Indian Journal of Information Sources and Services
ISSN: 2231-6094 (P) Vol.15, No.2, 2025, pp.83-90

DOI: https://doi.org/10.51983/ijiss-2025.IJISS.15.2.12

The Role of Natural Language Processing (NLP) in Translating Legal Documents with High Accuracy

Mavluda Ahmedshaeva[1], Nodirakhon Pulatova[2], Sherzod Khayitov[3], Dilafruz Sufiyeva[4] and Mohinur Nizomova[5]

[1]Professor of the Department of Theory of State of Law Tashkent State University of Law, Uzbekistan
[2]Senior lecturer of the Department of Theory of State of Law Tashkent State University of Law, Uzbekistan
[3]Associate Professor of the Department of Theory of State of Law Tashkent State University of Law, Uzbekistan
[4]Senior lecturer of the Department of Theory of State of Law Tashkent State University of Law, Uzbekistan
[5]Karshi state university, Uzbekistan
E-mail: [1]mavludaahmedshaeva@gmail.com, [2]dilanor@mail.ru, [3]xayitovsher085@gmail.com, [4]dilafruz.moj@gmail.com, [5]mohinurnizomova90@gmail.com
ORCID: [1]https://orcid.org/0000-0001-8144-0988, [2]https://orcid.org/0009-0004-8741-902X, [3]https://orcid.org/0000-0003-3804-7210, [4]https://orcid.org/0009-0009-8609-8515, [5]https://orcid.org/0009-0002-1872-3322
(Received 19 February 2025; Revised 30 March 2025, Accepted 15 April 2025; Available online 25 June 2025)

Abstract - **The accurate rendition of legal documents is vital to ensure effective communication and justice beyond linguistic boundaries. Given the complexities of legal texts — that involve terminologies with specialized jargon, intricate sentence structures, and culture-bound concepts — standard translation processes fall short of the desired precision. The development of Natural Language Processing (NLP) technologies in recent times has introduced innovative solutions for enhancing the quality and accuracy of legal translations. This paper explains the role of NLP in legal translation, specifically how it helps improve terminological consistency, performs automatic text analysis, and supports human translators. The study explains the marriage between machine translation, corpus-based approaches, and AI tools in the legal field with both the possibilities and limitations of the technology. Through a critical examination, the paper demonstrates how NLP assists in enhancing efficiency, accuracy, and reliability in legal document translation and how there remains a continued need for human expertise.**

Keywords: **Natural Language Processing (NLP), Legal Translation, Machine Translation, Legal Terminology, Accuracy in Translation, Artificial Intelligence, Computational Linguistics, Legal Texts, Quality Assurance, Multilingual Legal Communication**

I. INTRODUCTION

The globalization of businesses and other activities raises issues in terms of the localization of documents. For professional documents involving legal, commercial or other technical matters, translation is one possible option. While the translation between some languages has been facilitated, that is not the case for less commonly used languages. The translation of legal documents involves complex issues including the proper understanding and verbalization of the original lyrics and the use of appropriate specialized terminology. Translations must be done with the highest possible precision to be considered valid in a court of law (Mahari et al., 2023). The development of a generic Natural Language Processing (NLP) system capable of working for all the languages of the entire world in the same way as translation engines for the most prominent languages is a major challenging research task. However, given the imperious need, and as a first step, it is necessary that complex and consistent NLP tools be made available for more languages than those spoken in the major Western countries (Tohma & Kutlu, 2020).

Although the use of translation machines could contribute to a significant acceleration in the process, there are still some major difficulties in ensuring the accuracy of the translation. Language differences cause variations in the general structure of the sentence, as in word order, even in sentences with similar meaning. Terms used in legal texts, usually characterized by a very extensive vocabulary, are easily misinterpreted by conventional translation systems (Hemat, 2014). In addition, a lot of written lyrics or idiomatic expressions do not translate, in most cases, into an identical expression. Since most of the written legal documents are characterized by the complexity of their contents and the usage of specialized terminology, translating them from one language to another with high accuracy is even more difficult. In the legal domain, the literal translation cannot be respected because of the wide variety of legal concepts that exist in each legal system (Hajnasiri et al., 2015).

