IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 10, October, 2026, pp. 3357-3367
Journal homepage: http://iieta.org/journals/isi

Mapping Research on Legal Natural Language Processing and Large Language Models: A Bibliometric Study

Hanane Moustaoui[1*], Aidi Ahmi[2], Hicham Bensaid[1], Mouhsine Lakhdissi[3]

[1] SEEDS Team, Institut National des Postes et Télécommunications, Rabat 10112, Morocco
[2] Accounting Information System Research and Development Institute (AISRED), Tunku Puteri Intan Safinaz School of Accountancy, University Utara Malaysia, Sintok 06010, Malaysia
[3] Faculty of Sciences and Technology, Hassan I University, Settat 26000, Morocco

Corresponding Author Email: moustaouihanane@gmail.com

Copyright: ©2026 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/isi.311010

Received: 9 June 2026
Revised: 25 August 2026
Accepted: 1 September 2026
Available online: 31 October 2026

Keywords:
legal natural language processing, legal informatics, large language models, bibliometric analysis, keyword co-occurrence, multilingual legal natural language processing

ABSTRACT

This study maps research on natural language processing (NLP) and large language models (LLMs) in the legal domain using bibliometric methods. A corpus of 1,109 Scopus-indexed publications was analyzed for 2019–2025, with incomplete coverage for 2025. Records were screened and cleaned with OpenRefine and BiblioMagika; author-keyword variants were harmonized with assistance from an LLM and then checked by human reviewers. VOSviewer and Biblioshiny were used to examine publication output, citation patterns, keyword co-occurrence, thematic structure, and geographical collaboration. Annual output increased from 29 publications in 2019 to 347 in 2024, with the strongest growth occurring after 2021. The corpus includes 4,322 contributing authors and 6,156 citations, although citation indicators favor earlier publications because they have had more time to accumulate citations. China and India account for the largest national outputs, while English-language publications remain dominant. The keyword and thematic analyses point to four broad strands: legal information retrieval, NLP method development, transformer-based legal analytics, and semantic approaches, alongside growing attention to generative models and multilingual legal NLP. The reviewed literature remains concentrated on legal classification, case retrieval, judgment prediction, and named-entity recognition; work on synthetic data is visible but forms a smaller part of the corpus. These findings describe the distribution and thematic direction of indexed research rather than its practical effects on legal services. The analysis is limited by reliance on Scopus and by the sensitivity of keyword maps to screening, harmonization, and threshold choices.

1. INTRODUCTION

Bibliometric analysis is an indispensable tool for evaluating research, especially in the sciences, and has grown in popularity since the mid-1990s [1]. Initially originating from information science, these techniques have now been adopted in a broad spectrum of academic disciplines [2]. What makes them stand out is their capacity to offer an objective perspective on scientific contributions, allowing literature reviews to move beyond individual judgment [3]. In the dynamic field of legal natural language processing (NLP), these methods act as a dependable guide for monitoring field developments and evaluating influence. We are researching the ways in which NLP and large language models (LLMs) have transformed the legal field. Recently, powerful language models have played a major role in this advancement. For example, Bidirectional Encoder Representations from Transformers (BERT) marked significant progress in understanding word context and has led to improvements in search query accuracy [4]. GPT-3 is widely used in text generation tasks, chatbots, and other generative tasks [5]. Models such as Text-to-Text Transfer Transformer (T5) introduced a flexible text-to-text framework, enhancing translation and summarization tasks [6]. Summarization stands out: deep learning approaches are increasingly being deployed to distil court decisions and legal documents down to their essentials [7]. XLNet reads sentences from left to right and from right to left, which improves the question-answering accuracy since the sentence is parsed from both sides [8]. Bidirectional models have been adapted to legal question answering [9]. Meta AI's LLaMA has recently gained popularity and power over the new tech tools; it's an open-source LLM with a variety of sizes ranging from 7 billion to 65 billion parameters [10]. DeepSeek-V3 achieves competitive benchmark results comparable to leading models, requiring far less investment in training [11]. Legal NLP tasks come with significant gaps such as inconsistent document formatting, lack of training data, privacy and bias issues [10, 11]. Processing legal texts in under-resourced languages adds an additional layer of complexity [11]. This paper offers a bibliometric study of how NLP and LLMs are applied in the legal field, highlighting key trends, contributors, institutions, countries, topics, and collaboration patterns.

Several reviews have been published in legal NLP, but most

of them are narrative or systematic reviews; in those reviews, the material and results are organized based on the author's judgment. Greco and Tagarelli [12] showed that the transformer architecture will be the state of the art and widely used in legal artificial intelligence (AI) tasks. Our study maps the volume, geography, and thematic structure of legal NLP research from 2019 to 2025 with partial coverage of 2025. The result is a data-driven complement to existing reviews. It offers researchers and practitioners a concrete, evidence-based map of where the domain has concentrated its efforts and where it has yet to look.

In order to achieve this, this study answers the following questions:

RQ1: How has the application of NLP and LLMs in the legal and jurisprudential field changed over time?

RQ2: Which authors, institutions, and countries are the most influential contributors to this field?

RQ3: What are the most impactful keywords and topics, and in what way have they been transformed?

RQ4: What are the leading research clusters and collaboration networks forging this field?

RQ5: What future research directions can be inferred based on current bibliometric patterns?

2. DATA AND METHODOLOGY

Bibliometric methods were applied in this study to the scholarly literature on the application of NLP and LLMs. This workflow comprises ten steps: IDENTIFICATION, EXPORT, SCREENING, CORPUS, PREPARATION, CLEANING, PROPOSAL, HARMONISATION, VERIFICATION and ANALYSIS, as it is described in the Figure 1. This flow follows the workflow described in the biblioMagika® bibliometric toolkit [13]. In the first phase, we export and retrieve the data from Scopus. We use one database to avoid combining databases and reconciling records with divergent metadata, which may introduce matching and duplication errors. The total records retrieved here are 6969. We load both the truncated and untruncated files to be used in the biblioMagika tool [13]. This extraction was done on 17 December 2024 using the query shown in Figure 1; this query contains two blocks joined with AND. The first block focuses on the NLP and LLM methods and architectures, and the second one captures the legal field; since the two blocks are linked with the AND operator, a record that didn't cover both fields will not be retrieved. We carried out a screening phase; we added a relevance column in the Excel file for each record, analyzed the title, and marked the value of the cell as "Yes" if the work is relevant to the topic and "No" if not. The abstract was consulted whenever the title was not decisive. This first cleaning step was done to avoid including works under heterogeneous terminology; many items are homonyms with concepts outside the subject area: NLP also denotes neuro-linguistic programming, GPT is also used in biomedicine and economics, and "transformer model" occurs in electrical engineering. In the preparation step, we load the relevant data in biblioMagika® and biblioMagika®Split in order to clean multi-value fields authors, affiliations, countries, remove duplicate records using Scopus EID, and complete missing metadata. In this step, we end up with 1109 cleaned records, excluding 5860 irrelevant and duplicated research works; then we use the OpenRefine tool to reduce terminological variation in author names, affiliations, countries, and keywords. For each field, we generate a text facet and create clusters using Nearest Neighbor Search [14] and Levenshtein distance [15], with 1.0 as the radius and six as the blockchars. This cleaning improves the accuracy of authors, institution, and country-level indicators which will be fragmented across variant spelling of the same entity but don't change the composition of the corpus the total number of records remains 1109. The new step that we introduce is harmonization, where the LLM assists in cleaning the author keywords. As shown in Figure 2. The prompt used follows the structural components identified in the prompting literature [16]; no examples were supplied; therefore, the zero-shot technique was used [5]. A human verification step was added before moving to the analysis phase.

