

ClauseEase

AI based Contract Language Simplifier

ClauseEase

Transforming Legal Documents through Intelligent Automation

- ✓ AI-Driven Contract Understanding
- ✓ Multi-Format Uploads (PDF, DOCX, TXT)
- ✓ Simplify Complex Legal Language
- ✓ Clear Contracts. Confident Decisions

What is ClauseEase?

Project Statement:

Legal contracts are often written in dense, complex language that is difficult for non-lawyers to understand.

This project, **ClauseEase**, aims to build an **AI-powered system** that automatically extracts, analyses, and simplifies legal contract clauses without losing legal meaning or intent. The goal is to make contracts more **readable, accessible, and understandable** for all users.

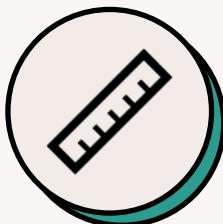
Objectives



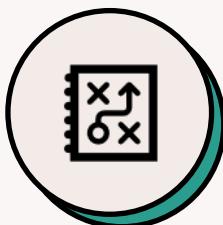
Automate contract text extraction and preprocessing.



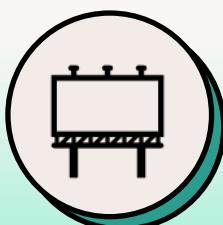
Identify and classify legal clauses using NLP and AI models.



Recognize complex legal terms and generate plain-English explanations.

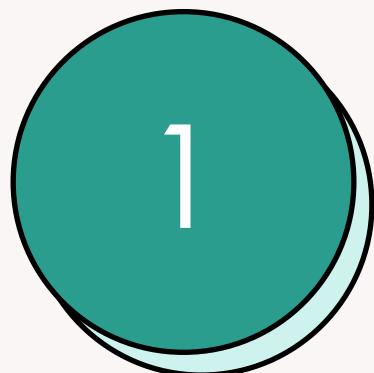


Simplify contract language while preserving legal accuracy.



Provide users with an intuitive interface for contract analysis.

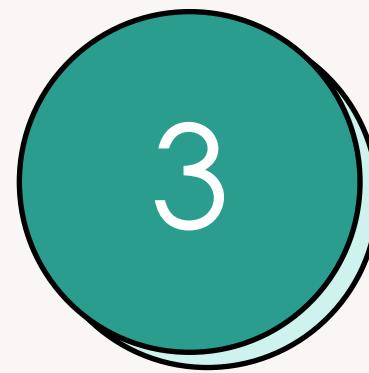
ClauseEase - Modules



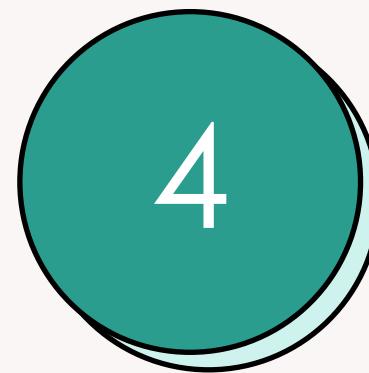
Document Ingestion
& Text Preprocessing



Legal Clause
Detection



Legal Term
Recognition



Language
Simplification



Context Linking &
Feedback

Requirement Analysis

Hardware:

- **OS:** Compatible with Windows, macOS, or Linux.
- **RAM:** Minimum **4GB** RAM required.
- **CPU:** Processor speed of Core i5 or equivalent/higher recommended (GPU recommended for higher evaluation models)

Basic Requirements:

- **Python 3.10** or higher
- **Web Browser:** Google Chrome / Edge / Firefox (latest version)
- **IDE / Code Editor:** VS Code / PyCharm / IntelliJ

Software Requirement Analysis

- **Flask**
- HTML5, CSS3, JavaScript
- **Jinja2** templating (Flask integrated)
- **Docker** Desktop (for deployment)

Integrations:

Model Name	Purpose
<code>prajjwal1/bert-mini</code>	Legal clause classification
<code>google/flan-t5-small</code>	Legal text simplification / paraphrasing
<code>en_core_web_sm (spaCy)</code>	Entity extraction & linguistic analysis

