



## Next Generation Legal FIR Analysis: AI-based Summarization and Section Identification

Loshini K<sup>1\*</sup>, Deepasangkini K<sup>2</sup>, Tharun S<sup>3</sup>, K.S. Janu<sup>4</sup>, D.Parameswari<sup>5</sup>

DOI:10.31033/IJEMR/16.2.2026.1854

<sup>1\*</sup> Loshini K, Department of Artificial Intelligence and Machine Learning, Jerusalem College of Engineering, Chennai, Tamil Nadu, India.

<sup>2</sup> Deepasangkini K, Department of Artificial Intelligence and Machine Learning, Jerusalem College of Engineering, Chennai, Tamil Nadu, India.


<sup>3</sup> Tharun S, Department of Artificial Intelligence and Machine Learning, Jerusalem College of Engineering, Chennai, Tamil Nadu, India.

<sup>4</sup> K.S. Janu, Department of Artificial Intelligence and Machine Learning, Jerusalem College of Engineering, Chennai, Tamil Nadu, India.

<sup>5</sup> D.Parameswari, Department of Artificial Intelligence and Machine Learning, Jerusalem College of Engineering, Chennai, Tamil Nadu, India.

The increasing number and complexity of first information reports (firs) have made manual legal analysis time-consuming and error-prone. Although firs are essential to criminal investigations, their unstructured format and varied language often make legal interpretation difficult. This paper introduces jurismate, an ai-based system that helps automate the analysis of firs. Jurismate uses the gemini 1.5 language model to extract and structure important details from fir documents and applies a bart-based zero-shot classification method to determine whether an fir is lawful, unlawful, or unclear without requiring labeled data. To support legal research, the system uses semantic embeddings stored in pgvector to retrieve relevant laws and past cases based on similarity. Experimental results show that jurismate improves efficiency, ensures consistent analysis, and provides better support for legal decision-making.

**Keywords:** FIR Analysis, Legal AI, Semantic Retrieval, Zero-Shot Classification

Corresponding Author	How to Cite this Article	To Browse
Loshini K, Department of Artificial Intelligence and Machine Learning, Jerusalem College of Engineering, Chennai, Tamil Nadu, India. Email: <a href="mailto:loshini.am2022@jerusalemengg.ac.in">loshini.am2022@jerusalemengg.ac.in</a>	Loshini K, Deepasangkini K, Tharun S, K.S. Janu, D.Parameswari, Next Generation Legal FIR Analysis: AI-based Summarization and Section Identification. Int J Engg Mgmt Res. 2026;16(2):12-19. Available From <a href="https://ijemr.vandanapublications.com/index.php/j/article/view/1854">https://ijemr.vandanapublications.com/index.php/j/article/view/1854</a>	

Manuscript Received  
2026-02-28

Review Round 1  
2026-03-18

Review Round 2

Review Round 3

Accepted  
2026-04-01

Conflict of Interest  
None

Funding  
Nil

Ethical Approval  
Yes

Plagiarism X-checker  
5.34

Note



© 2026 by Loshini K, Deepasangkini K, Tharun S, K.S. Janu, D.Parameswari and Published by Vandana Publications. This is an Open Access article licensed under a Creative Commons Attribution 4.0 International License <https://creativecommons.org/licenses/by/4.0/> unported [CC BY 4.0].



## 1. Introduction

Manual processing of legal documents is becoming increasingly challenging due to their sheer volume and complexity, particularly when it comes to processing First Information Reports (FIRs). The FIRs provide the foundation of a criminal investigation, however, because of the unstructured nature and inconsistency of language they often lead to delays in processing as well as misinterpretations of the FIR's contents due to the complexity of the applicable law. Manually examining these documents is not only cumbersome but also results in inconsistency, especially where the relevant authority is under stress because of a high volume of cases. Some recent advances in Artificial Intelligence (AI) and Natural Language Processing (NLP) are developing ways to alleviate these issues. This paper outlines a proposed system to process FIRs that uses an AI-based solution to automate the process of extracting text from FIRs, as well as applying zero-shot classification to classify FIRs according to applicable laws and precedents. By providing this type of solution, the proposed system will improve efficiency and accuracy, while allowing for easier and more efficient management of legal workflows.