II. LITERATURE REVIEW

The use of natural language processing (NLP) in various domains has been an area of great academic concern. Bobojonova et al. (2024) talked about the traditions and history of librarianship in Central Asia, particularly how technological advances, including NLP and digitalization, have transformed traditional library services and enhanced

access to scientific and cultural heritage. Goltz et al., (2019) addressed the intersection of AI and global legal systems, with a focus on the application of NLP for the processing of legal texts. The authors' article outlined how AI-driven systems, while offering great possibilities, also introduce questions of interpretive accuracy within the legal profession. Jabborova et al., (2024) traced the historical evolution of book publishing, from handwritten manuscripts to modern printing technologies. This evolution forms the basis of present-day NLP applications in digitizing and semantically processing historical texts. Karimov et al., (2024) and Khaydarova et al., (2024) demonstrated how NLP-based technologies are revolutionizing Uzbekistan's tourism and library sectors. They revealed how multilingual mobile applications and online libraries are increasingly helping to improve tourism experiences and the availability of national heritage. Kurbanazarova et al., (2024) also examined the use of NLP in language learning alone using speech recognition and intelligent learning systems. Through this, it is evident how NLP is revolutionizing interactive learning in wireless and mobile environments. Mahari et al., (2023) bridged the disciplinary gap between legal studies and NLP, arguing for closer collaboration to ensure more reliable AI-based legal text processing. Similarly, Vladika et al., (2024) discussed NLP applications in the German legal field, reiterating its necessity for legal research and documentation. Shterionov et al., (2020) and Torfi et al.'s (2020) research also added to the understanding of neural machine translation and post-editing, with immediate applications for NLP in technical and professional domains. Together, these studies cement NLP's place as a groundbreaking tool for legal, educational, and cultural domains.

III. Understanding Natural Language Processing (NLP) in Translation

Natural Language Processing (NLP) is an area of computer science and artificial intelligence concerned with the interaction between computers and humans in a natural and intelligible language fashion. It incorporates the computational study of language and the design of algorithms to help computers analyze and generate comprehensible text. NLP enables computers to process, understand, produce and learn human language. Developed upon concepts and principles from computational linguistics and language modeling, Natural Language Processing fundamentally concerns itself with teaching computers to understand and act upon human-based languages. The broad range of components like text and context understanding, language translation, language generation, question answering and summarization, sentiment analysis, and speech learning, enables the handling of a multitude of input languages and dialects (Mahari et al., 2023).

Translating language however, is filled with complexities which includes but isn't limited to the understanding of context and mood, implied meanings, and dialect. Legal translations are especially tougher as the text is a combination of regulations and archaic reverend language, often including references and citations (Kurbanazarova et al., 2024). Implementing a methodical approach to convert the written word from one language into another usually calls for NLP, where algorithms can be used to process text and convert it to another language with an understandable meaning. Unfortunately, this isn't a straightforward task and the translation of legalese is an established barrier in keeping the precision and context of the original text.

Translating legal documents accurately requires a customized NLP approach as the text found in such documents is largely complex, archaic and outdated, and due to the possible cultural and language differences of the two involved languages. The intricacy in translating legal documents generally denotes a different lexicon and sentence structure that requires a specialized approach distinct from more general uses of NLP, like neural machine translation and automated web translations. Process issues thus highlight the precision necessary to maintain the text contextuality and accuracy. Barriers to maintaining accuracy and context in legal documents translate into the misuse of jurisprudence or change to the meaning of the original text during the translation process. Owing to NLP's ability to automatically process and interpret data existing in linguistics, translations can be handled much faster, thereby enhancing the production rate and cutting expenses. However, the system's conception of the meaning of words can often be pragmatic and NLP's unfitness to interpret terminologies understandingly makes it hard to retain the source text's lawful import. However, as previously stated, as technology progresses Natural Language Processing does also, and so companies turn to more modern solutions (Shanmugapriya & Srinivasan, 2016).

Definition and Basics of NLP

Natural Language Processing (NLP) is a compound scientific discipline that investigates the problems in data processing, data accessing, and data manipulation, translating and interpreting the continuous and discrete natural language. Naturally, writing, speaking and drawing diagrams are representing language thought. Through the intelligent language process, it makes the computers as the real medium of mind divorced from human brains. Some of the main branches of scientific algorithmics, including complex: cognition theories, artificial intelligence, automatic control optimization, information theory computer hardware and software, and some developing sciences, such as linguistics, phonetics psychology, and neuropsychology are invoked in NLP. Syntax is comprised of the theory of grammar which decides the structure and style of sentences. It thus can code and generate correct and incorrect sentences. Semantics is the theory of sentence meaning after some sentence structure is decided. It can code, understand, infer, demand, supply, exchange, negotiate, and answer the meaning of sentences, texts or dialogs. Pragmatics, the upper theory, is the meaning combination of sentence, word, passage, diagrams, and actions in the overall situation. Conceptually, pragmatics contains syntax and semantics (Torfi et al., 2020). It's possible for example to use the theory of discourse, rhetoric, plan, scenario, and script to understand text and actions well.