- **Google Firebase (Cloud Firestore)** → NoSQL database

Py Libraries/ Dependencies	Purpose / Usage
<code>fitz (PyMuPDF)</code>	PDF text extraction
<code>python-docx</code>	DOCX text extraction
<code>re</code>	Regular expressions for cleaning text
<code>nltk</code>	Sentence tokenization (<code>sent_tokenize</code>)
<code>spacy</code>	Named Entity Recognition (NER)
<code>torch (PyTorch)</code>	Running Legal-BERT / Google t5 model
<code>transformers</code>	Pretrained AI models (Legal-BERT, t5-small)
<code>scikit-learn</code>	Basic ML utilities (label encoding, evaluation)
<code>numpy</code>	Numerical operations
<code>firebase-admin</code>	Firestore connection in Flask
<code>json / os</code>	File handling, data serialization

Data Flow Diagrams (DFDs)

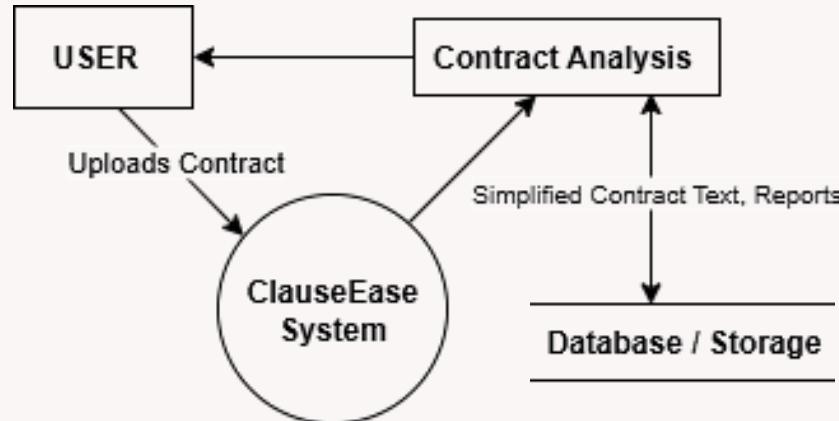


Fig.1: 0-level DFD

Note:

The output can be shown into **Basic**, **Intermediate** and **Advanced** paraphrases, according to user Query

