1.05mon>
- Moorkens, J., & Lewis, D. (2019). Copyright and the re-use of translation as data. In *The Routledge Handbook of Translation and Technology* (pp. 469–481). <https://doi.org/10.4324/9781315311258-28>
- Mordor Intelligence. (2023). *Language services market size & share analysis: Growth trends & forecasts (2023–2028)*. <https://www.mordorintelligence.com/industry-reports/language-services-market>
- Motiejūnienė, J., & Kasperavičienė, R. (2019). Translation project management: An ultimate skill for translators? *Current Trends in Translation Teaching and Learning E*, 6, 160–192.
- Nguyen, X.-P., Joty, S., Kui, W., & Aw, A. T. (2022). Refining low-resource unsupervised translation by language disentanglement of multilingual model. *Advances in Neural Information Processing Systems*, 35, 36230–36242.
- O'Hagan, M. (2020). *The routledge handbook of translation and technology*. Routledge.
- Ofcom. (2022). *What is 5G*. <https://www.ofcom.org.uk/phones-telecoms-and-internet/advice-for-consumers/advice/what-is-5g>

- Olohan, M. (2015). *Scientific and Technical Translation* (1st Editio). Routledge. <https://doi.org/10.4324/9781315679600>
- Olohan, M. (2021). Scientific translation. *Routledge Encyclopedia of Translation Studies*. <https://doi.org/10.4324/9781315678627-108>
- Omar, A., Khafaga, A. F., & Shaalan, I. E. N. A. W. (2020). The impact of translation software on improving the performance of translation majors: A case study of the Saudi universities. *International Journal of Advanced Computer Science and Applications*, 11(4), 287–292. <https://doi.org/10.14569/IJACSA.2020.0110439>
- PACTE group, Hurtado Albir, A., Galán-Mañas, A., Kuznik, A., Olalla-Soler, C., Rodríguez-Inés, P., & Romero, L. (2018). Competence levels in translation: Working towards a European framework. *The Interpreter and Translator Trainer*, 12(2), 111–131. <https://doi.org/10.1080/1750399X.2018.1466093>
- Papineni, K., Roukos, S., Ward, T., and Zhu, W. -J. (2002). Bleu: A method for automatic evaluation of machine translation. In *40th annual meeting of the association for computational linguistics*, 40(July), 311–318.
- Passban, P., Roosta, T., Gupta, R., Chadha, A., and Chung, C. (2022). Training mixed-domain translation models via federated learning. In *Proceedings of the 2022 conference of the north american chapter of the association for computational linguistics: Human language technologies*, 2576–2586. <https://doi.org/10.18653/v1/2022.naacl-main.186>
- Phrase. (2022). *Memsorce and phrase announce a new joint identity and a new localization Suite*. Slat—Press Releases. <https://slator.com/memsource-and-phrase-announce-a-new-joint-identity-and-a-new-localization-suite/>
- Pinto, G., Ferreira, C., Souza, C., Steinmacher, I., and Meirelles, P. (2019). Training software engineers using open-source software: The students' perspective. *Proceedings: 2019 IEEE/ACM 41st international conference on software engineering: Software engineering education and training, ICSE-SEET 2019*, 147–157. <https://doi.org/10.1109/ICSE-SEET.2019.00024>
- Quamar, A., Efthymiou, V., Lei, C., & Özcan, F. (2022). Natural language interfaces to data. *Foundations and Trends in Databases*, 11(4), 319–414. <https://doi.org/10.1561/19000000078>
- Raad, B. (2020). The role of machine translation in language learning. *International Journal of Academic Research*, 7(June), 60–67.
- Rico, C. (2017). The ePortfolio: Constructing learning in translation technology. *The Interpreter and Translator Trainer*, 11(1), 79–95. <https://doi.org/10.1080/1750399X.2017.1306995>
- Rienties, B., Giesbers, B., Tempelaar, D., Lygo-Baker, S., Segers, M., & Gijsselaers, W. (2012). The role of scaffolding and motivation in CSCL. *Computers and Education*, 59(3), 893–906. <https://doi.org/10.1016/j.compedu.2012.04.010>
- Riktors, M. (2018). Impact of corpora quality on neural machine translation. *Frontiers in Artificial Intelligence and Applications*, 307, 126–133. <https://doi.org/10.3233/978-1-61499-912-6-126>
- Risku, H. (2002). Situatedness in translation studies. *Cognitive Systems Research*, 3(3), 523–533. [https://doi.org/10.1016/S1389-0417\(02\)00055-4](https://doi.org/10.1016/S1389-0417(02)00055-4)
- Roosta, T., Passban, P., and Chadha, A. (2021). *Communication-efficient federated learning for neural machine translation*. <http://arxiv.org/abs/2112.06135>
- Rothwell, A., & Svoboda, T. (2019). Tracking translator training in tools and technologies: Findings of the EMT survey 2017. *The Journal of Specialised Translation*, 32, 26–60.
- Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and open-AI models: A preliminary review. *Future Internet*, 15(6), 24. <https://doi.org/10.3390/fi15060192>
- RWS Group. (2021). *Translation management system user guide*.
- Scheibengraf, M. (2023). *Technical translation: Definition, importance, and challenges*. phrase: Terminology management. <https://phrase.com/blog/posts/technical-translation-definition-importance-and-challenges/>
- Schroeder, C. (2018). Education in localization: How language service providers benefit from educational partnerships. *English Studies at NBU*, 4(2), 117–130. <https://doi.org/10.33919/esnbu.18.2.3>
- Seinen, W., & Van Der Meer, J. (2020). *Who owns my language data? Realities, Rules and Recommendations*.
- Sharma, A., & Bhatia, M. (2014). Quality assurance in localization. *International Journal of Information and Computation Technology*, 4(9), 891–896.
- Shi, Y., Shi, C., and Zhou, Z. (2019). Error types of machine translation of popular science text. In *Proceedings of the 2019 international conference on artificial intelligence and advanced manufacturing: AIAM 2019*, 1–4. <https://doi.org/10.1145/3358331.3358366>
- Shin, Y., Kim, D., & Song, D. (2020). Types and timing of scaffolding to promote meaningful peer interaction and increase learning performance in computer-supported collaborative learning

- environments. *Journal of Educational Computing Research*, 58(3), 640–661. <https://doi.org/10.1177/0735633119877134>
- Shuttleworth, M. (2015). Translation management systems. In C. Sinai (Ed.), *Routledge encyclopedia of translation technology* (p. 14). Routledge.
- Shuttleworth, M. (2017). Cutting teeth on translation technology: How students at University College London are being trained to become tomorrow's translators. *Tradução Em Revista*, 22, 18–38.
- Slator. (2023). *Slator 2023 language industry market report*. <https://slator.com/2023-language-industry-market-report/>
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L., and Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. *AMTA 2006: Proceedings of the 7th conference of the association for machine translation of the Americas: Visions for the future of machine translation*, 223–231.
- Sobral, S. R. (2021). Project based learning with peer assessment in an introductory programming course. *International Journal of Information and Education Technology*, 11(7), 337–341. <https://doi.org/10.18178/ijiet.2021.11.7.1532>
- Srinivasa, K. G., Deka, G. C., & Krishnaraj, P.M. (2021). A multi-step process towards integrating free and open source software in engineering education. In *Research anthology on usage and development of open source software* (pp. 389–397). IGI Global. <https://doi.org/10.4018/978-1-7998-9158-1.ch021>
- Strijbos, J. W. (2011). Assessment of (computer-supported) collaborative learning. *IEEE Transactions on Learning Technologies*, 4(1), 59–73. <https://doi.org/10.1109/TLT.2010.37>
- Strijbos, J. W. (2016). Assessment of collaborative learning. In G. T. L. Brown & L. Harris (Eds.), *Handbook of social and human conditions in assessment* (pp. 302–318). Routledge.
- Stuart, K., Botella, A., and Ferri, I. (2016). A Corpus-Driven Approach to Sentiment Analysis of Patient Narratives. In A. Moreno Ortiz and C. Pérez-Hernandez (Eds.), *8th international conference on corpus linguistics (CILC2016)* (pp. 381–395). EPiC Series in Language and Linguistics.
- Tarasenko, R. O., Amelina, S. M., & Azaryan, A. A. (2019). Features of the use of cloud-based translation systems in the process of forming information competence of translators. *CEUR Workshop Proceedings*, 2433, 322–335.