Practically, it joins together with lexicon grammar analyzers knowledge based systems, database, computational, electronic and communication system to automatic (semi-automatic) manipulation information using the machine that understands, generate and translate total text (Yang et al., 2024; Aziz, 2021).

Some incompatibility between meanings and forms, e.g., words ambiguous, many words to correspond to one meaning (concept) or vice versa meaning multinomial, figurative words or passage meaning incongruous as well as the intrinsic deficiency of the no formal language system in the computer, and so on, makes the NLP a high-level hard problem. In itself, the language processing is a non-automatic human mental activity that highly involves the individual abilities and social knowledge. Moreover, there are many kinds errors that happen when people early study the language, talking, writing, reading, drawing, presenting, etc. Even beyond the problem in the language processing itself, a vast of requirements about data input and output, easily change of problem or requirement, difficulties of the evidence collecting, can no foreseeable change, expectation or rule, and social controversial etc. are obstructive in the NLP realization. Matters would become worse when its application into computer system, such as the languages (branches) typically be designed becoming an appliance with functions and feelings of the computer or information science, the hush-hush of privacy (the state, the low, the secrets etc.) on the possible communication by the computers, the legal responsibility for a mistake computer processing that language (a special problem is presenting in the professional job of translating the law) the mistranslation of the form (technology) etc., all being a severe constriction on the NLP flourish. In despite of advances in NLP in recent years there were hardly some acceptable results (including practical systems). In any case, legal workers engaged in practice may take advantages of this or that tools, apart from professional individuals should have notion about that. Amateurs will help them avoid falling into a black box and can stimulate them into building the useful system or frame on close grounds. The understanding to this text, more knowledge and creative idea about that on the more general, direct, and state-of-art matters are pre-requirements.

Challenges in Translating Legal Documents

Legal translation, a specialized domain within translation studies, has translational products that need to meet a higher accuracy rate in translation tasks. Legal translators often coped with the difficulties in bridging different legal systems when incomparable legal concepts were founded. On the other hand, Natural Language Processing (NLP) tasks are mainly focused on the improvement of machine translation performance which may not be sophisticated enough for legal experts in legal translation due to poor quality of legal target text. Besides, they were also concerning about how accurately NLP tasks could help in the translating process to cope with the quality measure of translation product. High performance in accuracy rate will broaden the application in legal domain and further benefit the legal industry in the age of Information Technology. Legal language, as a type of professional jargons, is even more difficult to be understood by laypersons. Therefore, legalese is inevitable in the legal system but causes understanding difficulties to non-experts. In meeting the needs of precision and having legal effect, law became an abstruse and sophisticated language with no tolerance for variations. The elaboration of terminologies and sentence structures in legal texts of different legal systems is regarded as the foremost difficulty in translating legal documents (Mahari et al., 2023). Any misunderstanding or misinterpretation on legal texts may lead to enormous risks and damages in civil procedures, legislative matters, and criminal trials. In the internationalizing context, translational conflicts may arise due to discrepancies of cultural background and context in legal systems. Such discrepancies cannot be translated literally or automatically by machine in the transfer process of legal translation. Despite rapid progress in modern NLP technologies, none of these can exactly capture the cultural and contextual factors of different legal systems. In large part due to the legal idiosyncrasies of different legal systems, human experts in the legal domain are still needed in the translation process. This is why professional legal translators, upon the raw output from machines, are required to meticulously refine them. Further, trans- or cross-lingual expertise in more than one legal system is well beyond the capacity of current NLP technologies but is native to highly competent legal professionals across the globe. Nonetheless, the current practices in verifying and revising machine translation by human experts seem underspecificalized and more akin to posthoc error detection and correction. Due to the risks in the misinterpreting of source texts or the poor translation into other languages, ensuring accurate translation with high productivity is a challenge currently facing clients, clients' attorney, and legal professionals in Table I.