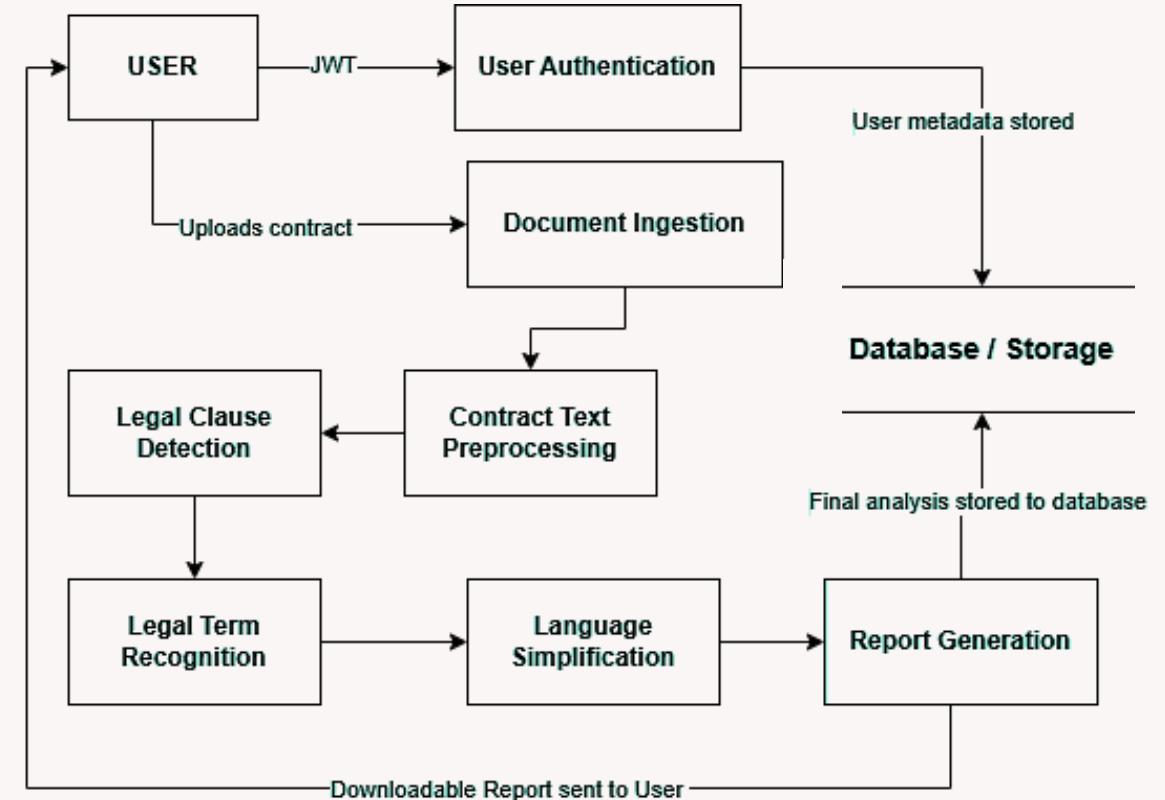
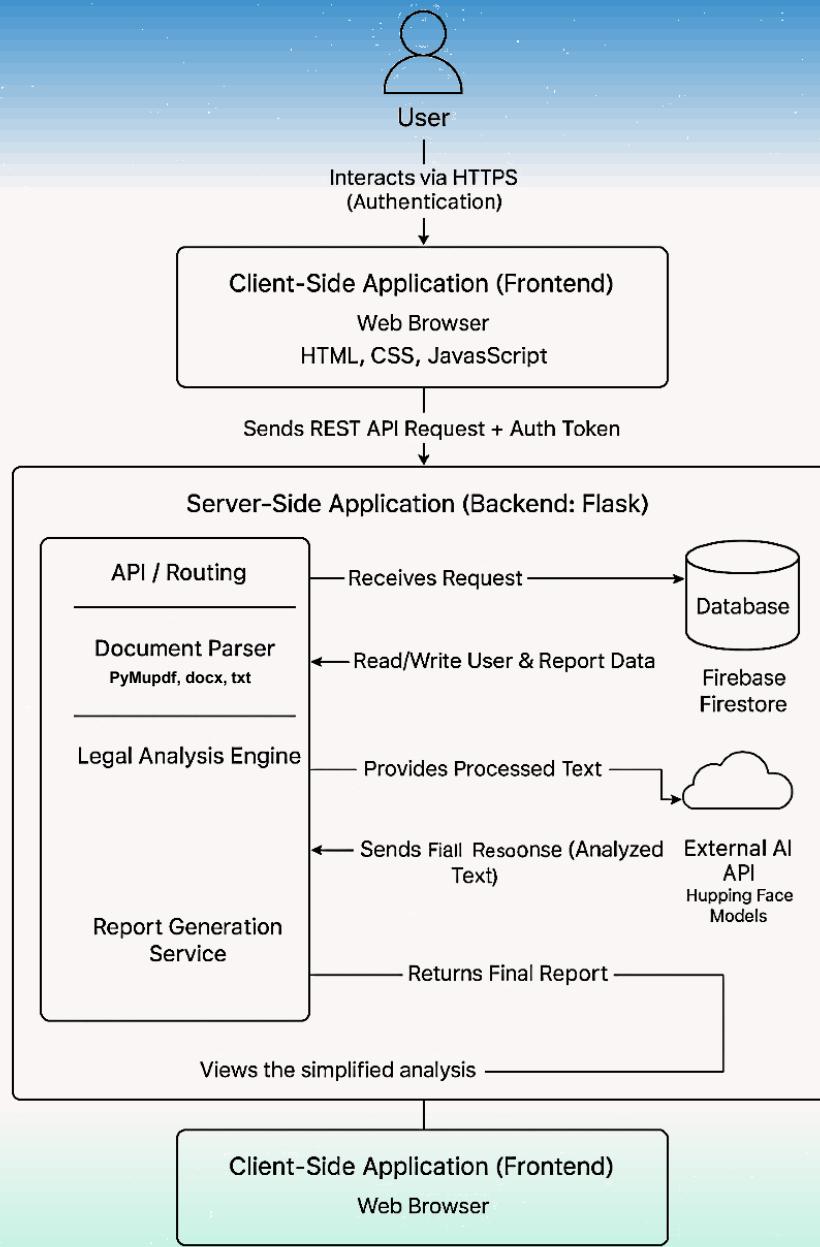


