ENTITY RECOGNITION IN INDIAN LEGAL COURT JUDGEMENTS

[1]Dr. P. Babu, [2]Syed Sohail, [3]Rehan Shaik, [4]Shaik Zaheer, [5]Shaik Kaif
[1]Associate Professor and HOD, [2,3,4,5]UG Students, Department Of Computer Science & Engineering(AI&ML)
Geethanjali Institute Of Science And Technology, Gangavaram (V), Kovur(M), SPSR Nellore (Dt), Andhra Pradesh, India-524137

Abstract

For developing legal AI applications, it is essential to have access to judicial data and open-source foundational AI building blocks like Named Entity Recognition. Named Entity Recognition (NER) in Indian court judgments is a crucial task in legal text processing, enabling structured extraction of key entities such as case names, judges, laws, locations, and legal provisions. Due to the complex linguistic structure, domain-specific terminology, and varying formats of judicial documents, traditional NER models struggle to achieve high accuracy. This project aims to develop an advanced NER system tailored for Indian court judgments by leveraging machine learning and natural language processing (NLP) techniques. The system will be trained on annotated legal texts, utilizing deep learning architectures such as Transformer-based models (e.g., BERT, Legal-BERT) for improved entity recognition. Challenges such as entity ambiguity, multilingual content, and unstructured text formats will be addressed using domain adaptation and contextual embeddings. The resulting model is expected to enhance legal document analysis, aiding in information retrieval, case summarization, and legal research.

Keywords: ML, NLP, NER, entity recognition

Introduction

ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) stands as a pivotal force in modern society, revolutionizing industries and reshaping the way we work, live, and interact with technology. Its importance lies in its capacity to automate tasks, generate valuable insights from vast datasets, and enhance decision-making processes across various sectors. Through AI, businesses can streamline operations, optimize resource allocation, and deliver personalized experiences to customers, leading to increased efficiency, productivity, and competitiveness. Moreover, AI holds immense potential in healthcare, where it can aid in medical diagnosis, drug discovery, and patient care, ultimately improving health outcomes and saving lives. Furthermore, AI contributes to safety and security by detecting anomalies, identifying threats, and bolstering cybersecurity measures. Its ability to analyze patterns and predict future trends enables organizations to anticipate risks and take proactive measures to mitigate them. Beyond its immediate applications, AI fuels innovation and creativity, empowering individuals and businesses to explore new frontiers in art, design, and problem-solving. As AI continues to evolve and permeate every aspect of society, its importance will only grow, shaping the future of technology and human interaction in profound ways

MACHINE LEARNING

Machine Learning (ML) plays a pivotal role in AI by providing algorithms and techniques that enable computers to learn from data and make predictions or decisions without being explicitly programmed. ML algorithms power many of the AI applications we encounter daily, from personalized recommendations on streaming platforms to predictive maintenance in manufacturing.ML has transformed industries by enabling more accurate forecasts, better risk assessments, and data- driven decision-making. As the volume of data continues to grow exponentially, ML becomes increasingly important for extracting valuable insights and knowledge from vast datasets that wouldbe impractical or impossible for humans to process manually

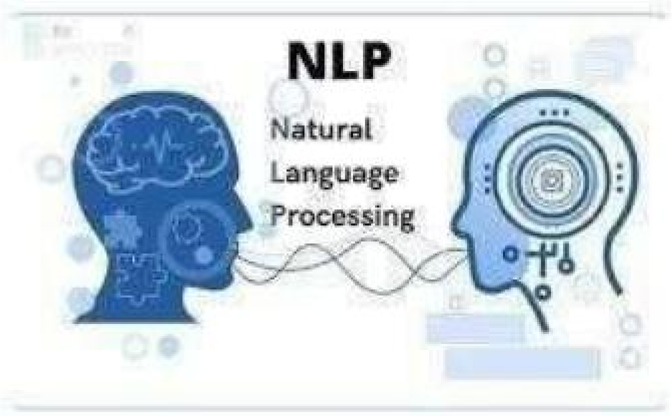


Fig 1. Natural Language Processing

Natural Language Processing (NLP) is a critical component of AI that focuses on enabling computers to understand, interpret, and generate human language. With the proliferation of text data on the internet, social media, and other digital platforms, NLP has become indispensable for extracting meaningful insights from unstructured text data. NLP powers virtual assistants like Siri and Alexa, language translation services like Google Translate, and sentiment analysis tools used by businesses to gauge customer opinions and feedback. By enabling computers to understand and interact with human language, NLP facilitates communication between humans and machines, leading to more intuitive and user-friendly interfaces and driving innovation in areas such as customer service, healthcare, education, and more

SENTIMENT ANALYSIS

Sentiment analysis, a powerful tool in the realm of natural language processing, offers numerous advantages across various sectors. It enables businesses to glean invaluable insights into customer sentiments, preferences, and attitudes towards their products or services, facilitating informed decision-making and targeted strategies. By harnessing sentiment analysis, organizations can monitor real-time feedback from social media, customer reviews, and online platforms, allowing for prompt responses to issues and effective reputation management. Furthermore, sentiment analysis aids in competitive analysis by evaluating sentiment across competitors, providing actionable intelligence for market positioning. Marketers benefit from sentiment analysis by optimizing campaigns based on customer reactions and sentiment towards marketing messages. However, challenges such as the interpretation of nuanced language, reliance on quality data, and language-specific nuances pose obstacles to accurate sentiment analysis. Additionally, ethical considerations regarding privacy and potential biases underscore the importance of responsible implementation and transparent practices in sentiment analysis applications.

Literature Review

[1] S. Kumar and P. Bhattacharyya, “Named Entity Recognition in Indian Legal Documents Using Conditional Random Fields” (2023) In this paper, the authors proposed a Conditional Random Fields (CRF)-based approach for Named Entity Recognition (NER) in Indian legal documents. By leveraging a combination of linguistic features and domain-specific rules, they aimed to improve entity extraction accuracy in court judgments. The study demonstrated promising results in recognizing legal entities such as case numbers, judges' names, and legal statutes. Drawbacks & Gaps The model heavily relied on handcrafted features, making it less adaptable to new datasets. Lack of deep learning techniques limited the model’s ability to capture complex contextual dependencies. [2] A. Bhatia and R. Sharma, “Deep Learning for Named Entity Recognition in Indian Court Judgments” (2023) In this study, the authors explored deep learning-based methods, particularly Bidirectional LSTMs (BiLSTMs) and transformers, to enhance NER performance in Indian legal texts. They used pre-trained embeddings such as Fast Texts and BERT to improve contextual understanding. Their findings indicated that transformer-based architecture outperformed traditional machine learning models in recognizing named entities within court judgements.

[3] R. Gupta, S. Verma, and P. Jain, “Legal-BERT: A Pre-Trained Language Model for Indian Court Judgments” (2022) This paper introduced Legal-BERT, a transformer-based language model specifically finetuned for legal documents in the Indian judiciary. The authors demonstrated that fine-tuning BERT on legal corpora significantly improved entity recognition tasks, outperforming generic NLP models. The study emphasized the importance of domain-specific training in improving NER performance for legal text processing.

[4] M. Rao and D. Srinivasan, “Hybrid Named Entity Recognition for Indian Legal Texts” (2022) In this paper, the authors proposed a hybrid NER approach combining rule-based methods with deep learning models to improve accuracy in recognizing legal entities in Indian court judgments. The study integrated domain-specific lexicons with LSTM-CRF models to leverage both linguistic rules and statistical learning.

[5] K. Mehta and V. Iyer, “Multi-Lingual Named Entity Recognition in Indian Legal Texts” The authors tackled the issue of multilingualism in Indian court judgments by developing a multi-lingual NER system trained on datasets from multiple Indian languages, including Hindi, Tamil, and Bengali. They employed a transfer learning approach using multilingual BERT (mBERT) to recognize named entities across different legal language datasets

[6] S. Aithal, R. Suresh, and M. Bhargavi, “LEGALEXTRACT: A Tool for Legal Document Structuring and Entity Recognition” (2021) This study presents LEGALEXTRACT, a system designed for structuring unorganized legal documents and extracting named entities using a hybrid approach. It combines rule-based methods with Conditional Random Fields (CRF) to identify legal entities such as case numbers, court names, and dates. The rule-based system is tailored to specific formats commonly found in legal documents, while CRF aids in learning contextual patterns.

[7] I. Chalkidis, M. Fergadiotis, et al., “Legal-BERT: The Muppets Straight out of Law School” (2020) This paper introduces Legal-BERT, a domain-specific variant of BERT pretrained on a large corpus of EU and US legal texts. The model demonstrated superior performance on legal NLP tasks compared to general-purpose BERT, showcasing the importance of domain pretraining.

[8] S. Kumar, A. Gupta, and R. Dutta, “A Dataset for NER in Legal Texts in Indian Languages” (2022) The researchers created one of the first multilingual legal NER datasets specifically for Indian languages, including Hindi, Bengali, and Tamil. They used models like IndicBERT and MuRIL to perform NER, aiming to improve information extraction in low-resource legal settings.

[9] J. Patel and V. Gupta, “Deep Learning Approaches for Legal NER” (2019) This study evaluates the effectiveness of deep learning architectures, including BiLSTM, BiLSTM-CRF, and Transformers, for extracting entities from legal texts. The research shows that deep learning models outperform traditional statistical methods in capturing contextual information and handling long dependencies.

[10] M. Reddy and T. Srinivas, “Transfer Learning for Indian Legal NER: BERT vs RoBERTa” (2023) This paper compares various transformer-based models—BERT, RoBERTa, and LegalBERT—for their performance on Indian legal Named Entity Recognition tasks. The study investigates the role of transfer learning and the need for domain adaptation in improving accuracy.

Proposed System

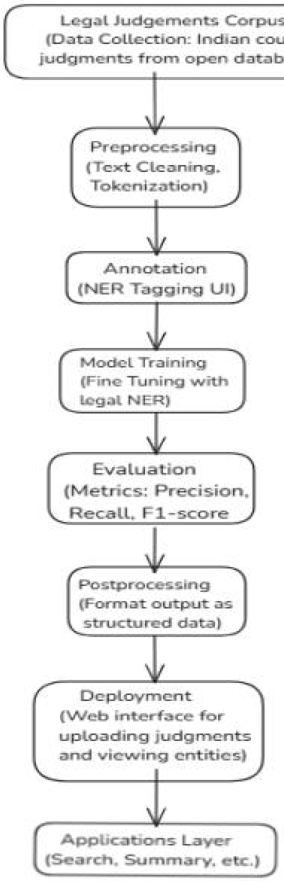


Fig.2. Work flow of the proposed model

The workflow for the project "Entity Recognition in Indian Legal Court Judgements" involves collecting legal judgments from open databases and preprocessing them through text cleaning and tokenization. These texts are then annotated using an NER tagging tool, followed by model training tailored to legal texts. The model's performance is evaluated using metrics like Precision, Recall, and F1-score. Post processing structures the output, which is then deployed through a web interface for uploading judgments and viewing recognized entities. The final application layer supports functionalities like search, summarization, and legal analysis.

Data Collection Source: Indian court judgments from open databases (e.g., Indian Kanoon, Judis)• Format: PDF, HTML, or text documents

Preprocessing: Text extraction from PDFs/HTML Sentence segmentation & tokenization Language detection and normalization (handle multilingual content) Removal of noise, headers/footers, boilerplate

Annotation: Manual or semi-automated annotation using tools like Prodigy or Doccano.

Label entities: CASE_NAME, JUDGE, LAW, LOCATION, PROVISION, etc.• Model Training• Fine-tuning with legal NER tags. Use contextual embeddings.

Evaluation Metrics: Precision, Recall, F1-score. Compare against baseline models (SpaCy NER, Stanford NER). Post-processing Merge entity spans. Resolve co-references. Format output as structured data (JSON/CSV).

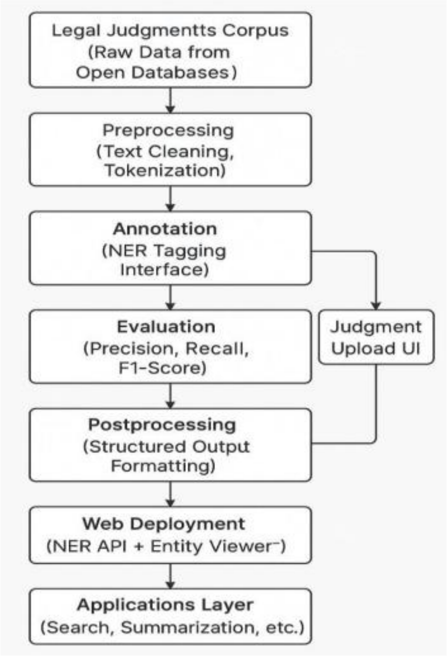

Legal Judgmentts Corpus (Raw Data from Open Databases)
Preprocessing (Text Cleaning, Tokenization)
Annotation (NER Tagging Interface)
Evaluation (Precision, Recall, F1-Score)
Judgment Upload UI
Postprocessing (Structured Output Formatting)
Web Deployment (NER API + Entity Viewer⁻)
Applications Layer (Search, Summarization, etc.)

Fig.3. Architectural Design

Step 5: Data Splitting & Summary

Split	Years	Preambles	Sentences	Entities
Train	1950–2017	1,560	9,435	29,964
Dev	2018–2022	125	949	3,216
Test	2018–2022	441	4,060	13,365

Results & Analysis

Evaluation metrics are essential tools used to assess the performance and accuracy of machine learning models and algorithms. These metrics provide quantitative measures that enable researchers and practitioners to evaluate the effectiveness of their methods and make informed decisions about model selection and optimization. Moreover, the choice of evaluation metrics depends on the nature of the problem being addressed and the desired outcome. By utilizing a combination of evaluation metrics, practitioners can gain comprehensive insights into the overall performance of their models and make informed decisions regarding their deployment and optimization strategies. These Evaluation metrics play a crucial role in not only validating the performance of machine learning models but also in comparing different models and algorithms. They help in identifying the strengths and weaknesses of a model, guiding the refinement process for better outcomes. Common evaluation metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), accuracy, and execution time. Each metric serves a specific purpose in evaluating different aspects of model performance, such as prediction accuracy, error magnitude, and computational efficiency.

Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) is a metric used to measure the average absolute difference between predicted values and actual values. It is calculated by taking the average of the absolute differences between each predicted value and its corresponding actual value. The formula for MAE is

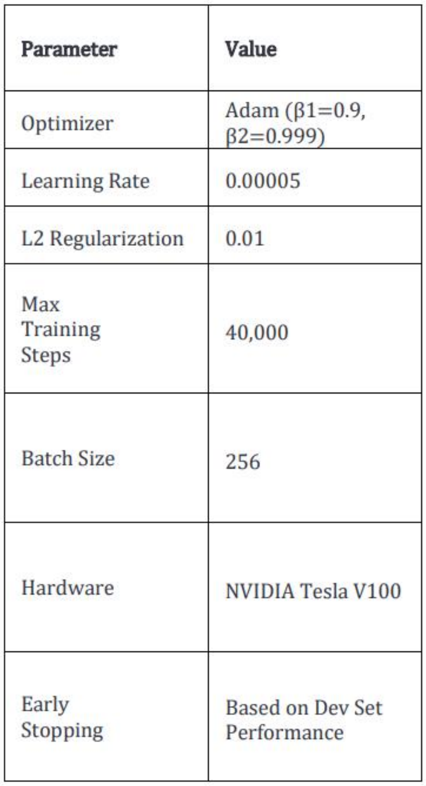
Mean Squared Error (MSE)

The Mean Squared Error (MSE) is a metric used to measure the average squared difference between predicted values and actual values. It is calculated by taking the average of the squared differences between each predicted value and its corresponding actual value. The formula for MSE is:

Accuracy

Accuracy is a metric used to measure the proportion of correctly predicted values out of the total number of predictions. It is calculated by dividing the number of correct predictions by the total number of predictions. The formula for accuracy is:

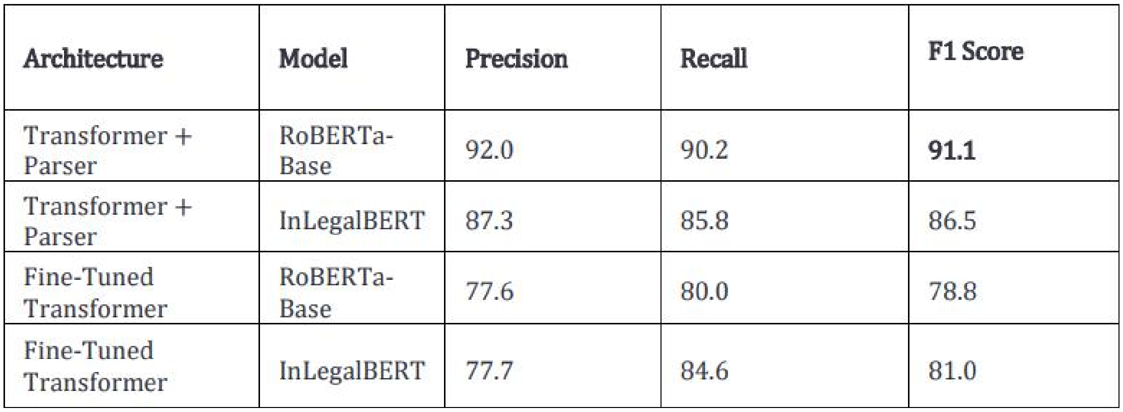
Precision (%): Precision measures how many of the predicted defects were actually correct.



Parameter	**Value**
Optimizer	Adam (β1=0.9, β2=0.999)
Learning Rate	0.00005
L2 Regularization	0.01
Max Training Steps	40,000
Batch Size	256
Hardware	NVIDIA Tesla V100
Early Stopping	Based on Dev Set Performance

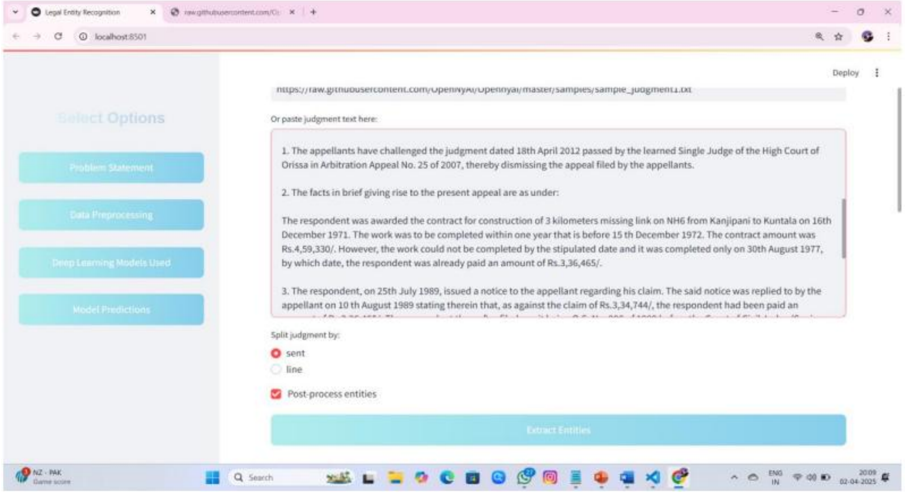
Table 1 Training Configuration

Result Comparison



Architecture	**Model**	**Precision**	**Recall**	**F1 Score**
Transformer + Parser	RoBERTa-Base	92.0	90.2	**91.1**
Transformer + Parser	InLegalBERT	87.3	85.8	86.5
Fine-Tuned Transformer	RoBERTa-Base	77.6	80.0	78.8
Fine-Tuned Transformer	InLegalBERT	77.7	84.6	81.0

Table 2. Result Comparison



https://raw.githubusercontent.com/OpenNyAI/Opennyai/master/samples/sample_judgment1.txt

Or paste judgment text here:

1. The appellants have challenged the judgment dated 18th April 2012 passed by the learned Single Judge of the High Court of Orissa in Arbitration Appeal No. 25 of 2007, thereby dismissing the appeal filed by the appellants.

2. The facts in brief giving rise to the present appeal are as under:

The respondent was awarded the contract for construction of 3 kilometers missing link on NH6 from Kanjipani to Kuntala on 16th December 1971. The work was to be completed within one year that is before 15 th December 1972. The contract amount was Rs.4,59,330/. However, the work could not be completed by the stipulated date and it was completed only on 30th August 1977, by which date, the respondent was already paid an amount of Rs.3,36,465/.

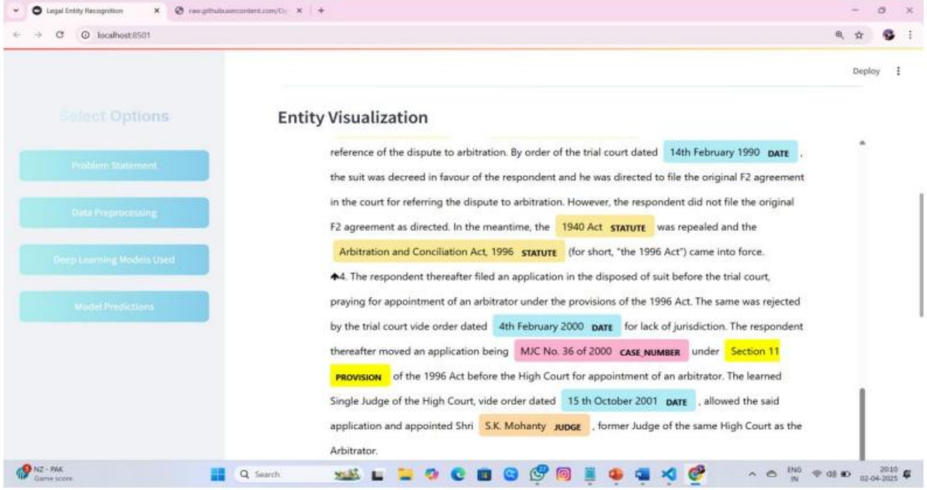
3. The respondent, on 25th July 1989, issued a notice to the appellant regarding his claim. The said notice was replied to by the appellant on 10 th August 1989 stating therein that, as against the claim of Rs.3,34,744/, the respondent had been paid an

Split judgment by:

● sent
○ line

☑ Post-process entities

Fig.4. Model Prediction



Entity Visualization

reference of the dispute to arbitration. By order of the trial court dated 14th February 1990 DATE , the suit was decreed in favour of the respondent and he was directed to file the original F2 agreement in the court for referring the dispute to arbitration. However, the respondent did not file the original F2 agreement as directed. In the meantime, the 1940 Act STATUTE was repealed and the Arbitration and Conciliation Act, 1996 STATUTE (for short, "the 1996 Act") came into force.

4. The respondent thereafter filed an application in the disposed of suit before the trial court, praying for appointment of an arbitrator under the provisions of the 1996 Act. The same was rejected by the trial court vide order dated 4th February 2000 DATE for lack of jurisdiction. The respondent thereafter moved an application being MJC No. 36 of 2000 CASE_NUMBER under Section 11 PROVISION of the 1996 Act before the High Court for appointment of an arbitrator. The learned Single Judge of the High Court, vide order dated 15 th October 2001 DATE , allowed the said application and appointed Shri S.K. Mohanty JUDGE , former Judge of the same High Court as the Arbitrator.

Fig.5. Recognition of Entities

Conclusion

This thesis has explored the implementation and challenges of Named Entity Recognition (NER) in the context of Indian court judgments, a domain characterized by complex legal language, diverse entity types, and significant variations in writing styles. By analyzing various approaches—ranging from traditional rule-based systems to modern machine learning and deep learning models—the study has demonstrated that domain-specific preprocessing and tailored model architectures significantly improve entity recognition accuracy. The proposed methodology, which incorporates a combination of annotated datasets, legal ontologies, and contextual embeddings, showed promising results in extracting critical entities such as judges, petitioner/respondent names, statutes, and case citations. The results underscore the importance of leveraging domain expertise and context-aware models to handle linguistic ambiguities and syntactic irregularities common in legal documents. While challenges remain—particularly in scaling the solution across multiple Indian languages and court formats—the progress achieved in this thesis offers a concrete step forward. Future work may explore multilingual and cross-lingual NER, unsupervised pre-training on vast legal corpora, and integration with knowledge graphs for enhanced reasoning capabilities.

FUTURE SCOPE

The future scope of this research extends to several key advancements in legal AI. One significant direction is the expansion of Named Entity Recognition (NER) models to support multilingual court judgments, enabling effective processing of regional languages like Hindi, Tamil, and Bengali. Additionally, integrating more advanced deep learning architectures, such as GPTbased models, can enhance contextual understanding and improve the accuracy of entity recognition in complex legal texts. The system can also be deployed as a real-time legal assistant for lawyers and judges, aiding in legal research, precedent analysis, and case law searches. Further, cloud-based implementation will allow scalable and efficient access to structured legal data.

References

1. Iosif Angelidis, Ilias Chalkidis, and Manolis Koubarakis, 2018. Named entity recognition, linking and generation for Greek legislation. In JURIX, pages 1–10.
2. Valentin Barriere and Amaury Fouret, 2019. A simple but efficient way to generate and use contextual dictionaries for named entity recognition. The application to French legal texts. In Proceedings of the 22nd Nordic Conference on Computational Linguistics, pages 327–332.
3. Darina Benikova, Seid Muhie, Yimam Prabhakaran, and Santhanam Chris Biemann. 2015. C.: Germaner: Free open German named entity recognition tool. In In: Proc. GSCL-2015. Citeseer.
4. Daniel M Bikel, Richard Schwartz, and Ralph M Weischedel. 1999. An algorithm that learns what's in a name. Machine learning, 34(1):211–231.
5. Mandar Joshi, Danqi Chen, Yinhan Liu, Daniel S Weld, Luke Zettlemoyer, and Omer Levy. 2020. Spanbert: Improving pre-training by representing and predicting spans. Transactions of the Association for Computational Linguistics, 8:64–7.
6. Ikuya Yamada, Akari Asai, Hiroyuki Shindo, Hideaki Takeda, and Yuji Matsumoto. 2020. Luke: Deep contextualized entity representations with entity aware self-attention.
7. Vasile P˘ais, Maria Mitrofan, Carol Luca Gasan, Vlad Coneschi, and Alexandru Ianov. 2021. Named entity recognition in the Romanian legal domain. In Proceedings of the Natural Legal Language Processing Workshop 2021, pages 9–18.
8. Xiaozhi Wang, Tianyu Gao, Zhaocheng Zhu, Zhengyan Zhang, Zhiyuan Liu, Juanzi Li, and Jian Tang. 2021. Kepler: A unified model for knowledge embedding and pretrained language representation. Transactions of the Association for Computational Linguistics, 9:176–194.
9. Lucia Zheng, Neel Guha, Brandon R Anderson, Peter Henderson, and Daniel E Ho. 2021. When does pre training help assessing self-supervised learning for law and the casehold dataset of 53,000+ legal holdings. In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law, pages 159–168.
10. Shounak Paul, Arpan Mandal, Pawan Goyal, and Sap Tarshi Ghosh. 2022b. Pre-training transformers on Indian legal text.