TABLE I CHALLENGES IN TRANSLATING LEGAL DOCUMENTS

Challenges	**Description**
Bridging Different Legal Systems	Legal concepts often lack direct equivalence across legal systems, causing difficulties in translation.
Poor Machine Translation Quality	NLP-based translations sometimes fail to meet the precision required for legal texts.
Legalese Complexity	Legal jargon and sentence structures are difficult for non-experts and machines to interpret correctly.
Cultural and Contextual Discrepancies	Different legal systems are rooted in varying cultural and contextual frameworks, making literal translation impossible.
Risk of Misinterpretation	Misunderstandings in legal translation can result in severe legal and financial consequences.
Limited NLP Capacity	Current NLP technologies cannot fully replace human expertise, especially in translingual legal contexts.
Post-editing Challenges	Human verification and refinement of machine-generated translations are essential but still underdeveloped in practice.

IV. Applications of NLP in Legal Document Translation

This section provides a structured overview of the use of Natural Language Processing (NLP) tools and explains the implications of processes for quality assurance in the context of legal translation workflows. On this basis, knowledge is shared about how NLP tools can be effectively adapted to the translation of legal texts, enabling Human Language Technology (HLT) companies to measure and maximize the added value of NLP tools when applied to this text type. As one of the first steps in such workflows, pre-processing also plays a vital role. To this end, state-of-the-art methods, as well as methodologies, on how to preprocess legal texts in a way that makes them optimal for the application of NLP technologies are shared. Accordingly, various paragraphs outline different ways in which NLP technologies are currently used in order to translate legal texts. They start by acknowledging that given the complexity of legal texts and the necessity for accurate interpretation, full automatic translation is rarely seen the best solution and judgment is usually made on a case-by-case basis concerning the need for further human post-editing (Vladika et al., 2024). Still, in an age of information overflow and with the sheer quantity of legal documents continuously growing, even small improvements in a system can be valuable. Since legal professionals are having to handle an overwhelming number of documents with strict deadlines, there is most certainly a demand for fast translation solutions, even if the accuracy occasionally has to be sacrificed for speed. In such scenarios, monolingual automated translation can come in handy, providing the gist of the text in any desired language. Unfortunately, current tools reach only moderate accuracy, especially when dealing with less common languages or specific jargon. With the enormous effort invested annually into the creation of new laws and regulation, legal texts are continuously growing in number and complexity. Furthermore, a lot of documents are merely amendments to previous ones, meaning that understanding new legislation often also requires the comprehension of an intricate network of older texts. Given these circumstances, it is not surprising that effective legal research and daily work in law firms can otherwise be quite challenging. Times of paper-based research are long gone and large-scale quantitative methods are increasingly pushing traditional ways of work aside. Yet, current methods are more or less restricted to the detection most frequently occurring texts and can lack in precision and versatility, meaning that more sophisticated tools are needed for helping legal professionals detect less conventional sources or information.

Automated Translation Systems

Globalization and the digital economy are increasing the demand for cross-border transactions, agreements, and contracts. At the same time, the penetration of the global internet is exposing people to more foreign language content. Understanding these legal texts, rights, and obligations is a challenge that centuries of technological advancement in machine translation have not adequately addressed (Goltz et al., 2019). The role of NLP (Natural Language Processing) in translating legal texts quickly, cost-effectively, and in a way that is close enough to an attorney's translation for most practical purposes - further closing the gap between the global accumulation of laws and the availability of artificial intelligence to make them transparent - is explored.

Since the 1980s the advent of the personal computer and software applications have had a significant cutback effect on law office methodologies that are solely based on legal doctrine (when the client, opponents, or its law firm are local) - the traditional way in which the legal profession is/was practiced worldwide. Nonetheless, practices still seem unable to compulsively accommodate it, which kept them bound to local/domestic services. That is, until the introduction of legal NLP advances. End-to-end solutions have become available to effortlessly transfer the lawyer's approach into the machine for the performance of case-specific tasks which drastically reduced the needs for paralegals. Nonetheless, the most prominent of such services are based on supervised machine learning that requires the profesional constant insights to establish good baselines. Translating text between natural languages requires a deep learning machine learning technology known as "machine learning", which is highly complex to develop. So, for the time being, the relevant technology stabilization and market penetration are not likely to be happen; and if/when it is, it is more likely to lead to job reduction in the legal profession rather than increased accessibility to justice and transparency of law in Table II.