Fig.2: Level-1 DFD

System Architecture



- ❑ ClauseEase uses a **Flask backend** integrated with AI models for clause detection and simplification.
 - The **frontend** provides an interactive user interface, while **Firebase Firestore** securely stores user data and processed clauses.
 - Requests flow from the **Frontend → Backend → AI HuggingFace models → Results** returned and saved for real-time access.

Fig. 3: System Architecture Diagram

Database Table Description

users (collection)

```

  |- user_id_1 (document) (PK)
    |   metadata:
    |   |   name: "Arnab"
    |   |   email: "arnab@test.com" (UK)
    |   |   password_hash: "..."
    |   |   created_at: "2025-10-08"

    |- analyses (subcollection)
      |- analysis_id_1 (document) (PK)
        |   uploaded_at: "2025-10-08 18:20"
        |   results_path: "/uploads/contract1.results.json"
        |   user_email: "arnab@test.com" (FK)
        |   details:
        |   |   file_name: "contract1.pdf"
        |   |   clause_count: 12 (PK)
        |   |   simplified_text_count: 12
  
```

Collection / Subcollection	Description	Key Fields / Documents
users	Main collection storing all registered user accounts and their profile metadata.	Document ID: unique user identifier (UID) Fields: name, email, password_hash, created_at
users → {user_id} → analyses	Subcollection under each user storing all uploaded contract analyses. Each document corresponds to one analyzed contract.	Document ID: unique analysis ID Fields: results_path (path to JSON result), uploaded_at, user_email, details (summary info like file name, clause count, etc.)
analyses → file_details	(If nested further) Holds specific details about each uploaded contract file, extracted clauses, and results metadata.	Fields: file_name, file_size, file_type, results_path, simplified_text_count, clause_count
legal_terms	A global collection storing predefined legal terms and definitions for recognition and glossary support.	term, definition, category

Index:

PK: Primary Key

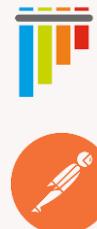
UK: Unique Key

FK: Foreign Key

Testing & Implementation

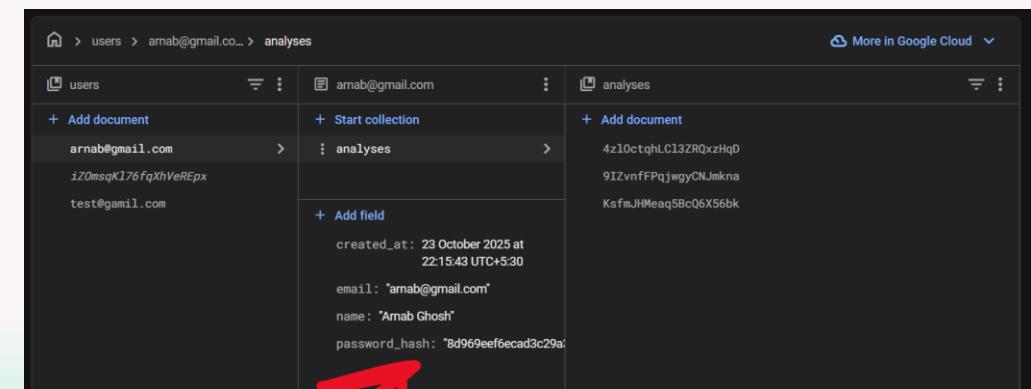
Testing Techniques

- ❑ **Unit Testing:** Tested each individual module:
 - *Text Extraction, Clause Detection, Simplification, and Glossary Mapping* modules.
- ❑ **Integration Testing:** Verified seamless communication between:
 - Flask backend, HuggingFace models, and Firestore database.
- ❑ **API Testing (Postman):**
 - All Flask routes (e.g. /api/analyze) tested using **Postman** for correct request/response handling. {json output}
- ❑ **User Interface Testing:**
 - Ensured responsive layout, smooth tab switching, and accurate result rendering in HTML frontend.
- ❑ **Performance Testing:**
 - Verified handling of large contracts (**PDF/DOCX**) and AI inference time optimization.
- ❑ **Security Testing:**
 - Password hashing and authentication verified for data safety.