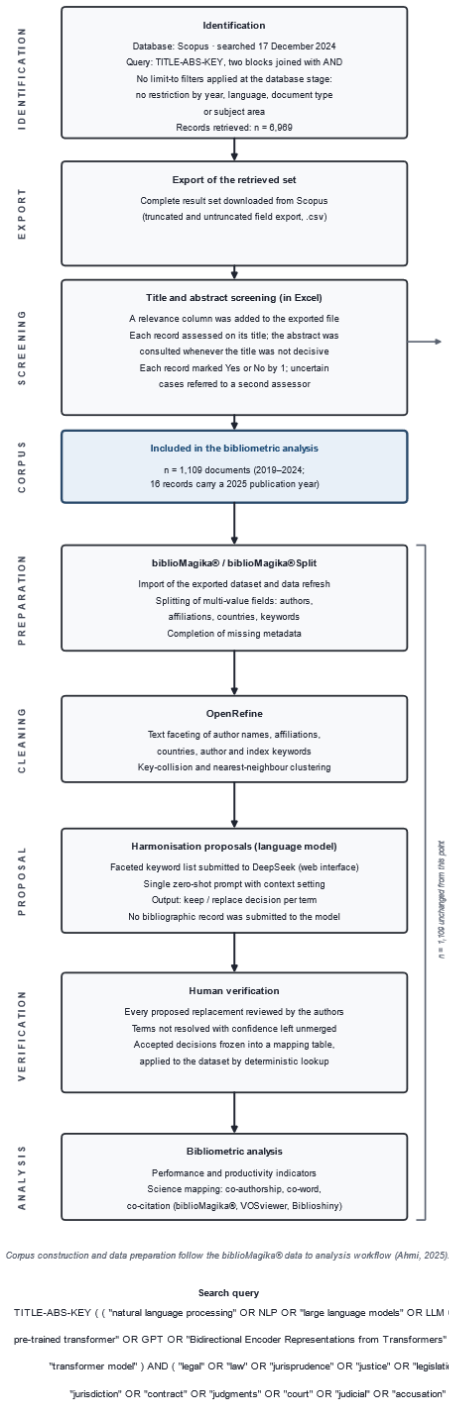

Figure 1. Methodological workflow applied in bibliometric analysis

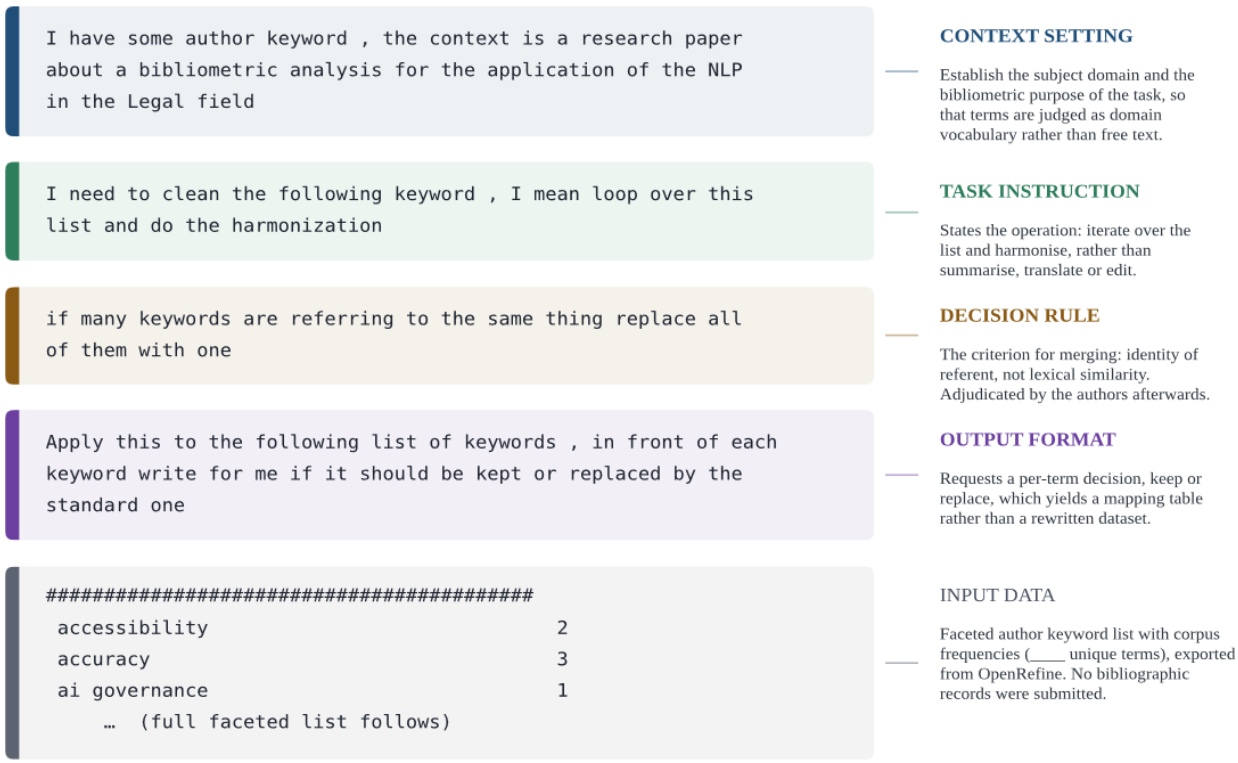

Figure 2. Anatomy of the keyword harmonization prompt

3. RESULTS

3.1 Analysis of bibliometric data (2019–2025)

Figure 3 shows that from 2019 to 2025, 1109 publications have been registered on the application of NLP in the legal field. Where 4,322 authors have contributed which reveal a strong collaboration. 6,165 citations were highlighted, with an average of 5.55 citations per paper and 9.08 citations per cited paper. With an h-index of 33, g-index of 56, and m-index of 4.714, this reflects the accumulation of citations across a growing body of work over the seven years. The main bibliometric indicators for the 2019–2025 period are summarized in Table 1.

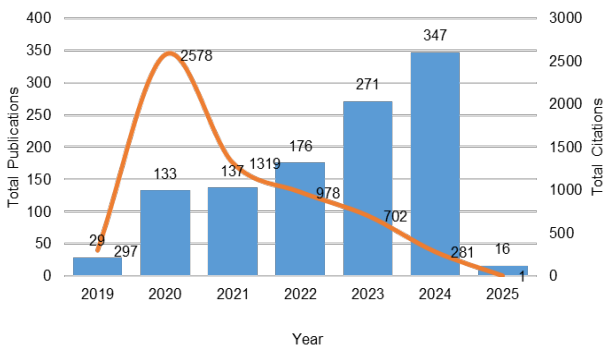

Figure 3. Year-by-year analysis (2019–2025) of publications and trends in natural language processing (NLP) legal research
Source: Generated by the author using BiblioMagika [13].