## 2. Literature Survey

Recent advances in the use of artificial intelligence (AI) and natural language processing (NLP) to help manage huge amounts of unstructured legal text have enabled the legal industry to more quickly and effectively process such data. One of the main types of legal documents that are used to lay the foundation for investigating criminal activities are First Information Reports (FIRs). The narrative structure of these documents, as well as the legal terminology used and the many different ways in which language can be used to convey the same information, make it incredibly complicated to use FIRs to investigate crimes that are based on these reports.

Ashley [1] introduced the HYPO system, one of the earliest case-based reasoning models for legal analysis, which evaluated disputes by comparing them with prior cases. This work demonstrated the feasibility of AI-assisted legal reasoning but relied heavily on structured representations and expert-defined dimensions.

Lipianina-Honcharenko et al.[2] have put forward a cyclical AI-based framework for analysing legal documents that includes the extraction of text from legal files, classifying them and then assessing their compliance. They have shown that machine learning models can yield highly accurate results when analysing legal documents; however, they have also pointed out that due to the complexity and evolving nature of legal systems, their approach needs to be flexible.

Kabir and Alam [3] examined the impact that AI has on legal research and decision-making, highlighting document analysis, prediction and efficiency improvements. The authors also note that even with the use of AI systems, ethical and legal responsibilities, such as the need for transparency and human oversight, should always be maintained.

According to Chalkidis et al. [4], large-scale classification of legal text using deep learning approaches has been proven to be significantly improved by utilising transformer-based models as compared to classic models, as well as to capture the legal context and semantics of the documents in question more effectively than classic modelling methods.

Zhong et al. [5] developed a framework for predicting legal judgements using NLP and neural network technology that illustrates the ability of AI to draw out useful features from legal narratives that can assist in making judicial decisions.

BERT, a transformer-based model for understanding text context, was introduced by Devlin et al. [6]. The architecture of BERT forms the basis of many models in the legal domain and also served as the foundation for zero-shot classification models applied within the field of law for categorizing legal documents.

Lewis [7] and colleagues propose a sequence-to-sequence denoising transformer model known as BART, which allows for zero-shot text classification. BART-based models are particularly useful in areas of the law where having a large amount of labeled training data is not possible.

The authors Reimers and Gurevych [8] introduced Sentence-BERT, which allows the rapid generation of semantically embedded sentences that can be used for similarity searches.

As a result, many current legal information retrieval systems use Sentence-BERT to determine how closely a particular document relates to another by comparing their meanings in context, rather than simply by matching keywords.

The foundation for semantic retrieval systems currently used with vector databases like PGVector was established by Johnson et al. [9], who showed the efficacy of vector similarity search using high-dimensional embeddings.

By fusing statistical learning with case-based methods, Ashley and Brüninghaus [10] expanded AI-based legal reasoning, focusing on explainability and consistency in legal decision support systems.

When taken as a whole, these studies demonstrate the shift from rule-based legal systems to data-centric, AI-driven strategies. However, the majority of current solutions place little emphasis on FIRs and instead concentrate on structured legal texts like contracts or judgments. The suggested system, which combines automated extraction, zero-shot classification, and semantic retrieval especially designed for FIR analysis, is motivated by this gap.

### 3. Proposed Methodology

The proposed system, JurisMate, is an AI-based framework that uses advanced natural language processing and semantic retrieval techniques to automate the analysis of First Information Reports (FIRs). The main goal of the system is to make FIR classification and legal information retrieval easier, more consistent, and more accurate.

#### A. Overall System Concept

JurisMate processes FIR documents digitally (in PDF format) and creates a structured, machine-readable representation. Rather than using traditional keyword or rule-based methods, this method leverages transformer language models and vector semantics to help understand the legal context so you can make sense of what it means based on the information contained within FIRs instead of just looking at keywords. As such, this is much better at handling the often unstructured format and varied terminology used in many FIRs.