Implementation

- ❑ **Frontend:** Built using HTML/CSS/JS with a clean UI following *layout.html* theme.
- ❑ **Backend:** Implemented using **Flask** for routing and model orchestration.
- ❑ **AI Processing:** BERT/NLP HF models handle clause detection and simplification logic.
- ❑ **Database:** User and analysis data securely stored in **Firestore (NoSQL)** collections.
- ❑ **Deployment:** Docker Desktop



Application Programming Interface (API) Testing

□ Testing Tools Used:

- ❖ **Postman:** API validation
 - ❖ **Pytest:** Backend module testing
 - ❖ **Firebase Console:** Firestore data integrity checks

HTTP <http://127.0.0.1:3000/api/analyze> Save

POST <http://127.0.0.1:3000/api/analyze> Send

Params Authorization Headers (9) **Body** Scripts Settings Cookies

none form-data x-www-form-urlencoded raw binary GraphQL

Key	Value	Description	...
<input checked="" type="checkbox"/> document	Agreement between owners and labour contractor for suppl... Upload		Bulk Edit
Key	Text	Description	

Body Cookies Headers (6) Test Results |

JSON Preview Visualize | 200 OK • 4 m 4.26 s • 49.42 KB •

```
8   "terms": {  
9     "arbitration": "A method of resolving disputes outside the courts.",  
10    "breach": "A violation of a law, duty, or other form of obligation.",  
11    "governing law": "The legal jurisdiction whose laws will apply to the agreement.",  
12    "jurisdiction": "The official power to make legal decisions and judgments.",  
13    "liability": "Legal responsibility for one's acts or omissions.",  
14    "termination": "The act of ending a contract or agreement."  
15  },  
16  "text": "*Disclaimer - This legal form and document is for reference only. Any document that you enter into, should be in consultation with an Advocate or a Solicitor.  
The Government will not be responsible for any claim arising out of the use of the following document. Agreement between owners and labour contractor for supply of  
labour This Agreement made this .... day of ..... at..... is entered into by and between....., a Company incorporated and existing in  
accordance with the Companies Act and having its Registered Office at ..... through Mr. ..... of the Company (hereinafter referred to as 'FIRST  
PARTY') which expression shall unless be repugnant to the context include its successors and assigns of ONE PART. AND ..... a Company incorporated and  
existing in accordance with the Companies Act and having its registered office at ..... through Mr. ..... of the Company (hereinafter  
referred to as 'SECOND PARTY') which expression shall unless be repugnant to the context include its successors and assigns of OTHER PART. WHEREAS .....  
At the very outset be it clearly understood and agreed that the following principles shall apply unless the context requires otherwise: (a) The singular includes  
the plural and conversely. (b) A gender includes all genders. (c) Where a word or phrase is given a particular meaning, other grammatical forms of that word or  
phrase have corresponding meanings. (d) If any day is appointed or specified by this agreement for the payment of any money on Sunday or other official bank  
holiday, the day so appointed or specified shall be deemed to be the next day which is not in turn a weekly off day or a day so appointed as a holiday. NOW THIS  
AGREEMENT WITNESSETH AS UNDER: 1. OFFER & ACCEPTANCE i. That Second Party hereby offers to provide work force for executing any Job assigned by First Party as per  
its requirements at its premises on contractual basis and First Party hereby accepts the same. 2. SCOPE OF SERVICES i. ..... 3. DEALING PERSONS i.  
..... 4. RIGHTS & OBLIGATIONS OF BOTH PARTIES i. That Second Party agrees to execute fulfil & discharge the work as may be assigned by First Party from time  
to time in the manner agreed to the entire satisfaction of First Party on all days as per the requirements of First Party during the period of this agreement. ii.  
That the estimated requirement of personnel as agreed may be increased or decreased depending upon the evaluation of the contractor required during course due to
```