Table 1. Main bibliometric information for natural language processing (NLP) in the legal field (2019–2025)

Main Information	Data
Publication years	2019–2025
Total publications	1,109
Citable year	7
Number of contributing authors	4,322
Number of cited papers	678
Total citations	6,156
Citation per paper	5.55
Citation per cited paper	9.08
Citation per year	879.43
Citation per author	1.42
Author per paper	3.90
Citation sums within h-core	3,623
h-index	33
g-index	56
m-index	4.714

Source: Generated by the author using biblioMagika [13].

Collaboration is noticeably a core component here since the papers are around 1,026 citations annually, signaling their improving visibility and relevance. The h-core alone gathered 3,623, emphasizing influential works that continue to guide today's academic research. This work shows that advanced NLP techniques address legal challenges and improve workflows. With 70% of work appearing as conference paper confirms how quickly things are moving. The fact that 97.57% of research is in English highlights a clear priority and the need to introduce other languages beyond English.

3.2 Analysis of total publications by year

Over time, the application of NLP in the legal system is expanding, reaching a peak of 347 papers published in 2024. Tasks like multiclass classification, summarization, and question answering [17] have seen huge progress. In 2020, the citation activity reached its maximum value, a sign that the earlier foundational work is still shaping the field today. As the field progresses, researchers should stay open and agile to new directions.

3.3 Top 10 papers

Table 2 presents the ten most influential papers that have marked NLP in the legal field identified through bibliometric analysis. Their impact is assessed using total citations and per-year averages. The 2020 contribution from Chalkidis and colleagues stands out, having reached 344 citations and a yearly average of 57.33 citations. This remarkable study introduced LEGAL-BERT, a language model personalized for legal texts. The stable high citation count indicates that this work has continued to be cited across the study period.

Table 2. Top 10 papers in natural language processing (NLP) applications in the legal field

No.	Author(s)	Title	Source Title	TC	C/Y
1	Chalkidis et al. [18]	LEGAL-BERT: The Muppets straight out of law school	Findings of the Association for Computational Linguistics: EMNLP 2020	344	57.33
2	Medvedeva et al. [19]	Using machine learning to predict decisions of the European Court of Human Rights	Artificial Intelligence and Law	210	35.00
3	Chalkidis et al. [20]	Neural legal judgment prediction in English	ACL 2019-57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference	180	25.71
4	Zhong et al. [21]	How does NLP benefit the legal system: A summary of legal artificial intelligence	Proceedings of the Annual Meeting of the Association for Computational Linguistics	167	27.83
5	Xiao et al. [22]	Lawformer: A pre-trained language model (PLM) for Chinese legal long documents	AI Open	125	25.00
6	Shao et al. [23]	BERT-PLI: Modeling paragraph-level interactions for legal case retrieval	IJCAI International Joint Conference on Artificial Intelligence	118	19.67
7	Chalkidis et al. [24]	Large-scale multi-label text classification on EU legislation	ACL 2019-57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference	115	16.43
8	ul Hassan and Le [25]	Automated Requirements Identification from Construction Contract Documents Using Natural Language Processing (NLP)	Journal of Legal Affairs and Dispute Resolution in Engineering and Construction	92	15.33
9	Chalkidis et al. [26]	LexGLUE: A Benchmark Dataset for Legal Language Understanding in English	Proceedings of the Annual Meeting of the Association for Computational Linguistics	86	21.50
10	Sydora and Stroulia [27]	Rule-based compliance checking and generative design for building interiors using BIM	Automation in Construction	78	13.00

Note: TC = total citations, C/Y = citations per year.
Source: Generated by the author using BiblioMagika [13].

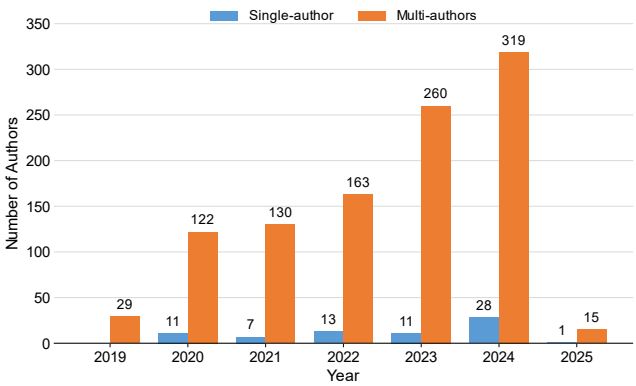

Figure 4. Distribution of single-author and multi-author publications in natural language processing (NLP) applications in the legal field

3.4 Authorship analysis

The authorship analysis shown in Figure 4 reveals that multi-author publications dominate research on NLP applications in the legal field, reflecting strong collaboration. While single-author contributions remain limited, they have shown steady growth. Promoting a collaborative environment, as well as supporting single-author projects, can provide valuable insights and new ideas, which are vital for moving this important interdisciplinary field forward.

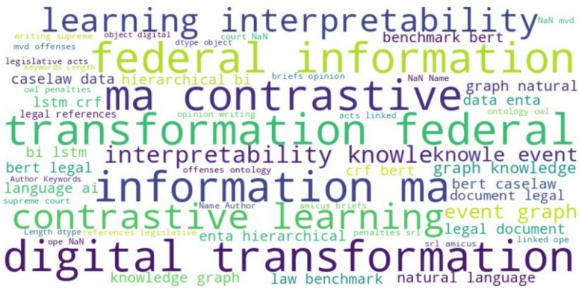

Figure 5. First look at key themes before semantic cleaning: word cloud of legal natural language processing (NLP) research

3.5 Word cloud analysis

The significance of the clean-up process for bibliometric data becomes apparent when comparing the word clouds produced before and after the harmonization process. The first word cloud (Figure 5) shows major discrepancies, including redundant terms, variations in case sensitivity, and fragmented

phrases (e.g., "biLSTM interpretability knowledge event"), which blur the topical orientation of the study field. After applying DeepSeek for text harmonization and redundancy removal using a well-defined prompt. Figure 6 shows a word cloud that paints a much cleaner and more consistent snapshot of key themes like NLP, legal document analysis, knowledge graphs, contracts, and blockchain. This cleaning step didn't just improve the clarity of the visualization but also reinforced the dependability of our later bibliometric work, supporting more precise trend extraction, as we can see in Figure 6.