TABLE II APPLICATIONS OF NLP IN LEGAL DOCUMENT TRANSLATION

Application Area	Description	Advantages and Challenges
Pre-processing and Text Normalization	Includes tokenization, sentence splitting, and formatting of legal documents to prepare them for NLP tasks.	Advantage: Reduces noise and improves input quality. Challenge: Legal texts may have non-standard formats.
Named Entity Recognition (NER)	Identifies legal entities such as parties, dates, laws, contracts, and case numbers within texts.	Advantage: Automates identification of key entities. Challenge: Variations in legal language reduce accuracy.
Legal Terminology Extraction	Extracts domain-specific terminology for constructing or updating legal glossaries.	Advantage: Enhances translation consistency. Challenge: Requires domain-specific corpora.
Machine Translation (MT)	Automates translation of legal documents between languages using neural machine translation (NMT) models.	Advantage: Reduces workload and time. Challenge: Limited accuracy, especially for low-resource languages.
Post-Editing of MT Output	Human revision of machine-translated documents to ensure legal and linguistic accuracy.	Advantage: Improves quality and compliance. Challenge: Time-consuming and

		requires legal expertise.
Information Retrieval	Locates and extracts relevant legal information from large corpora or databases.	Advantage: Speeds up legal research. Challenge: High dependency on accurate indexing and search algorithms.
Text Classification	Classifies legal texts by type (contracts, court decisions, legislation, etc.) or topic.	Advantage: Automates document organization. Challenge: Needs well-labeled training data.
Summarization	Generates concise summaries of lengthy legal documents.	Advantage: Saves time for legal practitioners. Challenge: Risk of omitting critical legal information.
Cross-lingual Search	Enables searching across multilingual legal corpora.	Advantage: Broadens accessibility. Challenge: Maintaining semantic equivalence across languages.

Quality Assurance and Post-Editing

The quality assurance of translated legal texts involves methods to ensure that the automated translated text exhibits high accuracy, consistency, and legal compliance. Post-editing is explored from different angles. Firstly, the interplay between automation and human expertise is discussed based on case studies: for instance, how the involvement of human experts in an automated translation process allowed them to detect and find solutions to incorrect translation and sentence analysis results not detected by language models and parsers. Also explored is how feedback can be looped back from human post-editors into the machine translation system to improve translation quality. Finally, best practices in manual post-editing and processes are explored based on examples of a legal text translation from German to English. The role of post-editing is exemplified; the right balance between speed and accuracy needs to be found, especially in time-critical scenarios such as legal translations. However, also when post-editing fast tracks the translation, an accuracy check is indispensable.

One of the critical qualities of translated legal texts is their correctness. Legal documents must be translated into a foreign language with high accuracy, consistency, and legal compliance. Any incorrect legal term may cause not only legal disputes but also financial loss. In order to ensure the translation quality of automated translation, it is necessary to carry out a quality assurance (QA) step for the translation results. As post-editing quality measurement, the post-editing error rate is used. The correlation between the characteristics of the sentence pairs and correction categories answering the six descriptors of the post-editing guidelines is explored. To this end, post-edited sentence pairs are first analyzed by post-editors and their error categories are annotated. Then, it is checked if specific sentence characteristics, as well as features extracted by machine learned classifiers, can be predictive of the error categories by correlating them with the annotations (Shterionov et al., 2020).

V. Benefits and Limitations of NLP in Legal Document Translation

There are benefits and limitations which come in using Natural Language Processing (NLP) in legal document translation. Using NLP can speed up the translation process of legal documents. By using NLP, legal document translation can be done more quickly and consistently than with traditional manual methods (Mahari et al., 2023). NLP can create high-quality translations of legal documents while preserving their meaning and intention, as well as avoiding any translation errors. NLP makes it easier to detect duplicate translations and copy and paste errors during the translation phase. NLP systems can automate the process of translating large volumes of legal data into different languages in a short amount of time. Additionally, NLP in the legal field can streamline the handling of legal data that requires intensive study to understand the context and find the relevant parts. NLP can help save the scope of work by automating this time-consuming process, resulting in more efficient workflows (Umarova et al. 2024).

