The use of translation technologies and generative artificial intelligence by terminology professionals in institutional settings: an exploratory study

Tanja Wissik

Tanja Wissik
Austrian Academy of Sciences;
tanja.wissik@oeaw.ac.at;
ORCID: 0000-0002-1631-4560

Abstract

This contribution explores the use of translation technologies, including machine translation and generative artificial intelligence (GenAI), by terminology professionals in institutional settings through the analysis of qualitative interview data collected in two phases (through synchronous online interviews in phase I and asynchronous email interviews in phase II). The study shows that terminology professionals in institutional settings use a variety of tools for terminology work. Nearly all have experimented with GenAI applications but few have fully integrated them into their workflow so far. Furthermore, the impact of new technologies on terminology workflows, such as the emergence of the new role of technology manager and the need to adapt revision procedures and quality assurance when using GenAI, is also discussed.

Keywords: translation technologies, generative artificial intelligence, terminology professionals, terminology work, institutional settings.

Resumen

Esta contribución explora el uso de tecnologías de la traducción, incluidas la traducción automática y la inteligencia artificial generativa (GenAI), por parte de profesionales de la terminología en entornos institucionales, a través del análisis de datos cualitativos procedentes de entrevistas realizadas en dos fases (entrevistas en línea sincrónicas en la fase I y entrevistas asincrónicas por correo electrónico en la fase II). El estudio muestra que los profesionales de la terminología en entornos institucionales utilizan una amplia variedad de herramientas para el trabajo terminológico. Casi todos han experimentado con aplicaciones de GenAI, pero pocos han integrado plenamente estas aplicaciones en su flujo de trabajo hasta el momento. Asimismo, se analizan el impacto de las nuevas tecnologías en los flujos de trabajo terminológicos, como la aparición del nuevo rol de gestor tecnológico y la necesidad de adaptar los procedimientos de revisión y aseguramiento de la calidad cuando se utiliza GenAI.

Palabras clave: tecnologías de la traducción, inteligencia artificial generativa, profesionales de la terminologia, trabajo terminológico, entornos institucionales.

Rebuda: 26 de abril de 2025| Acceptació: 6 de octubre de 2025 | Publicació: 15 de desembre de 2025

Cita recomanada: Wissik, Tanja (2025). «The use of translation technologies and generative artificial intelligence by terminology professionals in institutional settings: an exploratory study». Revista Tradumàtica. Tecnologies de la Traducció, 23, 165-192. <https://doi.org/10.5565/rev/tradumatica.498>

http://revistes.uab.cat/tradumatica
ISSN 1578-7559
Universitat Autònoma de Barcelona

Resum

Aquesta contribució explora l'ús de tecnologies de la traducció, incloses la traducció automàtica i la intel·ligència artificial generativa (GenAI), per part de professionals de la terminologia en entorns institucionals, a través de l'anàlisi de dades qualitatives procedents d'entrevistes realitzades en dues fases (entrevistes en línia sincròniques en la fase I i entrevistes asincròniques per correu electrònic en la fase II). L'estudi mostra que els professionals de la terminologia en entorns institucionals utilitzen una àmplia varietat d'eines per al treball terminològic. Gairebé tots han experimentat amb aplicacions de GenAI, però pocs han integrat plenament aquestes aplicacions al seu flux de treball fins ara. Tanmateix, s'analitzen l'impacte de les noves tecnologies en els fluxos de treball terminològic, com l'aparició del nou rol de gestor tecnològic i la necessitat d'adaptar els procediments de revisió i assegurament de la qualitat quan s'utilitza GenAI.

Paraules clau: tecnologies de la traducció, intel·ligència artificial generativa, professionals de la terminologia, treball terminològic, entorns institucionals.

1. Introduction

Since the launch of ChatGPT in late 2022, generative artificial intelligence (GenAI) applications have changed the language technology landscape due to their accessibility and innovative capabilities. AI is not new to the field of language technology. Many tools utilised in natural language processing (NLP) build upon machine learning approaches, neural machine translation perhaps being the most well-known example. Language technologies are information technologies specialised in handling human language, designed to process, understand, and even generate human language. Translation technologies are a specific subset of language technologies dealing with aspects of translation in various contexts (Alcina, 2008; Bowker & Corpas Pastor, 2015), such as terminology management systems, computer-aided translation (CAT) tools or machine translation. While GenAI can be seen as related to the area of language technology as it can be used to create, translate or extract information from texts, it can also be extended to other areas, depending on the end purpose of the system (e.g. generating images or code).

Furthermore, the use of language technologies in terminology work is well established, and several tasks in terminology workflows have been supported and facilitated by language technologies for some time.

While there is increasing interest in GenAI in the area of terminology, as shown by several papers on experiments with GenAI and terminology tasks (Reineke, 2023; San Martín, 2024; Heinisch, 2025), little is known about how well these new technologies are already integrated into practical terminology workflows by terminology professionals in institutional settings. Terminology professionals are persons involved in various steps in the terminology workflow. They “can have various roles such as terminology project manager, terminologist, computational linguist, translator or localization expert” (TerminOrgs, 2016). Furthermore, domain experts and IT experts (i.e. those who develop new tools to improve terminology workflows) might be among the professionals involved in these tasks as well (Chiocchetti et al., 2023).

This article aims to explore the impact of new technological trends, including GenAI, on practical terminology work in institutional settings by looking at the use of language technologies by terminology professionals. The exploratory study reported on in this article is based on qualitative interview data.

The article is structured as follows: After the introduction, Section 2 reviews the literature on the use of large language models (LLMs) and GenAI for terminology tasks from the perspective of NLP, while in Section 3 the use of language technologies (including LLMs) for terminology tasks is discussed from the perspective of terminology practice. In Section 4 the research method for the study is described, followed by an overview of the results in Section 5, which are then discussed in Section 6. Section 7 presents a few concluding remarks.

2. LLMs and terminology tasks from the perspective of NLP

The application of AI in research on language technologies (e.g. for machine translation tasks) dates back into the mid-1970s (Hutchins, 2001). But with the blossoming of GenAI and open-sourced LLMs, recent studies have experimented with the applicability of LLMs to terminology tasks that were previously performed (semi-)automatically with rule-based, statistical or machine-learning-based methods, such as term extraction, relation extraction, and term variation detection (see Lefever & Rigouts Terryn, 2024). Giguere et al. (2023) compared a statistical model with an LLM (GPT-4) for a term extraction task and the LLM model produced less noise than the statistical model. Furthermore, Giguere et al. (2023) also compared the two approaches in real projects from language service providers: less human effort was required to clean up and finalise the list of extracted terms produced by the LLM-based approach than the list of terms produced by the statistical model normally embedded into CAT tools. Tran et al. (2023) explored the applicability of open and closed-sourced LLMs in automatic term extraction. Their empirical tests and evaluation showed that prompting GenAI models is an approach that can be used for term extraction tasks, especially when little annotated training data is available. However, the results also showed that while the models had good coverage, their precision was poor. Therefore, Tran et al. (2023) argue that when a complete training dataset is available, a fully-supervised automatic term extraction system would be the optimal choice. Xu et al. (2025) also discuss LLM-based methods for cross-domain terminology extraction, i.e. extraction of terms from texts across different domains, and their potential to adapt to diverse domain-specific and cross-domain requirements. Synonym identification is another task that can be performed automatically. Thießen et al. (2023) tested the identification of scientific synonyms using different LLMs, with GPT-3 achieving the best results. An LLM approach has also been applied to detect and validate terminological neologisms. McCrae (2019) compared pretrained language models for the identification of neological adjective-noun phrases, and the BERT model gave the best results, although the results from the other models were very similar. Based on these findings, McCrae (2019) suggested that the use of pretrained models in general is helpful for the identification of neologisms. Hosseini-Kivanani (2025) used a hybrid approach to neologism validation by combining LLM-based semantic similarity with graph-based contextual verification using resources such as WordNet and Wikipedia. The hybrid approach showed increased precision and recall compared to frequency-based or rule-based techniques. Machine translation can also be applied in terminology work (Caffrey & Valentini, 2019) and the number of studies on LLM-based machine translation models is growing (Kim et al., 2024; Siu, 2024).

Another method relevant to terminology work worth mentioning here is retrieval-augmented generation (RAG). This is a method that enhances the performance of LLMs by combining them with an external information retrieval system (Lewis et al., 2020). This external information system could be Wikipedia, as in the experiments carried out by Lewis et al. (2020), or any other knowledge source, such as an institutional database, a corpus of documents or even a term base. This allows LLMs to access and incorporate information outside of their training data, resulting in more accurate, relevant, and up-to-date responses than results from LLMs that do not use RAG. However, for high-performing RAG systems, high-quality data that is relevant and up-to-date is needed. (Allahyari & Yang, 2023).

3. Language technologies, LLMs and terminology tasks from the perspective of terminology practice

Regarding the use of technology by terminology professionals in institutional settings, an interview-based study on terminology workflows was conducted over ten years ago by Chiocchetti et al. (2014), when GenAI and LLMs were yet to appear on the horizon. There are some publications on case studies regarding the use of technology in specific institutions that perform terminology work. Caffrey & Valentini (2019) reported on the use of machine translation, term extraction and machine learning to predict new concept relations in the context of terminology work carried out in the World Intellectual Property Organization and its public terminology database, WIPO Pearl. Zorrilla-Agut & Fontenelle (2019) reported on automatic term extraction and a recognition module integrated into IATE, the EU terminology management system. They described the method used to implement machine learning for data cleaning and improvement: for example, they used automatic concept disambiguation to identify missing domain labels in terminology entries or to correct inaccurate domain labels. Stefanik (2023) reported on the use of terminology management tools and quality assurance tools within the European Commission's Directorate-General for Translation. Furthermore, there are some case studies on the use of specific technologies, like semantic web technologies for legal terminology (Martín-

Chozas et al., 2023) or the use of automatic term extraction by terminology professionals in institutional settings (Wissik, 2025).

Regarding the use of GenAI and LLMs, some articles discussing experiments on how GenAI applications can be used for practical terminology work have been published. Reineke (2023) tested definition writing, concept system creation and the determination of synonyms with ChatGPT for German. Reineke (2023) states that even results that seem useful should be treated with caution, as the answers to identical queries can differ considerably from one another, and the resources provided are often hallucinated. Łukasik (2023) performed experiments with English and Polish for tasks like domain identification, definition writing, term extraction and concept system creation. In his experiments, GPT-4 performed better in English than in Polish. Varga (2024) performed practical terminology tasks in Spanish and in Romanian and the results were mixed: in the identification of equivalents, Romanian equivalents for Spanish terms were correct, whereas Spanish equivalents for Romanian terms contained errors. Sánchez-Gijón & Palenzuela-Badiola (2023) did some experiments with terminology extraction and equivalent identification using GPT-3 and GPT-4. While both versions of ChatGPT performed well in extracting terminology, GPT-3 failed at providing correct equivalents. GPT-4 provided better results, although with some small errors (Sánchez-Gijón & Palenzuela-Badiola, 2023).

Furthermore, Massion (2024) discussed how LLMs can be used to identify cross-linguistic equivalents in order to facilitate multilingual terminology management. San Martín (2024), meanwhile, discussed how GenAI can be used for definition writing. He also introduced the term "AI-assisted terminography".

Heinisch (2025) experimented with LLMs for multilingual terminology tasks (including term extraction, definition creation and assessment of concept relations) in system-bound domains, such as higher education systems in different language variations (Austrian German and British English). Heinisch (2025) argues that LLMs do not always enhance the productivity of terminologists as it can be time-consuming to create or refine prompts, evaluate the output and verify the sources.

Łukasik (2023), Reineke (2023) and San Martín (2024) also discussed the need for the integration of AI into existing terminology tools, or the creation of tailored AI tools for terminology work, since all experiments were done with general chatbot interfaces, which have limitations for terminology work (San Martín, 2024).

On the subject of the integration of AI into existing tools, the integration of LLMs in corpus analysis tools (Anthony, 2023) is worthy of mention, since corpora are a common resource used in terminology work. For example, in the corpus management software AntConc (Anthony, 2024a), there is an integrated function (ChatAI) that gives the user access to LLMs via an API through a chat-like interface. There, the user can interact with the LLM directly, or supply it with results generated by other AntConc tools (e.g. concordances or lists of n-grams) and ask it to utilise that information in its responses (Anthony, 2024b).

There are also articles on how GenAI can be integrated into terminology instruction (Heinisch, 2024), but there are no recent overviews of how new technologies, including

GenAI and LLMs, are used by terminology professionals, particularly in institutional contexts, and how GenAI and LLMs are integrated into existing workflows.

Given the lack of a systematic overview in this area, an interview-based study was carried out to gather information and reveal insights.

4. Method

The study described in this contribution is part of a project on terminology workflows in the 21st century (Wissik, 2024) and was carried out in two parts. Part I was completed through qualitative expert interviews conducted online via Zoom in 2023, while part II was completed via email interviews in 2025. The remainder of this section describes the sampling of the data and how it was collected and analysed. Since some aspects of the project have already been published, namely the impact of term extraction (Wissik, 2025) and sustainability aspects in terminology work (Wissik, 2024), reference will be made to these publications when discussing the methodology of part I.

4.1 Part I: synchronous online interviews

For part I, synchronous online expert interviews were conducted (Meuser & Nagel, 1991: 443): 15 terminology professionals working in academic and non-academic institutions were interviewed.[1] For the purposes of this study, terminology professionals are defined as individuals involved in various steps in the terminology workflow, covering roles like terminology coordinators, terminologists, domain experts and IT experts (TerminOrgs, 2016; Chiocchetti et al., 2023). A role not encompassed by previous definitions, but included for this study, is technology manager. A technology manager oversees and coordinates the technology strategy, as well as the implementation and maintenance of the technical infrastructure related to terminology work, and follows new technological developments in research and the industry. A technology manager also functions as a link between developers and users. The selection criteria for an institution's terminology professionals to qualify for participation in the study were a) that practical terminology work be carried out at the institution and b) that a publicly available terminology database (or dictionary) be maintained by the institution. The experts interviewed from these institutions had to a) be involved in the terminology workflow on a regular basis and b) cover diverse roles within the workflow (Wissik, 2025). An initial list of institutions was compiled using directories of terminology networks and terminology associations. Furthermore, international organisations, such as certain specialised agencies of the United Nations, were added to the list. The terminology experts at these institutions were then contacted. Out of the full list of experts contacted, 15 agreed to participate in the study[2] (Wissik, 2025).

[1] Terminology professionals working in the private sector were not included in this study.
[2] Some contacts also suggested other terminology professionals could have qualified for participation in the study.

The sample is rather small, which can be seen as a limitation of this study. However, the sample covers the most current approaches to terminology work, the most common scenarios where terminology work is performed in institutional settings, and different types of terminology professionals (Wissik, 2025). Furthermore, more interviews do not necessarily mean more data, as shown by Guest et al. (2006), where 12 qualitative interviews sufficed to achieve data saturation.

The individual interviewee profiles are described in Table 1.

Interview number	Role	Type of institution	Type of terminology work done in the institution	Less-resourced language(s) involved	Size of public database in entries[3]
INT 1	Developer / IT Expert	Academic / Research	systematic terminology work / ad hoc terminology work / preparatory work for standardisation	Yes	100,000 – 400,000 entries
INT 2	Technology Manager	Academic / Research	systematic terminology work / ad hoc terminology work / preparatory work for standardisation	Yes	100,000 – 400,000 entries
INT 3	Head of Terminology Unit / Terminologist	Academic / Research	systematic terminology work / ad hoc terminology work	Yes	40,000 – 100,000 entries
INT 4	Member of Terminology Committee	Academic / Research	systematic terminology work / ad hoc terminology work	Yes	100,000 – 400,000 entries
INT 5	Head of Terminology and Legal Translation Unit, Deputy Director for Development	Administration (national level)	translation-oriented terminology work / a posteriori terminology work / preparatory work for standardisation / coordination of terminology between the EU and the member state	Yes	> 400,000 entries
INT 6	Terminologist	Administration (regional level)	preparatory work for standardisation / ad hoc terminology work / systematic terminology work	No	< 10,000 entries

[3] Size of terminological database or dictionary database.

INT 7	Terminologist	Academic / Research	systematic terminology work / ad hoc terminology work	Yes	40,000 – 100,000 entries
INT 8	Head of Project Management Unit / Terminologist	Administration (regional level)	systematic terminology work / ad hoc terminology work	Yes	> 400,000 entries
INT 9	Terminologist	Academic / Research	systematic terminology work / ad hoc terminology work / preparatory work for standardisation	Yes	100,000 – 400,000 entries
INT 10	Terminology Coordinator / Terminology Manager	European Institution	translation-oriented terminology work / text-based terminology work	Yes	> 400,000 entries
INT 11	Head of Terminology Unit / Terminologist	International Organisation	translation-oriented terminology work / systematic terminology work (project-based) / ad hoc terminology work	No	10,000 – 40,000 entries
INT 12	Technology Manager	European Institution	translation-oriented terminology work / text-based terminology work	Yes	> 400,000 entries
INT 13	Head of Terminology Unit / Terminologist	International Organisation	translation-oriented terminology work / systematic terminology work (project-based)	No	40,000 – 100,000 entries
INT 14	Terminology Coordinator / Terminology Manager	Academic / Research	systematic terminology work	Yes	100,000 – 400,000 entries
INT 15	Head of Terminology Unit / Terminologist	Administration (regional level)	text-based terminology work / a posteriori terminology work / sometimes proactive terminology work / ad hoc terminology work	No	10,000 – 40,000 entries

Table 1: Individual profiles (Wissik, 2025)

The 15 semi-structured interviews were conducted online via Zoom, most of them in English (two in German). The questions for the interview protocol[4] included general questions about the institution and the background and role of the interviewee, as well as several questions regarding the use of technology in general and in terminology workflows in particular (see also Wissik, 2025).