Figure 6. Second look at key themes after semantic cleaning: word cloud of legal natural language processing (NLP) research

3.6 Network visualization of author keyword co-occurrence

Figure 7 presents a co-occurrence map that shows the interconnection between the keywords that have been selected by authors in legal NLP research. This analysis was performed using VOSviewer version 1.6.20 [28], which identifies tendencies and associations between research topics by visualizing the frequency with which keywords co-occur. Nodes represent individual keywords, with their size increasing proportionally with their occurrence frequency. Lines linking those nodes (edges) indicate co-occurrence between keywords; edges that are thicker reflect closer associations, implying tighter thematic relationships within the literature. This co-occurrence map was created using VOSviewer with the following configuration: the full counting method was employed, which counts each keyword occurrence equally without fractional attribution. A minimum threshold of seven occurrences per keyword was defined, filtering 1,939 author keywords down to 74 that met the criterion. We settled on this threshold because the keyword distribution is heavily asymmetrical: most of the terms appear only once or twice; retaining them would generate co-occurrence that would mirror single documents rather than stable thematic patterns. A minimum cluster size of 10 keywords was imposed to suppress tiny groupings. The default clustering resolution of 1.00 was left unchanged. The algorithm then produces four thematic clusters displayed in distinct colors; this was the output under the chosen settings, not a number chosen in advance.

Figure 7 shows four themes: blue and yellow one's centers on core NLP methods (deep learning, machine learning, text classification summarization). The red one groups task-oriented work on legal systems such as legal information retrieval, question answering, and named entity recognition (NER) [9, 29]. The green cluster reflects the recent shift toward generative models (LLMs, GPT) and an adjacent blockchain and smart contracts [30].

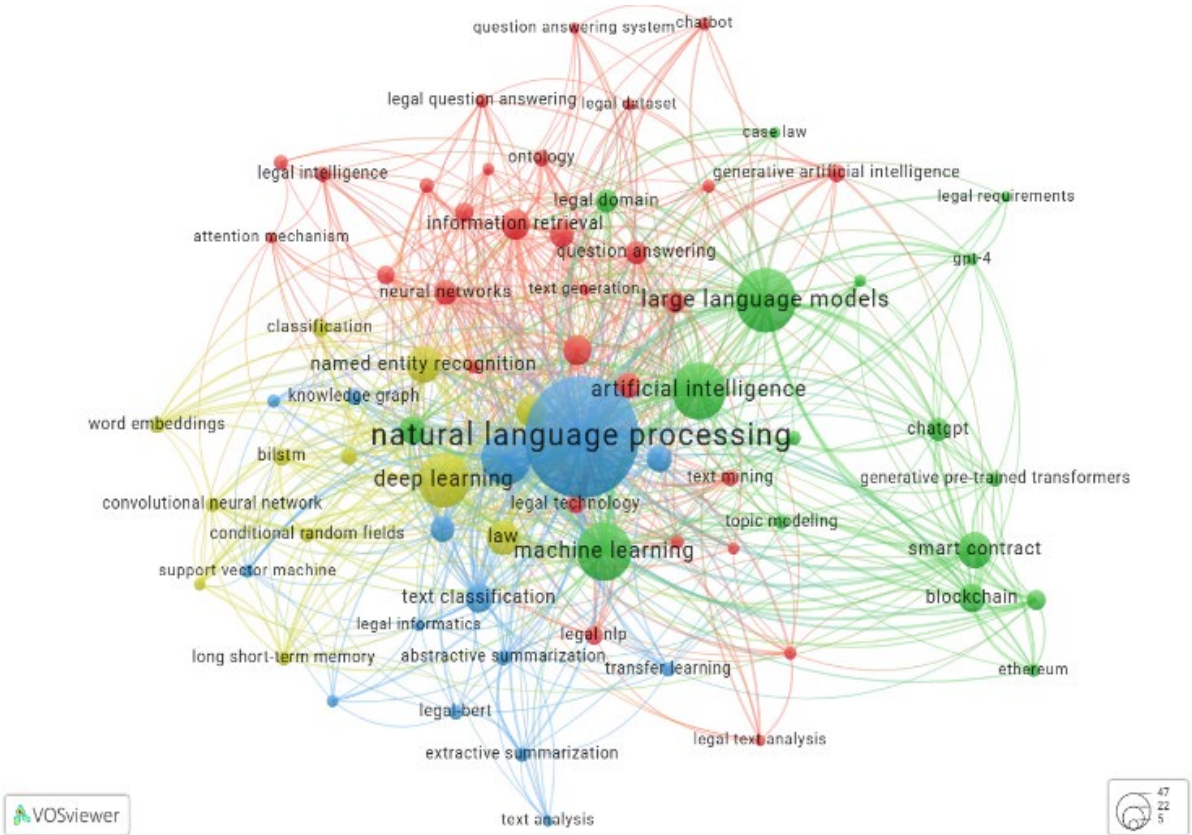

Figure 7. Network visualization of author keyword co-occurrence
Source: Generated by the author(s) using VOSviewer [18]. Dynamic View of the Network: https://bit.ly/4i7kaEg.

3.7 Thematic map analysis

In the thematic map shown in Figure 8, we have four quadrants based on the density and centrality. The density shows the topic development level, whereas centrality reveals the significance of the key terms in the entire field. In the major themes located in the right quadrant, we have terms related to the legal field, such as laws and legislation, legal domains, and legal cases. We can also see AI, since this quadrant groups well-developed and central topics that represent the core and driving field of this research area. Regarding the niche themes cited in the upper left having high density and low centrality, we found GPT, generative pre-trained models show a huge potential and applicability in the legal field: GPT-4 [30] report a score of approximately 297 points above the passing threshold in every jurisdiction but it makes errors in new statutes that he never seen during the training. Generative Pre-trained Transformer 3 (GPT-3) achieve results in the legal field that surpasses previously state of the art on statutory reasoning [31] using zero-shot, few-shot and chain-of-thought prompting, but this model makes some errors in new statutes that he never seen during the training. The previous two studies show that a general-purpose model can perform well on legal tasks without modification; GPT-3 was fine-tuned to classify legal rules encoded in legalRuleML, so the model will be more adapted to the legal domain [32].

In this niche theme, we found 3 foundations of legal NLP: NER models that extract main entities that structure a legal document and are used to identify parties, courts, statutes, and dates [33]; open legal data repositories also provide resources like ECHR-OD, an integrated open repository of European Court of Human Rights decisions used for machine learning applications [32]. For the basic themes in the lower right quadrant include transversal and foundational topics like question answering, NLP systems, deep learning, and LLMs: Question answering can be handled through neural text representations of legal knowledge, where the user query is matched with relevant, long, and complex statute law articles [9]. Or using a knowledge graph by converting unstructured legal text (law, court rulings, and contracts) into structured machine–readable nodes and edges [33]. Recently, LLMs have entered the field; inspired by the evolution of pre-trained language models (PLMs) in the general field, researchers try to use them in legal tasks [34]. Deep learning algorithms can also be used for legal text summarization [7]. Finally, we can see the emerging cluster that contains themes such as smart contracts, blockchain, and vulnerability detection. NLP, semantic web technologies, and text mining can be used to automate GDPR data compliance [35].

3.8 Geographical distribution of the publications

Figure 9 shows the geographic distribution of publications. The visualization describes the research output in several countries using a color gradient, where darker shades indicate fewer contributions. The visualization shows that China is the top publisher with 226 articles; India comes in second place, with 178 publications. European influence is substantial where Italy tops the list with 77 papers, UK and Germany is following at 60 and 50 respectively. European involvement is strengthened by France, Switzerland, and the Netherlands. Meanwhile, in the Asia-Pacific region, Japan stands out with 47 papers, followed by South Korea (21), Taiwan (20), and Australia (20). With 62 publications, Brazil dominates South America. However, coverage from Africa, Central America, and parts of the Middle East remains limited. Dominant countries maintain their high ranking; initiatives to collaborate and cooperate with the marginalized areas remain essential to ensure more balanced global advancement.

