

Is Artificial Intelligence Capable of Validating Language Policy Choices in the Translation of EU Legislation?

Réka Somssich*

Abstract

It is now commonplace that one of the professions seriously threatened to be taken over by the AI is that of the translators. Increasingly professionalised AI-based translation tools have, if not yet displaced translation, already transformed it. Translators are now more involved in post-editing work, i.e. checking the adequacy of the text translated by the AI for content and terminological consistency. This is no different in today's biggest translation hub, the European Union. But using AI-based translation tools also means that the primary linguistic decisions, including the choice of terms, would be made by the tool and not by the translator. However, this feature can pose significant challenges for EU law. On the one hand, because of the autonomy of EU law, new terms are still emerging and need to be properly reflected in the official languages, and on the other hand, in fast-developing fields, legislation is often adopted at EU level before it is drafted at national level, and the terms used in the European acts are therefore first settled there. Given that the language versions of EU law are official and authentic, the wording and terminology of the EU instruments automatically become part of that national language, so language choices of this kind are also language policy and language development choices taken at EU level but with serious national implications. The question is to what extent AI can help in making these choices.

Keywords

European language policy, national language policy, loanwords, adaptation of words, drafting language, neural translation tool

Submitted: 25.02.2025, accepted: 25.02.2025, published: 02.05.2025

*Réka Somssich: Eötvös Loránd University, Faculty of Law, somssichreka@ajk.elte.hu; This article is based on research presented at the 2024 conference on “Legal Translation in the European Union: What are the Challenges in Light of AI?”, organized by EUR FRAPP, ISIT and the University of Torino.

1. Introduction

The EU institutions are today's biggest translation centres: in 2022, the European Commission translated 2.6 million pages,¹ the European Parliament's overall translation output amounted 2,8 million pages² and the Council produced 1.3 million pages of translations.³ However, the resulting documents do not exist in isolation, but in conjunction with each other, embedded in the wider body of EU law. This is the case even if only a part of the translations is legislation, since in fact all EU documents are directly or indirectly related to legislation because such documents can in many cases be the precursor of binding legal acts, often being the first to contain the regulatory alternatives to a given question and thus the first to contain the terminology in a specific area.

This close, web-like interconnection of EU law and EU legal texts required terminological consistency from the outset and therefore could be the biggest winner and beneficiary of translation memories by the time they boomed. TRADOS has been used regularly by the Commission's translation service since the early 1990s, in 2007 the DGT developed its own translation memory the DG-MT. From 2017, a new era has begun, the EU launched its own neural translation tool, E-translation, powered by artificial intelligence (Foti, 2022). Moreover, since the end of 90's increasingly integrated terminological databases have been made available to translators and to the large public. The most significant from them is the IATE (Interactive Terminology for Europe) which is the EU's terminology management system. It was launched in 1999 "with the aim of providing a web-based infrastructure for all EU terminology resources, thus enhancing the availability and standardisation of their contents".⁴ In the case of IATE, the separate institutional databases have been integrated into a single common and public database, which today contains approximately 7 million terms.

The aim of this article is to identify the issues in which AI can help to make language policy decisions in EU law, according to the current state of science and technology. To do this, it first takes stock of the characteristics of EU law that influence language policy choices, in particular how a concept of EU law is presented in each official language (Chapter 2). It will be argued that both the specificities of EU legal acts and the types of concepts that appear in EU legal language are decisive for linguistic decisions. It will then identify where and at what levels language decisions are actually made in EU law, who is in a decision-making position and whether there are in fact conscious language policy choices behind these decisions (Chapter 3). We will then study the extent to which different languages use loanwords or create their own terminology in areas of law that

1 Annual activity report, 2022, European Commission, DGT. Available at commission.europa.eu/publications/annual-activity-report-2022-translation_en (accessed 24 Aug 2024).

2 Annual Activity Report, Directorate-General Translation, 2023. Available at europarl.europa.eu/cms-data/285562/9_TRAD-AAR-2023.pdf (accessed 24 Aug 2024).

3 Available at consilium.europa.eu/media/65104/2022-annual-activity-report-of-the-translation-service.pdf Available at (accessed 24 August 2024).

4 Available at iate.europa.eu/home (accessed 24 August 2024).

deal with new, rapidly evolving and not yet codified issues such as artificial intelligence, digitalisation, data protection, crisis management. In many cases, but not exclusively, the examples are drawn from Hungarian, whose solution is even more interesting because it is an isolated language (Chapter 4). Finally, we will look at the extent to which neural translation tools – in their current state of development – can be used to make EU language policy decisions (Chapter 5).

2. Characteristics of EU Law as an Inescapable Asset for AI-based Translation Tools

Consistency in terminology is one of the most important requirements for any legal language. However, in the case of EU law consistency is important not only for the internal coherence, but also because all official language versions of EU law are equally authentic and capable of producing the same legal effects, therefore the language of EU law must also be horizontally consistent, including at the level of language versions. The authenticity of EU law also means that the terms that appear in the language versions will be fixed in those languages which are at the same time official languages of the Member States. There are currently 24 official languages of EU law, five of which (English, German, French, Dutch, Greek) are official languages of several countries. The terms used in EU law are therefore integrated in these languages, creating thereby a kind of parallel legal language at EU level. However, the EU legal language and the national legal language are not two independent systems that operate side by side. Their relationship is not simple, the way in which they are linked is influenced by the type of the EU instrument.

The language of the EU regulations, because of their directly applicable nature, comes into direct contact with the language of national law, as it becomes part of the law applied at national instances. If the regulation must be accompanied by measures adopted at Member State level, they will also have to stick to the vocabulary of the regulation when referring back to it, except for the so-called “empty terms” (Robertson, 2011: 55)⁵ which are intentionally left vague in order to be concretised in national law. In this way, the terms used in the regulation (mostly domain specific terms) are automatically incorporated into the national legal language, sometimes refreshing it but sometimes, degrading it.⁶ Therefore, the terminology used in the regulations is not indifferent, as this will be a given thing for national law, which cannot be derogated from.

⁵ As an illustrative example for empty terms Robertson cites the *competent authority*.

⁶ The terminology of the recent Digital Single Market Regulations may have brought such an update to the legal languages of the Member States, as some of the terms they introduced have appeared as new words. Degradation however happens when the regulation in question contains technical terminology in a non-place.