However, legal documents contain legal nuances that are complex and difficult to interpret by NLP engines and can be ambiguous. A legal translation that ignores the legal nuances can result in the translation of the legal document that the meaning and intention of the legal document changes. It requires a high and accurate understanding of the proper use of legal terminology and the required legal tone in the translation, which is not something that NLP can do in a legal sense. While NLP models in the English language have reached state-of-the-art and are used in a wide variety of industries, results across other languages and legal systems are more variable and in many cases, are poor. Legal systems are highly variable across different languages and cultures, and as such, many English-based NLP tools cannot be simply repurposed. Only for selected languages and jurisdictions, there have been gradual improvements in relevant NLP tools, but in many cases, significant gaps exist. This might complicate the adaptability of NLP applications to different legal systems and languages, limiting its potential in this domain. Thus, more interdisciplinary and diverse collaborations and continuous research and development are needed in order to advance the state of the art in this domain. Given the complexity and importance of the rule of law, the language of legal documents, and the interpretive demands imposed by Common Law and Civil Law systems, there are ample opportunities for legal scholars to engage with this technology in ways that can yield insights or improvements beyond what computer scientists and NLP specialists might provide.

Advantages of NLP in Legal Translation

Advancements in machine learning algorithms and the proliferation of natural language processing tasks promise to revolutionize the way legal practitioners translate their legal documents. Because of this, there is an unprecedented level of interest in the potential of natural language technologies within the legal domain. The slow uptake of NLP in legal practice has often been attributed to a lack of trust in AI or a failure to understand how such technologies can be of use. However, recent advances in machine translation and the rapid democratization of standard translation tools suggest that the translation space, and particularly the professional subsector in law, is poised to undergo swift changes. While definite progress is yet to be made in applying machine translation tools to the complexity of legalese, other areas of recent advancements in NLP, such as the ability to parse and evaluate clauses, grounds, and judgements, present legal engineers with tangible and commercially exploitable tools. Legal experts with a good understanding of their own domain can leverage these NLP tools to track and interpret advances in the case law of jurisdictions they are following. The burgeoning field of computable contracts similarly promises practitioners tools to generate new agreement templates, and these can follow the style guides of the user's local jurisdiction. For other users, the democratisation of large language models arguably brings more invisible changes. The work of legal scholars is often directed by text work, and even a hierarchy layer glimpse behind the large models reveals insightful relations between clauses and terms that may be obvious to a domain expert, but are unprecedented in terms of the sheer volume of text ingested. Similarly, the transparency and searchability of large models purportedly levels the playing field between firms as it allows SMEs to efficiently navigate extensive legal corpora, whose sheer volume was previously dominated by a single group of Big Law collaborators (Mahari et al., 2023).

Limitations and Areas for Improvement

Since language is context dependent, linguistic expressions are not always subject to clear-cut rules (Mahari et al., 2023). Moreover, legal language is generally considered to be complex since laws, criminal-legal books, procedures, acts, and other legislative instruments have linguistic structures as well as expressions distinct from general language. Due to this situation, understanding legal texts and the extraction of information suitable for understanding and analysis of these documents require expertise knowledge in the field of law since it needs to be examined in the context of legally. The ability to understand, analyze, classify, retrieve, and question complex legal documents is inherently difficult due to the context of obvious legal terms and document structures. Fully automatic processing is challenging for state-of-the-art Natural Language Processing (NLP) approaches. Moreover, even if the necessary tools for legal NLP are available, extracting the correct meaning from legal documents may be difficult for human in case of interpretation of a dispute, and hence an analogous challenge may be even more so for algorithmic implementations. Mostly US based, text analytic will be used in summarization services for providing journalists with beneficial information and document investigation of long legal text instances and it will target text processing and analysis for better classifications and category controls.

With the current processing speeds and technology at hand, it is very difficult to create a general analysis mechanism capable of capturing the extent of situations and behaviours that occur in everyday communication. On the other hand, it is very simple for people to learn and to comprehend without much conscious effort, sounds and meaning evolving from recognition of the utterances in the actual world. NLP tries to deal with the grey area in the middle between these two scenarios. Different languages have different linguistic structures; hence designing an NLP application that has the same performance for all languages is non-trivial. Among 15 European languages, English is the best served in terms of language technology support. However, it still leaves more than 15 official European languages that are part of the EU without satisfactory language technology support.