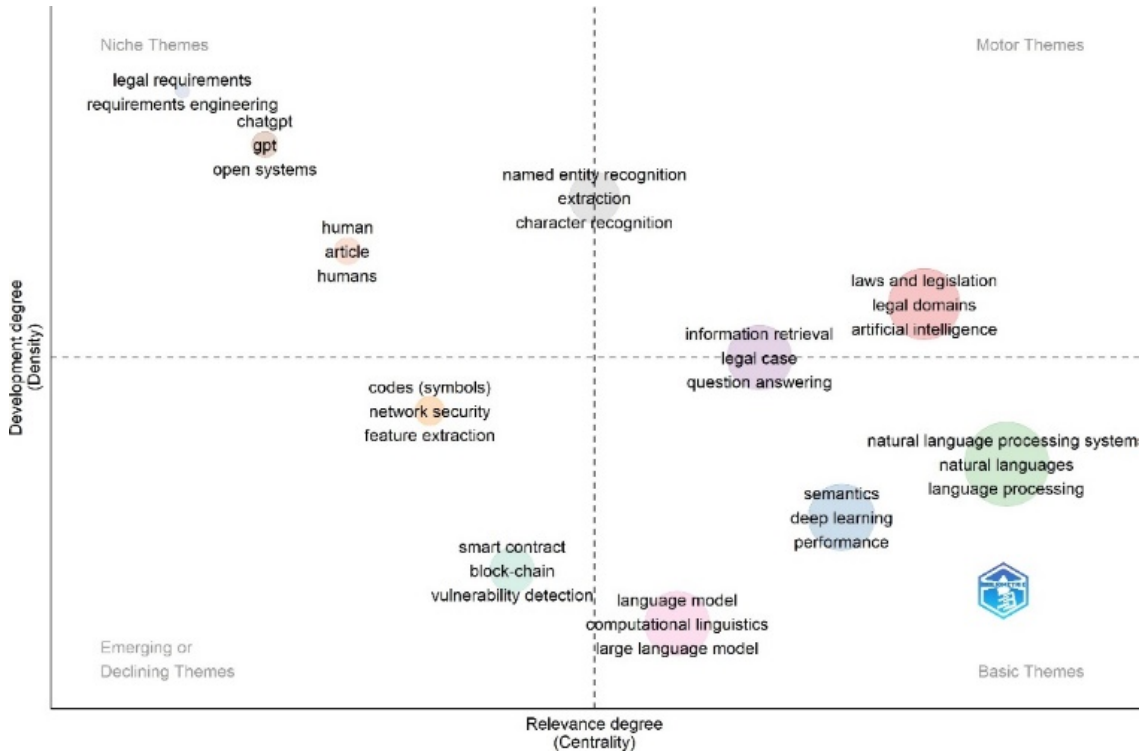

Figure 8. A thematic map of applications of natural language processing (NLP) in legal research after cleaning the data related to keywords

Source: Generated by the author(s) using Biblioshiny [36].

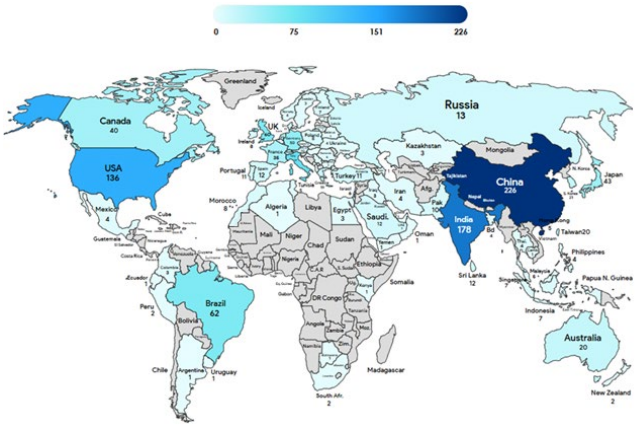

Figure 9. Geographical distribution of publications on legal natural language processing (NLP)
Source: Generated by the author(s) using iipmaps.com.

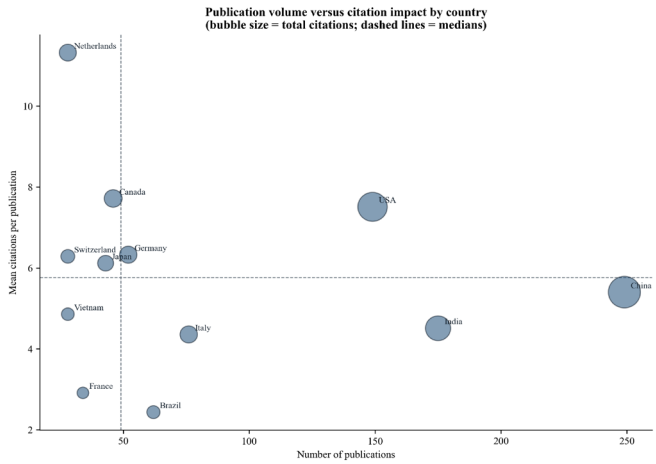

Figure 10. Publication volume versus citation impact by country

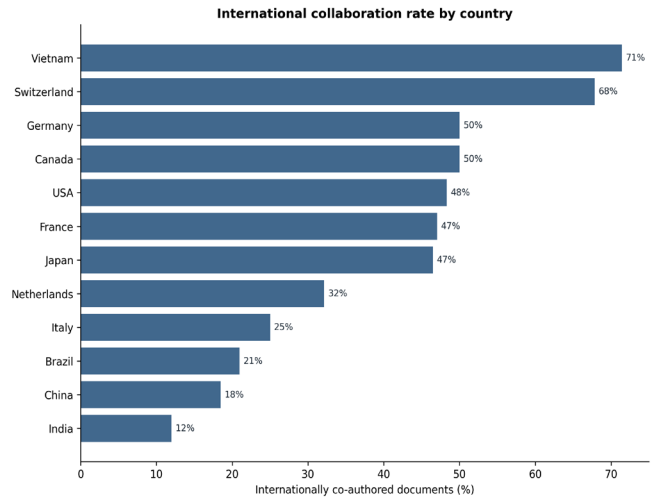

Figure 11. International collaboration rate by country

Figure 10 compares each country's publication volume against its citation impact. The figure shows that China and India, the highest-output countries, sit below the median in mean citations per publication, whereas small producers like the Netherlands achieve higher per-publication impact. Countries are ordered from the highest to the lowest. Figure 11 shows that publication volume and international collaboration do not move together in this domain. China and India are at the top for output but at the bottom for international co-authorship with just 18% and 12%; smaller contributors tell the opposite: Vietnam and Switzerland have high percentages with 71 M and 68%, while several European countries and the United States are in the middle range.