- TAUS. (2021). *Data Marketplace: For language data monetization and acquisition*. <https://datamarketplace.taus.net/>
- Tawmo, Adhikary, P. K., Dadure, P., & Pakray, P. (2022). An empirical analysis on abstractive text summarization. In *Computational intelligence in communications and business analytics, 4th international conference, CICBA 2022*, 280–287. https://doi.org/10.1007/978-3-031-10766-5_22
- Tehseen, I., Tahir, G. R., Shakeel, K., and Ali, M. (2018). Corpus based machine translation for scientific text. In *IFIP advances in information and communication technology* (Vol. 519, pp. 196–206). https://doi.org/10.1007/978-3-319-92007-8_17
- Toledo Báez, M. C. (2018). Machine translation and post-editing: impact of training and directionality on quality and productivity. *Tradumática: Tecnologías De La Traducción*, 16, 24. <https://doi.org/10.5565/rev/tradumatica.215>
- Toudic, D., and Krause, A. (2017). *European master's in translation competence framework 2017*. https://ec.europa.eu/info/sites/info/files/emt_competence_fw_k_2017_en_web.pdf
- UK Government. (2021). *5G programme findings*. <https://www.gov.uk/guidance/5g-programme-findings>
- Vagias, W. M. (2006). *Likert-type scale response anchors*.
- Vardaro, J., Schaeffer, M., & Hansen-Schirra, S. (2019). Translation quality and error recognition in professional neural machine translation post-editing. *Informatics*, 6(3), 41. <https://doi.org/10.3390/informatics6030041>
- Vela-Valido, J. (2021). Translation quality management in the AI Age. New technologies to perform translation quality assurance operations. *Tradumática: Tecnologías De La Traducción*, 19, 93–111. <https://doi.org/10.5565/rev/tradumatica.285>
- Verified Market Research. (2022). *Localization software market research report*.
- Vodă, A. I., Cautisanu, C., Grădinaru, C., Tănăsescu, C., & de Moraes, G. H. S. M. (2022). Exploring digital literacy skills in social sciences and humanities students. *Sustainability*, 14(5), 2483. <https://doi.org/10.3390/su14052483>
- W3C. (2007). *Localization vs internationalization*. <https://www.w3.org/International/questions/qa-i18n/>
- Wang, F., Yan, J., Meng, F., and Zhou, J. (2021). Selective knowledge distillation for neural machine translation. In *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: long papers)*, 6456–6466. <https://doi.org/10.18653/v1/2021.acl-long.504>

- Wang, L., & Wang, X. (2021). Building virtual communities of practice in post-editing training: A mixed-method quasi-experimental study. *Journal of Specialised Translation*, 36(36), 193–219.
- Wang, Q. (2009). Design and evaluation of a collaborative learning environment. *Computers and Education*, 53, 1138–1146.
- Weller, O., Marone, M., Braverman, V., Lawrie, D., and Van Durme, B. (2022). Pretrained Models for Multilingual Federated Learning. In *NAACL 2022–2022 conference of the north american chapter of the association for computational linguistics: Human language technologies, proceedings of the conference*, 1413–1421. <https://doi.org/10.18653/v1/2022.naacl-main.101>
- Wittkowsky, M. (2014). Integrating machine translation (MT) in the higher education of translators and technical writers. *Translating and the Computer*, 36, 100–108.
- Xia, X., Lo, D., Zhu, F., Wang, X., & Zhou, B. (2013). Software internationalization and localization: An industrial experience. In *2013 18th international conference on engineering of complex computer systems*, 222–231. <https://doi.org/10.1109/ICECCS.2013.40>
- Xu, C., & Li, Q. (2021). Machine translation and computer aided english translation. *Journal of Physics: Conference Series*, 1881(4), 042023. <https://doi.org/10.1088/1742-6596/1881/4/042023>
- Yiwen, W. (2017). A novel five-step data mining algorithm. *International Journal of Database Theory and Application*, 10(1), 119–126. <https://doi.org/10.14257/ijdta.2017.10.1.11>
- Yoleri, S., & Nur Anadolu, Z. (2022). Examination of digital literacy skills of undergraduate students according to various variables. *Advanced Education*. <https://doi.org/10.20535/2410-8286.262190>
- Zhang, X., & Vieira, L. N. (2021). CAT teaching practices: An international survey. *The Journal of Specialised Translation*, 36, 99–124.
- Zhang, Y., Zhang, F., Jin, Y., Cen, Y., Voronin, V., & Wan, S. (2023). Local correlation ensemble with GCN based on attention features for cross-domain person re-ID. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 19(2), 1–22. <https://doi.org/10.1145/3542820>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.