#### B. FIR Document Input and Text Extraction

FIRs are submitted to the system in the form of Digitally Formed PDF documents. The Gemini 1.5 Language Model is able to understand complex legal

terminology and document structure to extract text and to parse the structure of any submitted report. The data extracted from FIRs is segmented into logical text segments, which means that any legal-related information that is extracted from an FIR is preserved, with as little irrelevant data as possible. The system will not experience the errors associated with Optical Character Recognition (OCR) methods, allowing for better accuracy when processing extracted data.

#### C. Zero-Shot Legal Classification

By utilizing a zero-shot classification method utilizing the BART (facebook/bart-large-mnli) model to categorize FIRs without having labelled legal datasets, the FIR text extracted is determined as either Lawful, Unlawful, or Unclear, based upon pre-defined categories. The use of zero-shot learning makes it easy for the system to generalize its findings across all jurisdictions and types of documents that may exist within the same repository, and thus allows for flexibility and applicability in many different legal settings where fulcrums of labelled data do not exist.

#### D. Semantic Embedding Generation

After classification, the FIR data is transformed into vector embeddings with greater-than-3D dimensionality using sentence-based transformer like models, while maintaining semantic meaning and contextual relevance. As such, by the use of high-dimensional representation of the FIR contents, the FIR content can be better indexed for meaning compared to the traditional method of keyword signing.

#### E. PGVector-Based Semantic Retrieval

Embeddings that were generated by the process are stored within a Postgres extension called PGVector, allowing for easy retrieval of statutes, legal provisions and FIRs (previously processed) based on similarity. In order to identify legally relevant data by context, the vectors that were generated from each document are compared to determine the best matches. Thus, this new approach increases both the quality and relevance of sources that will be used in future cases where similar types of documents may need to be referenced.

#### F. Methodological Advantages

The proposed methodology has several advantages over existing approaches:

1. It eliminates the need for labeled training data by utilizing a zero-shot learning approach.
2. It utilizes a transformer-based model that understands context better than any of the previously used methods.
3. It provides semantic retrieval of the information through the use of vector databases.
4. It allows for the scaling of a very large volume of FIRs.

By integrating all of these elements into one comprehensive and scalable solution, this study presents a powerful platform to conduct intelligent analysis of FIRs rapidly and consistently leading to more accurate and timely decisions by legal experts.

### 4. System Architecture

The JurisMate solution’s system architecture provides a modular, layered structure that enables scalability, maintainability and the efficient integration of AI Processing. The system consists of four main layers: Frontend Layer, Backend Layer, AI Processing Layer, and Database Layer. Each of these four layers is responsible for a specific function and communicates with the other three layers via Secure, Well-defined interfaces. This separation of concerns allows for independent development and deployment and thus supports high system reliability.

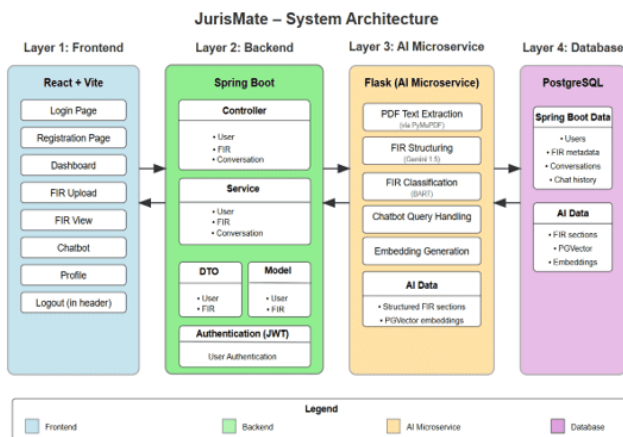


Figure 1: System Architecture

#### A. Frontend Layer

The Frontend Layer is User Focused, Built using React, this Layer offers the User a User interface (UI) to interact with the system. React provides the ability to build dynamic and responsive UIs based on Components that are focused on Document Uploading as well as visualising results.