VI. Case Studies and Examples

The case studies of Real Life Case Studies with NLP focus on the use of services in the cloud and the data and insights created by them. The use cases presented in the following have been chosen based on their different technological approaches and the insights derived from NLP-produced data. Due to the large number of NLP aspects and technologies relevant for the legal domain, priority has been given to use cases where tangible insights and outcomes are presented that can serve as a model for other scenarios. Moreover, the focus is primarily on NLP in the context of text data. This section aims to illustrate a broad variety of use case examples and derived insights in order to provide both an overview of the state of the art as well as to create a foundation for further research and discussion. The first case study is concerned with the use of state-of-the-art NLP to improve information retrieval in legal metadata catalogues. The second case study takes a more data-driven approach and introduces the analysis of generated NLP insights to show new possibilities of assessing legal text from a different perspective (Vladika et al., 2024). Lastly, the recent development of NLP technologies available in the cloud is scrutinized, followed by the update on the current and potential future use cases and the data and insights that are created by them.

Success Stories in NLP-Based Legal Document Translation

This section is dedicated to the successes in using Natural Language Processing (NLP) technologies for legal translation. The recent history and future of developing NLP-based legal document translation technologies and methodologies are first contemplated. As the case studies show, the methodologies applied in developing these technologies and technologies vary significantly, and there is no single best method or technology in this field. In most of the success stories, experts in NLP, especially developers of

the underlying technology, and legal experts worked together in professional settings. As the concrete trials and successes in the sphere give way to the use of these technologies and methodologies within legal settings, NLP professionals and legal professionals should collaborate. Such collaboration could lead to a more comprehensive and balanced development of these technologies and methods, and could maximize the utility of these technologies and methods in legal settings.

In the past decades, there have been impressive advances in machine translation. After the introduction of statistical methods and the following introduction of NLP technologies, the accuracy of machine translation has dramatically increased, especially in the translation of European languages. These advances have encouraged the application of machine translation to the translation of documents other than those requiring confidentiality, such as legal texts. However, for some legal settings, refined or professional translation is necessary, and for these documents, the accuracy of translation may be unsatisfactory. To further enhance the post-editing process, linguists and computer scientists specializing in NLP have already collaborated on the development of document-translating NLP technologies with a focus on linguistics. In the past decade, NLP-based translation technologies of documents with patents and technical fields that have developed rapidly within the task-intensive setting are discussed and analyzed. This discussion and analysis reveal that improved translation accuracy is one benefit of the use of machine translation technologies and methods. The technology and methods applied in the existing examples of machine translation technologies may not be the best and increasingly advanced other methods and technologies may already have been developed. Furthermore, the analysis of the successful development and application of technologies and methods within legal settings underscores the need for professional collaboration and the evaluation and refinement of primitive technologies and methods with the ongoing development and use on an empirical setting.

VII. Future Directions and Conclusion

Rapidly growing investments in Natural Language Processing (NLP) technologies are quickly making them versatile enough to accomplish tasks like translating legal documents easily and with high accuracy (Mahari et al., 2023). There are already those who have started collections of such tools and libraries on NLP preserving algorithms that could translate legal documents with astounding efficiency in a fraction of the time such work might require from a college-educated person proficient in two+ languages. Consequently, it likely won't be long before demand for 'language' or 'translation' classes will have evaporated, causing revenue prospects for translators and people-of-the-world jobs inherently dependent on a variety of languages to disappear. In the meantime, we can expect language-savvy people to develop an open-source movement in NLP aimed to counter what many perceive might be over-concentration of power and resources into the hands of organizations with very effective and proprietary NLP tools. The exact language of the law has never evolved comprehensively, due to the vast nature of law itself. That is, no official language exists for law, but rather spoken language has been used, with the educational translation of said being heavily charged. With NLP technology enhancing accuracy in fashioning and editing texts in languages other than the origin, legal systems will need to re-format their demands, as the malicious exploitation of ambiguity within the translation rights is most possible even in instances where the translation is based on transparency.

References

[1] Aziz, B. (2021). A note on the problem of semantic interpretation agreement in steganographic communications. *Journal of Internet Services and Information Security, 11*(3), 47-57. https://doi.org/10.22667/JISIS.2021.08.31.047

[2] Goltz, N. S., Cameron-Huff, A., & Dondoli, G. (2019). Rethinking Global-Regulation: world's law meets artificial intelligence. *Information & communications technology law*, *28*(1), 36-45.

