Project Synopsis

on

Assistance with Legal Documents using AI (LegalDocAI)

Submitted as a part of course curriculum for

Bachelor of Technology

in

Computer Science

Submitted by

Ananya Kushal (2200290120030)

Ankit Singh (2200290120033)

Kumari Chetana (2200290120093)

Aditya Madwal (2200290120011)

Under the Supervision of

Dr. Rishabh Jain

Associate Professor

KIET Group of Institutions, Ghaziabad

Department of Computer Science

Dr. A.P.J. Abdul Kalam Technical University

2024-2025

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Dr. Rishabh Jain (Associate Professor), Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his constant support and guidance throughout the course of our work. His sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his/her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Ajay Kumar Shrivastava, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature:
Student's Name:
Ananya Kushal (2200290120030)
Ankit Singh (2200290120033)
Kumari Chetana (2200290120093)
Aditya Madwal (2200290120011)

Guide Name & Signature
Dr.Rishabh Jain

Signature

Project coordinator

ABSTRACT

Our project **LegalDocAI** addresses a significant challenge in the legal industry: Managing and analyzing large volumes of legal documents while ensuring their security. Legal professionals often face time-consuming, error-prone processes when drafting, reviewing, and storing documents. Traditional methods of document storage are vulnerable to tampering, data breaches, and unauthorized access, making it difficult to ensure confidentiality and integrity. Moreover, analyzing complex legal documents manually can lead to oversights, especially when dealing with large contracts or agreements.

LegalDocAI is not only a solution for large legal firms but also provides significant benefits to small-scale law practices and middle-class individuals. Legal services are often expensive, particularly for document drafting, review, and secure storage, which can be burdensome for smaller firms and clients with limited resources. LegalDocAI democratizes access to advanced tools by offering an affordable, streamlined platform that reduces the need for manual legal work and lowers operational costs.

The solution offered by LegalDocAI automates the creation, analysis, and secure storage of legal documents, significantly reducing time and effort. By integrating AI, the platform can extract key clauses, generate summaries, and highlight important insights, assisting legal professionals in making quicker and more accurate decisions. Furthermore, LegalDocAI uses decentralized technology to store documents securely, ensuring that they are tamper-proof and accessible only to authorized users.

Use cases for LegalDocAI include contract management, where AI-driven analysis helps in identifying critical clauses; document verification, ensuring the integrity of sensitive legal files; and compliance auditing, where stored documents can be reviewed for regulatory compliance. This solution is ideal for law firms, corporate legal departments, and individual practitioners seeking to enhance document efficiency and security.

TABLE OF CONTENTS

CHAPTER 01 : INTRODUCTION

1.1 Introduction

The legal industry is heavily reliant on documents, from contracts to legal briefs. Manually drafting, analysing, and securely storing these documents is time-consuming and prone to human error. LegalDocAI addresses this by leveraging artificial intelligence for document creation and analysis while utilizing IPFS (Inter Planetary File System) for tamper-proof storage.

1.2 Problem Statement

Legal documentation can be a complicated and time-consuming process, especially for individuals and small businesses who may not have access to legal resources. In addition, the language and jargon used in legal documents can be difficult for non-lawyers to understand, which can lead to errors and misunderstandings.

Objective:

The primary objectives of LegalDocAI are:

1. Automate the process of legal document creation and analysis.
2. Ensure secure and decentralized storage of documents using IPFS.
3. Provide a user-friendly interface for legal professionals to manage and analyze documents effectively.

Potential Features:

1. User-friendly interface for inputting relevant information such as parties involved, terms of the agreement, and other necessary details.
2. AI-powered document generation that automatically drafts legal documents in plain language and using easy-to-understand terms.
3. Ability to customize legal documents based on the specific needs of the user.
4. Integration with existing legal resources and databases to ensure accuracy and completeness of the legal documents.
5. Option for users to seek legal advice from an expert in case of complex legal issues. Impact: The proposed solution can greatly benefit individuals and

small businesses in India, who often face challenges with legal documentation due to limited access to legal resources. By simplifying legal documentation, this solution can potentially save time, reduce errors, and increase access to justice. Data: Participants can use publicly available legal databases and resources to train the AI model for document generation.

1.3Scope

LegalDocAI is designed to be scalable, catering to law firms, legal departments, and individual practitioners. It will support features such as real-time document creation, AI-powered analysis, and decentralized file storage. The platform will be able to integrate with various blockchain networks and ensure compliance with data privacy regulations.

Chapter 2: Literature Review

- **AI-Powered Legal Document Management:**
  Authors - K.Meena, G.Renuka, Kooram Saraswati
  AI technology is transforming legal processes, improving efficiency in managing legal documents (Meena et al., 2024). By integrating machine learning models like neural networks and decision trees, AI assists legal professionals by automating tasks such as document analysis and prediction (Ashley, 2020). This automation reduces the time required for case preparation and enhances the quality of legal decisions (Noguti et al., 2020).

- **Indian Legal Text Summarization Using AI:**
  Authors - Satyajit Ghosh, Mousumi Dutta
  The growing backlog of over 4 crore pending cases in India has highlighted the need for AI-powered legal text summarization (Ghosh et al., 2022). Legal documents are complex, often lengthy, and contain domain-specific abbreviations, making manual summarization time-consuming. The authors experimented with domain-independent models like BART and PEGASUS for legal text summarization, utilizing a novel text normalization technique to make legal texts more suitable for machine learning models (Farzindar & Lapalme, 2004).

- **The Role of AI Technology in Legal Research and Decision Making:**
  Authors - Md. Shahin Kabir, Md. Nazmul Alam
  The increasing use of Artificial Intelligence (AI) in the legal profession has highlighted its potential to transform the way lawyers conduct research and make decisions (Kabir & Alam, 2023). The authors of this paper examine the evolving role of AI technology in the field of law, specifically focusing on legal research and decision making.

- **AI-Based Legal Document Generation**
  Authors: R. McNamara, A. Smith, D. B. Sanderson
  This research focuses on the development of an AI-based legal document generation tool that uses deep learning techniques to automate the drafting of standard legal documents, such as contracts and agreements. The system adapts to specific user inputs and context,

ensuring that generated documents meet the required legal standards and client needs.

- **Document Management System Generative AI and AWS Comprehend**
  Authors : Niharika Tiwari, Prashansa Srivastava, Dr. Sadhana Rana
  This paper discusses the potential of document management systems that integrate generative AI and cloud features to enhance user productivity and efficiency. The study highlights the benefits of generative AI in boosting creativity and quality, while cloud resources provide scalability and mobility. A cloud-based document management system offers users the freedom to access and analyze documents from anywhere, deriving important insights through built-in analysis features.

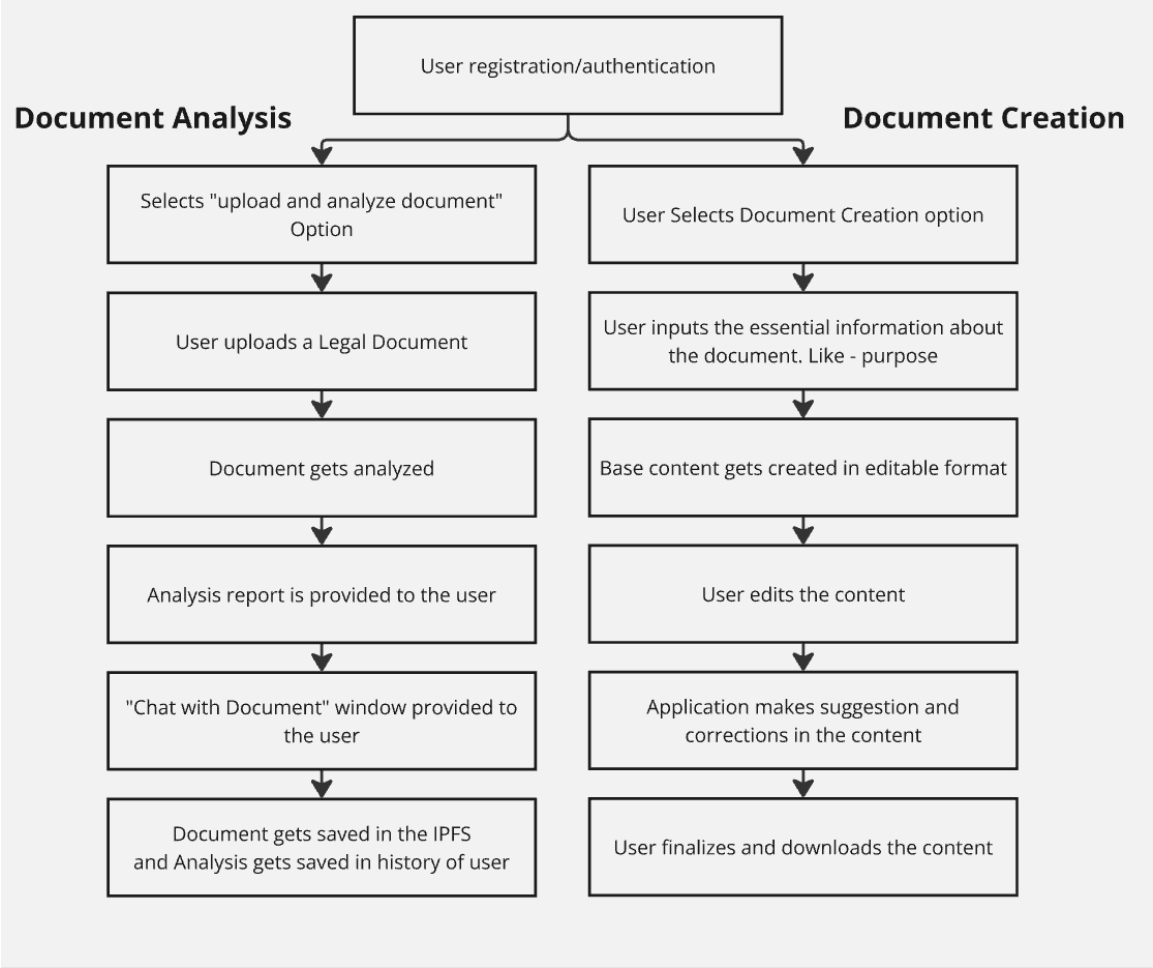
- **Document AI: A Document Content Explorer using Natural Language Processing**
  Authors : John Chris Baguhin, Edison Amatorio, Vanesa Bea Salvaleon, Gina Rose Quion, Darrel Cardana, Max Angelo Perin
  This paper presents a novel approach to addressing information overload in academic research through the development of Document AI, a document content explorer that leverages natural language processing (NLP), semantic indexing, and AI-driven assistance. The study highlights the challenges faced by researchers in navigating the vast amount of academic literature and proposes Document AI as a solution to enhance research efficiency. The paper discusses the design and methodology of Document AI, its usability, and its potential impact on the future of research exploration.

- **LEGALSEVA - AI-POWERED LEGAL DOCUMENTATION ASSISTANT**
  Authors : Rithik Raj Pandey, Sarthak Khandelwal, Satyam Srivastava, Yash Triyar, Mrs. Muquitha Almas
  This paper presents a novel AI-powered legal documentation assistant, LEGALSEVA, designed to simplify legal documentation for small businesses and individuals in India. The system utilizes Optical Character Recognition (OCR) technology and Custom Trained GPT to generate simplified versions of complex legal documents. The study demonstrates the effectiveness of LEGALSEVA in reducing errors, saving time, and promoting access to justice. The paper also discusses

the importance of ethical considerations, such as data privacy and AI training bias, in the development of LEGALSEVA. The results show that LEGALSEVA has the potential to revolutionize the legal documentation process, making it more accessible and user-friendly for non-lawyers.

Chapter 3: Proposed Methodology

3.1 Flowchart



3.2 Algorithm Proposed

"AI-Powered Legal Document Analysis and Secure Storage for LegalDocAI."

The LegalDocAI platform offers an end-to-end solution for legal document management, enabling users to create, analyze, and securely store legal documents with the help of AI and IPFS storage technology. The platform allows legal professionals to input documents, which are then processed by an AI model to extract key clauses, generate summaries, highlight important legal insights and make suggestions. Following the analysis, the documents are stored securely, ensuring that they remain tamper-proof and accessible only to authorized users. This process streamlines document handling, improves accuracy, and reduces the overall time spent on legal reviews.

Inputs (User Provides):

- **User Information**: Credentials for authentication.
- **Legal Document**: Uploaded document in PDF format.
- **Document Specifications**: Type of document (contract, agreement, etc.).

Outputs (User Receives):

- **AI-Generated Summary**: Concise overview of the document.
- **Key Clauses and Insights**: Highlighted important clauses, risks, or legal terms.
- **Secure Access Link**: Document link for future retrieval and verification.

Chapter 4: Technology Used

4.1 Frontend

- **React.js** : A powerful JavaScript library used to build a fast, dynamic, and responsive user interface for the LegalDocAI platform. It enables the creation of a seamless user experience where users can upload, analyze, and manage legal documents easily.

- **Tailwind CSS** : A utility-first CSS framework that enables the rapid development of a modern, responsive UI with minimal effort. Tailwind provides predefined classes that streamline the styling process, ensuring that the design is clean and mobile-friendly.

4.2 Backend

- **Django**: A high-level Python web framework that handles the backend logic, including managing the APIs, authentication, and database interactions. It ensures scalability, security, and efficient data handling.

- **Django REST Framework (DRF):** Extends Django by providing tools to create robust, flexible REST APIs, allowing the frontend (React.js) to communicate seamlessly with the backend. DRF is essential for handling API requests for document uploads, AI-powered analysis, and decentralized storage integration.

4.3 Database

- **PostgreSQL:** A powerful, open-source relational database system that handles the storage of user data, document metadata, and analytics results. PostgreSQL is known for its reliability and support for complex queries, making it ideal for managing large volumes of legal data.

4.4 AI Integration

- **OpenAI API:** This API powers the document analysis, generating summaries, extracting key clauses, and providing legal insights. By using natural language processing (NLP), it enhances the efficiency of

legal professionals by automating the review and understanding of complex documents.

- **NLTK (Natural Language Toolkit)**: A leading library for building Python programs to work with human language data.

- **Hugging Face Transformers**

- **Pretrained LLMs** (GPT-3, BERT)

- **spaCy**: An open-source NLP library in Python, offering features like tokenization, part-of-speech tagging, and named entity recognition.

- **TensorFlow**: An open-source library for machine learning and deep learning models, widely used in AI-powered applications.

4.5 Cloud Platform:

- **AWS (Amazon Web Services)**: A cloud computing service used for hosting, storage, and scalable backend services.

4.5 Decentralized Storage

- **IPFS (Inter Planetary File System):** A peer-to-peer decentralized storage network used for securely storing documents off-chain. IPFS ensures that files are stored across a distributed network, making them tamper-proof and immutable.

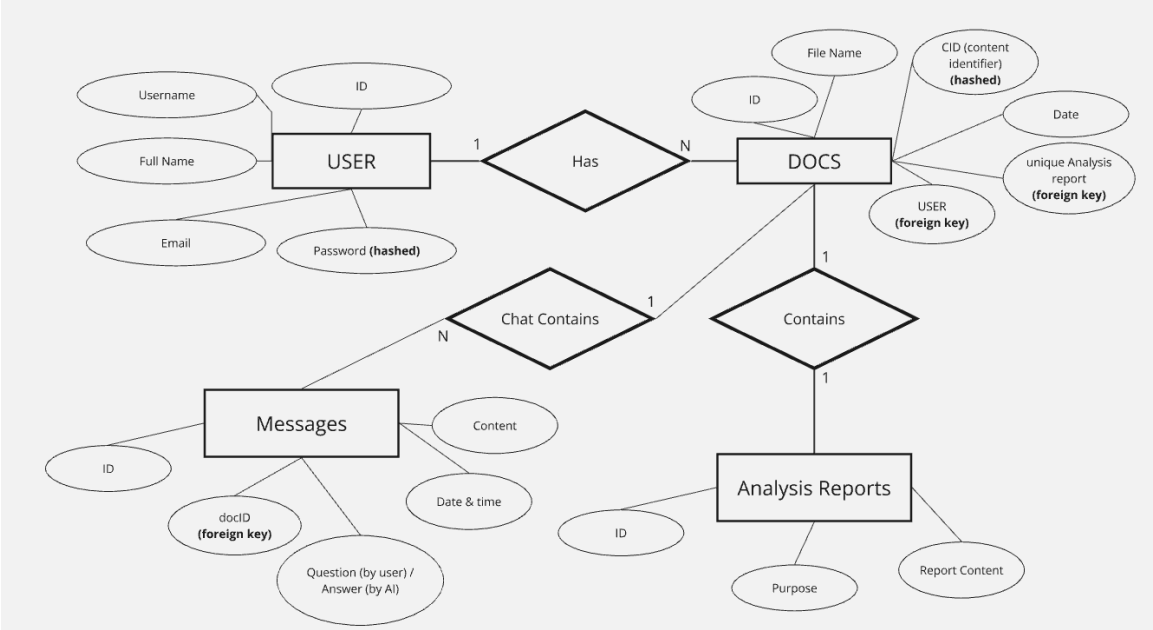
- **Pinata API:** This API is used to interact with IPFS, making it easier to pin and manage files on the IPFS network. Pinata ensures that the documents uploaded are reliably accessible, providing a secure link to each stored file.

4.6 Security

- **JWT (JSON Web Tokens)**: Used for secure user authentication and authorization. JWT ensures that only authenticated users can access

protected resources like document uploads, analysis, and retrieval. It generates signed tokens that allow secure communication between the client and server without exposing sensitive information.

Chapter 5: ER Diagram



Chapter 6: Conclusion

LegalDocAI offers an innovative solution to legal document management by combining **AI-driven analysis** with **blockchain-based security**. It enhances productivity, reduces human errors, and ensures the security and integrity of legal documents. The platform is scalable and can cater to legal professionals across different fields, improving document management workflows while ensuring data privacy and compliance.

References

- Kabir, M. S., & Alam, M. N. (2023). The Role of AI Technology for Legal Research and Decision Making. International Research Journal of Engineering and Technology, 10(7), 1088-1092.
- Ghosh, S., Dutta, M., & Das, T. (2022). Indian Legal Text Summarization: A Text Normalisation-based Approach. 2022 IEEE 19th India Council International Conference (INDICON). Adamas University, Kolkata, India.
- Meena, K., Tanusha, G., Renuka, G., Saraswathi, K., & Anuradha, S. (2024). Enhancing Legal Document Management Efficiency: An AI-Powered Solution Addressing Interpretation Challenges. Journal of Engineering and Technology Management, 72, 650-662.
- M. Steichen, B. Fiz, R. Norvill, W. Shbair and R. State, "Blockchain-Based, Decentralized Access Control for IPFS," 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canada, 2018, pp. 1499-1506, doi: 10.1109/Cybermatics_2018.2018.00253. keywords: {Blockchain;Access control;Smart contracts;Cryptography;Registers;Internet of Things}
- Imogen, Vimala & Sreenidhi, J. & Nivedha, V.. (2024). AI-Powered Legal Documentation Assistant. Journal of Artificial Intelligence and Capsule Networks. 6. 210-226. 10.36548/jaicn.2024.2.007.
- Arora, Jhanvi, Tanay Patankar, Alay Shah, and Shubham Joshi. "Artificial Intelligence as Legal Research Assistant." In FIRE (Working Notes), pp. 60-65. 2020.