The Frontend Layer is responsible for accepting FIR PDF from users, confirming the Upload via RESTful API, and forwarding that document to the Backend. Client Side Validation occurs to confirm that the correct File Formats are being sent to the Backend to avoid Unnecessary Processing.

#### B. Backend Layer

Utilizing Spring Boot to build the Backend Layer has allowed us to maintain centralized control over our system by controlling application processing logic, authenticating users, validating requests, and supplying a means of communication between the frontend and AI Processing Services.

REST APIs are exposed on the backend so that FIR submission and classification requests can be submitted and retrieved. Since this creates an abstraction layer that hides the complexity of the backend from the frontend, this enables the frontend to have a clean and secure interaction path.

#### C. AI Processing Layer

The AI Processing Layer is distinctively represented as an independent Flask-based microservice, used primarily for complex computing jobs relating to natural language processing. It provides FIR text extraction; Zero-Shot Legal Classification, and SEMANTIC EMBEDDING GENERATION for all related posts by creating text-based representations of the underlying documents with ease via transformer-based models allowing a greater understanding of the nuances associated with the documents involved. The Flask deployment model enables quick-scale deployment or decommissioning of AI modules, therefore offering highly flexible-base builds.

#### D. Database Layer

By using PostgreSQL aggregate-type data as well as using PGVector extension as part of this database layer, the foundation for accessing large quantities of data via an efficient "vectors" approach is created. Metadata for the System FIRs (Federally Insured) records are organized using traditional relational schema, while high dimensional vectors are stored as vector data. By using PGVector, Document Retrieval (DR) allows for effective similarity calculations as well as for determining most applicable legal documents/statutes based on the interpretation of the terms within the document rather than simple key word matching.

Unified storage provides consistency in the organization of the data, thus maximizing the performance and reliability of data management.

### E. Architectural Advantages

The multilayered structure affords various benefits, such as the ability to build on top of existing architectures to create new features, flexibility and ease of customization, and, best of all, the ability to introduce new technologies into the existing architecture with minimal hassle. The AI service's decoupled microservice architecture, combined with PGVector, provides the architecture with the ability to execute resource-intensive tasks separately from less resource-intensive operations and provides robust semantic retrieval of advanced data types. Thus, the architecture provides a scalable, efficient, and effective way to conduct FIR analysis to provide legal support for real-world applications.

## 5. Implementation Details

The JurisMate System has been built using a variety of current web-based frameworks, AI Models and Databases allowing for efficient processing, scalability and maintainability. Implemented in a Microservice model where Application Logic is separated from AI processing and Data Storage.

### A. Tools and Frameworks

React serves as the Frontend User Interface (UI) Framework for a responsive, modular application. On the back end, Spring Boot provides the foundation for creating RESTful services, validating requests and securing communication among the different components of a system. All AI features have been created as a separate Flask Microservice to allow for increased scalability and more efficient processing of CPU-intensive jobs. All communication between the microservices occurs through REST APIs using JSON formatted data.

### B. Text Extraction and Parsing

The FIRs accept digital documents in PDF form and analyze them with the Gemini 1.5 machine learning language model. Gemini 1.5 allows you to accurately extract and structurally parse the texts of legal documents, providing a clear display of contextual and semantic relationships and avoiding many features found with Optical Character Recognition (OCR) systems.

### C. Legal Classification Model

We use the BART (facebook/bart-large-mnli) transformer model to perform zero-shot legal classification of FIR text extracted from reports. In this method, extracted FIR text is classified into predetermined legal categories without the need for any training data labeled with those categories. This means you will always get valid classifications no matter where you use the model or what type of documents you use.

### D. Semantic Embedding Generation

The Sentence Transformer models convert the FIR documents into dense vector representations which provide the context for context-based retrieval by capturing semantic relationships among the legal texts. Thus, these vector representations can be utilized as a basis for similarity-based searches.