[3] Hajnasiri, H., Geranmayeh, M., Moghadam, Z. B., & Ghasemzadeh, S. (2015). An investigation of consultation effect on anxiety after legal abortion. *International Academic Journal of Innovative Research, 2*(1), 66–73.

[4] Hemat, M. A. (2014). A Review of Researches on Persian Text Mining. *International Academic Journal of Science and Engineering, 1*(2), 152–160.

[5] Jabborova, D., Tanieva, G., Isakov, B., Yaxshiyeva, M., Sulaymanova, N., Tillakhodjaeva, K., Akmalxonov A., & Sattorova, Z. (2024). From Manuscripts to Machines: The Evolution of Book Publishing Devices. *Indian Journal of Information Sources and Services*, *14*(4), 117–124. https://doi.org/10.51983/ijiss-2024.14.4.18

[6] Karimov, N., Kalandarova, M., Makhkamova, N., Asrorova, Z., Saydamatov, F., Ablyakimova, R., Karshiboeva, M., & Odilov, B. (2024). Impact of Mobile Applications on Tourism Development in Uzbekistan. *Indian Journal of Information Sources and Services*, *14*(4), 175–181. https://doi.org/10.51983/ijiss-2024.14.4.27

[7] Khaydarova, S., Khujamova, Y., Toshbaeva, M., Muhitdinov, D., Mamanazarova, G., Tukhtakulova, O., & Karimov, N. (2024). The Vital Role of Libraries in Enriching Tourism Experiences. *Indian Journal of Information Sources and Services*, *14*(2), 11–16. https://doi.org/10.51983/ijiss-2024.14.2.02

[8] Kurbanazarova, N., Shavkidinova, D., Khaydarov, M., Mukhitdinova, N., Khudoymurodova, K., Toshniyozova, D., ... & Alimova, R. (2024). Development of Speech Recognition in Wireless Mobile Networks for An Intelligent Learning System in Language Education. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *15*(3), 298-311. https://doi.org/10.58346/JOWUA.2024.I3.020

[9] Mahari, R., Stammbach, D., Ash, E., & 'Sandy' Pentland, A. (2023). The Law and NLP: Bridging Disciplinary Disconnects.

[10] Nazarova, S., Askarov, M., Karimov, N., Madraimov, A., Muminov, A., Abirov, V., & Shosaidov, A. (2024). The Role of Online Libraries in Advancing the Study of Uzbek Culture. *Indian Journal of Information Sources and Services*, *14*(3), 207-215. https://doi.org/10.51983/ijiss-2024.14.3.26

[11] Shanmugapriya, K. P., & Srinivasan, B. (2016). Markov Logic Network for Entity Disambiguation and Mining Sentiments from Songs Using Semi Hidden Markov Model. *International Journal of Advances in Engineering and Emerging Technology, 7*(1), 93–109.

[12] Shterionov, D., Carmo, F. D., Moorkens, J., Hossari, M., Wagner, J., Paquin, E., ... & Way, A. (2020). A roadmap to neural automatic post-editing: an empirical approach. *Machine Translation*, *34*, 67-96.

[13] Tohma, K., & Kutlu, Y. (2020). Challenges Encountered in Turkish Natural Language Processing Studies. *Natural and Engineering*

Sciences, 5(3), 204-211. https://doi.org/10.28978/nesciences.833188

[14] Torfi, A., A. Shirvani, R., Keneshloo, Y., Tavaf, N., & A. Fox, E. (2020). Natural Language Processing Advancements By Deep Learning: A Survey.

[15] Umarova, M., Karimov, N., Kulmetov, M., Kalandarova, G., Babadjanova, N., Toliboeva, N., Shamshiev, A., & Saparniyazova, M. (2024). The Impact of Online Databases on Enhancing Educational Quality. *Indian Journal of Information Sources and Services*, *14*(3), 241–250. https://doi.org/10.51983/ijiss-2024.14.3.31

[16] Vladika, J., Meisenbacher, S., Preis, M., Klymenko, A., & Matthes, F. (2024). Towards A Structured Overview of Use Cases for Natural Language Processing in the Legal Domain: A German Perspective.

[17] Yang, C. S., Lu, H., & Qian, S. F. (2024). Fine tuning SSP algorithms for MIMO antenna systems for higher throughputs and lesser interferences. *International Journal of Communication and Computer Technologies, 12*(2), 1-10. https://doi.org/10.31838/IJCCTS/12.02.01