

PATTERNS OF HUMAN-MACHINE INTERACTION IN LEGAL AND INSTITUTIONAL TRANSLATION: FROM HYPE TO FACT

Fernando Prieto Ramos¹

Centre for Legal and Institutional Translation Studies (Transius),

Faculty of Translation and Interpreting, University of Geneva

Abstract

After several years of intense technological adaptations in the translation industry, there is a need to take stock of their implications, especially as regards the integration of machine translation (MT) in work processes. This study presents the results of a large-scale survey on the use of machine-generated inputs, particularly through translation memories (TMs) and MT systems, in multiple international organizations. It first focuses on the relevance of legal translation in international institutional settings before comparing patterns of use of computer tools for the translation of legal documents as opposed to other texts in these settings, and how such patterns vary across organizations or depending on translators' profiles. The findings reveal a landscape of widespread "augmented translation", but with the prevalent use of TMs as suitable tools for verifying the relevance and reliability of the sources of previous translations, while MT suggestions are integrated and post-edited approximately two thirds of the time, or slightly less frequently in the case of legal translation. This points to a more cautious approach to human-machine interaction for the translation of legal texts, which is also reflected in several variations between institutions. From-scratch translation is limited to a minority of cases, while the scores per profile groups based on domain-specific backgrounds and experience levels were strikingly similar across institutions.

¹ ORCID: <https://orcid.org/0000-0002-4314-2813>; Email: fernando.prieto@unige.ch.

The factual overview provided serves to debunk some of the myths that have fuelled the hype about MT in recent years.

Keywords: machine translation, translation memories, legal and institutional translation, human-machine interaction, augmented translation, from-scratch translation, international organizations

Resumo

Após vários anos de intensas adaptações tecnológicas no sector da tradução, é necessário fazer um balanço das suas implicações, especialmente no que diz respeito à integração da tradução automática (TA) nos fluxos de trabalho. Este artigo apresenta os resultados de um inquérito em larga escala relativo à utilização de *inputs* gerados por máquinas, nomeadamente através de memórias de tradução (MT) e de sistemas de TA, em várias organizações internacionais. Após algumas considerações sobre a relevância da tradução jurídica em contextos institucionais internacionais, são comparados os padrões de utilização de ferramentas informáticas na tradução de documentos jurídicos, por oposição a outros textos nesses contextos, e a forma como esses padrões variam consoante as organizações ou os perfis dos tradutores. Os resultados revelam um cenário de “tradução aumentada” generalizada, mas com a utilização predominante de MT como ferramentas adequadas para verificar a relevância e a fiabilidade das fontes de traduções anteriores, enquanto as sugestões da TA são integradas e pós-editadas aproximadamente em dois terços dos casos, ou ligeiramente menos frequentemente no caso da tradução jurídica. Isto sugere uma abordagem mais cautelosa da interação homem-máquina para a tradução de textos jurídicos, o que também se reflete em variações entre instituições. A tradução “de raiz” limita-se a uma minoria de casos, ao passo que as pontuações por grupos de perfis de acordo com a formação em domínios específicos e com o nível de experiência são muito semelhantes em todas as

instituições. A panorâmica fornecida por estes dados serve para desmistificar alguns dos mitos que têm vindo a encorajar os exageros sobre la TA nos últimos anos.

Palavras-chave: tradução automática, memórias de tradução, tradução jurídica e institucional, interação homem-máquina, tradução aumentada, tradução de raiz, organizações internacionais

1. Introduction

The use of computer-assisted translation (CAT) tools has been a reality for many years in all segments of the translation sector, including legal translation and institutional translation services. The recent development of neural machine translation (NMT) has also permeated the translation industry rapidly, and has become a very popular tool for professional translators, in conjunction with previously-existing tools and resources, such as translation memories (TMs) and terminological databases. As a result, from-scratch translation is increasingly rare, while the reuse of data from previous translations is becoming commonplace.

Legal texts have often been regarded as less suited than other domains for machine translation (MT) due to the prominent culture-bound components of legal communication, the intricacies of resolving conceptual incongruities between legal systems and the legal implications and high risks involved (e.g., Nunes Vieira et al., 2021, pp. 1522-1525; Sánchez-Gijón & Kenny, 2022, pp. 85-86), especially in the case of binding texts and court settings. For the same reasons, legal translation is generally perceived as a field demanding the human expertise to meet high quality standards, and is reflected in the prioritization of legal texts for revision within quality assurance policies (DGT, 2015, pp. 5-9; Prieto Ramos, 2020a, pp. 471-473).

Recent studies on NMT performance in legal translation have yielded mixed results, including a wide range of issues found in machine output depending on the NMT system

employed, the language pair and the text sample translated (e.g., Killman, 2024; Quinci & Pontrandolfo, 2023; Roiss, 2021; Vigier Moreno & Pérez Macías, 2022; Wiesmann, 2019). Given the potential risks of relying on NMT and artificial intelligence (AI) more broadly in this field, the question arises whether patterns of NMT and CAT tool use for legal translation differ from other translation areas or specializations. We are especially interested in the evolution of practices in institutional translation services, and international organizations (IOs) in particular, not only because of the prominent role of legal translation in these settings, but also because of their advanced capacity to adapt workflows and tools to new technological developments.

Practices and results of NMT use in IOs have been reported in a few cases, led by the European Commission's Directorate-General for Translation (DGT) (on the use of the EU's eTranslation system, see Arnejšek & Unk, 2020; Stefaniak, 2020, 2023). For example, one of the most comprehensive of these studies in terms of language pairs and text types considered found a high level of variability between results per language, but also that the use of eTranslation for the translation of legal acts did not entail more risks than for other text categories with regard to terminology (Stefaniak, 2023). However, the field lacks a broader overview of NMT use for legal translation across IOs.

After several years of intense technological adaptations and significant “hype” about their consequences, there is a need to take stock of actual practices and their implications. To what extent are translators becoming post-editors in their daily work? Are there any significant differences in the use of CAT tools for the translation of legal documents as opposed to other texts in these settings? If so, can they be related to the specificities of legal texts or legal translation? Do patterns of machine-human interaction for legal translation, and for institutional translation more generally, vary across organizations or depending on translators' backgrounds or levels of experience?

Our study addresses these questions as part of a larger project on legal translation in international institutional settings.² It reports the results of a large-scale survey conducted in 2023 on the use of technology in IOs. 474 translation professionals from 24 institutions took part in the survey, including multiple EU institutions and intergovernmental organizations (IGOs). The survey collected data on the profiles of respondents, their working practices and current use of technological tools. We will first focus on the results related to legal translation and translator profiles within the international language services examined (Section 2) before presenting the findings on the use of tools, including the specific patterns identified for the translation of legal texts, as well as variations across organizations and profile groups based on domain-specific backgrounds and experience levels (Section 3).

2. Legal specialization in institutional translation services

In order to contextualize the analysis of patterns of technology use, and determine the relevance of practices and profiles in connection with legal translation in particular, we will first focus on the following indicators: the frequency of legal translation, and the main specializations, academic backgrounds and experience levels of the survey respondents. The breakdown of results per institution will include the large and medium-sized translation services that account for the most statistically-significant cohort groups of survey participants, i.e., within the EU institutions, the European Commission (EC), the European Parliament (EP), the Council of the EU (COEU) and the Court of Justice of the EU (CJEU), with between 53 and 77 respondents; and among the IGOs, the United Nations (UN) with 40 respondents, and the World Trade Organization (WTO) with 18. The remaining services accounted for fewer participants, generally in line with their respective sizes, and were grouped under “other EU institutions” or “other IGOs” according to the respondents’

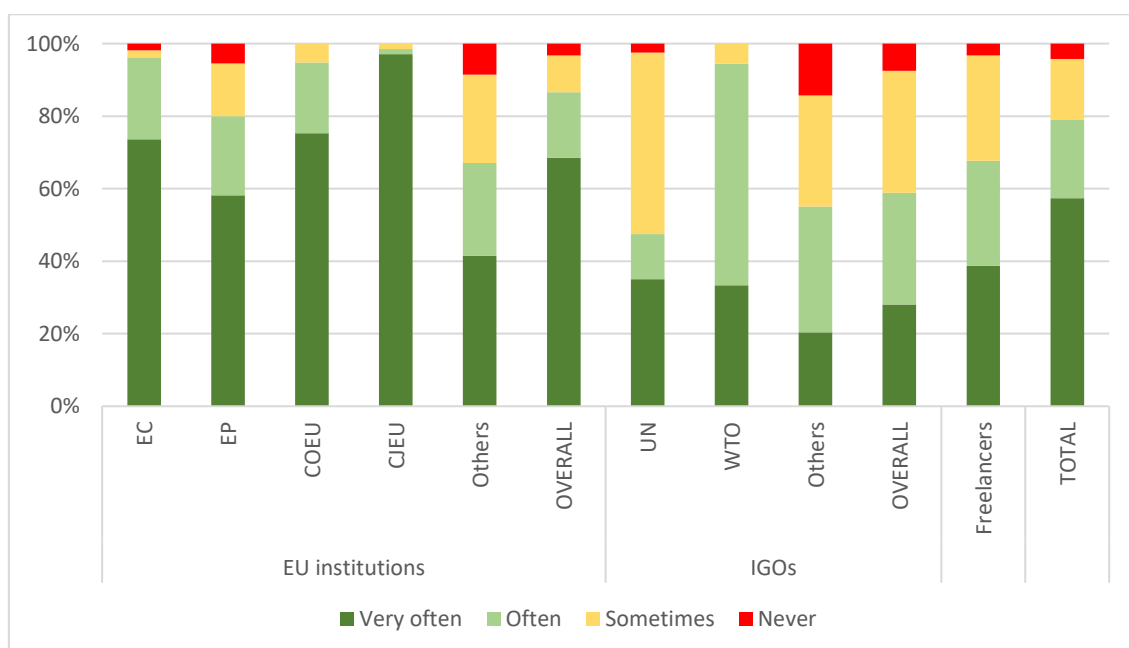
² “Legal Translation in International Institutional Settings: Scope, Strategies and Quality Markers” (<https://transius.unige.ch/letrint/>), led by the author and supported by the Swiss National Science Foundation through a Consolidator Grant.

affiliations. Finally, a cohort of 31 translators regularly working for IOs as temporary staff or freelance service providers, which make up a significant group within the institutional translation landscape, also participated in the survey and were classified as “freelancers”.

Table 1. Frequency of legal translation

	Very often	Often	Sometimes	Never
EC	73.58%	22.64%	1.89%	1.89%
EP	58.18%	21.82%	14.55%	5.45%
COEU	75.32%	19.48%	5.19%	-
<i>EU law-making institutions</i>	<i>69.73%</i>	<i>21.08%</i>	<i>7.02%</i>	<i>2.16%</i>
CJEU	97.10%	1.45%	1.45%	-
Other EU institutions	41.46%	25.61%	24.39%	8.54%
Overall – EU institutions	68.45%	18.15%	10.12%	3.27%
UN	35.00%	12.50%	50.00%	2.50%
WTO	33.33%	61.11%	5.56%	-
Other IGOs	20.41%	34.69%	30.61%	14.29%
Overall – IGOs	28.04%	30.84%	33.64%	7.48%
Freelancers (several inst.)	38.71%	29.03%	29.03%	3.23%
Total	57.38%	21.73%	16.67%	4.22%

Figure 1. Frequency of legal translation



According to their responses, the *frequency of legal translation* (see Table 1 and Figure 1), broadly defined in the survey as the translation of “texts of a legal or administrative nature, either binding or non-binding (e.g., legal acts, treaties, agreements, resolutions, model laws) or texts about compliance with rules, guidelines or obligations (e.g., documents used in monitoring procedures or court proceedings)”, is understandably higher at the CJEU, followed by the institutions involved in the EU’s ordinary legislative procedure (the COEU, the EC and the EP), the WTO (with an important dispute settlement function) and the UN (where monitoring procedures account for the largest proportion of translation volumes). Overall, almost 80% of translation professionals in IOs translate legal texts often (21.73%) or very often (57.38%), and only 4.22% are never confronted with legal texts.

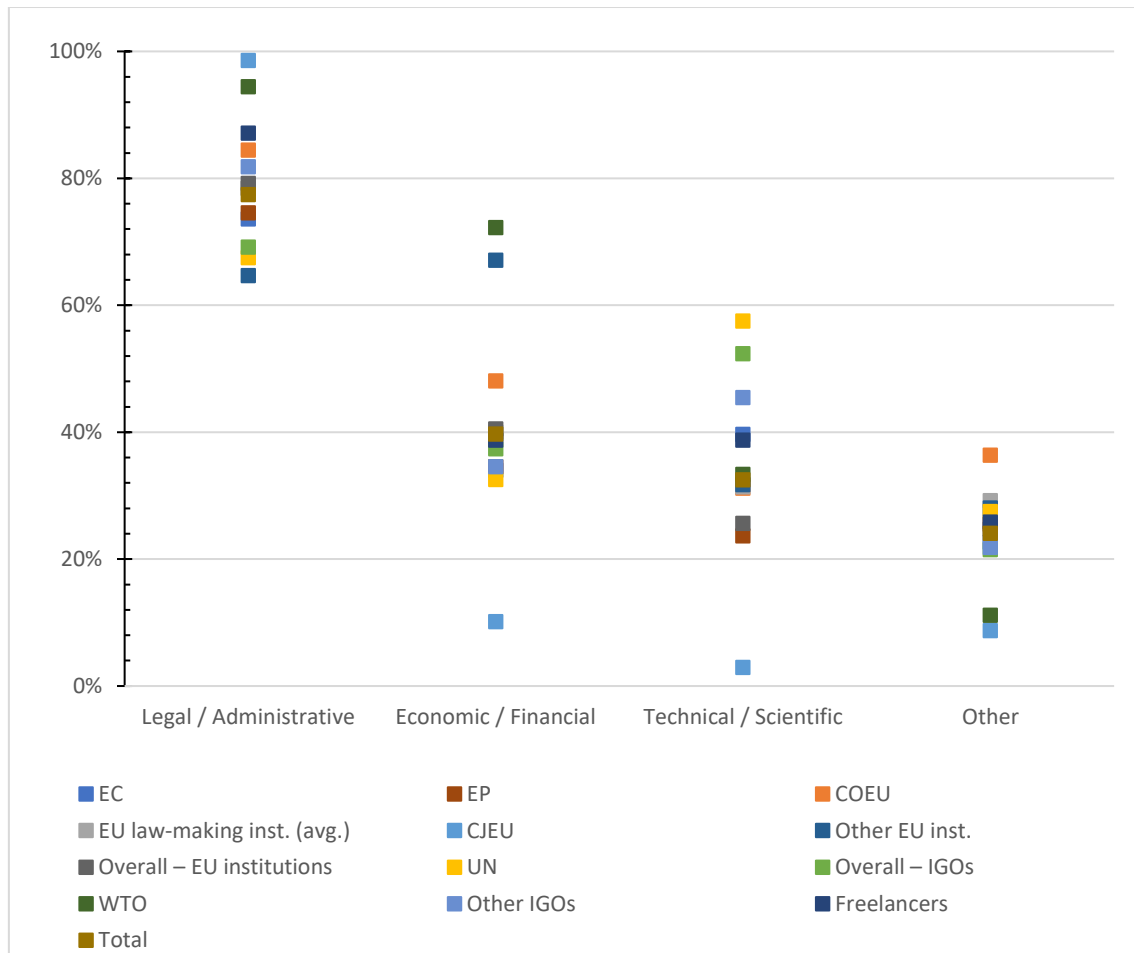
Table 2. Main areas of translation specialization

	Legal / Administrative	Economic / Financial	Technical / Scientific	Other
EC	73.58%	33.96%	39.62%	22.64%
EP	74.55%	34.55%	23.64%	25.45%
COEU	84.42%	48.05%	31.17%	36.36%
<i>EU law-making institutions</i>	<i>78.37%</i>	<i>40.01%</i>	<i>31.35%</i>	<i>29.19%</i>
CJEU	98.55%	10.14%	2.90%	8.70%
Other EU institutions	64.63%	67.07%	31.71%	28.05%
Overall – EU institutions	79.17%	40.48%	25.60%	24.70%
UN	67.50%	32.50%	57.50%	27.50%
WTO	94.44%	72.22%	33.33%	11.11%
Other IGOs	61.22%	28.57%	55.10%	20.41%
Overall – IGOs	69.16%	37.38%	52.34%	21.50%
Freelancers (several inst.)	87.10%	38.71%	38.71%	25.81%
Total	77.43%	39.66%	32.49%	24.05%

These results attest to the prominence of multilingual legal text production at IOs, and align with the findings of an earlier smaller-scale survey (Prieto Ramos, 2020b) and with the volumes of text genres translated in these institutional settings (Prieto Ramos & Guzmán,

2021). This data also reflects how the legal dimension of institutional work conditions the way in which organizations conduct their missions and ensure accessibility and implementation of rules and policies in multiple languages (Prieto Ramos, 2020a).

Figure 2. Main areas of translation specialization



As for the *legal specialization* among institutional translation professionals, the results are also in line with the frequency of legal translation practice. Legal and/or administrative translation was perceived as the main specialization by almost 80% of respondents across institutions (see Table 2 and Figure 2). This was also the proportion registered in EU law-making institutions, with the COEU slightly above the average, while practically all CJEU lawyer-linguists responsible for translating court documents were specialized in legal translation, and IGOs registered an average of two thirds, with the WTO remarkably above

the average (94.44%). This organization illustrates how the legal specialization is often combined with other domain specializations, mainly economic translation in this case (72.22%), also reflecting the thematic diversity of legal texts themselves. Technical and/or scientific translation was the third most frequent specialization overall, and second in the EC and, more clearly, in the UN and other IGOs (upwards of 50%).

In order to complete the analysis of profile specializations, respondents were asked to specify their main *academic backgrounds* and whether they had received training in law or legal translation in particular. Except for the CJEU, where a law degree is required for recruitment, the majority of respondents had completed university studies in translation, with proportions between approximately 60% in the EU law-making institutions and almost 90% at IGOs. Language degrees were the second most common (average of 41.56%), especially at the EU law-making institutions, followed by law degrees. By factoring in legal specialization as part of academic training in translation, participants were classified into four major groups with a view to subsequently analyzing patterns of technology use per profiles:

- LLs: translation professionals specialized in law as their main qualification, including those appointed as “lawyer-linguists” at the CJEU or holding other job titles;
- LTs: translation graduates with some training in legal translation and law ;
- Ts: translation graduates with no training in legal translation or law ; and
- T0s: translation professionals with no academic backgrounds in translation or law.

The largest group was that of LTs (see Table 3), except for the CJEU translation service (exclusively composed of LLs). Nonetheless, variations were significant, as LTs accounted for more than 50% of the total IGO cohort (and even more than 60% at the WTO), as opposed to less than 40% at the EU law-making institutions (except for the EP, at almost 42%), despite the important weight of legal translation, and legislative translation in particular, in the latter group of institutions. However, the number of T0s reached one third

in EU law-making institutions *versus* approximately 10% at IGOs, while the proportion of LLs was generally comparable between these major EU and IGO settings.

Table 3. Distribution of profile groups

	LL	LT	T	T0
EC	13.21%	37.74%	26.42%	22.64%
EP	7.27%	41.82%	16.36%	34.55%
COEU	6.49%	37.66%	15.58%	40.26%
<i>EU law-making institutions</i>	<i>8.64%</i>	<i>38.92%</i>	<i>18.92%</i>	<i>33.52%</i>
CJEU	100.00%	-	-	-
Other EU institutions	18.29%	25.61%	26.83%	29.27%
Overall – EU institutions	29.76%	27.68%	16.96%	25.60%
UN	12.50%	52.50%	25.00%	10.00%
WTO	27.78%	61.11%	11.11%	-
Other IGOs	2.04%	48.98%	34.69%	14.29%
Overall – IGOs	10.28%	52.34%	27.10%	10.28%
Freelancers (several inst.)	19.35%	58.06%	9.68%	12.90%
Total	24.68%	35.23%	18.78%	21.31%

Table 4. Experience in institutional translation

	0-5 years	5-15 years	≥15 years
EC	28.30%	32.08%	39.62%
EP	20.00%	30.91%	49.09%
COEU	9.09%	38.96%	51.95%
<i>EU law-making institutions</i>	<i>17.84%</i>	<i>34.59%</i>	<i>47.56%</i>
CJEU	18.84%	33.33%	47.83%
Other EU institutions	20.73%	29.27%	50.00%
Overall – EU institutions	18.75%	33.04%	48.21%
UN	17.50%	32.50%	50.00%
WTO	11.11%	44.44%	44.44%
Other IGOs	18.37%	30.61%	51.02%
Overall – IGOs	16.82%	33.64%	49.53%
Freelancers (several inst.)	19.35%	41.94%	38.71%
Total	18.35%	33.76%	47.89%

Finally, in terms of *experience in institutional translation*, the distribution of participants is very similar across institutions, with averages of almost one half of translation professionals with 15 or more years of experience, one third with 5 to 15 years of experience, and the rest (approximately 18%) with less experience (see Table 4). According to our institutional informants, this distribution is representative of the composition of translation teams at IOs, and a clear illustration of their gradual renewal. Only the COEU and the WTO cohorts diverged from these averages, with smaller proportions of junior staff and a higher rate of mid- or long-experience groups. The reverse applied to the EC, with above average figures for the least experienced and below average for the more experienced. More specifically, the average experience in legal translation varied between 14.9 years among CJEU lawyer-linguists, 13.1 in EU law-making institutions and 8.6 in IGOs. As mentioned above, this kind of information is, together with domain specialization, extremely relevant for the comparison of technology use among the various profile groups.

3. Patterns of human-machine interaction in legal and institutional translation

As part of the LETRINT project, multiple interviews with translation services managers and diverse exchanges with translation professionals pointed to a marked increase in the degree of human-machine interaction and the diversity of technological tools used since the advent of NMT in particular. However, these interactions were not homogeneous. While there is a growing trend to customize CAT environments by integrating access to the most relevant software, functionalities and resources in a single interface (e.g., eLUNa at the UN), not all institutions have completed such tailoring processes, and not all translators or all tasks rely on the same tools and data sources. Three main combinations of tools emerged as common practice:

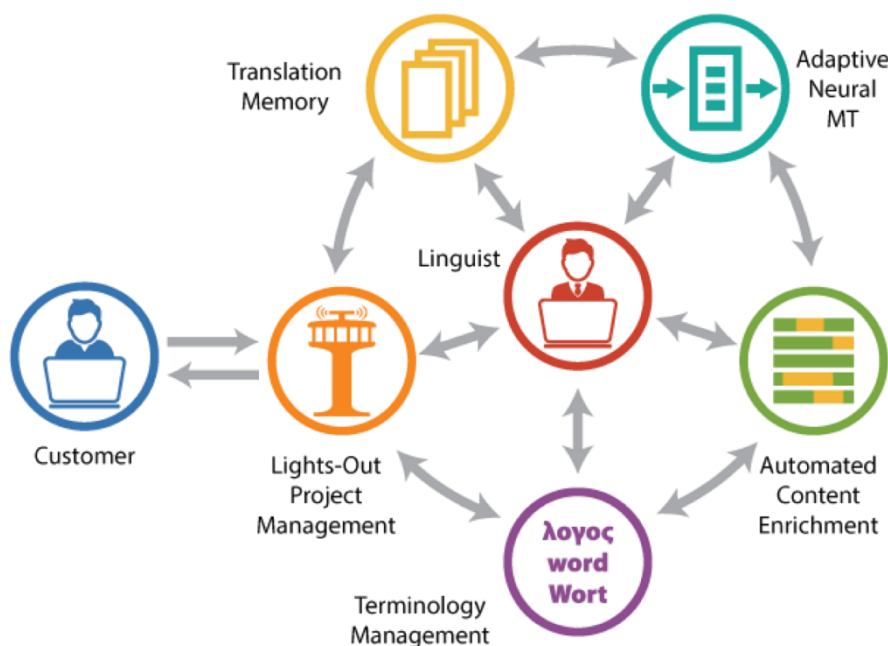
(1) The use of TMs to retrieve previously translated segments for potential reuse in a new translation based on the level of similarity, relevance and reliability of the matching

formulation retrieved (e.g., Euramis in the EU), with integrated or parallel access to other relevant tools and data, such as terminological databases (e.g., IATE in the EU) and text repositories (e.g., EUR-Lex for EU law), and also activating NMT systems³ as an add-on to suggest translations for specific text segments (e.g., eTranslation in the EU institutions; see, e.g., Stefaniak, 2023; Strandvik, 2020). Therefore, this working approach entails integrating post-editing (i.e., editing or reformulation of MT suggestions) in the translation process.

(2) The use of the same environment as in scenario (1), but without activating MT, i.e., only editing TM retrievals but not MT output.

(3) The use of CAT tools as in scenarios (1) or (2), but not relying on automated retrievals of TM matches or MT suggestions.

Figure 3. Augmented translation overview
(from Lommel, 2017, © Common Sense Advisory, Inc.)



The first combination of tools can be considered *augmented translation*, as denominated by Lommel (2017, 2018), since it leverages a wider range of technologies (see representation

³ For the purposes of this study, we will include translations generated through AI-powered large language models (LLMs) within the concept of MT, although their use at IOs has not been reported as frequent at the date of writing.

in Figure 3, including additional software used in institutional translation workflows) and involves a broader diversity of translational or quality assurance actions embedded in the translation process, including post-editing MT inputs. The second scenario can be regarded as a *TM-based intermediate approach*, with a focus on processing TM inputs and a reduced or less diversified human-machine interaction. Finally, the last scenario entails a lower level of machine-generated inputs and a higher dependence on human-driven searches for data management, rather than automated suggestions of text segments for the translation draft, so the approach corresponds more closely to machine-assisted *translation from scratch*.

All three approaches may co-exist in the same translation service or be adopted by the same translator depending on the translation project. Except for cases where MT or TM inputs are not relevant or not recommended because of the lack of suitable resources, international translation services do not generally impose the approach to be applied in every situation, as long as the final translation product is consistent with the applicable conventions and meets quality requirements. It is normally up to translation professionals, rather than the institution, to decide based on their expert assessment of TM or MT reliability for particular text types, as well as risk analysis or time efficiency considerations. Our survey was precisely conceived to gather data about fluctuations in the use of machine inputs for the translation of legal texts as opposed to other texts in institutional settings, and how these patterns converge or differ across organizations or depending on translators' domain-specific backgrounds or translation experience.

When asked about the *working method* that best described their most common practice, the overwhelming majority of respondents (average of 87.95%, see Table 5) chose "I integrate several machine-generated inputs (TM matches and/or MT suggestions) into the translation process and revise them along the way" (i.e., some degree of "augmented translation"). Only the remaining minority (average of 12.05%) selected "I translate or revise without using TMs or MT systems, but rely on other tools and resources, including searches

in relevant databases and text repositories” (scenario 3 above, with no automated segment retrievals). The only significant deviations were found (1) among EP translators (18.52%), which suggests that using machine-generated inputs is less efficient for part of their work⁴ more often than in other EU institutions; and (2) especially in “other IGOs” (average of 20.83%), mostly as a result of practices in the International Criminal Court (ICC), where MT in particular is not normally used due to the nature of the Court’s documents (see also Tomić & Beltrán Montoliu, 2013). Overall, however, from-scratch translation is rare in IOs.

Table 5. Integration of machine-generated inputs as main practice per institution

	Machine-generated inputs (TM / MT)	Only other resources
EC	92.45%	7.55%
EP	81.48%	18.52%
COEU	92.21%	7.79%
<i>EU law-making institutions</i>	<i>89.12%</i>	<i>10.86%</i>
CJEU	89.86%	10.14%
Other EU institutions	89.74%	10.26%
Overall – EU institutions	89.43%	10.57%
UN	86.84%	13.16%
WTO	100.00%	-
Other IGOs	79.17%	20.83%
Overall – IGOs	85.44%	14.56%
Total	87.95%	12.05%

The results per profile groups showed very few differences (Table 6), with fluctuations according to domain-specific background within a range of less than 4 percentage points, and variations per experience level within 2 percentage points. Those with a law background registered a slightly lower frequency of use of machine-generated inputs (85.32%), while LTs

⁴ As noted by our informants, this can be associated with the significant volume of EP amendments to draft legislation proposed in all the EU official languages (i.e., already translated) by the EC as part of the EU’s ordinary legislative procedure. In this context, apart from the actual amended text, the acts must use the wordings of the EC proposals, which is why it is not efficient to rely on MT in particular.

registered the highest frequency (89.63%); and less experienced translators (i.e., the younger generation of institutional translators) resorted to machine inputs slightly more frequently (89.41%) than their more senior peers (with scores very close to the overall average). These results suggest that academic backgrounds or experience levels are not crucial factors when it comes to adapting to technological changes in institutional settings. The findings rather corroborate that the fluctuations in working approaches can be associated to the relevance of tools and resources for specific text types or translation assignments in certain settings.

Table 6. Integration of machine-generated inputs as main practice per profile group

	Machine-generated inputs (TM / MT)	Only other resources
Background		
LL	86.32%	13.68%
LT	89.63%	10.37%
T	86.90%	13.10%
T0	88.00%	12.00%
Experience		
0-5 years	89.41%	10.59%
5-15 years	87.97%	12.03%
≥15 years	87.38%	12.62%
Overall	87.95%	12.05%

As for patterns of *use of TM and MT for legal translation and institutional translation* more generally, frequency levels were assigned a value within a scale of 0 (never) to 3 (more than two thirds of the time) (see Table 7). The average results point to a very prevalent use of TMs (i.e., adapting TM matches) as standard practice for the translation of all kinds of texts (average of 2.68) and specifically for legal texts (2.67) in IOs. This practice is combined with MT (i.e., post-editing MT suggestions for specific segments) just less than two thirds of the time, with an overall average of 1.96 for institutional translation in general and 1.81 for legal translation in particular. This difference primarily derives from the fact that almost 18% of

respondents stated that they never used MT for translating legal texts, as opposed to 12% for institutional translation in general, and the proportion of those who used it most of the time was almost 4% lower in the case of legal translation (39.12% *versus* 43.03% for institutional translation).

Table 7. Frequency of use of TM and MT inputs

	Translation	Never (0)	Up to 1/3 of the time (1)	Approx. 2/3 of the time (2)	More than 2/3 of the time (3)	Average
TM inputs	Institutional	2.44%	4.89%	14.43%	78.24%	2.68
	Legal	3.18%	5.13%	13.69%	78.00%	2.67
MT inputs	Institutional	11.98%	23.47%	21.52%	43.03%	1.96
	Legal	17.60%	22.49%	20.78%	39.12%	1.81

The use of TM retrievals of segments previously translated (from scratch or not) and/or validated by human translators in the same organization is particularly suited for institutional translation given the crucial role of translation precedents in ensuring consistency and adherence to institutional discourse conventions. These requirements are particularly critical in the case of legal texts, as intertextual consistency and semantic accuracy are conditions for legal certainty in the context of supranational law, as with any body of law and any jurisdiction. In practice, this entails verifying the relevance and reliability of sources of segments reused in translation processes, and these sources are indeed provided in TM retrievals. The quality or reliability level of an entire TM can be known to the translator and measured by the host institution depending on the relevance and validation of the data employed to feed the system. As noted by van der Meer and Joscelyne (2017, p. 30), “[t]ranslation memory matching, or leveraging, is certainly very productive in all cases where people work on updates of documents”, which is often the case in institutional translation.

In contrast, MT systems do not provide the sources for their specific suggestions and the reliability of their output is thus generally more unpredictable and cannot be assessed by verifying the sources or the level of similarity with previous translations directly in the output.

While MT engines can be trained for better performance using quality data, the connection between specific MT suggestions and their sources resembles a “black box” and requires special vigilance and cognitive effort of the translator, including, often, additional searches to double-check sources and established terminology, among other aspects. All in all, it is not surprising that MT may be deliberately activated or avoided for specific texts or segments based on the potential time gains and risks involved. In our survey, this emerged in the responses to open-ended questions about the tools and working methods applied,⁵ for example:

When we have to use a previous version of a proposal of a legal act (which is often), we tend not to use MT. (Comment 76)

I think there's quite a big difference between translation memory matches and MT system suggestions in that the latter were never previously revised and validated; they very often don't contain the validated terminology or follow internal rules; any references they feature weren't checked for correctness; and any mistakes that the original text may contain wouldn't be addressed. Therefore, human translations and translation memory matches are much more similar in that respect. I think MT system suggestions are completely different and require much more work from the editor/reviser/translator. (Comment 110)

Translation memory matches are the most reliable. (Comment 156)

With translation memory matches from my colleagues, I can assume a certain level of reliability, whereas the quality of the MT output can vary wildly without it necessarily being clear why – one sentence can be nigh on perfect, whereas the next sentence in the same text is utter garbage. (Comment 163)

⁵ Perceptions of the impact of MT use on working processes will be examined in a separate study.

For legal texts with lots of references and footnotes it is sometimes dangerous to rely too much on neural MT or even TMs as there are lots of quotes from other documents that have to be individually searched. A good translator should use the proper mix of CAT tools and personal research/intellect.
(Comment 188)

You just need to make sure you know exactly which parts of a legislative act can be dealt with by post-editing MT (e.g. explanatory memoranda producing no legal effects), which parts need project-specific translation memories and which parts need full-on human intervention based on previous legal references and terminology resources (e.g. enacting terms producing legal effects in Member States).
(Comment 211)

The *results per institution* point to some significant differences (see Table 8 and Figure 4). Within EU law-making institutions, the EC and the EP registered values above the overall averages for each tool (if we do not consider the rate of translation without machine-generated inputs at the EP –see Table 5–), and more particularly TMs, while the COEU scores were the lowest for MT among all the large institutions examined, with frequencies of MT input of just above one third of the time (1.37) for legal translation and between one and two thirds of the time (1.67) for translation work more generally. These rates can be related to the translation of COEU amendments to previously translated legislative drafts (for which MT is less useful –see also footnote 4–), as well as to the politically-sensitive content handled by this translation service in other texts, apart from legal acts (see Hanzl and Beaven, 2017). Nonetheless, TM use frequency at the COEU is close to the overall average, but less frequent in the case of legal texts, presumably, also because of the practical approach to translating certain amendments. Overall, the results point to a high popularity of Euramis TMs within EU translation services, and reflect the requirements of legislative drafting, in particular, the specificities of the EU’s ordinary legislative procedure.

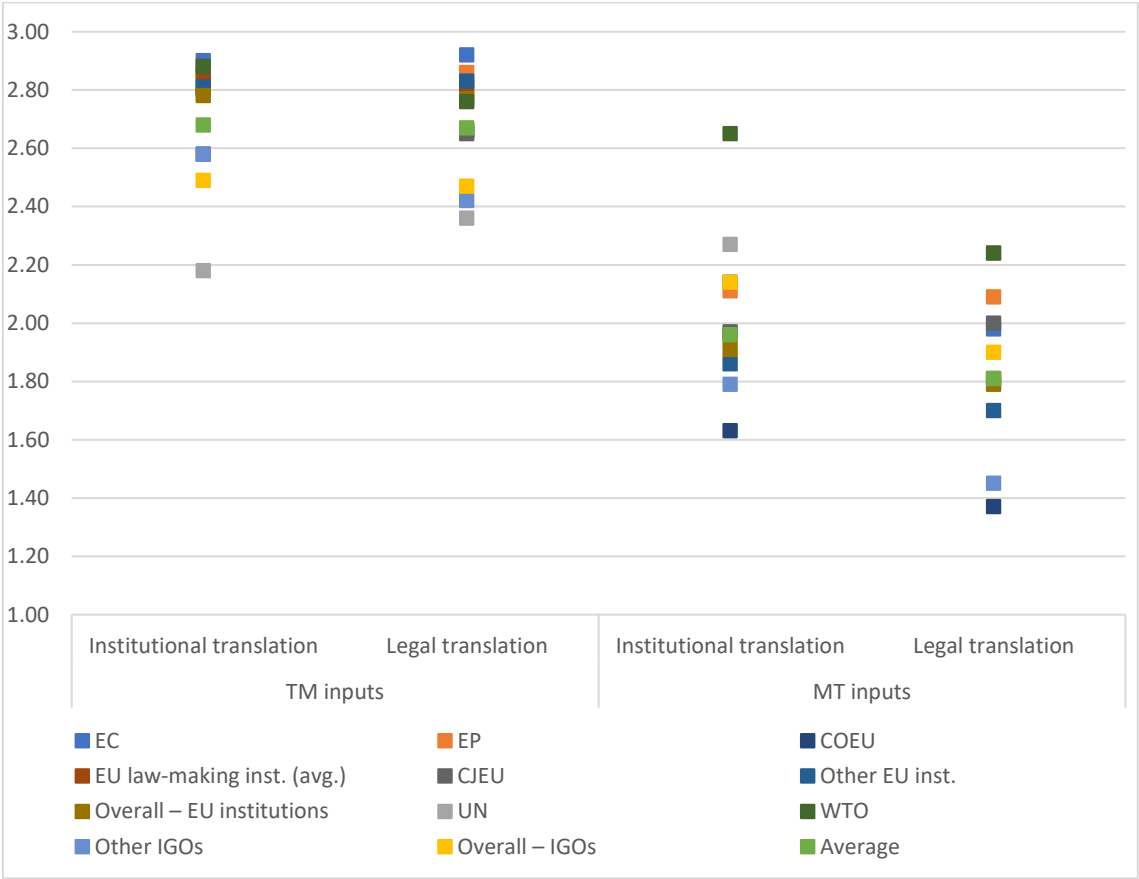
Table 8. Frequency of use of TM and MT inputs per institution

	TM inputs		MT inputs	
	Inst. translation	Legal translation	Inst. translation	Legal translation
EC	2.90	2.92	2.14	1.98
EP	2.86	2.86	2.11	2.09
COEU	2.80	2.66	1.63	1.37
<i>EU law-making institutions</i>	<i>2.85</i>	<i>2.81</i>	<i>1.96</i>	<i>1.81</i>
CJEU	2.58	2.65	1.97	2.00
Other EU institutions	2.81	2.83	1.86	1.70
Overall – EU institutions	2.78	2.77	1.91	1.79
UN	2.18	2.36	2.27	2.24
WTO	2.88	2.76	2.65	2.24
Other IGOs	2.58	2.42	1.79	1.45
Overall – IGOs	2.49	2.47	2.14	1.90
Average	2.68	2.67	1.96	1.81

The second lowest frequency score of MT use was found among “other IGOs” with medium or small language services (1.45 for legal translation and 1.79 for institutional translation). However, the lowest score for TM use was found at the UN, even if the frequency is still high at more than two thirds of the time (overall average of 2.18 and 2.36 for legal translation). The UN is also the only institution where the frequency values for TM and MT converged clearly, with an overall score of 2.27 for MT inputs in translation work in general. Yet, the pattern of more reliance on TM use in the case of legal texts also applies to this organization (frequency of 2.36 *versus* 2.24 for MT). The positive difference in TM use between legal translation and institutional translation in general at the UN, albeit limited, is actually the most marked of all institutions (+0.18). These findings suggest an effective combination of TM retrievals with suggestions generated by commercially available MT engines in the UN’s CAT tool, eLUNa; and perhaps more crucially, they can also be partly associated to the higher proportion of documents used in monitoring processes, including

multiple national reports, as compared to more legislative texts in the EU law-making institutions and case-law at the CJEU.

Figure 4. Frequency of use of TM and MT inputs per institution



The most frequent use of MT is found at the WTO, traditionally among the IGOs at the forefront of technological adaptations (see Pasteur, 2013). As opposed to the UN, however, the use of TM remains more frequent for all text types (score of 2.88 *versus* 2.65 for MT). The WTO thus illustrates well the trend towards augmented translation, while still aligning with the pattern of less frequent use of MT in the case of legal translation (frequency score of 2.24). The difference with MT use in institutional translation in general (-0.41) is actually the most marked for any organization, even if both values are above the average for all IOs. The only institution where MT use frequency is practically the same for the

translation of legal texts and all other texts is the CJEU, precisely because the bulk of the work of the Court's translation service is legal translation.

Table 9. Frequency of use of TM and MT inputs per profile group

	TM inputs		MT inputs	
	Inst. translation	Legal translation	Inst. translation	Legal translation
Background				
LL	2.58	2.60	1.98	1.89
LT	2.71	2.70	1.94	1.79
T	2.70	2.64	2.04	1.82
T0	2.75	2.69	1.89	1.76
Experience				
0-5 years	2.58	2.61	1.86	1.76
5-15 years	2.75	2.76	1.96	1.82
≥15 years	2.68	2.62	1.99	1.83
Overall	2.68	2.67	1.96	1.81

Finally, as opposed to institutional variations, the *results per profile group* were strikingly homogeneous, and did not deviate more than +/-0.10 from the overall averages in any cohorts based on academic background or experience level (see Table 9). Interestingly, LLs seem to integrate TM matches slightly less often than other groups, perhaps as legal profiles tend to double-check primary legal sources more regularly (see also Prieto Ramos, 2020b). As for experience levels, more recent recruits, while adopting machine-generated inputs most often (see Table 6), actually registered frequency scores slightly below the average for each tool. Although the difference is also insignificant, it might be related to early career adaptations to institution-specific tools, including the need to make more verifications to gain knowledge about house style and source reliability, and ultimately develop further confidence for dealing with the specificities of each dataset and machine input type.

4. Conclusions

The data obtained through our large-scale survey corroborates the prominent role and frequency of legal translation and legal specialization within translation services in IOs, thus providing an ideal professional ground to analyze patterns of human-machine interaction in legal and institutional translation. The trends elicited point to the widespread prevalence of “augmented translation” in these settings, including the use of TMs by default for most of the translation work and post-editing MT inputs along the way approximately two thirds of the time, or almost two thirds of the time in the case of legal translation. From-scratch translation is limited to specific cases and the main approach for an average of only 12.05% of translators.

This brings us to our first central question: *To what extent are translators becoming post-editors in their daily work?* Our findings reveal a landscape in which MT is part of a wider range of CAT tools and resources, and it is used or not depending on the needs and relevance for each particular translation assignment and setting. MT can thus be described as a “joker up the sleeve” (Strandvik, 2020, p. 466) that co-exists with TMs and other technologies, but is not always employed. Humans remain “in the loop” and decide when and to what extent post-editing is integrated in the translation process in light of several quality, risk and time-efficiency considerations.

This debunks the idea that post-editing is the new norm in institutional translation services. In fact, none of the institutions examined applies an approach exclusively based on post-editing of MT-generated drafts, but only MT as an add-on and optional post-editing embedded in translation work. According to our institutional informants, few organizations have recently adopted light post-editing (i.e., fixing the main issues in MT output) only for specific texts intended for limited use or involving low risk, and even fewer institutions have reported the use of unedited MT for informative or internal purposes that would not normally be processed through the institutional translation services anyway. In other words,

MT may assist, or in some cases complement, rather than compete with, human translators, whose skillset now comprises post-editing. This departs from some of the myths that have fuelled the hype about NMT improvements, including the notion that translators are becoming exclusively post-editors, losing their distinctive role or simply being replaced by machines.

Our findings about patterns of use of TM and MT inputs also serve to reply to our second research question: *Are there any significant differences in the use of CAT tools for the translation of legal documents as opposed to other texts in these settings? If so, can they be related to the specificities of legal texts or legal translation?* The similarities and differences found, as also illustrated by the respondents' comments, can be connected to source relevance and reliability considerations that are critical in legal and institutional translation. Because of the nature and the strict consistency requirements of translation in these settings, institutional translators need to verify the origin and similarity level of translation precedents retrieved by the machine, which is facilitated by using TM systems, but not MT. This TM functionality is crucial in contexts of reuse of documents previously translated and validated by translators in the same organization, including legal texts.

The frequency of TM use is accordingly higher than that of MT in the institutions under examination. The particular sensitivity, potential risks and translation difficulties associated with legal texts make them less frequent candidates for MT than other texts, even if the frequency scores do not differ dramatically. In turn, these results also point to a common core paradigm that is conditioned by quality requirements of consistency and accuracy shared between multilingual legal communication and institutional communication more generally.

Finally, the data gathered in connection with our third question (*Do patterns of machine-human interaction for legal translation, and for institutional translation more generally, vary across organizations or depending on translators' backgrounds or levels of experience?*) clearly suggests that

fluctuations in TM and MT use can be linked to institution-specific text types and workflows, rather than translators' preferences or habits based on individual backgrounds or experience levels.

The correlation between text types and tool suitability is particularly apparent in the case of the EU law-making institutions, where the popularity of TMs reflects the nature and great volume of legislative translation in these institutions, with high levels of intertextuality, and where machine-generated inputs are less relevant for translating EP and COEU amendments to legal act proposals previously translated at the EC. When read in conjunction, the EP's above-average rates of translation without machine-generated inputs (18.52%) and the COEU's comparatively low frequency scores of TM use and MT in particular for legal translation point to a similar pattern that is partly conditioned by the nature of legislative work in these institutions. The score for MT was only lower at the ICC, as MT was not used in this institution due to the apparent lack of suitability for its texts. In contrast, the WTO stood out with all its staff using machine-generated inputs (as opposed to the overall average of 12.05% of respondents not using them) and the most frequent integration of MT, followed by the UN, the only institution where TM and MT use frequencies converged significantly.

These fluctuations contrast with another major finding of the study: interinstitutional frequency scores per translator background and years of institutional translation experience yielded a remarkably homogeneous snapshot of practices, with no significant differences between profile groups. This means that institutional translators of all backgrounds and experience levels (and ages) seem to adapt to technological advances in line with general trends and institutional specificities, rather than based on attitudes towards technology. This finding challenges the common assumption that the younger generations, generally considered more technologically-savvy, are more proficient in using emerging technologies, or that more senior translators are less efficient in embracing technological change.

Overall, the stock-taking exercise presented here provides a unique factual overview of patterns of human-machine interaction in legal and institutional translation. The main commonalities and differences identified can serve as a basis for further analyses of needs, gaps and opportunities with regard to specific tools, contexts or text types. The impact of these advances deserves more attention, as technologies are evolving rapidly and are re-shaping the ways in which we understand and engage with legal and institutional translation.

Acknowledgments

I would like to thank Diego Guzmán for his assistance with data processing, as well as all institutional informants for their kind cooperation in the framework of the LETRINT project.

References

- Arnejšek, M. & Unk, A. (2020). Multidimensional assessment of the eTranslation output for English–Slovene. In A. Martins et al. (Eds.), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (pp. 383–392). European Association for Machine Translation.
- Directorate-General for Translation, European Commission (2015). *DGT Translation Quality Guidelines*. European Commission.
- Hanzl, J. & Beaven, J. (2017). Quality assurance at the Council of the EU’s Translation Service. In T. Svoboda, L. Biel & K. Łoboda (Eds.), *Quality Aspects in Institutional Translation* (pp. 139–153). Language Science Press. <https://doi.org/10.5281/zenodo.1048196>.
- Killman, J. (2024). Machine Translation in the Legal Context: A Spanish-to-English Comparative Product Study of Statistical vs. Neural MT Output. In M. Winters, S. Deane-Cox, & U. Böser (Eds.), *Translation, Interpreting and Technological Change: Innovations in Research, Practice and Training* (pp. 136–156). Bloomsbury.
- Lommel, A. (2017). *How AI Will Augment Human Translation*. Common Sense Advisory.

- Lommel, A. (2018). Augmented translation: A new approach to combining human and machine capabilities. In C. Cherry & G. Neubig (Eds.), *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 2: User Track)* (pp. 5–12). Association for Machine Translation in the Americas.
- Nunes Vieira, L., O'Hagan, M. & O'Sullivan, C. (2021). Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases. *Information, Communication & Society*, 24(11), 1515–1532. <https://doi.org/10.1080/1369118X.2020.1776370>.
- Pasteur, O. (2013). Technology at the Service of Specialized Translators at International Organizations. In A. Borja Albi & F. Prieto Ramos (Eds.), *Legal Translation in Context: Professional Issues and Prospects* (pp. 283–297). Peter Lang.
- Prieto Ramos, F. (2020a). Translation at International Organizations: The Legal and Linguistic Hierarchies of Multilingualism. In M. Ji & S. Laviosa (Eds.), *The Oxford Handbook of Translation and Social Practices* (pp. 455–477). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190067205.013.13>.
- Prieto Ramos, F. (2020b). The Use of Resources for Legal Terminological Decision-Making: Patterns and Profile Variations Among Institutional Translators. *Perspectives*, 29(2), 278–310. <https://doi.org/10.1080/0907676X.2020.1803376>.
- Prieto Ramos, F. & Guzmán, D. (2021). Examining institutional translation through a legal lens: A comparative analysis of multilingual text production at international organizations. *Target*, 33(2), 258–281. <https://doi.org/10.1075/target.21003.pri>.
- Quinci, C. & Pontrandolfo, G. (2023). Testing neural machine translation against different levels of specialisation: An exploratory investigation across legal genres and languages. *Trans-kom*, 16(1), 174–209.
- Roiss, S. (2021). Y las máquinas rompieron a traducir... Consideraciones didácticas en relación con la traducción automática de referencias culturales en el ámbito jurídico. *Trans. Revista de Traductología*, 25, 491–505. <https://doi.org/10.24310/TRANS.2021.v1i25.11978>.

- Sánchez-Gijón, P. & Kenny, D. (2022). Selecting and Preparing Texts for Machine Translation: Pre-Editing and Writing for a Global Audience. In D. Kenny (Ed.), *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence* (pp. 81–103). Language Science Press.
<https://doi.org/10.5281/zenodo.6759980>.
- Stefaniak, K. (2020). Evaluating the usefulness of neural machine translation for the Polish translators in the European Commission. In A. Martins et al. (Eds.), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (pp. 263–269). European Association for Machine Translation.
- Stefaniak, K. (2023). Machine Translation and Terminology: The Experience of the European Commission. In S. Castilho, R. Caro Quintana, M. Stasimioti & V. Sosoni (Eds.), *Proceedings of the New Trends in Translation and Technology Conference - NeTIT 2022* (pp. 134–141). INCOMA Ltd. Shoumen.
- Strandvik, I. (2020). Digital transformation and institutional translation – change and challenges. In J.-M. Della-Zuanna & C. Kurz (Eds.), *Translation Quality in the Era of Digitalization* (pp. 466–494). Elanders GmbH.
- Tomić, A. & Beltrán Montoliu, A. (2013). Translation at the International Criminal Court. In A. Borja Albi & F. Prieto Ramos (Eds.), *Legal Translation in Context: Professional Issues and Prospects* (pp. 221–242). Peter Lang.
- van der Meer, J. & Joscelyne, A. (2017). *Nunc est Tempus: Redesign your Translation Business, Now!* TAUS BV.
- Vigier Moreno, F. J. & Pérez Macías, L. (2022). Assessing neural machine translation of court documents: A case study on the translation of a Spanish remand order into English. *Revista de Llengua i Dret / Journal of Language and Law*, 78, 73–91.
<https://doi.org/10.2436/rld.i78.2022.3691>.
- Wiesmann, E. (2019). Machine translation in the field of law: A study of the translation of Italian legal texts into German. *Comparative Legilinguistics*, 37, 117–153.
<https://doi.org/10.14746/cl.2019.37.4>.