### E. Vector Storage and Retrieval

With PostgreSQL's ability to extend with PGVector, It supports loading semantic embeddings for the purposes of performing Vector Similarity Calculation on relevant statutes & Case law. This will also allow retrieval of FIRs that have yet to be processed but meet the same contextual criteria as those processed.

## 6. Experimental Results and Discussion

The evaluation of the JurisMate system, specifically its functionality for automating FIR processing and analysing Legal Documents, was completed through Functional Tests and Observational Analyses. The evaluation assessed the JurisMate's performance regarding System Behaviour, Processing Efficiency and the consistency of the Legal Classification carried out by JurisMate.

### A. System Behavior

The performance of the solution (FIR ingestion, the text extraction process, the classification of documents, and the semantic retrieval of documents) exhibited a consistent and dependable nature throughout the entire testing process. The system's ability to parse and structure digital FIR PDF documents without losing any contextual information is further substantiated by the modular architecture of the application that allowed for seamless communication between its frontend,

backend, AI micro-service, and data storage components, ensuring that no breaks occurred in the processing stages of a transaction from start to finish.

**B. FIR Processing Time**

The new system has dramatically decreased processing time when compared to manual FIR analysis. Activities typically requiring a great deal of human resources, such as document reading, legal analysis, and looking up references, are completed automatically in a few seconds or minutes. This allows for quicker case processing and a significant increase in operational efficiencies in high-volume legal settings.

**C. Classification Consistency**

Through the usage of the zero-shot classification technique, FIRs could be categorized consistently across multiple input sources using similar contextual data. The approach differs from that of manual interpretation, whereby interpretation is based on reviewer bias and opinions, whereas the use of artificial intelligence for classification removes individual subjectivity, thereby providing increased reliability.

**D. Accuracy and Observations**

The system's results showed an equivalence in accuracy and relevance between the user-created samples and the manually created samples of the same data set, despite the inability to quantify the results using the labeled dataset. Additionally, by using vector embeddings to perform semantic searches, I was able to retrieve legal articles and similar decisions relevant to my semantic studies through contextual understanding.

Overall, the experimental observations confirm that JurisMate improves efficiency, consistency, and reliability in FIR analysis, making it suitable for real-world legal and law enforcement applications.