3.9 Modeling paradigms in the CORPUS: A year-by-year view through the corpus

Figure 12 shows that most of the publications leaned on classical representation and embeddings; see large-scale multi-label classification of EU legislation [24] and recurrent architecture for tasks like legal document classification [37]. Encoder-based transformers arrived soon after, including domain-specific variants such as LEGAL-BERT [18], applications to case retrieval [23], and judgement prediction [38]. In 2023, researchers shifted to generative LLMs; classical machine learning and recurrent architectures were progressively displaced.

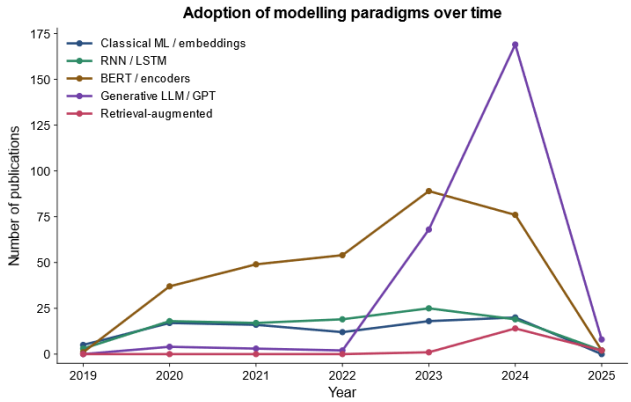

Figure 12. Evolution of modeling paradigms in the bibliometric data tracked through yearly term frequencies
Coverage of 2025 is partial (extraction in December 2024).

Figure 13 shows the evolution of synthetic data generation in the Legal field in the corpus between 2019 and 2024. The left image represents total number of publications per year mentioning synthetic data or data augmentation in their title, abstract or keywords whereas the right one shows the same thing but as a percentage of the annual output, both figures shows that we have absence in this topic in 2019 but such work rose to 11 publications in 2024 which represents around 3% of total annual publications. This topic has moved through clear stages: first, classical text augmentation [39] tests methods like synonym and mention replacement for German legal NER that show similar results to a bidirectional Long Short-Term Memory Network with a Conditional Random Field layer and a Transformer-based model consuming less time and under low-resource domain constraints. The next wave involves dedicated neural architectures [40] trained as an encoder-decoder with selective masked denoising objectives that exploit specialized legal vocabulary and study [41] pursue dsemantic augmentation of questions for legal QA. Recently, researchers have shifted their interest to LLMs [42], showing that with few examples and well-engineered prompts, LLMs can generate pseudo-labelled data; meanwhile, study [43] used GPT to create human-like annotated data and shows that this approach can enhance the accuracy of NER in the legal field and address data scarcity issues.

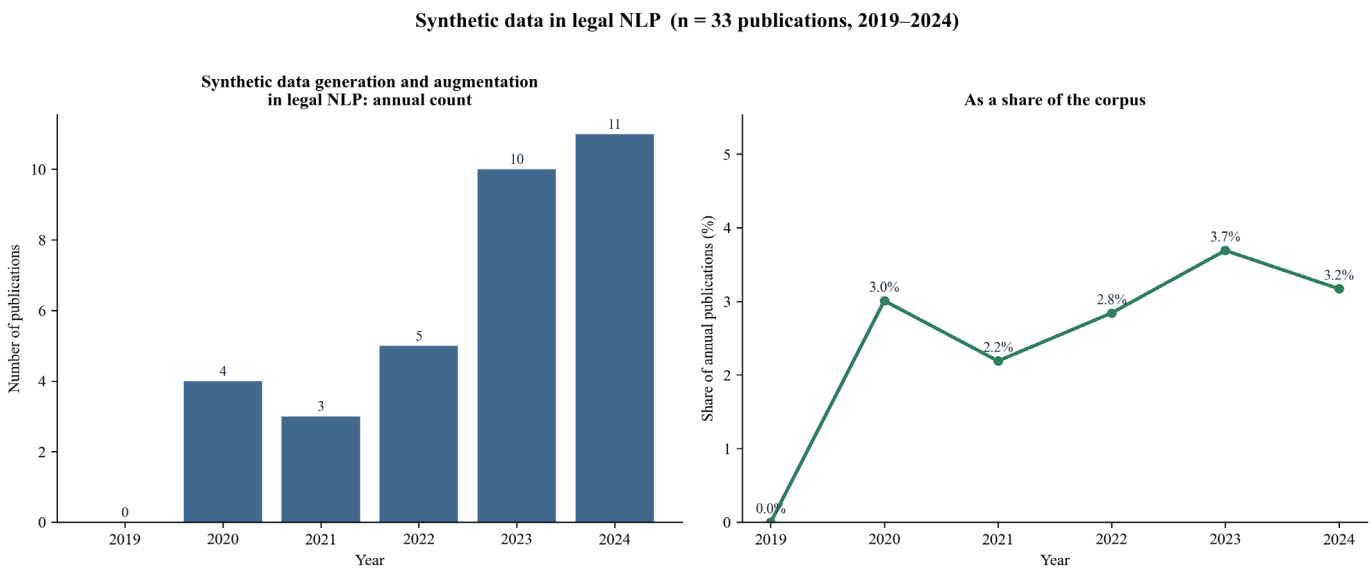

Figure 13. Evolution of synthetic data generation in legal natural language processing (NLP) (extraction in December 2024).

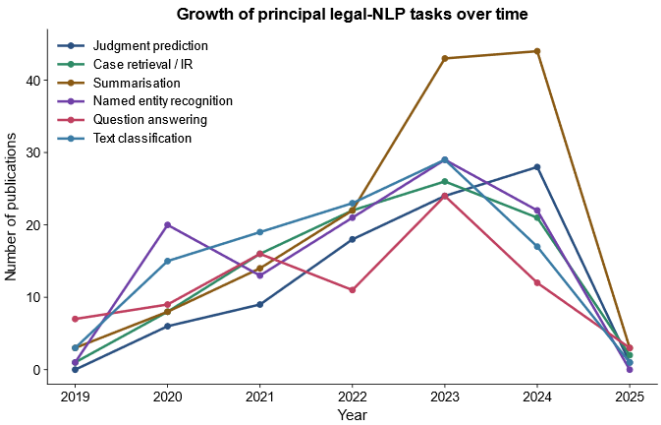

Figure 14. Evolution of legal-natural language processing (NLP) tasks over time from 2019 to 2024
Coverage of 2025 is partial.

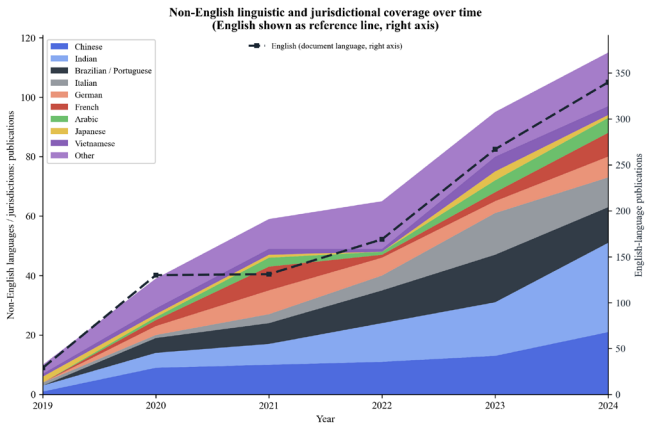

Figure 15. Yearly linguistic and jurisdictional coverage from 2019 to 2024

3.10 Legal-natural language processing tasks: Temporal trends in the literature

Figure 14 shows the growth of principal legal-NLP tasks over the period 2019 to 2024. Different legal-NLP tasks grew at very different speeds. Text classification and case retrieval were established early and remain major threads. Anchored by well-known work on large-scale multi-label legislation classification [44] and paragraph-level case retrieval modeling [23]. Judgment prediction became especially active and one of the most crowded areas thanks to benchmark corpora for English and Indian law [20, 45]. NER also developed a multilingual effort across Romanian, Brazilian, and Indian legal systems [29, 46]. Summarization and question answering existed earlier but took off in 2023 with the rise of generative models.

Figure 15 shows the temporal diversification of non-English language and jurisdiction coverage (stacked bands, left axis) counted as term frequencies in titles, abstracts, or keywords. The reference line tracks English-language publications by Scopus-documented language (right axis). English is displayed separately with a different scale to prevent masking diversification. The figure shows English-language publications rising from 29 in 2019 to 340 in 2024, maintaining their dominance throughout. On the other hand, Chinese and Indian law lead the non-English growth, yet Brazilian, Portuguese, Italian, German, French, Arabic, and others are increasing, so no single jurisdiction dominates the non-English band.

4. CONCLUSIONS

This bibliometric study is based on 1109 documents indexed in Scopus. The aim of this work is to characterize how the field has developed from 2019 to 2024; the coverage of 2025 is partial. This study uses rigorous bibliometric techniques like keyword pairing and theme mapping. Three patterns stand out: 3esearch output expanded rapidly from 29 papers in 2019 to 347 in 2024, with the sharpest growth occurring after 2021. The field is splitting along linguistic and jurisdictional lines. English still leads by volume; a broadening array of national legal systems, including Chinese, Indian, Brazilian, Italian, and German, are developing dedicated corpora and resources. Thematically, the literature has shifted from encoder-based transformers to generative LLMs. Researchers also give more focus to judgment prediction, case retrieval, classification, and NER. These findings track where

scholarly attention has gone, not its consequences for legal practice. The approach carries several limitations since we relied on Scopus alone, which means workshop papers, preprints, and legal scholarship from non-indexed venues are largely missing; citation-based metrics advantage earlier publications, which have had longer to accumulate references, so citation counts skew toward earlier work.

Data scarcity in the legal field is one of the big blocking issues in this domain especially legal data with a high quality. There is currently a real doubt about whether today's data generation methods can really produce legal texts that hold up scrutiny. Synthetic data generation helps in two ways: data privacy protection and providing researchers with the rich and varied dataset they need in order to train and test their models properly. Data isn't the only problem; interdisciplinary collaboration and holding the line on solid and rigorous methods are a must. These pieces mean NLP could improve how law works everywhere. Where we are facing the issue of legal data scarcity, new approaches to generating and augmenting data will help to overcome this gap. As a result, legal NLP tools will be available to many more legal systems around the world.

Future research could extend this analysis to additional databases such as Web of Science and the ACL Anthology to see how much of the picture Scopus is missing, follow-up analysis covering 2025, and develop more rules or scores to evaluate the synthetic legal data generated using LLMs.

REFERENCES

[1] Ellegaard, O., Wallin, J.A. (2015). The bibliometric analysis of scholarly production: How great is the impact? Scientometrics, 105(3): 1809-1831. https://doi.org/10.1007/s11192-015-1645-z

[2] Ellegaard, O. (2018). The application of bibliometric analysis: Disciplinary and user aspects. Scientometrics, 116(1): 181-202. https://doi.org/10.1007/s11192-018-2765-z

[3] Zupic, I., Čater, T. (2015). Bibliometric methods in management and organization. Organizational Research Methods, 18(3): 429-472. https://doi.org/10.1177/1094428114562629

[4] Devlin, J., Chang, M.W., Lee, K., Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, Minnesota, pp. 4171-4186. https://doi.org/10.18653/v1/N19-1423

[5] Brown, T., Mann, B., Ryder, N., et al. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33: 1877-1901. https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf.

[6] Raffel, C., Shazeer, N., Roberts, A., et al. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140): 1-67.

[7] Anand, D., Wagh, R. (2022). Effective deep learning approaches for summarization of legal texts. Journal of King Saud University-Computer and Information Sciences, 34(5): 2141-2150. https://doi.org/10.1016/j.jksuci.2019.11.015

[8] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R.R., Le, Q.V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. In Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, pp. 1-11. https://proceedings.neurips.cc/paper_files/paper/2019/file/dc6a7e655d7e5840e66733e9ee67cc69-Paper.pdf.

[9] Kien, P.M., Nguyen, H.T., Bach, N.X., Tran, V., Le Nguyen, M., Phuong, T.M. (2020). Answering legal questions by learning neural attentive text representation. In Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain, pp. 988-998. https://doi.org/10.18653/v1/2020.coling-main.86

[10] Jang, M.E., Stikkel, G. (2024). Leveraging natural language processing and large language models for assisting due diligence in the legal domain. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Mexico City, Mexico, pp. 155-164. https://doi.org/10.18653/v1/2024.naacl-industry.14

[11] Krasadakis, P., Sakkopoulos, E., Verykios, V.S. (2024). A survey on challenges and advances in natural language processing with a focus on legal informatics and low-resource languages. Electronics, 13(3): 648. https://doi.org/10.3390/electronics13030648

[12] Greco, C.M., Tagarelli, A. (2024). Bringing order into the realm of transformer-based language models for artificial intelligence and law. Artificial Intelligence and Law, 32(4): 863-1010. https://doi.org/10.1007/s10506-023-09374-7

[13] Ahmi, A. (2024). Bibliometric Analysis Using biblioMagika®. 2nd ed. Alor Setar, Kedah, Malaysia: Academic Research Society of Malaysia.

[14] Gu, X., Akoglu, L., Rinaldo, A. (2019). Statistical analysis of nearest neighbor methods for anomaly detection. In Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, pp. 1-11. https://proceedings.neurips.cc/paper_files/paper/2019/file/805163a0f0f128e473726ccda5f91bac-Paper.pdf.

[15] Al-Bakry, A., Al-Rikaby, M. (2016). Enhanced Levenshtein edit distance method functioning as a string-to-string similarity measure. Iraqi Journal for Computers and Informatics, 42(1): 48-54. https://doi.org/10.25195/ijci.v42i1.83

[16] Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM Computing Surveys, 55(9): 1-35. https://doi.org/10.1145/3560815

[17] Allam, A.M.N., Haggag, M.H. (2016). The question answering systems: A survey. International Journal of Research and Reviews in Information Sciences, 2(3): 211-221.

[18] Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., Androutsopoulos, I. (2020). LEGAL-BERT: The muppets straight out of law school. In Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 2898-2904. https://doi.org/10.18653/v1/2020.findings-emnlp.261

[19] Medvedeva, M., Vols, M., Wieling, M. (2020). Using machine learning to predict decisions of the European

Court of Human Rights. Artificial Intelligence and Law, 28(2): 237-266. https://doi.org/10.1007/s10506-019-09255-y

[20] Chalkidis, I., Androutsopoulos, I., Aletras, N. (2019). Neural legal judgment prediction in English. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, pp. 4317-4323. https://doi.org/10.18653/v1/P19-1424

[21] Zhong, H., Xiao, C., Tu, C., Zhang, T., Liu, Z., Sun, M. (2020). How does NLP benefit legal system: A summary of legal artificial intelligence. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 5218-5230. https://doi.org/10.18653/v1/2020.acl-main.466

[22] Xiao, C., Hu, X., Liu, Z., Tu, C., Sun, M. (2021). Lawformer: A pre-trained language model for Chinese legal long documents. AI Open, 2: 79-84. https://doi.org/10.1016/j.aiopen.2021.06.003

[23] Shao, Y., Mao, J., Liu, Y., et al. (2020). BERT-PLI: Modeling paragraph-level interactions for legal case retrieval. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20), pp. 3501-3507. https://doi.org/10.24963/ijcai.2020/484

[24] Chalkidis, I., Fergadiotis, E., Malakasiotis, P., Androutsopoulos, I. (2019). Large-scale multi-label text classification on EU legislation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, pp. 6314-6322. https://doi.org/10.18653/v1/P19-1636

[25] ul Hassan, F., Le, T. (2020). Automated requirements identification from construction contract documents using natural language processing. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 12(2): 04520009. https://doi.org/10.1061/(ASCE)LA.1943-4170.0000379

[26] Chalkidis, I., Jana, A., Hartung, D., et al. (2022). LexGLUE: A benchmark dataset for legal language understanding in English. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland, pp. 4310-4330. https://doi.org/10.18653/v1/2022.acl-long.297

[27] Sydora, C., Stroulia, E. (2020). Rule-based compliance checking and generative design for building interiors using BIM. Automation in Construction, 120: 103368. https://doi.org/10.1016/j.autcon.2020.103368

[28] Van Eck, N.J., Waltman, L. (2014). Visualizing bibliometric networks. In Measuring Scholarly Impact: Methods and Practice, Springer International Publishing, Cham, pp. 285-320. https://doi.org/10.1007/978-3-319-10377-8_13

[29] Correia, F.A., Almeida, A.A., Nunes, J.L., et al. (2022). Fine-grained legal entity annotation: A case study on the Brazilian supreme court. Information Processing & Management, 59(1): 102794. https://doi.org/10.1016/j.ipm.2021.102794

[30] Katz, D.M., Bommarito, M.J., Gao, S., Arredondo, P. (2024). GPT-4 passes the bar exam. Philosophical transactions. Series A, Mathematical, Physical, and Engineering Sciences, 382(2270): 20230254. https://doi.org/10.1098/rsta.2023.0254

[31] Blair-Stanek, A., Holzenberger, N., Van Durme, B. (2023). Can GPT-3 perform statutory reasoning? In Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law, New York, NY, USA, pp. 22-31. https://doi.org/10.1145/3594536.3595163

[32] Liga, D., Robaldo, L. (2023). Fine-tuning GPT-3 for legal rule classification. Computer Law & Security Review, 51: 105864. https://doi.org/10.1016/j.clsr.2023.105864

[33] Sovrano, F., Palmirani, M., Vitali, F. (2020). Legal knowledge extraction for knowledge graph based question-answering. In Legal Knowledge and Information Systems: JURIX 2020: The Thirty-third Annual Conference, Brno, Czech Republic, pp. 143-153. https://doi.org/10.3233/FAIA200858

[34] Moon, S., Lee, G., Chi, S. (2022). Automated system for construction specification review using natural language processing. Advanced Engineering Informatics, 51: 101495. https://doi.org/10.1016/j.aei.2021.101495

[35] Mahindrakar, A., Joshi, K.P. (2020). Automating GDPR compliance using policy integrated blockchain. In 2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity): IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), Baltimore, MD, USA, pp. 86-93. https://doi.org/10.1109/BigDataSecurity-HPSC-IDS49724.2020.00026

[36] Aria, M., Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4): 959-975. https://doi.org/10.1016/j.joi.2017.08.007

[37] Noguti, M.Y., Vellasques, E., Oliveira, L.S. (2020). Legal document classification: An application to law area prediction of petitions to public prosecution service. In 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, pp. 1-8. https://doi.org/10.1109/IJCNN48605.2020.9207211

[38] Dharma, P.Y., Pratama, A.R. (2023). Legal judgment prediction: A systematic literature review. In 2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA), Surabaya, Indonesia, pp. 691-696. https://doi.org/10.1109/ICAMIMIA60881.2023.10427855

[39] Erd, R., Feddoul, L., Lachenmaier, C., Mauch, M.J. (2022). Evaluation of data augmentation for named entity recognition in the German Legal Domain. In AI4LEGAL/KGSum@ ISWC, pp. 62-72.

[40] Ghosh, S., Evuru, C.K.R., Kumar, S., et al. (2023). DALE: Generative data augmentation for low-resource legal NLP. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Singapore, pp. 8511-8565. https://doi.org/10.18653/v1/2023.emnlp-main.528

[41] Li, B., Fan, S., Huang, J. (2024). CSAFT: Continuous semantic augmentation fine-tuning for legal large language models. In International Conference on Artificial Neural Networks, Springer Nature Switzerland, Cham, pp. 293-307. https://doi.org/10.1007/978-3-031-72344-5_20

[42] Sheik, R., Siva Sundara, K.P., Nirmala, S.J. (2024). Neural data augmentation for legal overruling task: Small deep learning models vs. large language models. Neural Processing Letters, 56(2): 121.

https://doi.org/10.1007/s11063-024-11574-4
[43] Zin, M.M., Nguyen, H.T., Satoh, K., Nishino, F. (2024). Addressing annotated data scarcity in legal information extraction. In JSAI International Symposium on Artificial Intelligence, Springer Nature Singapore, Singapore, pp. 77-92. https://doi.org/10.1007/978-981-97-3076-6_6
[44] Tuggener, D., Von Däniken, P., Peetz, T., Cieliebak, M. (2020). LEDGAR: A large-scale multi-label corpus for text classification of legal provisions in contracts. In Proceedings of the Twelfth Language Resources and Evaluation Conference, Marseille, France, pp. 1235-1241. https://aclanthology.org/2020.lrec-1.155/.
[45] Malik, V., Sanjay, R., Nigam, S.K., et al. (2021). ILDC for CJPE: Indian legal documents corpus for court judgment prediction and explanation. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, pp. 4046-4062. https://doi.org/10.18653/v1/2021.acl-long.313
[46] Păiș, V., Mitrofan, M., Gasan, C.L., Coneschi, V., Ianov, A. (2021). Named entity recognition in the Romanian legal domain. In Proceedings of the Natural Legal Language Processing Workshop 2021, Punta Cana, Dominican Republic, pp. 9-18. https://doi.org/10.18653/v1/2021.nllp-1.2