Find and replace Console Runner Start Proxy Cookies Vault Trash

http://127.0.0.1:3000/api/analyze

POST http://127.0.0.1:3000/api/analyze

Save

Send

Params Authorization Headers (9) Body Scripts Settings Cookies

none form-data x-www-form-urlencoded raw binary GraphQL

Key	Value	Description	Bulk Edit
document	Agreement between owners and labour contractor for suppl...	File	Attachment
Key	Text	Description	

Body Cookies Headers (6) Test Results

(JSON Preview Visualize)

200 OK 4 m 4.26 s 49.42 KB

```
1 {
2   "filename": "Agreement_between_owners_and_labour_contractor_for_supply_of_labour.docx",
3   "message": "Analysis completed successfully",
4   "results": [
5     {
6       "index": 1,
7       "simplified": "Disclaimer - This legal form and document is for reference only. Tout document que vous entrez ou utilisez doit \u00eatre en consultation avec un d\u00e9fenseur ou un solliciteur. Des Government will not be responsible for any claim arising out of the use of the following document. ...., at ..... is entered into by and between ..... Company incorporated and existing in accordance with the Companies Act and having its registered office at ..... (hereinafter referred to as \"FIRST PARTY\") which expression shall unless be repugnant to the context include its successors and assigns of ONE PART. ...., a Company incorporated and existing in accordance with the Companies Act and having its registered office at ..... through Mr. .... of the Company (hereinafter referred to as \"SECOND PARTY\") which expression shall unless be repugnant to the context include its successors and assigns of OTHER PART. At the very outset be it clearly understood and agreed that the following principles shall apply unless the context requires otherwise: a) The singular includes the plural and conversely. b) Un sexe englobe tous les sexes. c) Lorsqu'un mot ou phrase a un certain meaning, d'autres grammatical forms of that word or phrase have corresponding meanings. d) Si un jour est appointed or specified by this agreement for the payment of any money on Sunday or other official bank holiday, the day so appointed or specified shall be considered to be the next day which is not in turn a weekly off day or a day so appointed as a holiday. NOW THIS AGREEMENT WITNESSETH AS UNDER: 1. i. That Second Party hereby offers to provide work force for executing any job assigned by First Party as per its requirements at its premises on contractual basis and First Party hereby accepts the same. 2. SCOPE OF SERVICES i. 3. i. ...., b) BOTH PARTIES I. Que Second Party agrees to execute fulfill & discharge all work as may be assigned by First Party from time to time and Second Party agrees to the entire contract offered by First Party on all the terms and conditions mentioned in this Agreement. II. Que l'estimation des besoins en personnel tel qu'il convient, peut faire augmente ou diminuer en fonction des exigences des services re\u00e7us, un domm\u00e9t 7 jours notice a Second Partie. Die Zweit Party muss die zus\u00e4tzliche Arbeitsk\u00e4rte f\u00fcr die erforderlichen l\u00f6dungen bereitstellen. Die Zahl der erh\u00f6rten Arbeitsk\u00e4rte ist 100. Second Party shall maintain proper register at the aforementioned premises of First Party as well as with them in
```

Close and replace People

Printer Print Current Document Copy Paste Search Find

Simplified text output

Legal Terms recognized from the document

Snapshots

Upload Document

Choose File No file chosen

Intermediate (Innauer)

Analysis Results

Uploaded: employment_agreement.pdf

Simplification level: Intermediate

Generated Visualizations

Intermediate (Innauer)

Pipeline Modules

- 1. Document Ingestion
- 2. Preprocessing
- 3. Clause Detection
- 4. Legal Term Recognition
- 5. Language Simplification

Clause 1 – ('Confidentiality', 0.4767265319824219)