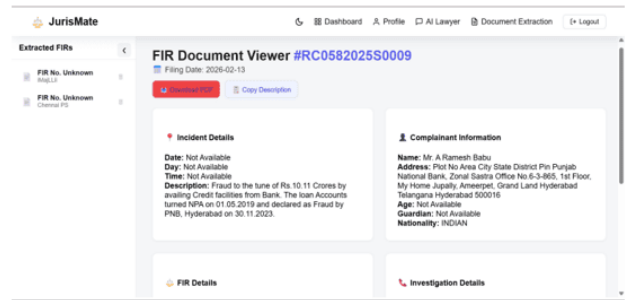


Figure 3: Document Extraction

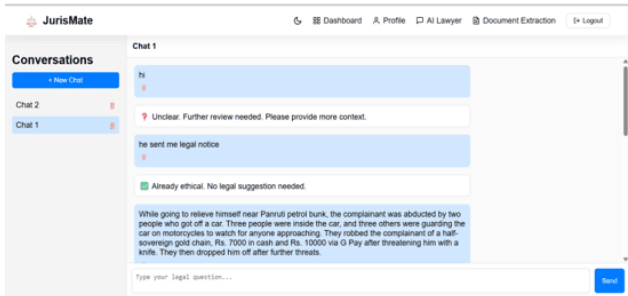


Figure 4: AI-Legal chatbot

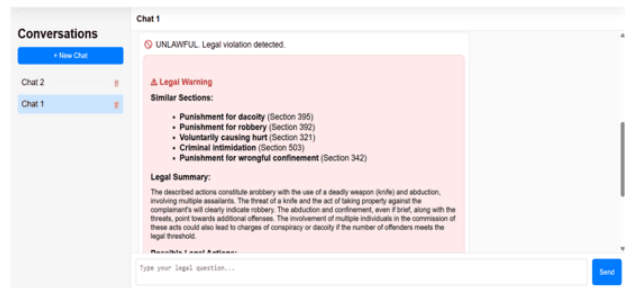


Figure 5: Classify as unlawful



Figure 6: List of Acts

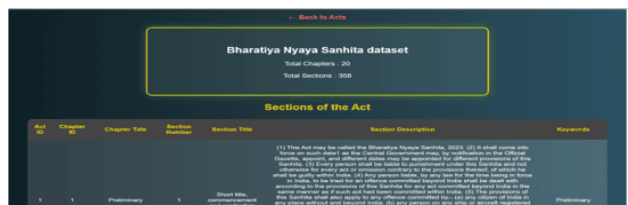


Figure 7: View Sections

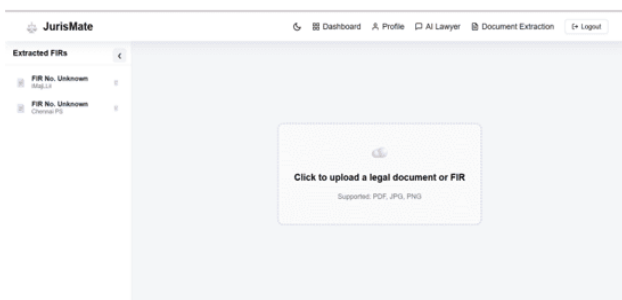


Figure 2: Upload FIR Document

**7. Recent Research and Advances**

Recent advancements in the fields of natural language processing (NLP), large-scale language modeling (LLM), and semantic information retrieval (SIR) have propelled forward the development of

AI-driven legal document analysis, creating a seamless connection between traditional, rule-based legal systems and innovative forms of decision-support technology that can analyze any kind of unstructured legal text, not just FIRs, but also other forms of case law.

### **A. Advances in Legal Text Understanding and Learning Paradigms**

Language models based on transformers have shown an excellent aptitude to understand legal language, contextual issues and vocabulary. Their usage of zero-shot and few-shot learning techniques has gained considerable interest due to limited amounts of labeled training data for legal documents. As a result, these models can apply to the classification of documents across various domains.

### **B. Semantic Retrieval and Vector-Based Search Innovations**

With the advent of semantic embedding (or context) based retrieval of case, statute and legal precedent through the use of vector based databases, has helped to eliminate keyword-based searches. Semantic Embedding models allow for a more efficient similarity search of large UK legal datasets; this advances the search process for retrieving relevant and accurate legal information.

### **C. System Architectures and AI Integration Frameworks**

Legal AI systems recently increasingly adopted microservices-based architectures to allow for flexible composing of scalable and modular applications. With RESTful APIs being used for AI services and cloud-ready designs being built from the ground up, these systems can easily connect to traditional legal technology platforms while delivering stable and high-performance services.

### **D. Towards Practical and Scalable Legal AI Systems**

The combination of these achievements has moved legal AI from research to the field through practical use. Current Research & Development trends in the legal AI space include scalable FIR analysis, multilingual legal processing, and hybrid systems that incorporate both symbolic legal rules and data-based learning models. Future R&D directions include predictive legal analytics,

cross-jurisdictional knowledge integration, and human-in-the-loop validation systems to provide legal accountability.

The advancement of these technologies solidifies the importance of the Development of an artificial intelligence driven framework like JurisMate (a System for Modernizing FIR Processing and Legal Decision Support) supports the current research trends in the area of AI through the adoption of transformer-based learning, Zero-Shot Classification and Semantic Retrieval for the purposes of modernizing FIR Processing and Legal Decision Support.

## **8. Limitations**

While the JurisMate system has many benefits, there are also some limitations present in the system. For starters, the JurisMate system is limited to FIRs that use digital document formats. The JurisMate system does not support FIRs that are scanned or handwritten as the technology relies on digital document formats and does not have Optical Character Recognition (OCR) capability, therefore it will not be viable in those countries or areas where FIRs are not yet adapted to fully utilize digital documentation.

Second, limitations in language and interpretation are evident in how the system has been developed:

- The system is designed primarily to accommodate FIRs created in specific, narrow languages.
- Regional languages or FIRs that have been written using a mixture of languages or use unusual developments of law will likely impact the system's ability to correctly extract texts and classify them.
- The continuing improvement of language or the capability of processing multiple languages through multilingual processing techniques will also be required.

Also, though the AI models offer consistent and contextually aware legal taxonomies, the system cannot supplant the legal judgement of attorneys. Attorneys must verify that final interpretations and decisions accord with the laws and procedures of a given jurisdiction.

Finally, the input FIR will significantly impact the performance of the system as well. If the input FIR are incomplete, vague or unclear, then classification accuracy and the semantic retrieval of results will be affected. Therefore, in order to deploy the system effectively in the real-world, limitations that arise from the quality and completeness of the input FIR must be resolved.

## 9. Conclusion and Future Work

This paper has introduced JurisMate which is an artificial intelligence-based system that uses advanced natural language processing and semantic retriever systems to automate the analysis of FIRs (First Information Reports). The JurisMate System combines transformer-based models for text extraction and zero-shot classification with vector-based similarity searching in order to overcome the major limitations of relying on manual processes for analysing FIR Data: inefficiency, inconsistency and limited Capacity for Scale. The Microservices Design Approach and Modular System Architecture that support JurisMate are suitable for developing reliable, consistent and Efficient Legal Document Analyses for real-world applications in both the Legal and Law Enforcement Environment.

The experimental evidence indicates that JurisMate has decreased FIR processing time, increased consistency of classification, and improved the ability to retrieve related information using context. By automating repetitive analytical processes, JurisMate facilitates legal business processes, while still allowing for interpretability and customisation of the process.

The next step for this system is to improve it with the ability to recognize text from scanned documents, allow for translations between many languages, and provide more sophisticated analytics for case outcomes. All of these enhancements will include fine-tuning based on specific legal topics, expanding the current knowledge base for each jurisdiction, and providing more access to and efficiency in court systems across North America and Europe.

## References

[1] K. D. Ashley. (1991). Reasoning with cases and hypotheticals in HYPO. *International Journal of Man-Machine Studies*, 34(6), 753–796.

[2] K. Lipianina-Honcharenko, O. Honcharenko, & Y. Savytskyi. (2024). A cyclical approach to legal document analysis: Leveraging AI for strategic policy evaluation. In: *CEUR Workshop Proceedings*, 3612, pp. 1–12.

[3] M. S. Kabir, & M. N. Alam. (2023). The role of artificial intelligence technology for legal research and decision making. *International Research Journal of Engineering and Technology (IRJET)*, 10(5), 1450–1456.

[4] I. Chalkidis, I. Androutopoulos, & N. Aletras. (2019). Neural legal judgment prediction in English. In: *Proc. 57th Annual Meeting of the Association for Computational Linguistics*, pp. 4317–4323.

[5] H. Zhong, C. Xiao, C. Tu, T. Zhang, Z. Liu, & M. Sun. (2018). Legal judgment prediction via topological learning. In: *Proc. 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 3540–3549.

[6] J. Devlin, M.-W. Chang, K. Lee, & K. Toutanova. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proc. NAACL-HLT*, Minneapolis, pp. 4171–4186.

[7] M. Lewis, Y. Liu, & N. Goyal, et al. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In: *Proc. ACL*, pp. 7871–7880.

[8] N. Reimers, & I. Gurevych. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In: *Proc. EMNLP-IJCNLP*, pp. 3982–3992.

[9] J. Johnson, M. Douze, & H. Jégou. (2021). Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3), pp. 535–547.

[10] K. D. Ashley, & S. Brüninghaus. (2009). Automatically classifying case texts and predicting outcomes. *Artificial Intelligence and Law*, 17(2), 125–165.

Disclaimer / Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Journals and/or the editor(s). Journals and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.