The situation is different for directives. In the case of directives, the national legislator has some linguistic leeway when, in the transposition phase, it adapts the terms used by the directive to its own legal language, by means of an intra-linguistic translation (Ruiz-Cortés, 2019; Roberston, 2011; Somssich, 2003).⁷ Of course, this is not always possible. The national legislator can only make use of this possibility if either the directive deliberately uses a term which is intended to cover a wide range of national law specific concepts under an umbrella term or if the term is not used at the same time in directly applicable EU acts, such as regulations or decisions with which the directive is interrelated. Decisions, when they appear as regulatory instruments, are closer to regulations in terms of their linguistic effect. Thus, if a directive is accompanied by a decision at EU level, the national legislator's linguistic room for manoeuvre in the transposition phase is almost lost. Soft law instruments, as we shall see, if they are precursors to later legislation, do not necessarily fix terms in a definitive way. On the contrary, they provide an opportunity for reflection as to whether the legislation ultimately uses this version.

It is clear from the above that, due to the strong and direct impact of EU law on national laws, language policy decisions have in some sense and in some areas been shifted from national to EU level. Terms used or sometimes even created by EU law will become part of the national languages. This is why it is not indifferent by whom and on the basis of what criteria such decisions are taken. However, before going through what language policy decisions are made in the current system of EU law, at what level, and how AI can be useful in this system, it is worth looking at the types of conceptual categories of EU law that need to be represented by appropriate terminology.

The concepts of EU law can basically be divided into two major groups, the group of legal concepts and the group of concepts used by law. By legal concepts we mean those concepts which, in a legal context (and usually based on law or legal doctrine), have a fully separable, specific content and meaning. Some of them, when used outside the legal context, will also carry this characteristic when used in the vernacular, others however may be damaged when affected by incorrect usage. This is usually due to the fact that the terms representing a legal concept have been taken over from the vernacular by the legal language, where they have fixed their meaning, usually with a narrower content but in the public use this precision is not always reflected. Concepts used by law are concepts expressed through colloquial or technical terms retaining their colloquial or technical meaning and whose attachment to the law is not so strong that they become legal concepts in their own right.

This can be common in the case of agricultural or processed foodstuffs, chemicals, etc. and often leads to corrigenda.

⁷ Typical examples of such translation are often found in private law. The various cases of termination of contract or warranty rights are often referred to in the directives by general terms, and it is up to national law to transpose them in a way similar to domestic law.

EU law has its own legal concepts, which were not part of Member States' laws, but were created by EU law, and therefore require a specific presentation at the level of terminology, often with completely artificially created words, word combinations or recalibration of existing words. For national legal language, these terms will clearly convey the EU character. *Framework decision, flexicurity, advocate-general, preliminary ruling procedure, European semester, conditionality mechanism* are good examples for such terms. In the case of these concepts – due to their novelty – the choice of terms is an important moment, since in many cases word creation and language development is actually taking place, and wrong choices cannot be corrected later. The second category of legal concepts of EU law are those concepts that has already existed in national laws and have a similar but eventually modified meaning in EU law but appear in the same terminology. Examples might be concepts that take their precise maybe varying meaning from the national and EU regulatory context such as *company, warranty, contract, invalidity* etc. Although an autonomous legal system,⁸ EU law also imports concepts from national laws where it takes over legal solutions from these legal systems. However, in the EU context, these concepts may no longer have quite the same meaning as in the “law of origine”. Therefore, for the law of origine, the term might become polysemic, for other legal languages – in the case of an unfamiliar legal concept – it is the naming of the concept that poses the challenge. *Timesharing, legal separation, reasonableness* might be raised as examples. Finally, the last group of legal concepts of EU are those concepts that have exactly the same meaning as in all the national laws. This includes basic legal concepts such as *scope, entry into force, publication*. From a translation perspective, these concepts should not be a problem.

As regards the group of concepts used by law, EU law is interesting in two respects. On the one hand, the Court of Justice cases on the interpretation the Common Customs Code have shown that even the content covered by perfectly ordinary colloquial expressions (such as nightwear) does not necessarily coincide in content in different languages.⁹ On the other hand, EU law covers an increasingly wide range of areas, often relying on very specific terminology, where the misuse of a technical term – for example in the case of a directly applicable regulation – can lead to serious financial damage or loss for economic operators. This is the case, for example, when a product is incorrectly named or classified in the wrong place in the Common Customs Tariff¹⁰ or when a product has to be relabelled by the force of EU law with a term not otherwise used in technical jargon (European Commission, 2010: 76). Under EU legislation, the identification of technical terminology is also more difficult because, unlike in national legislation, experts who are familiar with the national terminology of the specific field concerned are not involved in the adoption of the legislation, or only tangentially, and the translators

⁸ See Case C 26/62, *van Gend & Loos* of 5 February 1963 (ECLI:EU:C:1963:1).

⁹ See the famous pyjama cases (Case C-395/93, *Neckermann Versand* of 9 August 1994 (EU:C:1994:318), Case C-338/95, *Wiener SI GmbH* of 20 November 1997 (ECLI:EU:C:1997:552).

¹⁰ See Case C-74/13 *GSV* of 9 April 2014 (ECLI:EU:2014:423).

are primarily responsible for finding the right term. With regard to technical terminology, it is also important to emphasise that in technically fast developing and rapidly changing areas, the relevant terminology is often not developed quickly or quickly enough in national languages and regulation at EU level is sometimes developed sooner or, in some cases, the terminology needed for regulation is created by EU law. We can witness such phenomena in the field of artificial intelligence, digital market regulation or cyber security.

3. Language Policy Choices at EU Level

In view of the specificities of EU law described above, it is therefore necessary to determine, for the purposes of our topic, what language policy decisions are taken at EU level, who is involved in making them, and to what extent these decisions are characterised by awareness. A language policy decision may arise in relation to terms that do not yet have an established and usable equivalent in the language concerned, either because the concept to be covered is not yet known, or because the term is not yet fixed for the concept, or because of the need to distinguish, at linguistic level, between an EU concept and a similar but not identical national legal concept. In these cases, the freedom and responsibility of the translator of the text is opened up. In other words, this is where human judgement can play a greater role. In all other cases, existing EU or national legal and technical vocabulary should be used in order to establish the various language versions of EU law.