EMPLOYMENT AGREEMENT This EMPLOYMENT AGREEMENT made in _____ on this _____. Between _____ (hereinafter referred to as "Employer"), having its registered office at _____ IN AND _____, S/O _____ (father), residing at _____ hereinafter referred to as the "Employee". (The Company and the Employee are hereinafter jointly referred to as the "Parties" and individually as a "Party"). NOW THIS AGREEMENT WITNESSETH AND IT IS HEREBY AGREED BY AND BETWEEN THE PARTIES HERETO AS UNDERSIGNATOR The Employee is hereby appointed and shall hold office as _____ of the Company on the terms and conditions contained hereunder. 2. PROBATIONARY PERIOD The probationary period will be _____, and the company will evaluate the new hire's qualifications, skills and "fit" within our Company. The given time frame does not constitute an obligation on the part of the Company to retain the employee until the end of the probationary period. This period involves special orientation activities plus closer and more frequent performance evaluations.

Img1: Analysis Report

ClauseEase

Contract Language Simplifier

Welcome to ClauseEase

ClauseEase is a lightweight, intelligent web application that ingests contracts and other legal documents, automatically detects clauses, recognizes key legal terms, and provides plain-English simplifications.

Key Features

- Document ingestion and preprocessing for accurate parsing
- Automated clause detection and classification
- Legal term recognition with concise explanations
- Language simplification for non-legal readers

Img2: Home page

ClauseEase

Contract Language Simplifier

Transforming Legal Documents through Intelligent Automation

About ClauseEase

ClauseEase is designed to simplify the complex world of legal documents. By blending intelligent rule-based parsing with machine learning insights, it helps users rapidly identify critical clauses, legal terms, and obligations – translating legal jargon into clear, plain English.

The goal of ClauseEase is to empower lawyers, legal researchers, startups, and business professionals to review contracts efficiently and make informed decisions without getting lost in dense legal text.

How It Works:

- Upload your document – supported formats include PDF, DOCX, and plain text.
- Automated text extraction – the system preprocesses the file and segments it into clauses.
- Intelligent analysis – each clause is categorized and scanned for legal terminology.
- Readable output – you'll see the original clause, identified legal terms, and a simplified explanation side by side.

ClauseEase runs on a **FaaS-based backend** with modular design principles – making it easy to extend, integrate, and enhance. The current demo showcases how legal tech automation can make contracts more transparent and user-friendly.

Img4: About page

Contract Language Simplifier

Admin Panel Analyses

List of all uploaded analyses across users.

User	Filename	Uploaded At	Simplification	Actions
arnab2@gmail.com	employment_agreement.pdf	2025-11-04 17:35:10.56607...	Advanced	Download PDF Download JSON
arnab2@gmail.com	employment_agreement.pdf	2025-11-04 17:12:47.926503...	Advanced	Download PDF Download JSON
arnab2@gmail.com	employment_agreement.pdf	2025-11-04 13:46:43.58413...	Intermediate	Download PDF Download JSON
arnab5@test.com	Deed_of_absolute_sale.pdf	2025-11-05 11:55:33.071456...	Advanced	Download PDF Download JSON

Img6: Admin panel

Logged out successfully.

ClauseEase

Transforming Legal Documents through Intelligent Automation

AI-Driven Contract Understanding
Multi-format Uploads (PDF, DOCX, TXT)
Simplify Complex Legal Language
Clear Contracts. Confident Decisions

Login Register

Email: your@email.com

Password:

Sign In

Don't have an account? [Register](#)

Img3: Login/ Signup Page

ClauseEase

Contract Language Simplifier

Transforming Legal Documents through Intelligent Automation

Your Analysis History

File Name	Download PDF	Download JSON
CBSE-Class-3-Mathematics----4-Digit-Numbers.pdf	Download PDF	Download JSON
ClauseEase_documentation.pdf	Download PDF	Download JSON
employment_agreement.pdf	Download PDF	Download JSON

Img5: See history page

Conclusion

Summary

ClauseEase successfully simplifies complex legal contracts using AI and NLP. It reduces the effort required to interpret legal documents and improves accessibility for users without legal expertise.

Future Enhancements

- Integration of multilingual support for global use.
- Real-time clause comparison across multiple contracts.
- Improved summarization using advanced LLMs.
- