



PATRIZIA GIAMPIERI

LLMs as smart concordancers for legal translation practice

There are several literature-driven studies exploring how intelligent machines and humans can interact to streamline and improve legal translation processes. Nonetheless, no scholarly research has focused on the employment of GenAI/LLMs as concordancers in the legal translation workflow. This paper aims at filling this epistemic lacuna by considering Notebook LM. For this purpose, an extract of an Italian certificate of limited partnership (*società in accomandita semplice*) is translated into English by leveraging Notebook LM and querying it as a corpus concordancer. Hence, 45 certificates of partnerships drawn up in English are uploaded to the LLM system and used as reference documents for the translation task. Prompts are formulated with a view to extracting collocates, disambiguating words in context, verifying word usages, obtaining concordances, and clarifying relevant legilinguistic aspects. To complete the analysis, the US *Uniform Partnership Act 1997 – Last Amended 2013* (“UPA”) is also consulted via a separate folder of Notebook LM. The findings indicate the importance of formulating prompts as clearly and completely as possible in order to retrieve translation-salient information. The results also show that LLMs can be seamlessly integrated into the translation workflow: prompts can feature simple language and are articulated in a user-friendly environment, thus enabling accessibility by non-ICT specialists; responses are, most of the times, straightforward and coherent, thereby facilitating the reading and interpretation of corpus data. However, the upload of targeted and field-relevant documents, as well as expert knowledge are necessary to draft ad hoc queries, eschew unrelated responses, and interpret results consistently.

GenAI in legal translation – LLMs in legal translation – Legilinguistics – Corpus linguistics – Legal translation

Impiego di Large Language Models come concordanti nella traduzione giuridica

Diversi studi scientifici hanno analizzato la possibilità di impiegare la IA per semplificare e migliorare la traduzione giuridica. Tuttavia, la ricerca non si è dedicata all'impiego di LLM (Large Language Models) come concordanti. Il presente articolo si propone di colmare tale lacuna epistemica considerando Notebook LM per scopi traduttivi. A tal fine, un estratto di un atto costitutivo di società in accomandita semplice redatto in lingua italiana è tradotto in inglese utilizzando Notebook LM. Documenti corrispondenti in lingua inglese sono caricati sul sistema e utilizzati come riferimento ai fini della traduzione. Si formulano prompt con l'obiettivo di estrarre collocazioni, disambiguare parole nel contesto, verificare l'uso di termini, ottenere concordanze e chiarire aspetti legilinguistici. A completamento dell'analisi, si consulta separatamente lo *Uniform Partnership Act 1997 – Last Amended 2013* (“UPA”). I risultati dello studio sottolineano l'importanza di formulare prompt in modo chiaro e completo al fine di ottenere informazioni rilevanti per la traduzione. Si dimostra quindi che i sistemi LLM possono essere integrati nel processo traduttivo; i prompt possono essere formulati in un linguaggio semplice ed intuitivo, rendendoli accessibili anche a chi non è esperto di informatica o di analisi dei corpora. Le risposte del sistema sono, nella maggior parte dei casi, dirette e coerenti, facilitando così la lettura e l'interpretazione dei dati. Tuttavia, si rendono necessari l'impiego di documenti-fonte mirati, nonché la conoscenza specialistica del linguaggio giuridico al fine di ottenere risultati pertinenti ed interpretarli in modo coerente.

IA nella traduzione giuridica – LLM nella traduzione giuridica – Linguistica dei corpora – Traduzione giuridica

The Author is Associate Professor of English Language, Translation and Linguistics at the University of Perugia

SUMMARY: 1. Introduction. – 1.1. Integrating automation into the translation process. – 1.2. Automating legal services and legal translation practices. – 2. Aim of the paper and research questions. – 3. Methodology. – 4. Analysis. – 5. Discussion. – 6. Conclusions.

1. Introduction

1.1. Integrating automation into the translation process

In 2020, Allen and Hogan¹ envisaged the integration of machines in the translation workflow and described a model according to which automatic post-editing (APE) could be performed within MT (Machine Translation) processes. This was achieved by training machine learning algorithms that detected and automatically corrected errors on the basis of previous human post-editing of machine-generated texts. Hence, a machine learning pipeline was developed to leverage source-MT output-human post-edit (SRC-MTO-HPE) triplets, thereby obtaining post-editing (PE) rules and re-use them in other texts to translate. Such a new paradigm was, however, not without inherent failures. Do Carmo et al.², for instance, revealed a host of overcorrections and overlooked errors. Therefore, scholars found that although APE could enhance MT-driven workflows, human proofreading was necessary³.

For these reasons, Niu and Carpuat⁴ implemented a triplet-based training model that was capable of detecting missing or wrong variables automatically, thus allowing for automated instant corrections without necessitating human intervention. At

the same time, the authors aimed at enhancing the level of formality of target texts while maintaining the same quality output. To prove the validity of their system, Niu and Carpuat involved human evaluators in the assessment of the output texts. The evaluators had to verify meaning preservation and formality improvement. The results of their study point to the achievement of a high standard of translation quality and formality enhancement.

Escribe and Mitkov⁵ adopted a different perspective and envisaged a human-machine-interconnected scenario. The authors found that human-machine interactions were still limited because the human post-editing process (considered necessary by the authors) was not back-fed into the translation workflow. In this way, automated systems failed to make changes and corrections in the database. For this reason, Escribe and Mitkov leaned towards improving human post-editors' training and assistance rather than rendering the PE process fully automated.

Other scholars suggested corpus-based systems integrated into APE toolkits. Already in 2018, Negri et al.⁶ proposed an APE tool that could be trained on the basis of a virtual collaborative environment where human post-editors could interact. For this purpose, they developed a bilingual parallel corpus from which the system would learn and be trained. Such a corpus featured both source texts and hu-

1. ALLEN-HOGAN 2000.

2. DO CARMO-SHTERIONOV-MOORKENS et al. 2020.

3. SHTERIONOV-DO CARMO-MOORKENS et al. 2020.

4. NIU-CARPUAT 2020.

5. ESCRIBE-MITKOV 2023.

6. NEGRI-TURCHI-BERTOLDI-FEDERICO 2018; NEGRI-TURCHI-CHATTERJEE-BERTOLDI 2018.

man-translated target texts. An MT tool was anchored to the corpus in the post-translation phase. Namely, it firstly produced a target text, then it employed human-driven corpus evidence as a form of PE to correct and reintegrate target data. The authors later adopted a web-based approach, which was more satisfactory than the initial APE paradigm.

Huge strides forward have been made after the advent of GenAI and LLMs. Anthony⁷ shows how AI technology can be incorporated into corpus-based analyses by leveraging API (Application Programme Interface) keys and a cloud-based LLM interaction. More precisely, the author describes how concordancers such as AntConc⁸ are combined with LLMs to offer an easy way to navigate corpus data by relying on the precision and accuracy of corpus evidence. For example, users can interact with the integrated LLM tool both before and after the formulation of corpus-based queries. In the initial stage, they can prompt the AI interface to tailor (or assist in the tailoring) of corpus queries (for example, by listing a particular term functioning as both a noun and a verb in the corpus⁹). In the post-prompt phase, users can ask to elucidate corpus results or to furnish further, or similar, language evidence from the corpus. Despite the user-friendliness approach and the expeditiousness of the process, the whole paradigm is not free of issues. Firstly, Anthony¹⁰ highlights that API keys can be rather costly, especially when several corpus results are needed and, hence, diverse queries must be inputted. Secondly, privacy can be compromised or violated, as queries are transmitted to providers' servers, so users have no control over the information flow.

1.2. Automating legal services and legal translation practices

In view of the huge advancements in translation technologies, legal linguists and scholars, as well

as law firms, have gradually incorporated AI into their workflows¹¹. In 2023, 80% of the law firms surveyed by Thomson Reuters deemed chatbots as potentially employable in the legal profession. In 2025, Thomson Reuters actually found that in-house legal departments declared to be more proactive in the use of GenAI than similar outsourced services. In 2026, according to new Thomson Reuters' findings, GenAI has become a reality for everyone: it is a tool that helps reducing costs for end-users, and it is a way for law firms to offer services to a wider range of clients while asking for higher fees in view of the increased quality of their business¹². The downside of this patterning is that clients may decide to reduce their reliance on such premium services as they prefer conversing directly with chatbots to solve their legal issues¹³. This, however, is not risk-free: users may not formulate prompts adequately, thus risking being exposed to biased outputs or utterly wrong legal advice¹⁴.

In legal translation practice, Bajcic and Golenko¹⁵ report an increasing number of publications addressing LLMs applied in legal translation. The authors acknowledge that NMT (Neural Machine Translation) systems have not been employed satisfactorily in the legal field for several reasons. Amongst others, scholars have always lamented terminological and lexical issues, omissions, repetitions, and word-order oddities. According to the authors, this justified the fact that NMT was not completely feasible in the legal field in view of the peculiarities, intricacies, and system specificities of legal discourse. The advent of GenAI, however, has changed the legal translation landscape and its requirements. In this respect, Bajcic and Golenko acknowledge the high adaptability of such systems and their capacity to learn and be trained in real time. This implies that LLMs and GenAI perform better than early machine translation tools such as Google Translate.

7. ANTHONY 2025.

8. ANTHONY 2026.

9. ANTHONY 2025, p. 12.

10. ANTHONY 2025.

11. THOMSON REUTERS 2023; THOMSON REUTERS 2025; THOMSON REUTERS 2026.

12. THOMSON REUTERS 2026.

13. *Ibidem*.

14. THOMSON REUTERS 2025.

15. BAJČIĆ–GOLENKO 2024.

Nonetheless, several scholars have pinpointed the pitfalls of AI-driven automated translation¹⁶. To test the accuracy and reliability of AI in the legal translation sector, Altakhaineh et al.¹⁷ carry out a comparative analysis of human vs GenAI-driven legal translations in the English/Arabic language pair. The source text was mainly composed of contract clauses, e.g., indemnification, termination, governing law clauses, and those concerning the non-disclosure of confidential information. The authors addressed three main aspects: the use of correct legal terminology; the comprehensibility of the target text, and compliance with system-bound principles. Their findings reveal that human translators outperformed GenAI, showing more terminological preciseness and clarity, along with higher levels of compliance with legal Arabic principles. The major shortcomings resulting from intelligent machines' outputs revolved around the incapability of discerning terms in legal contexts and rendering articulated legal constructions.

A similar examination is carried out by Li and Wang¹⁸, who compare legal translations performed by humans and GPT-4. They prompt human translators and chatbots to translate the Chinese Contract Book of the Civil Code into English and notice that the GPT version was less complex and more readable. Simplification, however, was at the cost of precision and accuracy, thus producing a target text that could lead to misunderstandings and ambiguities. Additionally, the automated translation often featured literal renditions, thus disregarding legal conventional patterning. The human translation, conversely, was more rigorous and precise, resorting to various translation strategies to accurately render source terms and phrases in context.

Likewise, Badah et al.¹⁹ scrutinise human vs chatbot-powered translations. They focus on the English/

Arabic language combination where English was the target language. For their study, they consider a special power of attorney and a judicial pleading. They involved 27 students in Translation Studies and used two chatbots (namely, ChatGPT and DeepSeek). Similarly to Li and Wang²⁰ and Altakhaineh et al.²¹, the authors find that intelligent machines strive to produce consistent target terminology and often give rise to poor renditions of system- or culture-bound terms. On the other hand, AI outperformed the students in textual coherence and terminological accuracy (as long as no culture-, context-, or system-specific terminology was involved).

With a view to scrutinising the quality of machine translations, Briva-Iglesias et al.²² assess the legal translation performances of an LLM and an NMT system in four language combinations. In their language experiment, the authors consider English as a source language and Brazilian Portuguese, Catalan, Spanish, and Turkish as target languages. The legal text addressed was a contract of over 500 words. They used Google Translate and ChatGPT, which was prompted to translate the source text into each of the target languages indicated. They involved human professional translators with over five years' experience in their project. The authors find that GPT-4 produced reliable target texts with context-appropriate wording, terminological consistency (e.g., source terms and phrases were translated consistently in the target text when appearing multiple times), and better fluency.

In a related vein, Giampieri²³ scrutinises the quality of MT-based vs AI-driven legal translation (from Italian into English) by focusing on an extract of an informative article on mortgage contracts. To this aim, the outputs produced by DeepL, ChatGPT, Copilot, and Notebook LM²⁴ were evaluated. To assess the automated target texts,

16. ALTAKHAINEH–ALGHATHIAN–JARRAH 2025; GIAMPIERI 2025; LI–WANG 2025; BRIVA-IGLESIAS–CAVALHEIRO CAMARGO–DOGRU 2026.

17. ALTAKHAINEH–ALGHATHIAN–JARRAH 2025.

18. LI–WANG 2025.

19. BADAH–KHALAF–DWAIKAT–ALQBAILAT 2026.

20. LI–WANG 2025.

21. ALTAKHAINEH–ALGHATHIAN–JARRAH 2025.

22. BRIVA-IGLESIAS–CAVALHEIRO CAMARGO–DOGRU 2026.

23. GIAMPIERI 2025.

24. Notebook LM: <https://notebooklm.google/>.

Giamperieri (an expert legal translator with other 10 years' experience) consulted legal dictionaries, the case law, and literature-driven studies in the legal sub-genre addressed. Differently from Briva-Iglesias et al.'s²⁵ findings, the results of Giamperieri's study point to unsatisfactory texts generated by both the MT and AI tools. Namely, automated outputs contained language oddities, grammatical shortcomings, and word-order issues. As highlighted by Altakhaineh et al.²⁶, (intelligent) machines were not capable of disambiguating word meanings in context, and they obviously lacked the knowledge of legal system specificities.

System peculiarities are even more compelling in countries such as Canada, where two official languages and two different legal systems coexist. For these reasons, Girard et al.²⁷ compare MT with AI-driven legal translations and highlight how humans must remain central in the legal translation process in order to guarantee the achievement of high professional standards.

Despite the several strands of the literature examining the accuracy and reliability of intelligent machines' target texts in the legal field, there is still a paucity of study on the effective integration of GenAI/LLM and corpus consultation into the legal translation workflow. More precisely, no literature-driven investigation has considered the possibility to leverage chatbots as concordancers to provide answers on sectoral language patterning on the basis of the sources provided (i.e., texts). In particular, by uploading documents and prompting to produce source-related and source-consistent responses, chatbots could become accurate and user-friendly language investigation and language translation tools. This paper is aimed at unravelling this possibility.

2. Aim of the paper and research questions

In view of the huge strides forward made by GenAI/LLMs, their extensive applications in the legal domain, and the increasing research in legal translation, this paper is aimed at examining the possibility and feasibility of harnessing an LLM sys-

tem as a concordancer for legal translation purposes. Namely, documents addressing a specific legal sub-genre (i.e., certificates of partnership) are uploaded to Notebook LM, which is then queried as a concordancer for translation purposes. By doing so, the LLM platform is employed to produce language patterns and assist users in the translation process.

The questions posed by this paper are as follows: 1) To what extent can an LLM system provide accurate responses when a set of documents is uploaded and the system is queried as a corpus concordancer?; 2) Does the LLM system generate linguistic- and translation-relevant data?, and 3) Is the translation of a certificate of partnership possible by employing an LLM tool in such manner?

3. Methodology

As mentioned, this paper tests the possibility to translate a legal document by employing an LLM system as a corpus concordancer. For the purpose of this study, an extract of an Italian certificate of partnership (*Atto Costitutivo di Società in Accomandita Semplice*) was sourced from the Internet²⁸. The text is 115 word long.

Parallel texts drafted in English (i.e., certificates of limited partnerships) were retrieved from the US Security Exchange Commission portal (www.sec.gov), which is a database featuring authentic legal documents drafted in English. The retrieved parallel documents, namely 45, were uploaded to Notebook LM and used as sources – this is the term used by the LLM to refer to the document(s) provided. The system was then prompted to extract linguistic- and translation-relevant data, such as concordances, collocations, and word usages in context. In this way, Notebook LM could be harnessed as a legal translation resource.

To complete the legal and linguistic analyses, statutory documents were also included in the language resources. The US *Uniform Partnership Act 1997 – Last amended 2013* ("UPA")²⁹ was uploaded to another Notebook LM folder. By doing so, two sets of documents (or corpora) were queried: the first

25. BRIVA-IGLESIAS–CAVALHEIRO CAMARGO–DOGRU 2026.

26. ALTAKHAINEH–ALGHATHIAN–JARRAH 2025.

27. GIRARD–LEHOUX–JOBIN–POMERLEAU 2026.

28. See: <https://www.impresafutura.it/costituzione-e-perizie/fac-simile-atto-costitutivo-societa-in-accomandita-semplce>.

29. Uniform Partnership Act 1997, last amended 2013, available at www.uniformlaws.org.

one featuring 45 certificates of limited partnerships (“CLP”), and the second one containing only one document, i.e., UPA. CLP was sourced from the literature³⁰. As mentioned, the corpus was composed manually by selecting certificates of partnerships from the US SEC portal. CLP featured 6,769 word types (203,390 tokens).

Therefore, for the purposes of this study, authentic parallel documents (namely, CLP) are harnessed to provide samples of language, whereas a statute (i.e., UPA) is used to verify and account for legal principles, as well as to apply the corresponding adequate terminology. Additionally, the English sub-section of the Bononia Legal Corpus (BoLC)³¹ is consulted to verify language and lexicological consistency, as well as the frequency of occurrences of some expression proposed by the LLM system. Furthermore, the Italian Civil Code is consulted; in particular, the articles concerning the regulation of the *Società in Accomandita Semplice* (articles 2313-2324). In such manner, terminological discrepancies can be addressed.

Notebook LM is preferred over other similar LLM systems in view of its unrestricted free access to end-users. Therefore, the replicability of the activities is ensured.

4. Analysis

Example 1 showcases the Italian source text.

Example 1. Italian source text

ATTO COSTITUTIVO DI SOCIETÀ IN ACCOMANDITA SEMPLICE

La società ha per oggetto: [...]

Potrà inoltre compiere tutte le operazioni commerciali, industriali, finanziarie (in modo non prevalente non nei confronti del pubblico ed esclusivamente al fine di realizzare l'oggetto sociale), mobiliari ed immobiliari che saranno ritenute dall'amministrazione necessarie ed utili per il conseguimento dell'oggetto sociale, ed assumere, direttamente od indirettamente, interessenze e partecipazioni occasionali in altre Società od Imprese, sia in Italia che all'estero, aventi oggetto analogo, affine o connesso al proprio, nelle sole ipotesi in cui l'acquisizione risulti strumentale per il conseguimento dell'oggetto sociale, escluso in ogni caso il fine di collocamento presso terzi ed il potere degli amministratori di agire nei confronti del pubblico.

As can be noticed, the source text consists of a very long sentence of 102 words featuring parenthetical information (*in modo non prevalente [...]* oggetto sociale), near-synonyms (*necessarie ed utili; affine o connesso; società od imprese*), various syntactic discontinuities (in parentheses or between

commas), technical terms (*oggetto sociale, interessenze, partecipazioni, collocamento presso terzi*), and formulae (*escluso in ogni caso*). Rendering such peculiarities appropriately might be challenging for an automated system, especially in view of the high specificities of the terms and the complex structure of the long sentence. Consequently, it is worthy exploring whether an LLM-based tool harnessed via targeted queries can produce relevant and consistent terminology and wording.

For reasons of comprehensibility, the Italian source text is divided into several paragraphs, each reported in a table, where automated output is scrutinised. Therefore, the next tables display the reference documents (i.e., CLP or UPA) (first column), the prompt inputted in Notebook LM (second column), and the answers obtained (third column). It will be noticed that prompts are formulated with an increasing level of complexity. In this way, it will be possible to grasp how more comprehensive and flexible wording and language can give rise to different alternative responses. For reasons of space, it will not be possible to display how ineffective or incomplete prompt formulation engenders wrong or inaccurate answers.

As regards the responses furnished by Notebook LM, only the most relevant aspects are displayed

in the tables for reasons of space. It is worthwhile mentioning that the LLM system automatically identifies associations between words based on minimum distance values (e.g., adjacent or close words). In the answers, partnerships' names are anonymised by using “XX”.

30. GIAMPIERI–HARPER 2022.

31. ROSSINI FAVRETTI–TAMBURINI–MARTELLI 2007.

Ref.	Prompt	Response
UPA1	Does the source mention “partnership object”?	The exact phrase “partnership object” is not used as a formal or defined term.
UPA2	Does the source mention “partnership purpose” or “partnership aim” or similar words?	General partnership must have a business purpose .
CLP	Search for “object” or “objects” and tell me the phrases where they appear.	A) XX L.P. is formed for the object and purpose of engaging in any lawful act or activity for which limited partnerships may be formed under the Act and engaging in any related activities. B) The objects of the Partnership [...] include acquiring, managing, and trading various securities and evidences [sic] of indebtedness.

TAB. 1 — *La società ha per oggetto*

Table 1 is concerned with the first line: *la società ha per oggetto*. The term *oggetto* is focused on. Relevant legal-linguistic elements featured in the responses are marked in bold.

As can be gathered from Table 1, “object” is focused on being a calque of the Italian *oggetto*. The first prompt is formulated in UPA (see UPA1) and asks to inform whether the noun phrase “partnership object” is found in the source document. Notebook LM informs that the noun phrase does not appear in UPA. For this reason, a second query is written (i.e., UPA2) prompting whether “partnership purpose” or “partnership aim” (or other similar words) are featured in the document. As a response, the system suggests “business purpose” (see UPA2).

Additionally, the term “object(s)” is queried in CLP by requesting to list the phrases where “object(s)” is mentioned. As can be seen from the response column, phrases such as “object and purpose” and “the objects of the partnership” are found. Both the expressions “object(s) and purpose” and “object(s) of the company/partnership” are retrieved from BoLC, although the former is more frequent.

The queries and the related results shown in Table 1 underscore the importance of formulating accurate queries; the fact that “partnership object” is not featured in the corpus (see UPA1), does not imply that other similar expressions should be ruled out. This suggests that the initial query in UPA must be re-formulated to include more options, as exhibited in UPA2.

Table 2 deals with *potrà inoltre compiere tutte le operazioni commerciali, industriali, finanziarie* and addresses the noun phrase *operazioni commerciali, industriali, finanziarie*. As such a noun does not entail or imply any particular legal aspect, only CLP is queried. As concerns the verb *compiere*, a possible equivalent was featured in the results shown in A of Table 1 (i.e., “engage in”). Therefore, it is not addressed any further.

As observable from Table 2, the prompt requires the listing of “operations” or “activities” (considered synonymous) collocating with possible translations of *commerciali, industriali, finanziarie*. From the response, it is evident that *operazioni* corresponds to “activities” in CLP and it collocates

Ref.	Prompt	Response
CLP	Search for “operations” or “activities” collocating with “business”, and/or “industrial”, and/or “commercial”, and/or “financial” and tell me the phrases or sentences that you obtain.	XX L.P. was formed to engage exclusively in specific business and financial activities , such as developing and operating a hydroelectric generating plant.

TAB. 2 — *Potrà inoltre compiere tutte le operazioni commerciali, industriali, finanziarie*

Ref.	Source term/phrase	Prompt	Response
CLP1	<i>In modo non prevalente</i>	Search for“(non) prevailing”, “secondary” or synonyms referring to any “activity” or “business” and tell me the phrases or sentences that you obtain.	A) The character of the business for XX L.P. includes (iv) conducting Business Activities Ancillary to the businesses set forth in clauses (i), (ii) and (iii). B) Activities [...] convenient, desirable, or incidental to its stated business goals. C) To conduct any ancillary business necessary to accomplish those purposes.
CLP2	<i>Nei confronti del pubblico</i>	Tell me if “towards the public” or “towards the general public” or similar phrases are found in the sources.	A) “Revenue Procedures” are defined as statements of procedure from the IRS that are issued to the general public . B) Information secret does not apply to material that has become generally available to the public.
CLP3	<i>Nei confronti del pubblico</i>	Tell me if and how “the public” or “the general public” or similar phrases collocate with “operation”, “activities” or “business” (or similar).	A) Activity undertaken by a person to raise capital for themselves or another from any public or private source. B) Available to the public.

TAB. 3 — (*In modo non prevalente non nei confronti del pubblico ed esclusivamente al fine di realizzare l'oggetto sociale*)

with “business” and “financial”. Therefore, a fully consistent response is provided.

Table 3 displays prompts and responses related to the parenthetical phrase (*in modo non prevalente non nei confronti del pubblico ed esclusivamente al fine di realizzare l'oggetto sociale*). The following words and word pairs collocating with *oggetto sociale* are investigated in CLP: *in modo non prevalente*; *pubblico*, and *realizzare*. Given the multiple searches, Table 3 features an additional column with the specific phrases or words queried.

The query in CLP1 is aimed at verifying whether *non prevalente* can be rendered as “non prevailing”, “secondary” or any other synonymous word or expression when collocating with “activity” or “business”. As noticeable, the query is rather broad in scope, as it includes possible synonyms. The responses clearly produce “ancillary to” and “incidental to”. This is another case where the LLM system proves to be capable of retrieving and proposing adequate formulae and legal language conventions from the documents provided. Indeed, by searching for “non prevailing” in BoLC, no hits are found. Conversely, the word “secondary” generates 2,416 results, whereas “ancillary” and “incidental” appear 1,590 and 1,690 times, respectively.

Both “ancillary” and “incidental” collocate more frequently with “business” or “activities” than “secondary”. Therefore, as indicated by Notebook LM, “ancillary” and “incidental” are more adequate collocates of “activities” or “business”. In addition, as showcased in CLP1 C, a possible rendition of *realizzare (l'oggetto sociale)*, is “accomplish”.

In CLP2, *nei confronti del pubblico* is queried as “towards the (general) public”. Notebook LM, however, does not provide consistent responses (e.g., “information secret [...] available to the public”), although it mentions “to the general public” (see A and B in CLP2). To obtain more fine-tuned results, the prompt is reformulated so that the system is asked to verify whether “(general) public” collocates with “operation”, “activities”, “business”, or any other synonym (see CLP3). The answers received are slightly inconsistent, as they show “activity undertaken by a person to raise capital [...] from any public or private source”, thus lacking any reference to “activities”, “business” or “operation”. Nonetheless, the responses provided in CLP2 and CLP3 can be considered partly satisfactory for the translation purposes. For example, CLP3 showcases the verb “undertaken”, which is another possible rendition of *realizzare*, whereas CLP2 (as men-

Ref.	Prompt	Response
CLP	Tell me if “movable” and/or “immovable” (or any synonymous terms) collocate with “operations”, or “activities” (or synonyms).	A) Dealing in any manner whatsoever with any real property [...] and any business related thereto [...] alongside any and all other activities . B) All partnership property , whether real or personal , tangible or intangible.

TAB. 4 — Operazioni [...] mobiliari ed immobiliari

tioned) displays “the general public” corresponding to *il pubblico*. Table 4 deals with the modifiers *mobiliari ed immobiliari* pairing with *operazioni*.

As can be noticed, the prompt is particularly flexible and all-encompassing. Namely, it includes “movable/immovable” (and similar modifiers) pairing with “operations” or “activities” and any other possible collocate. Notebook LM provides evidence of “real property” and “personal property” collocating with “business” and “activities”, thus providing coherent and straightforward answers (see A and B in Table 4).

Table 5 illustrates queries and the results related to the Italian relative clause *che saranno ritenute dall'amministrazione necessarie ed utili per il conseguimento dell'oggetto sociale*. In this statement, English correspondences of the term *amministrazione* are searched for in UPA to retrieve correct terminology. Then, *necessarie ed utili* is investigated, although Table 3 already showed “convenient”, “desirable” (see CLP1, B) and “necessary” (see CLP1, C) as possible translation options.

It should be pointed out that the LLM’s response retrieved from UPA is partly inaccurate, as

the paragraph quoted is not (f), but (e). Apart from such slight inconsistency, the information provided is correct. As a matter of fact, the source (i.e., UPA) mentions several times that the business of a partnership is “managed” by its partners. For example, § 306(d) quotes “the failure of a limited liability partnership to observe formalities relating to the exercise of its powers or management of its business is not a ground for imposing liability on a partner for a debt”. This implies that *amministrazione* is best rendered as “management” rather than as the calque “administration”.

Once this salient legal and linguistic aspect is clarified, the query in CLP is aimed at assessing whether and how “necessary”, “useful” or any synonymous adjectives collocate with “operation”, “activities” or “business”, and “managers” or “management” (or similar words). As can be seen, queries have become slightly more complex, thus including various terminological searches and word pairing. The results are very convincing and provide different language solutions, such as “required or appropriate for the management of the business” (see A); “necessary to effectuate the purposes of

Ref.	Prompt	Response
UPA	Are “partners” the “administrators” of a partnership?	UPA §401(f): “Each partner has equal rights in the management and conduct of the partnership business.” This provision establishes that general partners are the managers of a general partnership.
CLP	Tell me if and how “managers” or “management” or similar words collocate with “operation”, “activities” or “business” (or similar) AND “necessary”, “useful” or similar adjectives.	A) All specific powers required or appropriate for the management of the business. B) To do all things necessary to effectuate the purposes of the Partnership. C) The General Manager is empowered to take all actions on behalf of the Partnership that it deems necessary or advisable .

TAB. 5 — Che saranno ritenute dall'amministrazione necessarie ed utili per il conseguimento dell'oggetto sociale

Ref.	Prompt	Response
UPA	Does the source mention “interest” as a partner’s “quota” or “share” or “membership”?	Section 1101 (16): “Interest” means: (A) a share in a business corporation; (B) a membership in a nonprofit corporation; (C) a partnership interest in a general partnership; (D) a partnership interest in a limited partnership; (E) a membership interest in a limited liability company.
CLP1	Tell me if and how “interest” collocates with any verb meaning “buy” or “obtain”.	Purchase [...] of its Partnership Interest
CLP2	Is there the phrase “either locally or abroad” or similar expressions?	A) Inside or outside the State of Delaware. B) Within or without the State of Texas. C) Whether as to the U.S. or any non-U.S. Country. D) In the United States or outside of the United States. E) On a regional, national or international basis.

TAB. 6 — *Ed assumere, direttamente od indirettamente, interessenze e partecipazioni occasionali in altre Società od Imprese, sia in Italia che all'estero*

the Partnership” (where “purposes of the Partnership” corresponds to the Italian *per il conseguimento dell'oggetto sociale*) (see B), and “necessary or advisable” (see C).

The next line to address is rather intricate because of the word *interessenze*. The text reads as follows: *ed assumere, direttamente od indirettamente, interessenze e partecipazioni occasionali in altre Società od Imprese, sia in Italia che all'estero*. UPA is consulted via the LLM system to obtain clarifications on possible renditions of *interessenze*, then CLP is queried as regards the verb *assumere* collocating with *interessenze* and to investigate the expression *sia in Italia che all'estero* (see Table 6).

The term “interest” is investigated in UPA because it is a calque of *interessenze*. The word “share” is under scrutiny because of translation options provided by bilingual dictionaries³², whereas “quota” is searched for because it is a calque (or a cognate) of the Italian *quota*. Art. 2317 of the Italian Civil Code, in fact, mentions that the limited members of a *Società in Accomandita Semplice* are liable up to the amount of their quotas. The response offered by Notebook LM on “interest” is very insightful and clarifies that *interessenze* corresponds to a “partnership interest” in the specific case in point (see UPA D).

In CLP1, the query is aimed at unveiling possible verbs collocating with *interessenze* signifying “buy” or “obtain”. It is, hence, supposed that *assumere* (literally “to assume”) has to be interpreted as “to buy” or “to obtain”, given that such interests are derived (i.e., taken or bought) from other companies or entities. As visible, Notebook LM furnishes “purchase” as a possible verb.

The question in CLP2 explores translations of *sia in Italia che all'estero*, possibly excluding the word *Italia*. Indeed, the prompt requires the system to verify whether “either locally or abroad” or similar expressions are mentioned in the source document. As evident, Notebook LM provides various useful renditions.

The next expression to address is *aventi oggetto analogo, affine o connesso al proprio*. In this phrase, the near-synonyms *analogo*, *affine* o *connesso* are dealt with. As there are no peculiar legal implications or legal principles involved, only CLP is queried (see Table 7).

As can be observed in Table 7, the query aims to verify whether the calque “with a similar, analogous or related purpose” (or any similar expression) is sourced from CLP. Notebook LM indicates two main related solutions, i.e., “similar or related objects” (A), and “relating to and compatible with

32. See https://dizionari.corriere.it/dizionario_inglese/Italiano/I/interessenza.shtml.

Ref.	Prompt	Response
------	--------	----------

CLP	Tell me if there are phrases similar to “with a similar, analogous or related purpose”.	A) Similar or related objects B) Relating to, and compatible with , the purposes.
-----	---	--

TAB. 7 — *Aventi oggetto analogo, affine o connesso al proprio*

Ref.	Prompt	Response
------	--------	----------

CLP	Tell me if “only when the purchase” (or similar phrases) collocates with “is necessary for the purposes of the partnership” (or similar).	A) Necessary or desirable in furtherance of the purposes of the Partnership. B) Incidental to or necessary, suitable or convenient for the accomplishment of the purposes specified. C) So far as such powers and actions are necessary or convenient to . D) Provided that .
-----	---	---

TAB. 8 — *Nelle sole ipotesi in cui l’acquisizione risulti strumentale per il conseguimento dell’oggetto sociale*

the purposes” (B). It can be argued that the responses produced by the LLM system are targeted and accurate. For example, both phrases “relating to the purposes” and “compatible with the purpose(s)” are retrieved from BoLC.

Table 8 addresses the next line that is *nelle sole ipotesi in cui l’acquisizione risulti strumentale per il conseguimento dell’oggetto sociale*. In this sentence, the prepositional phrases *nelle sole ipotesi in cui* and *per il conseguimento (dell’oggetto sociale)* are explored. As no particular statutory definitions are involved, only CLP is investigated.

Table 8 displays Notebook LM’s clear answers to the question. As observable, “only when” (which should translate *nelle sole ipotesi in cui*) is not featured in the parallel documents provided. Conversely, these contain expressions such as “so far as” and “provided that” (see C and D). Also, the queried expression “is necessary for” (which renders *risulti strumentale per*) finds several synonyms, amongst which are “desirable in furtherance of” and “suitable for” (see A and B).

This is another example where Notebook LM proves to be insightful for the retrieval of legal formulaic language. For instance, “only when” is featured 249 times in BoLC, whereas “so far as” appears 12,449 times and “provided that” is found 3,076 times.

The next line to deal with is *escluso in ogni caso il fine di collocamento presso terzi*. In this case, the

expression *escluso in ogni caso* and *collocamento presso terzi* are examined. The noun phrase *collocamento presso terzi* refers to *interessenze* and probably denotes the sale of partnership interests to third parties, i.e., the public. To better frame this from a legal perspective, UPA is consulted along with CLP (see Table 9).

As observable in Table 9, UPA does not provide evidence when prompted to inform whether interests could be offered to the public or third parties. Given the lack of data by UPA, the same question is posed in CLP (see CLP1). The answers are “offer and sell to the public” and “distribute and issue”, which are proposed as verb phrases collocating with “interests” (see A and B in Table 9). In BoLC, the lemmas “offer” and “issue” collocate with “interests”.

CLP2 aims to verify whether *escluso in ogni caso* can be rendered with the calque “excluding in any case” or any similar expression. The system proposes several solutions, amongst which “other than” and “exclusive of” are considered the most adequate (see C and E). Indeed, by searching for “excluding” followed by “in any case” in BoLC, no hits are obtained; by searching for “other than” and “exclusive of”, conversely, 15,157 and 145 occurrences appear, respectively. Therefore, the LLM tool proves to be a valuable resource to retrieve conventional legal formulae and multi-words.

Ref.	Prompt	Response
UPA	Tell me if there are possibilities to offer interests to the public (or third parties) and the expressions that are used.	[None]
CLP1	Tell me if there are possibilities to offer interests to the public (or third parties) and the expressions that are used.	A) Raising Capital through Public Offerings: The General Partner may be authorized to raise capital for the Partnership by offering and selling to the public a specific number of interests. B) Distributing and issuing the Interests.
CLP2	Tell me if there is anything similar to and more formal than “excluding in any case”.	A) Provided that none of such actions will be considered. B) Not including. C) Other than. D) None of which shall be. E) Exclusive of.

TAB. 9 — *Escluso in ogni caso il fine di collocamento presso terzi*

Ref.	Prompt	Response
CLP1	Tell me if I can say “the power of the managers (or partners) to act towards the Public” or anything related, according to the sources.	A) Right or authority to act for or to bind the Partnership. B) Enter into and terminate agreements and contracts with third parties.
CLP2	Is there anything related to “perform actions/acts” “for/to/with the public”?	Offering and selling to the public.

TAB. 10 — *Ed il potere degli amministratori di agire nei confronti del pubblico*

The next and last phrase to tackle is *ed il potere degli amministratori di agire nei confronti del pubblico*, where the verb phrase *agire nei confronti del pubblico* is inquired (see Table 10).

CLP1 explores whether a calque of the Italian phrase (e.g., “the power of the managers to act towards the Public”) appears in the source documents. The LLM tool provides consistent output (e.g., “right or authority to act for [...] the Partnership”, see A, and “enter into [...] agreements [...] with third parties”, see B).

Given that the results of the query in CLP1 do not furnish evidence of “public”, CLP2 aims to examine whether “acts” or “actions” collocate with “public”. Unfortunately, the response is not satisfactory, as no relevant collocations are shown. Conversely, the BoLC generates concordances with “towards the public”.

In light of the analyses carried out in the sections above, Table 11 illustrates the query-based

translation. Alternative renditions (sourced and obtained from the uploaded documents) are written in squared brackets. For reasons of comprehensibility, the Italian text is also reported and both texts are divided into sub-paragraphs. For summarising purposes, Appendix 1 displays a summary of all prompts formulated in this study.

As visible, Notebook LM has allowed the rendition of the source text by providing multiple translation solutions in formal legal language, as also attested by BoLC.

5. Discussion

This paper has tested the possibility for LLMs to be seamlessly integrated into the legal translation workflow. By leveraging on Notebook LM, it has explored the feasibility of uploading parallel documents to the system and of formulating queries in simple language. As a matter of fact, as can

Italian text	LLM-assited translation
<i>La società ha per oggetto: [...]</i>	The Partnership is formed for the object of The objects of the Partnership include [encompass] [...].
<i>Potrà inoltre compiere tutte le operazioni commerciali, industriali, finanziarie (in modo non prevalente non nei confronti del pubblico ed esclusivamente al fine di realizzare l'oggetto sociale), mobiliari ed immobiliari</i>	Furthermore, the Partnership may engage in business, industrial and financial activities, dealing with real or personal property, that are incidental to its stated goals [that are ancillary to its principal businesses/activities], that are undertaken for the general public, are necessary to accomplish its purposes
<i>che saranno ritenute dall'amministrazione necessarie ed utili per il conseguimento dell'oggetto sociale,</i>	and that are deemed as advisable and appropriate to effectuate the Partnership's purposes [and that the management deems required and appropriate to accomplish the purposes of the Partnership].
<i>ed assumere, direttamente od indirettamente, interessenze e partecipazioni occasionali in altre Società od Imprese, sia in Italia che all'estero, aventi oggetto analogo, affine o connesso al proprio,</i>	The Partnership may also purchase interests and shares, either directly or indirectly, form other companies or businesses, either on a national or international basis [inside or outside the State], with similar or related objects [with related and compatible purposes]
<i>nelle sole ipotesi in cui l'acquisizione risulti strumentale per il conseguimento dell'oggetto sociale,</i>	and so far as such purchase is necessary or desirable in furtherance of the purposes of the Partnership [and provided that it is necessary or convenient for the accomplishment of the Partnership purposes],
<i>escluso in ogni caso il fine di collocamento presso terzi ed il potere degli amministratori di agire nei confronti del pubblico.</i>	exclusive of [other than] offering [selling] interests to the Public and without granting the partners [managers] the authority to transact any business with the Public [transact any business on behalf of the Partnership].

TAB. 11 — Italian text and LLM-assited translation

be noticed from the sections above, the queries were written in standard language, thus allowing for non-ICT specialists to access and use such resources. The responses, however, needed expert knowledge, as well as the interpretation of the results in context, and the capability to distinguish related from unrelated contents. Hence, the ability to read and understand contextual and co-textual information was pivotal.

Most of the times, the output produced by Notebook LM proved to be consistent and translation-salient. This was visible, for example, in Table 1, when proposing ad hoc renditions of *oggetto della società* (i.e., “object and purpose” or “objects”).

This study has also foregrounded the pivotal role played by clear and inclusive prompt formulation. Prompts which, for example, did not account for the listing of possible synonyms did not yield consist-

ent results or enough information. For example, the straightforward (and rather constrained) query “does the source mention ‘partnership object?’” in UPA1 of Table 1 did not provide compelling responses. The LLM feedback, in fact, only revolved around a negative response (i.e., “the exact phrase ‘partnership object’ is not used”). Hence, this study highlighted the importance of prompting the listing of synonyms in responses. One of the formulations, in fact, featured “tell me if there is anything similar to and more formal than ‘excluding in any case’” (see CLP2 in Table 9). This more inclusive query allowed for the retrieval of a host of possible translation options, amongst which were “other than” and “exclusive of”.

The system proved to be capable of foregrounding concordances, although not in classical linguistic terms. When prompted to generate phrases and sentences with the words queried, in fact, Note-

book LM aptly produced related contents. For example, in Table 2 the query was formulated in a way to search for “operations” or “activities” collocating with “business” and/or “financial” and list the phrases or sentences obtained. The results were satisfactory, as the system showcased a linguistically consistent target expression such as “engage exclusively in specific business and financial activities”. By doing so, Notebook LM framed relevant strings of text featuring the word(s) searched for, as well as collocations, and word usages in context, although not in a standard computational linguistics fashion.

Clearly, the documents uploaded to Notebook LM functioned as corpus files, thus allowing for the navigation of authentic data in a user-friendly and less technical fashion. Indeed, queries were aimed at listing collocations, as in the prompt “tell me if ‘only when the purchase’ (or similar phrases) collocates with ‘is necessary for the purposes of the partnership’ (or similar)” (see Table 8), or “tell me if and how ‘interest’ collocates with any verb meaning ‘buy’ or ‘obtain’” (see CLP1 in Table 6). In this sense, not only collocations were explored, but also legal thesauri.

To make the searches more legally consistent, as mentioned, UPA was consulted, thus enabling the compliance with legal principles and compelling statutory definitions. For instance, harnessing UPA was pivotal to shed light on the fact that the “administrators” of a partnership are the “partners” themselves acting as “managers”. In this way, calques from the Italian *amministratori* (i.e., “administrators”) or *amministrazione* (i.e., “administration”) could be ruled out (see Table 5). This also underscores that users’ acquaintance with the legal subject-matter addressed is necessary to perform more informed queries and searches. Inexperienced users, in fact, may consider and search for incorrect data as plausible translation options, without questioning the possibility for more adequate and legally consistent terminology to be a better fit. This also corroborates that critical analyses and reasoning are necessary when employing automate translation tools. Hence, informed decisions as to the best translation options allow users to better construct legal patterning by adhering to consolidated legilinguistic conventions.

Queries were also formulated to foreground patterns of formal and formulaic language, as in the case of CLP2 in Table 9 (i.e., “tell me if there is

anything similar to and more formal than ‘excluding in any case’”). Similarly, in Table 3, prompts required the system to search for and propose ad hoc synonyms of “secondary” or “non prevailing”. In this case, Notebook LM skilfully retrieved “ancillary to” and “incidental to” by providing word usages in context (i.e., concordances, although not in the strictest sense). Another case in point concerned the retrieval of translations of *nelle sole ipotesi*. Notebook LM indicated “so far as” and “provided that”, which proved to be particularly frequent in BoLC. Therefore, the system was able to extract and propose conventional legal patterning as formulated in the documents provided.

Sometimes, however, the tool did not yield satisfactory results. Apart from non-complete (or non-inclusive enough) prompts (i.e., the prompts that did not envisage the listing of any possible synonyms or related words/phrases in the responses), the system at times produced unrelated outputs. In one instance, Notebook LM quoted a wrong statutory paragraph. This occurred in UPA, Table 5, where the letter “f” of § 401 was cited in lieu of the correct letter “e”. Albeit seeming a trivial error, referring to a wrong statutory paragraph actually gives rise to misunderstandings, confusions and a probable incorrect application of legal principles. It is self-evident that such circumstances must be eschewed. In another circumstance, Notebook LM did not deliver answers. This was the case of Table 9, when the system was prompted to scan UPA and confirm whether the phrase “possibilities to offer interests to the public (or third parties)” (or similar) was featured in the corpus. The lack of an answer, however, could be due to the fact that UPA itself did not contain such a piece of information, which, conversely, was retrieved from CLP (see CLP1 in Table 9). Such infelicities underscore the importance of consulting multiple reliable and targeted resources.

Other times, the LLM tool did not produce consistent language patterning. When prompted to verify whether “the public” or “the general public” collocated with “operation”, “activities” or “business” in the uploaded documents, the system proposed unrelated sentences that only displayed “any public or private source” or “the public”. Another case of a lack of consistent feedback was visible in Table 10, where the prompt “is there anything related to ‘perform actions/acts for/to/with the public?’” was answered inconsistently (i.e.,

with the phrase “offering and selling to the public”). This implies that queries probably have to be reformulated and/or documents have to be searched for differently, as well as that cross-verification must be performed.

6. Conclusions

This study has proved that an LLM system can be harnessed as an effective translation resource that lists collocations, word usages in context, thesauri, and alternative translation options. Queries must be formulated consistently and ought to envisage the possibility for the tool to display synonyms, in case the specific word or phrase is not featured in the document(s) provided. Indeed, when targeted documents are uploaded to LLMs, the systems can show authentic patterning retrieved from the sources, thus allowing users to tap into field-relevant language.

One of the major advantages of such tools is that prompts need not be formulated in technical language, such as CQL (Corpus Query Language), which is required in several corpus platforms in order to obtain highly targeted results. This implies that queries can feature simple language, thus allowing non-ICT experts to access corpus data. In this way, LLMs make digital resource more affordable and available.

This paper has also proved that the consultation of statutory documents is necessary to guarantee terminological consistency and adherence to established legal principles. Such consultation, however, can still be carried out via an LLM toolkit, which can provide insightful responses, thus expediting the translation process and fine-tuning the search for the most accurate translation candidates. Nonetheless, acquaintance with the legal subject-matter dealt with is a prerequisite for quality assurance. Indeed, lacking sufficient knowledge of legal principles could give rise to inconsistencies, wrong interpretations of words or phrases, and the incapacity to grasp nuanced legal meanings.

Another aspect that needs addressing and should not be underestimated is the representativeness of the documents uploaded to the LLM platform. This signifies that the files where information and data are tapped into must be representative of the sub-domain or sub-genre addressed, otherwise intelligent machines would not be able to provide sensitive and coherent responses that can be harnessed for translation purposes.

Considering the advantages and challenges described above, the results of this study frame LLMs as potentially useful resources that can be employed and consulted as concordancers for technical translations. Hence, they have the capability of assisting users in the translation workflow, provided that users employ critical thinking and are able to discern word meanings, usages, and legal principles in context.

The questions that this paper posed at the beginning were the following ones: 1) To what extent can an LLM system provide accurate responses when a set of documents is uploaded and the system is queried as a corpus concordancer?; 2) Does the LLM system generate linguistic- and translation-relevant data?, and 3) Is the translation of a certificate of partnership possible by employing an LLM tool in such manner? The first question has already been answered: as far as the results of this study are concerned, most of the times Notebook LM has furnished coherent responses. Such responses were often accurate and employable as satisfactory translation options. Therefore, the answer to the second question is affirmative: the LLM-driven output was generally relevant and adequate for translation purposes. Only in a very few cases were responses inaccurate or unrelated to the query formulated or to the context provided. The answer to the third question is also affirmative, i.e., the translation of a certificate of partnership is possible by employing LLMs, as long as users are familiar with the subject-matter dealt with, they consult other reliable sources (such as statutory documents, and/or technical dictionaries), and employ critical thinking. This means that they must carefully ponder each response provided by the AI system and verify whether it actually fits in the target context by also adhering to target system specificities.

The implications of this study for the scientific community are manifold. As mentioned, the analyses carried out in this paper has shown that LLMs can be compared to and harnessed as user-friendly corpus platforms. In this way, by formulating queries in simple language, users can interact freely and continuously with the AI-powered platform, thus fashioning and fine-tuning prompts on the basis of their knowledge and technical background. This has, indeed, opened up new possibilities and widened cost-effective accessibility to digital resources.

This study is not without inherent epistemic lacunas. One of the major limitations lies in the fact that it only addressed the translation of a short extract of a legal text, i.e., a certificate of partnership. By considering an array of legal documents in their integrity (e.g., contracts, notary deeds, judicial proceedings, judgments, statutory documents, etc.), more insightful results could be gathered. However, this initial investigation may pave the

way for further studies, probably exploring other avenues of legilinguistic research. For example, scholars may consider uploading legal dictionaries in LLMs, along with parallel documents and statutes. In this way, all prompts would be formulated against one set of documents only, thus including more legilinguistic resources. Hence, responses could be more complete, coherent, and/or sensible.

References

- J. ALLEN, C. HOGAN (2000), *Toward the development of a post editing module for raw machine translation output: A controlled language perspective*, in “Proceedings of the 3rd International Controlled Language Applications Workshop (CLAW-00)” (Seattle, Washington, USA), 2000
- A.R.M. ALTAKHAINEH, G.A. ALGHATHIAN, M.M. JARRAH (2025), *A Comparative Study of Accuracy in Human vs. AI Translation of Legal Documents into Arabic*, in “International Journal of Language & Law”, vol. 14, 2025
- L. ANTHONY (2026), *AntConc (Version 4.4.2)* [Computer Software], Waseda University, Tokyo, 2026, available at: <https://www.laurenceanthony.net/software.html>
- L. ANTHONY (2025), *Integrating AI technology into corpus-based language learning through ChatAI*, in “Computer Assisted Language Learning”, 2025
- A.M. BADAH, C.N. KHALAF, F.J. DWAIKAT, N. ALQBAILAT (2026), *The usage of artificial intelligence in legal translation: Bridging the gap between law and language*, in “Ampersand”, vol. 16, 2026, 100248
- M. BAJČIĆ, D. GOLENKO (2024), *Applying large language models in legal translation: The state-of-the-art*, in “International Journal of Language & Law”, vol. 13, 2024
- V. BRIVA-IGLESIAS, J.L. CAVALHEIRO CAMARGO, G. DOGRU (2026), *Large Language Models “Ad referendum”: How good are they at machine translation in the legal domain?*, in arXiv, 2402.07681, 2026
- F. DO CARMO, D. SHTERIONOV, J. MOORKENS, J. WAGNER, M. HOSSARI, E. PAQUIN, D. SCHMIDTKE, D. GROVES, A. WAY (2020), *A review of the state-of-the-art in automatic post-editing*, in “Machine Translation”, vol. 35, 2020
- M. ESCRIBE, R. MITKOV (2023), *Applying Incremental Learning to Post-editing Systems: Towards Online Adaptation for Automatic Post-editing Models*, in J. Pan, S. Laviosa (eds.), “Corpora and Translation Education: Advances and Challenges”, Springer Nature, 2023
- P. GIAMPIERI (2025), *Assessing the quality of AI and MT in legal translation*, in “Altre Modernità”, vol. 33, 2025
- P. GIAMPIERI, M. HARPER (2022), *Legal translation: From dictionary to corpus*, Le Penseur, 2022
- M.H. GIRARD, E. LEHOUX-JOBIN, M. POMERLEAU (2026), *The use of MT and AI for translation by legal professionals in Canada. A survey-based study*, in “Digital Translation”, 2026
- X. LI, G. WANG (2025), *Application and Limitations of Large Language Models in the Translation of the Contract Book of the Civil Code*, in “Law and Humanities”, vol. 1, 2025, n. 3
- M. NEGRI, M. TURCHI, N. BERTOLDI, M. FEDERICO (2018), *Online neural automatic post-editing for neural machine translation*, in “Proceedings of the 5th Italian Conference on Computational Linguistics (CLIC-IT 2018)” (Torino, Italy), 2018

- M. NEGRI, M. TURCHI, R. CHATTERJEE, N. BERTOLDI (2018), *eSCAPE: A large-scale synthetic corpus for automatic post-editing*, in “Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)” (Miyazaki, Japan), 2018
- X. NIU, M. CARPUAT (2020), *Controlling Neural Machine Translation Formality with Synthetic Supervision*, in “Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)”, 2020
- R. ROSSINI FAVRETTI, F. TAMBURINI, E. MARTELLI (2007), *Words from Bononia Legal Corpus*, in W. Teubert (ed.), “Text corpora and multilingual lexicography”, John Benjamins, 2007
- D. SHTERIONOV, F. DO CARMO, J. MOORKENS, M. HOSSARI, J. WAGNER, E. PAQUIN, D. SCHMIDTKE, D. GROVES, A. WAY (2020), *A roadmap to neural automatic post-editing: An empirical approach*, in “Machine Translation”, vol. 34, 2020, n. 2
- THOMSON REUTERS (2026), *2026 State of the US Legal Market*, in thomsonreuters.com, 2026
- THOMSON REUTERS (2025), *2025 Generative AI in Professional Services Report*, in thomsonreuters.com, 2025
- THOMSON REUTERS (2023), *ChatGPT and Generative AI within Law Firms*, in thomsonreuters.com, 2023

Appendix 1

Prompt summary

Does the source mention “partnership object”?
Does the source mention “partnership purpose” or “partnership aim” or similar words?
Search for “object” or “objects” and tell me the phrases where they appear.
Search for “operations” or “activities” collocating with “business”, and/or “industrial”, and/or “commercial”, and/or “financial” and tell me the phrases or sentences that you obtain.
Search for “(non) prevailing”, “secondary” or synonyms referring to any “activity” or “business” and tell me the phrases or sentences that you obtain.
Tell me if “towards the public” or “towards the general public” or similar phrases are found in the sources.
Tell me if and how “the public” or “the general public” or similar phrases collocate with “operation”, “activities” or “business” (or similar).
Tell me if “movable” and/or “immovable” (or any synonymous terms) collocate with “operations”, or “activities” (or synonyms).
Are “partners” the “administrators” of a partnership?
Tell me if and how “managers” or “management” or similar words collocate with “operation”, “activities” or “business” (or similar) AND “necessary”, “useful” or similar adjectives.
Does the source mention “interest” as a partner
Tell me if and how “interest” collocates with any verb meaning “buy” or “obtain”
Is there the phrase “either locally or abroad” or similar expressions?
Tell me if there are phrases similar to “with a similar, analogous or related purpose”.
Tell me if “only when the purchase” (or similar phrases) collocates with “is necessary for the purposes of the partnership” (or similar).
Tell me if there are possibilities to offer interests to the public (or third parties) and the expressions that are used.

Tell me if there is anything similar to and more formal than “excluding in any case”.
Tell me if I can say “the power of the managers (or partners) to act towards the Public” or anything related, according to the sources.
Is there anything related to “perform actions/acts” “for/to/with the public”?