What are the language policy choices open to those involved in the production of the language versions when they have the opportunity to make real choices? The primary question is whether the given language follows in national context an approach of linguistic purism, i.e. is resistant to foreign expressions and linguistic borrowing and whether this approach should be followed even at European level what is not necessarily the case.¹¹ Yet it seems that languages that consistently avoid the use of foreign words are often resistant to their use, even in the EU context (Fischer, 2023). The issue is complicated when the same language is used by several countries and they may have different approaches, duplicating (or even multiplying) the ways in which these languages appear. Even in the case of an express or conscious language policy, the question is whether the given language has a structure and character that is easily receptive to foreign words. Slavic or Finno-Ugric languages are by their very nature not among these languages. In the Finno-Ugric languages, and especially in Hungarian, as we will see

¹¹ For example, a strategy document issued prior to Croatia's accession for the translation of EU legislation explicitly recommends the adoption of words of Latin and Greek origin, or of foreign origin in general, which, in the direction of internationalisation, guarantee that the term concerned will certainly mean the same thing (Robertson & Mac Aodha, 2023).

later, there is a clear tendency, even at EU level, to avoid the influence of the drafting language in the case of newly created terms. Between Indo-European languages there is obviously greater interoperability given the fact that the drafting (source) language of EU text is English.

Time is also a crucial factor in finding the right linguistic expression and making a language policy decision. In the case of any language, it is useless for the translator to strive for linguistic purity and avoid foreign words if the foreign language term is already in actual use in the vernacular or in the professional community, and there is little chance that EU law – despite its binding nature – can supplant it or replace it.¹²

Another important policy decision is word creation, which occurs when a completely new concept appearing in a new term is created and, in some cases, the translator of the language version concerned does not want to use it as loanword or borrowing is not possible anyway. In such cases, an important aspect should be how the newly created words would be integrated into the language concerned. To denote new concepts, some languages can even reuse terms that have fallen out of the legal language and give them new meanings (European Commission, 2010: 83).¹³

But language policy decisions should not only be seen from the perspective of the national language. EU law, and indeed the EU itself, can have its own language policy objectives. For example, in certain cases, clarity of language, interoperability of languages and unambiguous recognition of symbolic words may be desirable. This may also justify, for example, in the case of Latin-based words (i.e. words that do not reflect the influence of the drafting language), promoting word borrowing where possible even for languages that follow a stronger policy of language purism. In many cases, this is more feasible because such languages are less resistant to Latin words (which have often been so accommodated over the centuries) than to English terms. Moreover, it is submitted that in the case of loanwords adapted to the spelling and pronunciation of the receiving language, the foreign origin of the word is no longer obvious (van der Sijs, 2004: 16).

Under the current system of EU law, who makes these language policy decisions and how? To answer this question, we need to look at the process by which EU law is created. From a linguistic point of view, the creation of an EU legal norm must be monitored from the drafting of each document until it becomes official. The official nature of the document means that the official language versions are equally authentic, i.e. they are equivalent in content. This equivalence does not therefore imply that each language version is a translation, but that the language versions are presumed to have been produced

¹² The term monitoring is a good example for Hungarian language as by the time an equivalent was proposed in order to replace the English term, the term was so widespread that the new term could not be admitted.

¹³ In Hungarian, for example, the word “*törvényszék*” was used to refer to courts in the period before World War II, but after that it disappeared from the legal language. As an unused term it could be brought back to distinguish the General Court. Of particular interest is that after 2010, Hungarian courts were renamed back to their pre-war traditional names, so that the unique character of the word ceased to exist and it became a term in use again.

independently and simultaneously. However, practice shows that parallel drafting of the 24 language versions of EU law is only a fiction, and that in fact it is through translation from a source language (usually English as drafting language) to other language versions (Robertson, 2011).

In this system, the producers of the source text (those who prepare the original draft and those who take part at the legislative process) have a particular responsibility for the choice of terms and the definition of the individual concepts as this determines the room of manoeuvre for translators of other language versions. However, in EU law-making, drafting is not a single stage during which the full text of the legal act is produced. This might be true for acts adopted by the Commission, but even there the text of the act is subject to the comitology procedure also changes during the procedure. But it is especially not true in the ordinary legislative procedure, where the draft act prepared by the Commission changes significantly in Council working groups (and eventually at higher instances at the COREPER or Council meetings) and before the European Parliament's committees, and is immediately followed by a translation phase in all institutions and at all phases. The participants at this stage are the following: at the Commission level, the experts who draft the proposed legislation in a language other than their mother tongue, the native-speaking proofreaders who linguistically check the draft in English and the translators who produce the first versions of the proposal in each official language. While those belonging to the first two categories make linguistic decisions only in relation to the original text, although important ones, the translators at the Commission must make the first real language policy decisions. At Council level, we have the experts involved in the Council working groups, who can help to get the terminology right, and the translators, who work from the version already produced by the Commission but can change it. In addition, in all legislative institutions, lawyer-linguists ensure that the legal content of the text is the same at the level of language. Their role is far most important at the Council and Parliament level given the fact that, as a general rule, the Commission only checks the legal and linguistic aspects for the legislation it adopts itself not the proposals submitted for legislative procedure. In the parliamentary stage, apart from the lawyer-linguists, it is mainly translators who play a role in language decisions. The EU's multilingual vehicle is therefore based on a mixed system where drafting, translating activities and ensuring the legal-linguistic consistency alternate (European Commission 2010, 34; Ringe, 2022). However, the system is more than a mere 'source-text—subsequent translations' exercise, where an original unchangeable text is to be transposed as such into other languages. The source text might have to be modified retroactively according to other language versions if these reveal errors or ambiguities in the original (Gallas, 2006: 124).

We cannot, of course, ignore the role of non-institutionalised relationships in making particular linguistic choices: the extent to which, for example, translators and lawyer-linguists of the same language at the different institutions communicate with each

other on a particular problematic term, whether there is a permanent informal cooperation between them, whether they maintain contact with national government bodies. The level of the involvement in language decisions of these latter is not irrelevant either. Government bodies can have an impact on the text of their language version either through informal cooperation with translators or in the formulation of modification to the text in Council working groups or at Council meetings, or ultimately by expressing linguistic reservations. This approach may be systemic, but it may also be *ad hoc*, affecting only certain texts considered more important by the government. Let's recall the declaration made by three Member States on the spelling of the euro and annexed to the Treaty.¹⁴

Language policy decisions therefore in the current EU decision-making mechanism are not taken by a single person, but by a chain of participants. Although the participants in the process are not all translators, as there are also experts, lawyers and lawyer-linguists, all of them are required to have the intercultural competences that are essential for translators (Károly, 2007: 59). In this regard, intercultural competence means not only being able to correctly interpret the legal vocabulary of the languages concerned, but also being able to apply it to the whole of EU law as a separate, autonomous body of law. And the recognition and interpretation of the appropriate context is inherently important in legal translation, and even more so in the EU context.

4. To What Extent is There a Consistent Language Policy Behind the Language Choices in Existing EU Law?

To understand the types of decisions that arise from a linguistic and, more broadly, a language policy perspective, it is worth looking at a few examples to see whether a consistent approach can be demonstrated for certain languages. The examples concern primarily, but not exclusively, the Hungarian language. Hungarian is also special compared to other official languages because, despite being a member of the Finno-Ugric language family, it is an isolated language. The connection between Hungarian and Finnish or Estonian is very remote. In our analysis we looked for examples mainly in areas that are newly regulated or rapidly developing and which, because of these characteristics, did not necessarily have their equivalents in all languages, or where EU law itself could create new concepts. So, we will basically look at terms related to the COVID-19 pandemic, artificial intelligence, digitalisation, data protection. In addition to the relevant secondary legal sources, our research was based on the Union's terminology database, IATE, the EU's legislative database Eurlex and the Union's neural translation tool,

¹⁴ Declaration No 58. by the Republic of Latvia, the Republic of Hungary and the Republic of Malta on the spelling of the name of the single currency in the Treaties.

E-translation. We will also see that the E-translation is not (yet) fully in line with the secondary legislation in force and sometimes proposes a different version for the linguistic equivalent of certain terms than the official version of the EU act. Through our examples, we will investigate whether there is any consistency across languages in the adoption of loanwords, even in an adapted form, whether Latin-based or English.

4.1. The Reception of Terms of Latin Origin

Our first word is *pandemic*. The English versions of Eurlex database currently stores 10 419 documents containing this term. The majority of these were created exclusively after 2020.¹⁵ For example, while there were 2060 documents using this term in 2020 and 2745 in 2021, there were only 35 in 2019. Understandably, *pandemic* became a leading term in legal acts with the COVID-19 crisis. *Pandemic* is a word of Latin origin which all the official languages knew and used either in its original form or in a form adapted to their own language, although most had also their own equivalent, not derived from the Latin version. The use of the term in EU legal documents has been a policy choice for these languages. If we look at the IATE database, we can see that all languages, with the exception of one, have opted for the Latin word and actually use it in legal documents, in addition to the fact that, secondarily, the database also indicates its own equivalent for eleven languages, all of which are translated loanwords of *global outbreak*.¹⁶ The only language that does not use the Latin-based equivalent is Hungarian,¹⁷ where the Latin version is only second in the database and is not used in legal documents at all. It should also be added that the translated term has not been able to displace the borrowed form of pandemic (*pandémia*) in the vernacular from the Hungarian language.¹⁸ However, the Hungarian language is apparently not consistent in its rejection of Latin-based words. Another important word for pandemic, which has become a key term for crisis response, is *resilience*. In all languages, including Hungarian, the Latin term is used in the EU databases in an adapted form, no language has its own version.

At the same time however, we can identify cases – where the Hungarian version could have preferred its own equivalent and other languages have done so – but it has opted for the Latin term instead. An example of this can be found in the field of artificial intelligence law. The word *predictive justice* was first used in a Commission Communication

15 Eurlex Available at: eur-lex.europa.eu/search.html?scope=EURLEX&text=pandemic&lang=en&type=quick&qid=1744654571603 (accessed 23 Aug 2024).

16 Rozšíření onemocnění (CZ), brote de COVID-19 (ES), ülemaailmne koroonaviiruse puhang (ET), maailmanlaajuinen koronavirusepidemia (FI), izbijanje bolesti COVID-19 (HR), взрив от COVID-19 (BG), pasaulinis COVID-19 protrūkis (LT), tifqigħa tal-COVID-19 (MT), 19izbruch COVID-19 (SI), globálne šírenie ochorenia COVID-19 (SK), coronavirus outbreak (EN).

17 The Hungarian equivalent is „világjárvány”.

18 With a simple google search we have more than 849 hits among Hungarian sites for the adapted form of pandemic and 650 000 for the translated term.

in 2018¹⁹ and is still not integrated in the general legal language, we only can find it in a few documents. However, the concept is an important emerging concept of AI law which refers to the use of algorithms and artificial intelligence to predict judicial outcomes. The keeping of the word *predictive* being of Latin origin was at stake for this word combination. According to IATE, 18 of the 24 official languages use the Latin loanword and only German, Slovenian, Estonian, Finnish, Polish, Irish translate it.²⁰ In this case, even the Hungarian language, which is otherwise resistant to loanwords, has taken the Latin term. Note however the endeavour of the Irish language to have its own version.²¹ As it is a new word, found in a total of six Eurlex documents as of August 2024, we looked at what the EU's neural translation tool, E-translation, offers for one of the languages which according to IATE has its own equivalent. For this we chose the German language, for which IATE proposes the term *vorausschauende Justiz*. In E-translation however we found *vorausschauende Gerechtigkeit* instead. Which one is the correct version and which one should be used? In this specific case it seems, it is the AI which needs human correction. *Justiz* and *Gerechtigkeit* mean different things in German, the first referring to the structure of the judiciary, the second to the quality and fairness of judicial decisions. In English, the word *justice* can mean both, depending on the context, but E-translation – still lacking the relevant training and corpus in that regard – has not recognised the context in this case.

Finally, on the subject of Latin terms, the last example shows when a new non-Latin-based term of EU law is reflected by a Latin-based word in one of the language versions. The word is a term used to describe one of the defining concepts of the famous GDPR,²² which is in English *personal data breach*. According to the Regulation “personal data breach means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed”.²³ The term has not been a challenge for most languages, as its elements have been translated literally. A surprising exception to this is the Hungarian language, which denotes the term by *adatvédelmi incidens*, using the adapted form of the foreign word *incident*. In addition, the specific term was introduced in a late version of the draft Regulation, only in the Council's first reading,²⁴ while the original Commission

¹⁹ Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Coordinated Plan on Artificial Intelligence, COM/2018/795 final.

²⁰ *Vorausschauende Justiz* (DE), *napovedno pravosodje* (SI), *prognoosiv õigusemõistmine* (ET), *ennakoiva oikeudenkäyttö* (FI), *maszynowa sprawiedliwość* (PL).

²¹ *ceartas tuarthach* (GA).

²² Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

²³ Article 4 (12) of the Regulation.