The interview data was recorded, transcribed and anonymised by removing names of institutions and departments, terminology databases, in-house tools and (research) projects.[5] Furthermore, names of specific languages were removed, because the mention of a particular language could identify an interviewee or their home institution. Rather than naming specific languages, they were categorised as well-resourced or less-resourced languages (for more details see Wissik, 2025). Thematic qualitative text analysis was chosen to analyse the data (Kuckartz, 2014). The categories for encoding the data were developed through a combination of inductive and deductive approaches. Encoding was done manually with the CATMA software (Gius et al., 2023).

4.2 Part II: asynchronous email interviews

When the expert interview study was being prepared in late 2022 and early 2023, the first version of ChatGPT had launched relatively recently (end of November 2022), and so the topic was fairly new and its impact on the field had yet to be fully understood. Therefore, specific questions regarding GenAI were not included. Part II of the study was conducted to fill this gap. For part II, the same sample as for part I was used. Therefore, the 15 participating terminology professionals (Table 1) were invited to take part in a follow-up study on the use of GenAI in practical terminology work via asynchronous email interviews (Ratislavová & Jakub Ratislav, 2014; Dahlin, 2021). In Translation Studies, email interviews are very new (Künzli & Gile 2021), even though they have been used for some time in other disciplines (Dahlin, 2021). Email interviews are based on asynchronous online communication between a researcher and a research participant via email over a period of time (Ratislavová & Ratiskav, 2014; Künzli & Gile, 2021). An advantage of email interviewing is that it is time and cost-efficient. For example, the data is already in written form and does not need to be transcribed (Dahlin, 2021), and more recent developments can be reflected and integrated into data collection, such as the publication of guidelines while interview data is being collected. Furthermore, research participants can answer the questions in their own time and reflect on their answers (Dahlin, 2021). This prolonged answering time may also have a negative effect on the data in that participants present a commonly shared standpoint in the community under research rather than "spontaneous answers" (Künzli & Gile, 2021). However, this can also occur in synchronous interviews.

---

4 The list of questions for part I is available on Zenodo via the following link: https://doi.org/10.5281/zenodo.11144968

5 In the consent form, it was agreed that the full interviews, even in anonymised form, were not to be published.

There were two sets of questions: i) for experts with greater involvement with the content aspect of terminology workflows (question set A, Annex 1) and ii) for IT experts (question set B, Annex 1). The questions addressed (i) the existence of institutional AI policies or guidelines concerning the use of GenAI or LLMs, (ii) whether the participating experts had experimented with GenAI applications for terminology or programming tasks, and (iii) the extent to which GenAI or LLMs play a role in daily terminology-related work.

Answers could be submitted in writing via email. At the time of writing this article, 13 out of the 15 original participants from the previous study had answered the follow-up questions (INT 4 and INT 13 are missing in the follow-up study). All answers were anonymised and encoded as described for the interview transcriptions. The interviewee numbering used was the same as in part I (see Table 1).

5. Results

The results from part I and part II of the study are presented separately in this section. In Sections 5.1 to 5.3 the results from part I are presented, while the results from part II are presented in Section 5.4.

5.1 Overview of the language technologies used by terminology professionals

This overview of the tools used summarises the results from part I of the study, so GenAI applications are not explicitly integrated into Figure 1.

The tools mentioned in the first part of the interview study can be roughly divided into two large categories, tools for data management and tools for communication purposes. Since this study focuses on translation technologies, tools used for communication purposes with different stakeholders (e.g. online discussion platforms, email) are not dealt with in this contribution. In the context of this study, data management is to be understood in its broadest sense, which subsumes accessing data. Data can refer to terminology data, corpus data or translation data (see Figure 1). Regarding terminology data, the following tools were mentioned: terminology management systems, dictionary writing systems, term extraction tools, data conversion scripts, quality control scripts, and tools to check terminological consistency in texts. Databases were also mentioned, mostly as a backend for terminology management or dictionary writing systems. However, terminology is sometimes also stored in spreadsheets, which are also used for data cleaning operations and for importing and exporting data from a database, or are shared with domain experts to review terminological data. APIs were mentioned as an alternative method of access for terminology databases, besides the primary user interface of terminology management systems, and data repositories were mentioned in the context of long-term archiving of terminological data. Machine translation is categorised in Figure 1 under translation, because it was mentioned in the context of translation-oriented terminology work and the pre-translation of texts. However, it is also categorised under terminology, because it is used to translate terms and terminological definitions (see Section 5.3).

Since several terminology professionals listed corpora (large collections of machine-readable texts) among the resources they use (Wissik, 2024b), it is no surprise that corpus management tools and annotation tools (e.g. part-of-speech tagger, lemmatiser) were mentioned.

The tools mentioned were either stand-alone tools or tools that could be integrated into other resources. Workflow tools were only mentioned as modules integrated into other tools, such as terminology management tools, therefore they do not appear as a separate category in Figure 1.

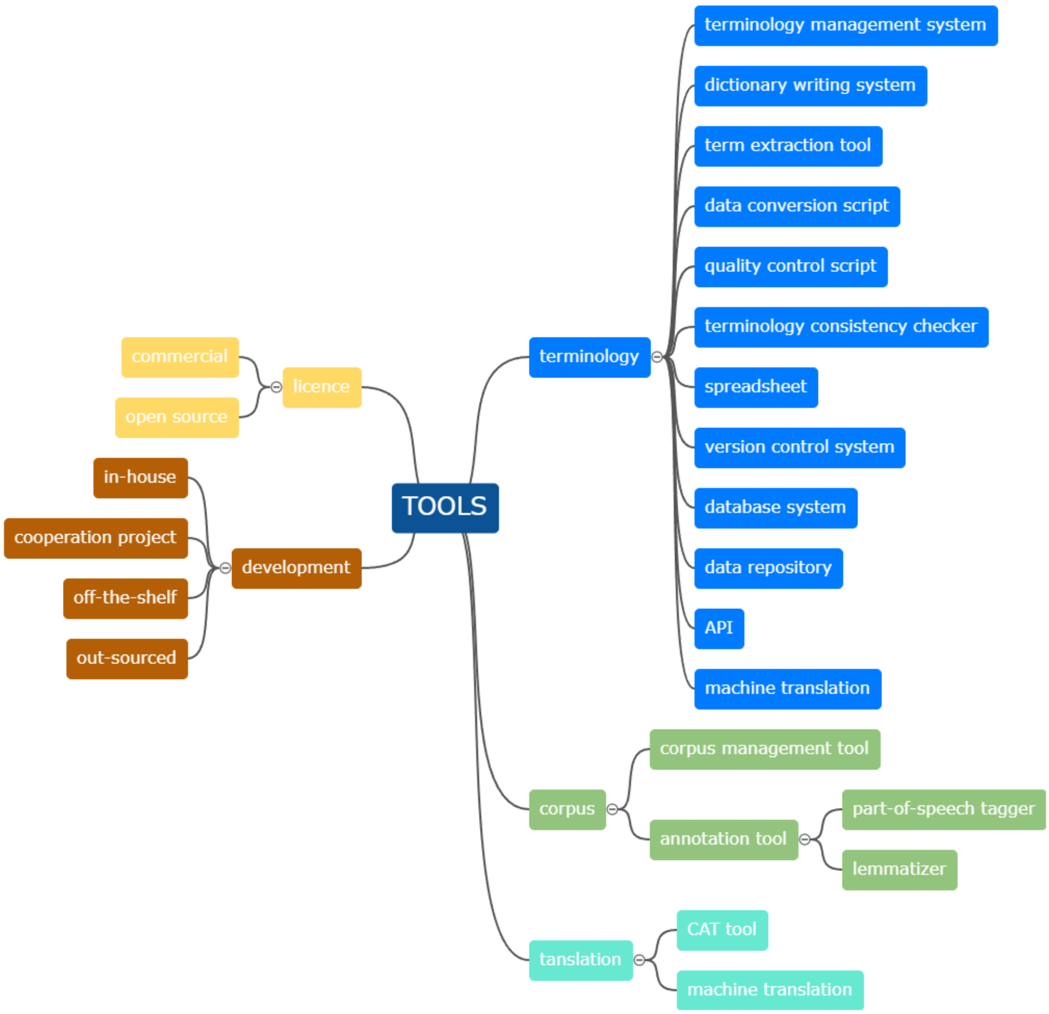

Figure 1: Tools used for handling terminology, corpus and translation data

Figure 1 shows the wide variety of tools used in terminology work. The fact that terminology professionals' use of technology in their work is steadily increasing is illustrated by the following quote:

> "[W]e now have a lot more to do with language technology in general anyway. So, I mean, when I started it was really just about making entries and now we take care of the whole thing, so now we take care of the licenses for the electronic dictionaries, we take care of

> the tool, we take care of the machine translation and then probably soon also ChatGPT and these things." (INT 15, translated by the author)

In the following sections we will go into greater detail regarding terminology management systems, term extraction, machine translation and GenAI applications.

5.2 Terminology management systems and term extraction

According to ISO 26162-2:2019, a terminology management system (TMS) is a "software tool specifically designed with a metadata structure for collecting, maintaining, and accessing terminological data." Dictionary writing systems (DWS) serve a similar purpose in the context of publishing terminological dictionaries. Abel (2022: 1) defines dictionary writing systems as software tools for "creating and editing dictionary entries, ensuring the long-term storage and reusability of data, as well as better support for smooth and consistent management throughout the whole lexicographic process."

All of the interview participants used terminology management systems or dictionary writing systems. The majority of participants (11: INT 1, INT 2, INT 3, INT 4, INT 7, INT 8, INT 10, INT 11, INT 12, INT 14, INT 15) used systems developed in-house or systems that were custom-made by external developers. In-house systems are not necessarily open-source; on the contrary, open source is relatively uncommon in the field of terminology management systems and in translation technology in general (Wissik, 2024a; Sandrini, 2019). However, one participant did describe adopting an open-source approach:

> "[W]e have been developing our own internal kind of in-house terminology management system and to [have] gone through several iterations of that and then at one point we decided that we want [the] next iteration to be something that we can share with outsiders, you know, we want to develop this as an open-source project and make it available not only to ourselves but also to others, to anybody out there who wants to use it as a hosted kind of software-as-a-service, you know, or people can actually get the source code and install it on their own servers and things like that." (INT 1)

Three participants (INT 5, INT 6, INT 9) used a commercial off-the-shelf terminology management system, specifically Multiterm,[6] for creating their terminological entries.

Sometimes, there are also additional systems used to manage the whole terminological database: "We create our terminological dictionaries using Multiterm and we have a system for managing our [terminological database]" (INT 9).

One participant (INT 13) reported that at their institution they used both a commercial (Muliterm) and an in-house tool to cater to the needs of different types of users.

Alongside terminology management systems, term extraction tools are often used by terminology professionals. Since an analysis of the use of term extraction tools was published in a previous article (Wissik, 2025), it will only be briefly summarised here. Based on Wissik (2025), three different groups of term extraction tools used by terminology professionals have been identified: 1) existing open-source or commercial

[6] https://www.trados.com/product/multiterm/

tools, developed by others; 2) tools developed in-house; and 3) tools developed in cooperation projects. Terminologists working with less-resourced languages more frequently used in-house term extraction tools or tools developed in cooperation projects due to a lack of technological resources for their working languages. This is in line with the findings of Secara et al. (2025): in their report based on interview data from representatives of the translation industry, they highlight the need for better translation technology support for less-resourced languages. Not all the terminology professionals used term extraction tools on a regular basis. Reasons for not using these tools included the results achieved not being satisfactory and always requiring a human review process. Furthermore, the quality of the results was found to greatly depend on usage scenarios and other constraints, such as language or domain. Other negative aspects included the tools not being integrated into the workflow, whether due to issues with information security and data protection rules or simply because of a lack of specific tools for the languages involved (Wissik, 2025).

5.3 Machine translation

There was a specific question regarding the role of machine translation in practical terminology work, because little is known about if and how machine translation is used by terminology professionals, apart from some case studies on translation memories and machine translation (Zorrilla-Agut, 2014; Caffrey & Valentini, 2019).

The application scenarios mentioned in the interviews range from projects for the evaluation of the neural machine translation of terminology in a specific low-resource language (INT 13), to using machine translation systems to draft translations of definitions, which are then reviewed by terminologists (INT 6). Other interview participants reported that they used machine translation to translate terms: "[T]he [results of our internal machine translation system, which is trained on in-house data] also provide the basis for us to work because we can look at the terms that the machine translation suggests and then try to, you know, see whether it's correct or not and try to do some search, so it's a preliminary, it's [a] basis, a good basis" (INT 13).

Another application scenario for machine translation is to identify entries where certain target languages are missing:

> "[A]nother way in which, in the past, we've used also the machine translation is [...] to find out basically records that in which we don't have a certain target language [...]. So typically, we've done this by doing carrying alternate extraction, for example, in [name of well-resourced language], seeing which were the most frequently used terms but that were not yet in the term base and then machine translating those in [name of well-resourced language] and then matching those English machine translated terms to the English terms that are already existing in the term base to try to find, you know, correspondences. [...] So, we've also used it like that for prioritising completion in some target languages." (INT 13)

An unsolved problem with machine translation mentioned for both in-house and externally sourced machine translation models was the inconsistent use of terminology:

“So we are working in close cooperation with the unit that has developed [our machine translation system]. But they and we know equally well how difficult it is to force this tool to use any terminology, so it’s still not solved” (INT 10).

Therefore, collaboration between terminologists and those who train and implement machine translation systems is important. Several interview partners were involved in such collaborations:

> “At present, what we do is we help [those] people who [are] in charge of [name of less-resourced language] spell checkers and [name of less-resourced language] translators to feed them with terms. We communicate the terms that can be updated. We provide them with [a] list of terms. Sometimes it is the other way round, they ask us for collaboration, they provide us with a list of terms that users suggested and we have to check them and decide which [ones] are correct, which [ones] are not, and give them that information.” (INT 7)

So far, we have discussed the more ‘traditional’ technologies used by professionals in terminology. In the next section we will look into the use of GenAI applications.

5.4 GenAI applications

In this section, the results of part II are described. First, we will discuss the availability of guidelines and policies, followed by experiments with GenAI applications, before finally tackling the questions of if and how GenAI is already integrated into the institutional terminology workflow. Even though AI-enhanced machine translation platforms such as DeepL are not GenAI applications per se, they are included in this section and not in the previous one because they were mentioned by the interview participants in part II of the study.

5.4.1 AI guidelines and policies

Seven out of the 13 interview participants mentioned that their institutions have guidelines or a policy document regarding the use of AI. Six participants said that they did not yet have such a document. Some mentioned that there were guidelines in development and that other information regarding the use of AI had already been shared through their intranet (e.g. warnings to not blindly trust AI results, to not enter any personal data, etc. (INT 6 and INT 11)). Furthermore, some participants mentioned that they had guidelines regarding the use of machine translation (INT 5) or online resources in general (INT 15), which needed to be extended to AI as well. In the context of these questions, one participant said that their institution has a dedicated AI portal, where not only guidelines but also learning resources and AI tools are made available (INT 10). One interview participant mentioned that their institution has an AI task force, which includes the terminology unit (INT 11), while another noted that their institution is involved in a project to develop and adapt an LLM for a less-resourced language (INT 3).

5.4.2 Experiments with GenAI applications

All of the interview participants except INT 14 had tried GenAI applications such as ChatGPT for terminology tasks. Two interviewees (INT 10, INT 12) mentioned that they were only allowed to use the in-house GenAI system and could not use publicly available chat interfaces for work-related tasks. Therefore, they reported on their experiences with their in-house GenAI systems. Some of the participants more involved in the technical side (INT 1, INT 2) mentioned that they had not tried out GenAI applications to help them write code (e.g. Copilot) but were planning to integrate GenAI-based support features into the terminology management systems that they were developing. Specifically, they were planning to include AI-based features for definition writing and translation, as well as term extraction.

Interviewees had tested GenAI with several terminology tasks: e.g. term extraction (INT 1, INT 2, INT 9, INT 10, INT 12), definition writing (INT 1, INT 2, INT 3, INT 6), definition translation (INT 2, INT 6, INT 10,[7] INT 12[8]), equivalent search (INT 3, INT 5, INT 6, INT 7), searching for etymology (INT 5, INT 7), domain suggestion (INT 11, INT 10), creating concept systems (INT 5), concept delimitation (INT 7), drafting responses to inquires in terminology consultation (INT 3), term relations (INT 5), finding or establishing synonyms (INT 11), searching for abbreviations (INT 11), searching for term use on a specific website (INT 11), verification of (translation) equivalence (INT 11), data cleaning (INT 10), and more general tasks like searching for specialised sources (INT 8), finding source documents in a different language (INT 11), document preprocessing (e.g. changing data formats) (INT 7, INT 8), suggesting text improvement (INT 8) and summarising long texts (INT 8).

The experiences of interview participants varied greatly, from "At first glance, it looks promising, but when you check the facts, there are a lot of inconsistencies, errors and even hallucinations" (INT 3) to "[W]e at the Terminology Team have been using ChatGPT for a few months now and we absolutely love it! Of course, it does make mistakes, and one has to be really careful with it and double check results using other tools/sources, but it's become one of our favourite and indispensable tools" (INT 11). Experiences vary greatly based on working languages, the tasks to be performed, and the GenAI application used. A recurring theme was that the results need human validation.

The limitations of GenAI applications such as ChatGPT were also discussed during the interviews. All participants reported that applications like ChatGPT give good results if the terminological task is in English, but if it is in a less-resourced language the proposed answers are rarely satisfying: "[I]f we bear in mind that we deal with terminology in a low-resource language, we find data provided were untrue, confusing or simply made-up" (INT 7). Another participant's answer, regarding the specific task of equivalent search, is along similar lines: "Since the [name of less-resourced language] is less developed in digital space and AI platforms, the proposed equivalents rarely provide acceptable options.

[7] Neural machine translation rather than GenAI.
[8] Neural machine translation rather than GenAI.

However, they often serve as a source of inspiration for possible [name of less-resourced language] terms" (INT 5). INT 12 mentioned that "bigger languages offer better quality than smaller ones in general."

Another interview participant shared their thoughts on language-specific LLMs: "We think that LLMs being developed for the [name of less-resourced language] (there is currently a national campaign to collect texts, with the aim of collecting 40 billion words) will be more reliable for our language" (INT 3).

There were also different experiences depending on the tasks performed. One interview participant rated term and definition extraction from source texts as good, but all the results required human validation: "The experience with these experimental [AI-based] features can be rated as good, but they require human validation. Particularly useful is the terms and definitions extraction from [...] legislative texts, which was a very manual and lengthy task and now it is highly automated" (INT 12). Another participant reported that tasks supporting data cleaning in the term base were inefficient: "The possibility to use AI for cleaning [name of term base] was explored and the conclusions were that it was not the right way to go. It did not bring added value to the cleaning procedure and other means were more efficient" (INT 10).

Furthermore, concerns regarding data sharing with GenAI applications were expressed, as in a statement from one of the interview participants: "However, we seldom [use AI applications] because we are concerned about sharing information and how our data could be used" (INT 7).

5.4.3 Integration of GenAI applications into the terminological workflow

For the majority of interview participants, AI did not (yet) play a role in their daily terminology work (INT 3, INT 5, INT 7, INT 8, INT 9, INT 14, INT 15), as illustrated by INT 3: "No, AI and LLMs do not play a significant role in our terminology work. The tools we use on [a] daily basis are not based on AI or LLMs" (INT 3). Furthermore, when asked about why AI had yet to be integrated into workflows, one interview participant referred to the limitations of GenAI, such as lower performance for less-resourced languages, hallucinations or simply incorrect results (see Section 5.4.2): "Unfortunately, we [have] refrained from integrating AI and LLMs in our unit due to the many limitations we [have] experienced" (INT 7).

Most participants did not use GenAI or LLMs for terminology management tasks or the creation of terminological entries: "We do not use tools with integrated AI modules for terminology management" (INT 8). However, GenAI applications were used for researching and searching for concepts and sources, as well as for data processing tasks:

> "[W]e have found that the following uses in general [of] ChatGPT can be an effective aid in some cases: researching difficult-to-document concepts, searching for specialized sources, suggesting text improvements, summarizing long texts, and changing data formats. These are

> tasks related with terminological work that take place before the creation of a term record and are carried out under the supervision of a terminologist." (INT 8)

Only three participants reported using GenAI applications regularly in their terminology work (INT 10, INT 11, INT 12). A separate case is the regular use of DeepL, an AI-enhanced machine translation tool that, however, is not a GenAI tool. It is used for definition and term translation, as reported by INT 6, precisely because it is not actually a GenAI tool: "I regularly use DeepL in my work to translate definitions and terms. The experience is basically good, but the translations have to be checked in any case, especially when they contain complex sentence structures. The advantage is that you can upload your own glossaries so that your own terminology is used in the translations" (INT 6, translated by the author). INT 11 used GenAI applications for terminology research and INT 10 explained that at their institution they used AI for definition drafting based on a source document: "We use our own definition [writing] prompt" (INT 10). Furthermore, INT 11 reported using ChatGPT for revision tasks (namely reviewing a glossary if all terms are domain and institution-specific) and then reviewing the results manually: "So at the end it helped me to remove irrelevant terms and I was quite happy with the results" (INT 11).

Regarding integration with other tools, only INT 10 and INT 12 reported the use of AI terminology support features integrated into their terminology management system. Furthermore, even though AI-based features may be integrated into a terminology management system, it does not necessarily mean they are integrated into the workflow, remaining optional additions: "We don't currently have any AI elements in our established workflow. Users have access to the existing tools and it is up to them to decide whether to use them or not" (INT 10).

6. Discussion

In this section, we will discuss the results from both part I and part II. This study has combined data from multiple actors in the terminology workflow (terminologists, terminology managers, (language) technology managers, domain experts, IT experts). A combined dataset bringing together a similar variety of actors within the terminology world was described as a research desideratum for translation workplace research by Rogle & Risku (2024). While the present study has some limitations, especially the small size of the sample, it gives insights into the use of language technology, including GenAI applications, by terminology professionals in institutional settings. However, the results presented are from an exploratory study and it must be stressed that the objective is certainly not to generalise the findings to the wider sector of terminology work. This study has focused exclusively on gathering information on terminology professionals in institutional settings and is not intended to be representative of terminology professionals in general, but rather a description of some of the current working situations in these settings.

The study has shown that terminology professionals in institutional settings use a variety of translation technologies, ranging from more traditional tools like terminology management systems, term extraction tools, corpus management tools and machine translation systems to GenAI applications and LLMs. Some interview participants stressed that more and more tools are being used in terminology work. This is also reflected by the emergence of a new role among terminology professionals, that of technology manager, sometimes also called a tool manager. A technology manager oversees the technology used for terminology work, develops and oversees the technology strategy of the terminology unit (sometimes also of the translation unit) and coordinates its development. If there are in-house developers, the technology manager coordinates their work and functions as a liaison between the (in-house or external) developer(s) and the end users (e.g. terminology managers, terminologists, translators).

Regarding the use of technology, all of the participants used some type of terminology management system or dictionary writing system. Most of them were custom-made tools, developed in-house or by an external developer. Three participants used a commercially-available tool, MultiTerm, and one participant mentioned using both an in-house system and a commercially available system. Open-source tools played a minor role, as already mentioned by Sandrini (2019) and Wissik (2024a).

This study has also shed light on the use of machine translation by terminology professionals, a topic neglected in the literature so far. Machine translation is used for various terminology tasks like term and definition translation or to identify records where certain target languages are missing. Terminology professionals also often work closely with developers of machine translation systems in order to give feedback and improve the performance of those systems.

Regarding the use of GenAI, only seven out of the 13 interview participants mentioned that their institution has guidelines or a policy document on AI. One participant said that their institution provides not only guidelines but also training materials on AI. One participant mentioned an AI task force, which includes representatives from the institution's terminology unit. Another stated that their institution is involved in the development of LLMs for their less-resourced language. All but one of the interview participants involved in the content-related side of terminology work had experimented with GenAI applications. The tasks that were carried out with GenAI ranged from specific terminology tasks like term extraction, definition writing, definition translation, equivalent search, searching for etymology, domain suggestion, creating concept systems, concept delimitation, identifying synonyms, and verification of (translation) equivalence, to more general tasks like searching for specialised sources, changing data formats, suggesting text improvements, and summarising long texts. The results were highly dependent on the working language, the tasks performed, and the specific GenAI application used. In general, the results of tasks in well-resourced languages are better than those of tasks in less-resourced languages. While nearly all of the participants had experimented with GenAI, the majority of participants did not (yet) use GenAI in their daily terminology work. In all of the institutions surveyed, the use of GenAI features was still optional and terminology professionals could decide to use them or not. Only two interview participants reported GenAI being integrated into existing tools, and two participants were planning to integrate GenAI features into their terminology management tool. However, the integration of AI features into existing tools does not automatically mean they will be adopted by professionals, as there is no obligation to use them.

Some institutions have internal rules on permissible technologies, prohibiting the use of publicly available GenAI chat interfaces for work-related tasks and allowing access to in-house GenAI tools only, and this may be specific to institutional settings. While they limit terminology professionals' opportunities to experiment with diverse tools and critically compare and evaluate their outputs, such restrictions may be deemed necessary in institutional settings because of issues with data security and publicly available AI tools.

A very interesting result from this study is that some interview participants who rarely use tools in their terminology work (except terminology management systems) reported very positive experiences when using GenAI applications (INT 13) and AI-powered machine translation systems (INT 6). These participants reported using these tools regularly for certain terminology tasks. This is in line with the results of an experiment involving an information-seeking task, using Google-like search engines and ChatGPT, where the users in the ChatGPT group reported significantly better experiences in terms of usefulness, enjoyment, and satisfaction (Xu et al., 2023). The ease of use of LLMs with conversational interfaces (Dingemanse, 2025) might contribute to a more positive user experience compared to other terminology tools.

The interview participants also discussed the limitations of technology in general, and GenAI in particular. One limitation that was mentioned is the lower quality of results in less-resourced languages. Evaluation studies such as Lai et al. (2023) and Alam et al. (2024) highlighted the performance-related limitations of current LLMs for less-resourced languages, since most current LLMs rely on predominantly English training data. This limitation affects the effective integration of LLMs into terminology workflows for less-resourced languages, which is common in the terminology sector. However, there are already some initiatives to adapt or (pre-)train LLMs for less-resourced languages (e.g. PLLuM Consortium, 2025; Etxaniz et al., 2024; Holdt et al., 2025). Other limitations mentioned are hallucinations or simply incorrect results. These limitations have implications for quality control and assurance, as discussed below. However, the use of RAG systems could limit hallucinations and incorrect results, especially for emerging domains that have yet to be included in training sets.

It was mentioned several times that GenAI results require human validation, because "AI-generated inaccuracies tend to be subtle; hence, deep knowledge of the subject-matter and of the source terminology is an unyielding prerequisite", as stated by Giampieri (2024: 355). This means that quality management measures will become increasingly important. Quality control and assurance requirements will need to become more stringent and far-reaching when GenAI is integrated into workflows. Thus, procedures will have to be adapted, since the results are more volatile and opaque (Strauß, 2024).

Concerns regarding data protection were also mentioned as a factor that discourages terminology professionals, particularly in institutional settings, from using (public) GenAI applications.

Another important issue is AI literacy (Long & Magerko, 2020), i.e. the knowledge and skills needed to understand, use, and critically evaluate AI technologies. Such knowledge can be shared through guidelines, and skills can be acquired through training materials and sessions. The data also highlights a need for awareness of the full capabilities of language technologies, as well as of which contexts and tasks are suitable for automation, when human intervention is needed, and when it is simply more efficient for a human to perform the task. For other terminology-related tasks, such as term extraction, training measures are important to achieve the best possible results from a specific tool (Wissik, 2025), but they are even more so when using GenAI.

Since both synchronous online and asynchronous email interviews were used in this study, it is worth reflecting on the influence the shift in interview format between the two phases may have had on the nature and richness of participants' responses. The live online format enabled real-time interaction, allowing for spontaneous answers, immediate clarification, and the observation of tone and nonverbal cues. However, such immediacy may have limited the time participants had to fully reflect on their answers. In contrast, the email interviews offered participants greater time for reflection and the opportunity to craft more deliberate, more structured, and sometimes more detailed responses. Even though the live interviews were conducted with a prepared set of questions, the structure of the interviews varied depending on how the conversation flowed. In contrast, the email interviews produced more structured responses that differed less between respondents because they all received the same questions in the same format. This also made analysis of the responses easier. However, organic digressions, which appeared in the live interviews and yield unexpected insights, were not found in the asynchronous email interviews.

The difficulty of establishing rapport in email interviews is mentioned in the literature (Oates et al., 2022). However, this was not the case in this study, because the participants had all participated in a live online interview before receiving the interview questions in phase II. In general, the two-phase approach was well suited to this study, and email interviews were an appropriate choice for a constantly evolving topic like GenAI, making it possible to integrate recent developments into data collection.

7. Conclusions

The objective of this study was to explore the use of translation technologies, including GenAI and LLMs, by terminology professionals in institutional settings. The data for the study was gathered in two steps: through 15 synchronous online interviews and 13 asynchronous email interviews conducted with terminology professionals in 2023 and 2025 respectively.

The analysis showed that terminology professionals are using more and more translation technologies, including GenAI applications. The increased use of technology is also reflected by the emergence of a new role — the technology manager — who oversees the technology strategy and the implementation of technology, as well as acting as a liaison between developers and end users.

Among the tools used by professionals there are more traditional ones, like terminology management systems and term extraction tools, as well as machine translation and GenAI applications. The study's results show that nearly all interview participants had experimented with GenAI applications for terminology tasks, and their experiences were highly dependent on factors like the working language(s) involved, the tasks to be performed, and the GenAI tool used.

One recurrent topic was that GenAI results require human validation. Due to the subtle nature of GenAI errors, revision and quality management need to be adapted when integrating GenAI into a terminology workflow.

Furthermore, there is a need for institutions to provide targeted AI training and to develop clear usage guidelines for GenAI in institutions in general and for terminology work in particular. Such guidelines should be regularly updated since the field is undergoing rapid change.

For future work, it would be interesting to complement the results of this study with quantitative data and to collect data from terminology professionals in commercial settings (companies or language service providers), so as to compare different settings and receive responses from a larger number of participants.

Acknowledgments

The author would like to thank all the interviewees. Without their support this study would not have been possible.

References

Abel, Andrea (2022). Dictionary writing systems. In: Patrick Hanks; Gilles-Maurice De Schryver (eds.). *International Handbook of Modern Lexis and Lexicography*. Berlin [etc.]: Springer, pp. 1-15. <https://doi.org/10.1007/978-3-642-45369-4_111-1>. [Accessed: 20250905].

Alcina, Amparo (2008). Translation Technologies: Scope, tools and Resources. *Target,* v. 20, n. 1, pp. 79-102. <https://doi.org/10.1075/target.20.1.05alc>. [Accessed: 20250905].

Allahyari, Mehdi; Yang, Angelina Yang (2023). *A Practical Approach to Retrieval Augmented Generation Systems*. <https://mallahyari.github.io/rag-ebook/>. [Accessed: 20250905].

Anthony, Laurence (2023). *Corpus AI: Integrating Large Language Models (LLMs) into a Corpus Analysis Toolkit.* Presentation given at the *49th Annual Conference of the Japan Association for English Corpus Studies (JAECS), Kansai University, Osaka*, Japan. <https://osf.io/srtyd/>. [Accessed: 20250905].

Anthony, Laurence (2024a). *AntConc (Version 4.3.1): A freeware corpus analysis toolkit for concordancing and text analysis* [Computer software]. <https://www.laurenceanthony.net/software/antconc/>. [Accessed: 20250905].

Anthony, Laurence (2024b). *Help for AntConc (Version 4.3.1)* https://www.laurenceanthony.net/software/antconc/releases/AntConc431/help.pdf. [Accessed: 20250905].

Bowker, Lynne; Corpas Pastor, Gloria (2015). Translation Technology. In: Ruslan Mitkov (ed.). *The Oxford Handbook of Computational Linguistic*s. 2nd edition. Oxford: Oxford University Press, pp. 871-905. <https://doi.org/10.1093/oxfordhb/9780199573691.013.007>. [Accessed: 20250905].

Caffrey, Colm; Valentini, Cristina (2019). Applications of technology in the Patent Cooperation Treaty (PCT) Translation Division of the World Intellectual Property Organization (WIPO). In: Minako O'Hagan (ed.). *The Routledge Handbook of Translation and Technology*, v. 1. Oxfordshire: Taylor & Francis , pp. 127-147. <https://www-taylorfrancis-com.uaccess.univie.ac.at/chapters/edit/10.4324/9781315311258-10/applications-technology-patent-cooperation-treaty-pct-translation-division-world-intellectual-property-organization-wipo-colm-caffrey-cristina-valentini>. [Accessed: 20250905].

Chiocchetti, Elena; Lušicky, Vesna; Wissik, Tanja (2023). Multilingual Legal Terminology Databases. Workflows and Roles. In: łucja Biel; Hendrik J. Kockaert (eds.). *Handbook of Terminology,* v. 3., Amsterdam [etc.]: John Benjamins, pp. 458-484. <https://doi.org/10.1075/hot.3.mul1>. [Accessed: 20250905].

Chiocchetti, Elena; Ralli, Natascia; Wissik, Tanja (2014). Terminology workflows in theory and practice. In: Vesna Lušicky; Gerhard Budin (eds.). *Proceedings of the 19th European Symposium on Languages for Special Purposes: Languages for Special Purposes in a Multilingual, Transcultural World.* University of Vienna, pp. 525-535. <https://phaidra.univie.ac.at/o:359149>. [Accessed: 20250905].

Dahlin, Emma (2021). Email Interviews: A Guide to Research Design and Implementation. *International Journal of Qualitative Method*, v. 20, pp. 1-10. <https://doi.org/10.1177/16094069211025453>. [Accessed: 20250905].

Etxaniz, Julen; Sainz, Oscar; Miguel, Naiara; Aldabe, Itziar; Rigau, German; Agirre, Eneko; Ormazabal, Aitor; Artetxe, Mikel; Soroa, Aitor (2024). Latxa: An Open Language Model and Evaluation Suite for Basque. In: Lun-Wei Ku; Andre Martins; Vivek Srikumar (eds.). *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers). Association for Computational Linguistics, pp. 14952-14972. <https://aclanthology.org/2024.acl-long.799/>. [Accessed: 20250905].

Giampieri, Patrizia (2024). The use of AI in the translation of legal documents: A critical analysis. *Trans-kom,* v. 17, n. 2, pp. 343–358. <https://www.trans-kom.eu/bd17nr02/trans-kom_17_02_05_Giampieri_AI.20241128.pdf>. [Accessed: 20250905].

Giguere, Julia; Pareja, Antonio; Mitkov, Ruslan (2023). Leveraging Large Language Models to Extract Terminology. In: *Proceedings of the First Workshop on NLP Tools and Resources for Translation and Interpreting Applications*, *Varna, Bulgaria.* Shoumen: INCOMA, pp. 57-60. <https://aclanthology.org/2023.nlp4tia-1.9/>. [Accessed: 20250905].

Gius, Evelyn; Meister, Jan Christoph; Meister, Malte; Petris, Marco; Schumacher, Mareike; Gerstorfer, Dominik (2023). *CATMA 7 (Version 7.0)* [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.1470118>. [Accessed: 20250905].

Guest, Greg; Bunce, Arwen; Johnson, Laura (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods,* v. 18, n.1, pp. 59-82. <https://journals.sagepub.com/doi/10.1177/1525822X05279903>. [Accessed: 20250905].

Heinisch, Barbara (2025). Large Language Models for Terminology Work: A question of the right prompt? *Journal for Language Technology and Computational Linguistics*, v. 38, n. 2, pp. 13-30. <https://doi.org/10.21248/jlcl.38.2025.280>. [Accessed: 20250905].

Heinisch, Barbara (2024). Large language model prompting in a university course on terminology. *Terminology Science & Research = Terminologie: Science et Recherche,* v. 27, pp. 28-51. <https://journal-eaft-aet.net/index.php/tsr/issue/archive>. [Accessed: 20250905].

Holdt, Špela Arhar; Antloga, Špela; Munda Tina; Pori, Eva; Krek, Simon (2025). From Words to Action: A National Initiative to Overcome Data Scarcity for the Slovene LLM. In: Špela Arhar Holdt; Nikolai Ilinykhi; Barbara Scalvini; Micaella Brunton; Iben Nyholm Debess; Crina Madalina Tudor (eds.). P*roceedings of the Third Workshop on Resources and Representations for Under-Resourced Languages and Domains (RESOURCEFUL-2025*). pp. 130-136. <https://aclanthology.org/2025.resourceful-1.27/>. [Accessed: 20250905].

Hosseini-Kivanani, Nina (2025). Neologism Validation with Large Language Models and Lexical Graphs. In: Federica Vezzani; Ana Ostroški Anić; Ana Salgado; Giovanni Tallarico (eds.). *Proceedings of the 1st International Workshop on Terminological Neologism Management (NeoTerm 2025*). <https://ceur-ws.org/Vol-3972/paper2.pdf>. [Accessed: 20250905].

Hutchins, W. John (2001). Machine translation over fifty years. *Histoire. Épistémologie, Langage,* v. XXII, n. 1, pp. 7-31. <https://www.persee.fr/doc/hel_0750-8069_2001_num_23_1_2815>. [Accessed: 20250905].

ISO 26162-2:2019 (2019). *Management of terminology resources-Terminology databases-Part 2: Software*. International Standards Organisation.

Kim, Sejoon; Sung, Mingi; Lee, Jeonghwan; Lim, Hyunkuk; Gimenez Perez, Jorge Froilan (2024). Efficient Terminology Integration for LLM-based Translation in Specialized Domains. In: Barry Haddow; Tom Kocmi; Philipp Koehn; Christof Monz (eds.). *Proceedings of the Ninth Conference on Machine Translation. Association for Computational Linguistics*, pp. 636-642. <https://doi.org/10.18653/v1/2024.wmt-1.51 >. [Accessed: 20250905].

Kuckartz, Udo (2014). *Qualitative Text Analysis: A Guide to Methods, Practice and Using Software.* London: SAGE Publications.

Künzli, Alexander; Gile, Daniel (2021). The impact of ICTs on surveys and interviews in translation and interpreting studies. *Parallèles,* v. 33, n. 2, pp. 18-34. <https://doi.org./10.17462/para.2021.02.02>. [Accessed: 20250905].

Lefever, Els; Rigouts Terryn, Ayla (2024). Computational Terminology. In: Yuhong Pen; Huihui Huang; Defeng Li (eds.). *New Advances in Translation Technology: Applications and Pedagogy.* Singapore: Springer, pp. 141-159. <https://doi.org/10.1007/978-981-97-2958-6_8>. [Accessed: 20250905].

Lewis, Patrick; Perez, Ethan; Piktus, Aleksandra; Petroni, Fabio; Karpukhin, Vladimir; Naman Goyal, Naman; Küttler, Heinrich; Lewis, Mike; Wen-tau Yih, Wen-tau; Rocktäschel, Tim; Riedel, Sebastian; Kiela, Douwe (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In: Hugo Larochelle; Marc'Aurelio Ranzato; Raia Hadsell; Maria-Florina Balcan; Honghao Lin (eds.). *Proceedings of Advances in Neural Information Processing Systems 33 (NeurIPS 2020).* <https://papers.nips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf>. [Accessed: 20250905].

Long, Duri; Magerko, Brian (2020). What is AI Literacy? Competencies and Design Considerations. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).* Association for Computing Machinery, pp. 1-16. <https://doi.org/10.1145/3313831.3376727>. [Accessed: 20250905].

Łukasik, Marek (2023). Terminology Work in the Era of AI-generated Content: A Case of ChatGPT: and end-user perspective. Presentation given at *5th International Conference on Terminology: Scientific, Administrative and Educational Dimensions of Terminology, 19-20 Oct. 2023.* <https://www.researchgate.net/publication/378297882_Terminology_work_in_the_era_of_AI-generated_content_A_case_of_ChatGPT_an_end-user_perspective>. [Accessed: 20250905].

Massion, François (2024). Terminology in the Age of AI: The Transformation of Terminology Theory and Practice. *Journal of Translation Studies,* v. 4, n. 1, pp. 67-94. <https://doi.org/10.3726/JTS012024.04>. [Accessed: 20250905].

McCrae, John P. (2019). Identification of adjective-noun neologisms using pretrained language models. In: Agata Savary; Carla Parra Escartín; Francis Bond; Jelena Mitrović; Verginica Barbu Mititelu (eds.). *Proceedings of the Joint Workshop on*

Multiword Expressions and WordNet (MWE-WN 2019), pp. 135–141. <https.//doi.org/10.18653/v1/W19-5116>. [Accessed: 20250905].

Meuser, Michael; Nagel, Ulrike (1991). Experteninterviews: vielfach erprobt, wenig bedacht. In: Detlef Garz; Klaus Kraimer (eds.). *Qualitative-empirische Sozialforschung. Konzepte, Methoden, Analysen*. Opladen: Westdeutscher, pp. 441-471.

Martín-Chozas, Patricia; Montiel-Ponsoda, Elena; Rodríguez-Doncel, Víctor (2023). The role of Semantic Web technologies in legal terminology. In: Łucja Biel; Hendrik J. Kockaert (eds.). *Handbook of Terminology,* v. 3. Amsterdam [etc.]: John Benjamins, pp. 542-567. <https://chooser.crossref.org/?doi=10.1075%2Fhot.3.rol1>. [Accessed: 20250905].

Oates, Maggie; Crichton, Kyle; Cranor, Lorrie; Budwig, Storm; Weston, Erica J. L.; Bernagozzi, Brigette M.; Pagaduan, Julie (2022). Audio, video, chat, email, or survey: How much does online interview mode matter? *PLoS One,* v. 17, n. 2. <https://doi.org/10.1371/journal.pone.0263876>. [Accessed: 20250905].

PLLuM Consortium (2025). *PLLuM: Polish Large Language Model*. <https://pllum.org.pl/>. [Accessed: 20250905].

Ratislavová, Kateřina; Ratislav, Jakub (2014). Asynchronous email interview as a qualitative research method in the humanities. *Human Affairs*, v. 24, pp. 452–460. <https://doi.org/10.2478/s13374-014-0240-y>. [Accessed: 20250905].

Reineke, Detlef (2023). Terminologiearbeit mit ChatGPT & Co. *Fachzeitschrift Für Terminologie,* v. 19, n. 1, pp. 25-28. <http://dttev.org/images/edition/ausgaben/edition-2023-1- e-version.pdf>. [Accessed: 20250905].

Rogl, Regina; Risku, Hanna (2024). Cognitive artefacts and boundary objects: On the changing role of tools in translation project management. In: Marion Winters; Sharon Deane-Cox; Ursula Böser (eds.). *Translation, Interpreting and Technological Change: Innovations in Research, Practice and Training*. London [etc.]: Bloomsbury Publishing, pp. 13-36.

San Martín, Antonio (2024). What Generative Artificial Intelligence Means for Terminological Definitions. In: Federica Vezzani; Giorgio Maria Di Nunzio; Beatriz Sánchez Cárdenas; Pamela Faber; Melania Cabezas García; Pilar León Araúz; Arianne Reimerink; Antonio San Martín, Antonio (eds.). *Proceedings of the 3rd International Conference on Multilingual Digital Terminology Today (MDTT 2024)*. <https://ceur-ws.org/Vol-3703/paper1.pdf>. [Accessed: 20250905].

Sánchez-Gijón, Pilar; Palenzuela-Badiola, Leire (2023). Analysis and Evaluation of ChatGPT-Induced HCI Shifts in the Digitalised Translation Process. In: Constantin Orasan; Gloria Corpas Pastor; Ruslan Mitkov; Johanna Monti (eds.). *Proceedings of Hit-IT: The International Conference on Human-Informed Translation and Interpreting Technology*. Naples: Incoma, pp. 227-67. <https://acl-bg.org/proceedings/2023/HiT-IT%202023/pdf/2023.hitit2023-1.21.pdf>. [Accessed: 20250905].

Sandrini, Peter (2019). *Translationspolitik für Regional-oder Minderheitensprachen. Unter besonderer Berücksichtigung einer Strategie der Offenheit*. Berlin: Frank und Timme.

Siu, Sai Cheong (2024). Revolutionising Translation with AI: Unravelling Neural Machine Translation and Generative Pre-trained Large Language Models. In: Yuhong Peng; Huihui Huang; Defeng Li (eds.). *New Advances in Translation Technology = New Frontiers in Translation Studies*. Singapore: Springer, pp. 29-54. <https://doi.org/10.1007/978-981-97-2958-6_3>. [Accessed: 20250905].

Secara, Alina; Rivas Ginel, Isabel; Toral, Antonio; Guerberof, Ana; Ciobanu, Dragoș; Brockmann, Justus; Plieseis, Claudia; Chereji, Raluca; Rossi, Caroline (2025). *LT-LiDER Language Technology Map: Technologies in Translation Practice and their Impact on the Skills Needed*. Zentrum für Translationswissenschaft. <https://doi.org/10.25365/phaidra.641>. [Accessed: 20250905].

Stefaniak, Karolina. (2023). Terminology management and terminology quality assurance in the European Commission's Directorate-General for Translation., In: Łucja Biel; Hendrik J. Kockaert (eds.). *Handbook of Terminology*, *v. 3: Legal Terminology.* Amsterdam [etc.]: John Benjamins, pp. 458–484. <https://doi.org/10.1075/hot.3.ter12>. [Accessed: 20250905].

Strauß, Stefan (2024). *CAIL-Critical AI Literacy (Endbericht*). Wien: Institut für Technikfolgen-Abschätzung (ITA). <https://epub.oeaw.ac.at/ita/ita-projektberichte/ITA-2024-07.pdf>. [Accessed: 20250905].

TerminOrgs (2016). Recommended skills for terminology professionals. <https://www.terminorgs.net/downloads/Terminology_skills.pdf>. [Accessed: 20250905].

Thießen, Freya; D'Souza, Jennifer; Stocker, Markus (2023). Probing Large Language Models for Scientific Synonyms, In: Jennifer D'Souza; Anisa Rula; David Chaves-Fraga; Mersedeh Sadeghi; Shahrom Hosseini Sohi; Julián Andrés Rojas; Pieter Colpaert; Edlira Vakaj; Sanju Tiwari; Sahar Vahdati; Anna Lisa Gentile (eds.). *SEMANTiCS 2023 EU: Workshops Proceedings Compound Volume 2023: 19th International Conference on Semantic System, September 20-22, 2003, Germany….* <https://ceur-ws.org/Vol-3510/paper_nlp_2.pdf>. [Accessed: 20250905].

Tran, Hanh Thi Hong; Martinc, Matej; Caporusso, Jaya; Doucet, Antoine; Pollak, Senja (2023). *The Recent Advances in Automatic Term Extraction: A survey*. <https://arxiv.org/abs/2301.06767>. [Accessed: 20250905].

Varga, Christina (2024). El uso de ChatGPT en el trabajo terminológico puntual y sistemático en rumano y en castellano. *Quaestiones Romanicae* v. XI, n. 2, pp. 377-396. <https://doi.org/10.35923/QR.11.02.32>. [Accessed: 20250905].

Wissik, Tanja (2024a). Dimensions of sustainability in terminology practice in institutional settings. *Terminology Science & Research = Terminologie: Science et Recherche*, v. 27, pp. 93-116. <https://journal-eaft-aet.net/index.php/tsr/article/view/7284/7481>. [Accessed: 20250905].

Wissik, Tanja (2024b). An infrastructural approach to terminology work. In: Vincent Vandeghinste; Thalassia Kontino (eds.). *CLARIN Annual Conference Proceedings, 15-17 October 2024, Barcelona, Spain.* pp. 1-5. <www.clarin.eu/sites/default/files/CLARIN2024_ConferenceProceedings_final.pdf>. [Accessed: 20250905].

Wissik, Tanja (2025). Impact of automatic term extraction on terminology work: A qualitative interview study in institutional settings. T*erminology,* v. 31, n. 1, pp. 110–136. <https://doi.org/10.1075/term.00085.wis>. [Accessed: 20250905].

Xu, Kang; Feng, Yifan; Li, Qiandi; Dong, Zhenjiang; Wei, Jianxiang (2025). Survey on terminology extraction from texts. *Journal of Big Data,* v. 12, n. 29. <https://doi.org/10.1186/s40537-025-01077-x>. [Accessed: 20250905].

Zorrilla-Agut, Paula (2014). When IATE met LISE: LISE clean-up and consolidation tools take on the IATE challenge. In: Vesna Lušicky; Gerhard Budin (eds.). *Proceedings of the 19th European Symposium on Languages for Special Purposes: Languages for Special Purposes in a Multilingual, Transcultural World. University of Vienna*, pp. 536-545. <https://phaidra.univie.ac.at/o:359149>. [Accessed: 20250905].

Zorrilla-Agut, Paula; Fontenelle, Thierry (2019). IATE 2. Modernising the EU's IATE terminological database to respond to the challenges of today's translation world and beyond. *Terminology*, v. 25, n. 2, pp. 146-174. <https://doi.org/10.1075/term.00034.zor>. [Accessed: 20250905].

Annex 1

Questions from part II of the study

Question set A

1. Does your institution/unit have a policy/guidelines regarding the use of applications based on (generative) AI or LLMs? Or is such a document currently being developed?
2. Have you tried to use ChatGPT, Gemini or other similar generative AI applications for some of the steps in terminology work? What was your experience with them? How would you rate the quality of the results for the languages you are working with?
3. Do AI and large language models (LLMs) play a role in your daily terminology work? Do you use tools based on AI or LLMs or that have integrated AI modules for terminology management, data cleaning, term extraction, term translation, definition writing or other tasks?

   a) Please name and describe the tool(s) and the task(s) you use them for.
   b) How was your experience with the tool(s) (pros and cons)?
   c) How do they integrate into your established workflow and with the other tools you use?
   d) Did you experience differences in quality for different languages?

Question set B

1. Does your institution/unit have a policy/guidelines regarding the use of applications based on (generative) AI or LLMs? Or is such a document currently being developed?
2. Do (generative) AI and large language models (LLMs) play a role in your daily work in developing terminology and linguistic tools? Do you use tools based on (generative) AI or LLMs for developing terminology management systems or other linguistic tools? Do you integrate (generative) AI or LLM-based applications into the tools you are developing? If so, for which tasks have you integrated AI? If so, do you already use existing LLMs out of the box, do you finetune existing LLMs, or do you train your own?
3. Have you gotten requests from users to integrate (generative) AI or LLM-based applications into the systems you develop? If so, for which tasks?