²⁴ ST 5419 2016 REV 1 - 2012/011 (OLP).

proposal²⁵ still contained the same wording as the other language versions, that is *personal data breach*,²⁶ as did the European Parliament's resolutions of 2012 and 2013.

The loanword *incidens* used by the Hungarian text does not appear in any of the other linguistic versions, it seems to be unique, so it is worth examining how and for what reason it became part of the core terminology of the Regulation, otherwise at a relatively late stage in the process leading to the adoption of the Regulation. In the Hungarian version of EU law, we first encounter the term *adattvédelmi incidens* (*data protection incident* if translated) in the context of the PNR agreement with the United States,²⁷ although it should be noted that it is used there as the English (American) equivalent of *privacy incident* and not of *personal data breach*. In the PNR agreement, moreover, not only Hungarian but several other official languages use *incident* as part of their own vocabulary: *incident* (FR), *incidente* (IT). It seems, however, that only in the case of the Hungarian language has the term from American law had a lasting effect.

It can be assumed that the Hungarian term integrated in the Hungarian legal language is a loanword from the American legal language which eventually became the designation of the EU concept too. Examining the national legislation, it can be seen that the term *adattvédelmi incidens* appears in Hungarian law in the 2015 amendment of Act CXII of 2011 on the Freedom of Information²⁸ – i.e. before the GDPR was adopted, at the time of the Council's reading, otherwise with a similar content to that defined in the GDPR. In this case, we can therefore see that the functional equivalent, which had been created shortly before the GDPR, superseded the general wording of the term in the drafting language, presumably at the suggestion of national experts in the Council working groups. The term itself as part of the directly applicable GDPR and due to its widespread use became a general and frequently used term of the Hungarian legal language.

4.2. The Impact of the Drafting Language – Terms of English Origin

We will now look at how terms of English origin, often used in their original English version by professionals and experts, appear in EU legal documents. First of all, we will look at the language versions of the term *geo-blocking*. The European Parliament and the Council adopted the so-called Geo-blocking Regulation in February 2018.²⁹ The primary

²⁵ COM/2012/011 final.

²⁶ In Hungarian: személyes adatok megsértése.

²⁷ 2012/472/EU: Council Decision of 26 April 2012 on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security. Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Record data to the United States Department of Homeland Security.

²⁸ Act CXXIX of 2015, § 1.

²⁹ Regulation (EU) 2018/302 of the European Parliament and of the Council of 28 February 2018 on addressing unjustified geo-blocking and other forms of discrimination based on customers' nationality, place

aim of the Regulation was to prevent discrimination based on customers' nationality, place of residence or place of establishment in cross-border transactions between a trader and a customer. Such a discrimination is where traders operating in one Member State block or limit access to their online interfaces, such as websites and apps, by customers from other Member States wishing to engage in cross-border transactions. This practice is called by the Regulation geo-blocking. The term was first used in EU law in a Communication on content in the Digital Single Market in 2012.³⁰ The terminological uncertainties of some of the language versions suggest that the linguistic equivalents of the concept were not yet fully developed, even in larger languages. For example, in the German version of the 2012 Communication, the term *geografische Sperre* is used, whereas this term does not appear in any other EU document and in the 2016 draft Regulation, the English term is used in the German language version, without hyphen, written as one (*Geoblocking*). It should be noted here that, apparently, the German literature also uses a descriptive term for the phenomenon covered by the Regulation (*territoriale Begrenzung audiovisueller Inhalte im Internet*),³¹ but the Regulation has finally chosen to borrow the English term. As far as the other official languages are concerned, most of them have reformulated the pair of words that make up the term by breaking up the English word combination and adapting them to their own linguistic rules, leaving them formally essentially intact.³² In those languages where similar forms of the words concerned are standard expressions (French, Spanish, Italian), this does not cause any problems of reception, whereas in other languages (especially Slavic languages), such a solution may be more difficult to accept. Interestingly, only the German and the Dutch language did not conjugate or transform the term according to their own linguistic rules and kept the English ending, treating it as a full-fledged loanword. These are two languages that certainly had or could have had their own terms for *geo-blocking*, but presumably for reasons of simplicity and identifiability they did not opt for it, just as they refused to conjugate and adapt the loanword according to their own rules. It is also striking that only Estonian, Hungarian and Finnish languages have used different equivalents, which are not adaptations of the original, but descriptive translations or loanword translations, in the case of Hungarian a very much similar description to the German one³³. Of these, however, Finnish is special, as it still consistently and repeatedly uses its

of residence or place of establishment within the internal market and amending Regulations (EC) No 2006/2004 and (EU) 2017/2394 and Directive 2009/22/EC.

³⁰ COM/2012/0789 final.

³¹ See for example Carlo Theiß (2016), *Geoblocking. Die territoriale Begrenzung audiovisueller Inhalte im Internet*, Jura.

³² bloqueo geográfico (ES), geoblokering (DA), blocage géographique (FR), geografiskog blokiranje (HR), gheobhlócaíl (IRL), blocchi geografici (IT), ģeogrāfisko bloķēšanu (LV), geografinio blokavimo (LT), tal-imblukkar ġeografiku (MT), blokowania geograficznego (PL), geoblocare (RO), bloqueio geográfico (PT), geoblokering (SV), geografické blokovanie (SK), geografisko blokiranje (SI), блокиране на географски принцип (BG), γεωγραφικός αποκλεισμός (EL).

³³ asukohapõhise tõkestusena (ET), területi alapú tartalomkorlátozás (HU), maarajoitukset (FI).

own linguistic equivalent in the text of the Regulation,³⁴ but offers a loanword (*geoblokkaus*) in the IATE database as recommended term. However, the term recommended seems to be used only in a limited number of non-legislative communications and European Parliament documents³⁵ and E-translation is not using the loanword either. Its existence still indicates that in Finnish, the own equivalent, although it exists and is used, has not been able to take full root.

In contrast to *geo-blocking*, the term *roaming* was certainly known in national languages long before the so-called Roaming Regulation was adopted in 2007.³⁶ Moreover, it is a term with a wider range of users than those who may have come into contact with the phenomenon of geo-blocking and hence with the term itself. Because of the prevalence of mobile use, *roaming*, in its English form, was already an established and used part of the vernacular speaker's vocabulary in most of the languages when EU legislation came into picture. If we look at the official language versions of the Regulation, we can see that many languages retained the English equivalent,³⁷ while others use their own translation,³⁸ presumably influenced by the extent to which the English equivalent had been adopted in the vernacular – and not just in the official regulatory language – by the time the EU legislation was drafted. The Hungarian language is an exception to this rule, as it also uses a double name in the title of the Regulation: in addition to its own language version, it also includes the English term in brackets.³⁹ However, in the text of the Regulation we only come across the own equivalent, which is the right thing to do, as it would be unfortunate to keep the double name throughout the act, making the text cumbersome. It can also be seen, however, that Finno-Ugric languages – Finnish, Estonian and Hungarian language versions – all preferred their own equivalent to the version of the drafting language. These two examples alone show that the Finno-Ugric languages clearly have a much harder time with the introduction of English loanwords than with Latin words.

But timing of language policy decisions is not irrelevant either. This is well observed in the case of the term *monitoring*. Here, the term emerged in EU law long before, for example, the EU legal vocabulary of the languages of the countries that joined the EU in 2004 was fixed. It can be seen that the languages of the old Member States, such as Danish, German, Spanish, French, Swedish, Portuguese and Finnish, have been able to take

³⁴ maarajoitukset.

³⁵ In Eurlex we have 15 hits for *geoblokkaus*. Available at: eur-lex.europa.eu/search.html?scope=EURLEX&text=geoblokkaus&lang=fi&type=quick&qid=1744654465177 (accessed 23 Aug 2024).

³⁶ Regulation (EC) No 717/2007 of the European Parliament and of the Council of 27 June 2007 on roaming on public mobile telephone networks within the Community and amending Directive 2002/21/EC.

³⁷ The Czech, the Danish, the German, the Italian, the Maltese, the Dutch, the Polish, the Croatian, the Slovak, the Swedish and the Romanian version all use the word *roaming*.

³⁸ Itinerancia (ES), rändlus (ET), itinérance (FR), fánaíocht (GA), itinerância (PT), gostovanje (SI), Verkkovierailu (FI).

³⁹ Barangolás (roaming).

root in their own linguistic versions,⁴⁰ while in the case of Polish, Hungarian, Slovak, Romanian, Bulgarian and Maltese, the English word or a borrowed form of it is used, probably because by the time the own linguistic version could have been developed, the English word was already in professional use.

Another example similar to *monitoring* is *know-how*. *Know-how* became part of not only professional but also everyday language in its original English form by the time the EU legislator adopted a directive on know-how in 2016.⁴¹ This was the case even when national laws tried to introduce a specific term for it. If we look at the language versions of the directive, we see that nine languages have adopted the English equivalent unchanged,⁴² while the others have produced their own translations. It is important to note, however, that in this case, unlike the other terms analysed so far, the term in question is not a term of a Regulation but a term of a Directive, where the national legislator has had the possibility to convert the vocabulary at the level of national legislation. This is exactly what happened with the Hungarian language. The Hungarian version of the Directive, unlike the other Finno-Ugric languages, does not use its own equivalent, but the English term. The transposing national legislation, however, already uses its own term⁴³ and only refers to the English word in brackets, which, given its widespread use, it could certainly not have avoided.

The difficulty caused by English as a drafting language is also apparent in other respects. This is particularly noticeable in the case of EU-level programmes, initiatives or bodies which, thanks to the characteristics of the English language, often have a catchy, easy-to-use name, usually consisting of short but colloquial word combinations.⁴⁴ This kind of flexibility is not common to all languages, however, and these languages are usually faced with the challenge of whether to give the initiative or programme a descriptive, less attractive name, or to retain the English version while ensuring interoperability between languages (European Commission, 2010: 95).

An example for such a translation challenge is a term – AI-Watch – again from the field of artificial intelligence. AI Watch – an initiative of the Commission aiming to review national AI strategies⁴⁵ – is used by most of the official languages in its original version, only some languages – Hungarian, Portuguese, Slovak – have their own mostly descriptive, longer and complicated version, in the case of Slovak with the English name

⁴⁰ Overvågning (DA), Überwachung (DE), seguimiento (ES), suivi (FR), övervakning (SV), acompanhamento (PT), seuranta (FI).

⁴¹ Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

⁴² The Bulgarian, the Czech, the Danish, the German, the Hungarian, the Italian, the Dutch, the Polish, the Romanian, the Slovak and the Swedish version.

⁴³ Védett ismeret (HU).

⁴⁴ See for example *Eurojust*, *NextGeneration EU*, *Clean sky*, *women on board*.

⁴⁵ Available at: ai-watch.ec.europa.eu/index_en (accessed 24 August 2024).

in brackets.⁴⁶ If we look at these descriptive names, one can see that they are definitely more unfortunate, more complicated and harder to remember than the English version. It can be seen that the other Finno-Ugric languages, apart from Hungarian, eventually retained the English name. It is interesting to note that E-translation is lagging behind in this respect as compared to IATE, as it recommends the use of the English name for both Portuguese and Hungarian languages.

However, keeping the English version in some languages is not always a solution. In the case of neologisms denoting new EU concepts, official languages often try to find their own linguistic equivalent, even if it is a challenge for the language in question. Challenges are often due to the fact that English is very innovative and flexible in creating new terms by merging existing terms denoting different concepts into one word and thus creating compound nouns indicating a new concept (European Commission 2010, 95). Specifically such a newly coined word is *flexicurity*. The term indicates a policy strategy the aim of which is to enhance, at the same time and in a deliberate way, the flexibility of labour markets, work organisations and labour relations on the one hand, and security – employment security and income security – on the other. The word is therefore a contraction of the words *flexibility* and *security*. In not all languages can this kind of word combination work in such a way that the resulting word really fits organically into the vocabulary and structure of the given language. According to IATE, German, Danish and Swedish languages also use the original English equivalent in their own languages. Nine other languages have either adapted English to their own linguistic rules and used it as a loanword, or could perform the same word combinations in their own language that English does.⁴⁷ Other languages, however, could not make use of such word combinations, so they either translated the two terms separately or used a circumlocution.⁴⁸ The Lithuanian and Polish languages use the longest of these descriptions (with the English term in brackets in the case of Polish), but we can also see that none of the three Finno-Ugric languages has adopted the English equivalent or a version adapted to its own language.

4.3. New Concepts – Old Terms?

However, the biggest challenge in language policy is not necessarily the use of words of foreign origin, but the linguistic representation of concepts that, due to the autonomy

⁴⁶ mesterségesintelligencia-figyelő (HU), Observatório da Inteligência Artificial (PT), monitorovací centrum pre umelú inteligenciu (SK).

⁴⁷ Flexikurita (CS), flexiguridad (ES), joustoturva (FI), flexicurité (FR), flessicurezza (IT), elastdrošība (LV), flessigurtà (MT), flexizekerheid (NL), flexissegarança (PT), flexicuritate (RO), flexiistota (SK).

⁴⁸ гъвкавост и сигурност (BG), ευελιξία και ασφάλεια στην απασχόληση (EL), turvaline paidlikkus (ET), solúbthacht agus socracht (GA), rugalmas biztonság (HU), darbo rinkos lankstumo ir užimtumo garantijų pusiausvyra (LT), model elastycznego rynku pracy i bezpieczeństwa socjalnego (flexicurity) (PL), prožna varnost (SI).

of EU law, are intended to be broad enough to cover all the specific solutions of 27 legal systems. Recognising such concepts requires contextual interpretation by the translator and a case-by-case decision as to whether the use of a national legal language's own term in an EU-level document would be restrictive or whether, despite the restrictive effect, the use of the term would still be the best choice to make. Such broad EU concepts are for example: *long-term care*, *self-employed person*, *certificate of conformity* which in many languages are translated by a different construction than the national term indicating the Member States' specific appearance of the concept. It is therefore a conscious decision whether the translator or drafter of the text uses the technique of foreignising or domesticating in such cases (Fischer, 2023).

These few examples alone show that a consistent language policy is not entirely visible for each language, but trends can be identified. All languages are naturally more receptive to words of Latin origin, as these words have been infiltrating these languages for centuries, while some languages, especially Finno-Ugric languages, are resistant to words of English origin. However, the prevalence of professional or colloquial language use always overrides language policy and linguistic purism. Every linguistic decision requires an independent assessment, since it is also necessary to consider how well the given word fits into the language in the given context. Borrowing might raise inter-linguistic equivalence but borrowed words might also be a difficulty in intra-language use, conjugation. Developing or using own terms will avoid the language fitting problem, it will however increase the possibility of misunderstanding.

5. How Can AI Be of Help?

The nature and type of the above decisions determine how AI can be used at its current level of development for furthering of terminology of EU law in the official language versions. In this context, it is worth distinguishing between the usefulness of AI as a research tool and AI-based (neural) translation tools. It is safe to say that the use of AI in the research phase can make the work of translators much easier if it can provide documented sources and incidence rates, which can help translators make decisions without contacting national experts personally. This finding is supported by most research (Tan et al., 2020). The main question, however, is whether neural translation tools can provide solutions to the types of dilemmas presented in the previous chapter.

The first important constraint is that neural translation tools operate with a fixed vocabulary, the vocabulary the tool is trained on. Words (or split words in the case of character-based systems) that the tool does not know are not translated or are translated incorrectly. Neural translation tools operate poorly in open-vocabulary settings (Tan et al., 2020: 12). Over the years, many suggestions have been made on how the neural translation tools should handle unknown words, such as using synonyms instead after the tool

was trained on a database of synonyms of words (Turganbayeva & Tukeyev, 2021; Li et al., 2016) but this is clearly not a solution for creating new expressions of EU law in different languages. These tools – apparently at least – cannot create words. Or if they do, this must be a system error. Arnejšek and Unk, looking at E-translation in the context of Slovenian, found that the neural translation tool is highly innovative in creating new words, using non-existent words in place of their existing equivalents in the event the engine encounters words not included in the data sets used in its training (Arnejšek & Unk, 2020: 387). However, these are not consciously generated words, but mistakes. The neural translation tool could not have for instance offered a non-English equivalent in the other official languages for *monitoring* or *flexicurity* that is understandable that fits into the language and is likely to be used, because the tool was not facing at the time of translating these words a fully-used language version and could not have been able to invent one especially with due regard to the above criteria. These decisions can only be made for the time being by the translators, national experts and lawyer-linguists.

It should still be added that in such cases if the neural translation tool leaves the foreign language term untouched in the translated text (instead of indicating the unknown word by a special character), another interesting question arises. The translator might be tempted when doing post-editing of the text translated by the neural translation tool not to look for a linguistic equivalent of the term and to leave the English term as it is raising thereby the number of loanwords.

The same risk threatens when a term does not yet have a mature equivalent in a given language, but there is an existing, first-used version that already appears in some soft law documents. If, with the ever-expanding corpus of EU law, the neural translation tool is constantly being taught, the tool will use this version, the translator producing the first version of the text may not care whether the proposed term is actually the one that should be used when producing the already binding document. Most empirical studies have concluded that translators often overlooked terminological errors or uncertainties at the post-editing stage and preferred to accept the tool's proposal and spotted errors (Sosoni, O'Shea & Stasimioti, 2022). In the case of *geo-blocking*, for example, the German version might have been retained in the text of the Regulation as it was already used in the previous Communications.

The next question is how the neural translation tool can draw on the vocabulary of national laws, if it is trained to do so. The neural translation tool would presumably not have given the Hungarian equivalent *adatvédelmi incidens* of the term *personal data breach* when drafting the GDPR, since the counterpart of the Hungarian term in the national legal corpus is the US legal-based *privacy incident*. The tool would presumably have translated *personal data breach*, as in other languages, element by element. True, in this case, such solution might have been better than the forced, otherwise earlier not widely used national equivalent and would have been more in line with the other linguistic equivalents, ensuring interoperability between them.

A further difficulty with training the tool on the national legal corpus is that this is only possible if the national legislation or part of it is translated. Even in this case, there are still difficulties in the use of specific, emerging concepts of EU law or concepts that have a slightly different meaning in the EU and in the national legal system. Recognising this requires knowledge of the intercultural background, which is a much stronger capacity for context recognition and interpretation than the translation tool is capable of today given the fact that a number of shortcomings of the neural translation tools can still be identified now days. According to Arnejšek and Unk such shortcomings are in the case of E-translation for instance that the neural machine translation includes polysemic misinterpretations and that it often uses generic words instead of domain specific terms (Arnejšek & Unk, 2020: 3868–389). As a conclusion the authors submit that in legal texts terminological errors and inconsistency causes the biggest problem with neuronal machine translation (Arnejšek & Unk, 2020: 389).

The same finding is reinforced in relation to terminological challenges by another empirical study conducted by Sosoni and others where human translations and post-edited machine translations of legal texts (although not EU texts) were compared in the case of Greek language (Sosoni, O’Shea & Stasimioti, 2022). One of the participants of this project said that they would have preferred using computer assisted translation tools instead of neural translation tool (Sosoni, O’Shea & Stasimioti, 2022: 103). Moneus and Sahari arrived at similar conclusions in the case Arabic and English legal texts (Moneus & Sahari, 2024). They found that AI translations lacked legal terms and used alternative simple words instead (Moneus & Sahari, 2024: 75). This shortcoming may of course also be due to poor or inadequate training, but also to a lack of contextual awareness. Skilled human translators will be able to deeply understand the languages and cultures, their judgment and knowledge which makes them able to “make informed decisions about translating idiomatic expressions, metaphors, and other language-specific features” (Moneus & Sahari, 2024: 76). The emphasis here is on informed decisions.

However, there is one segment of the above language choices where neural translation tools may already be useful. For example, if language policy trends can be identified for a particular language, the tool can be trained to follow them. In such an event the tool will suggest own linguistic equivalent for languages avoiding foreign terms, if available. It can also be trained to deal with certain recurring language policy decisions in the same way, for example, to recognise that when naming EU programmes and initiatives with a particularly good sounding, catchy original name, they should not use descriptive translations or own linguistic versions. In this respect, with the help of AI, even greater consistency in certain language policy decisions can be achieved than before.

6. Conclusions

Like many other fields, translation has been revolutionised by AI. Translators are already more post-editing. But not only that. At the current level of development of neural translation tools, language policy decisions in the EU translation mechanism also remain largely in the hands of translators, lawyer-linguists and other participants in the legislative process. The examples above showed that even in languages where there is a tendency towards either word borrowing or domestication, there is no consistency at the level of individual terms. This can be explained by two things. Either because the right language policy decisions can only be made on the basis of *ad hoc* cases, or because there is no coherent language policy at all. In this, AI could bring uniformity, provided that such uniformity is needed. Although not at the current level of sophistication of the tools, it can also help in time to create and suggest new words, offers alternatives. But the choice of what should be the authentic and official text is one of the people, and the responsibility lies with the institutions, ultimately the with EU. The paper tried to point out that even today, in the sixth decade of its existence, EU law is still confronted with serious language policy choices, the correctness of which, whether based on human or non-human decisions, will be judged by time.

References

- Arnejšek, Mateja & Unk, Alenka (2020). Multidimensional assessment of the eTranslation output for English–Slovene. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, Lisboa, Portugal*. European Association for Machine Translation, 383–392. Available at: aclanthology.org/2020.eamt-1.41 (accessed 24 August 2024).
- European Commission: Directorate-General for Translation, Várnai, Judit; Bérczi, Anna & Somssich, Réka, *Study on Lawmaking in the EU Multilingual Environment*, Publications Office, 2010. Available at: data.europa.eu/doi/10.2782/36562 (accessed 24 August 2024).
- Gallas, Tito (2006). Understanding EC law as ‘diplomatic law’ and its language. In Pozzo & Jacometti (Eds.), *Multilingualism and the Harmonisation of European Law* (pp. 119–128). The Netherlands: Kluwer Law International BV.
- Fischer, Márta (2023). *Eurolektusok*. Budapest: Akadémiai Kiadó.
- Foti, Markus (2022). eTranslation. Le système de traduction automatique de la Commission européenne en appui d’une Europe numérique. *Traduire – Revue française de la traduction*, 246, 28–35. DOI: [10.4000/traduire.2793](https://doi.org/10.4000/traduire.2793).
- Károly, Krisztina (2007). *Szövegtan és Fordítás*. Budapest: Akadémiai Kiadó.
- Li, Xiaoqing; Zhang, Jiajun & Zong, Cheqinq (2016). Towards Zero Unknown Word in Neural Machine Translation. *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*, 2852–2858.
- Moneus, Ahmed Mohammed & Sahari, Yousef (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Helyon*, 10(6). DOI: [10.1016/j.helyon.2024.e28106](https://doi.org/10.1016/j.helyon.2024.e28106).

- Ringe, Nils (2022). *The Language(s) of Politics. Multilingual Policy-Making in the European Union*. Ann Arbor: University of Michigan Press.
- Robertson, Colin (2011). Multilingual legislation in the European Union: EU and national legislative-language styles and terminology. *Research in Language*, 9(1), 51–67. DOI: [10.2478/v10015-011-0011-3](https://doi.org/10.2478/v10015-011-0011-3).
- Robertson, Colin & Mac Aodha, Máirtín (2023). Legal Terminology of the European Union. In Kockaert & Steurs (Eds.), *Handbook of Terminology Online* (pp. 244–270). Amsterdam: John Benjamins.
- Ruiz-Cortés, Elena (2019). Language in Supranational and National Law-Making – The Case of Directives and their Transposition into National Law. *International Journal of Language and Law*, 8, 34–49.
- van der Sijs, Nicoline (2004). The role of purism in language development – historical and political aspects. In ó Riagáin & Stolz (Eds.), *Purism – Second Helping* (pp. 1–24). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Somssich, Réka (2003). A jogfogalmi megfeleltetés problémái a közösségi jogban az irányelvek átültetésének szintjén – a jogi „fordítás” sajátos formája. *Magyar Jog*, 12, 746–753.
- Sosoni, Vilemini; O’Shea, John & Stasimioti, Maria (2022). Translating law: A comparison of human and post-edited translations from Greek to English. *Revista de Llengua i Dret/Journal of Language and Law*, 78, 92–120. DOI: [10.2436/rld.i78.2022.3704](https://doi.org/10.2436/rld.i78.2022.3704).
- Tan Zhixing; Wang, Shuo; Yang Zonghan; Chen Gang; Huang, Xuancheng; Sun Maosing & Liu, Yang (2020). Neural machine translation: A review of methods, resources, and tools. *AI Open*, 1, 5–21. DOI: [10.1016/j.aiopen.2020.11.001](https://doi.org/10.1016/j.aiopen.2020.11.001).
- Turganbayeva, Aliya & Tukeyev, Ualsher (2021). The solution of the problem of unknown words under neural machine translation of the Kazakh language. *Journal of Information and Telecommunication*, 5(2), 214–225. DOI: [10.1080/24751839.2020.1838713](https://doi.org/10.1080/24751839.2020.1838713).

Note: JLL and its contents are Open Access publications under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).



Copyright remains with the authors. You are free to share and adapt for any purpose if you give appropriate credit, include a link to the license, and indicate if changes were made.

Publishing Open Access is free, supports a greater global exchange of knowledge and improves your visibility.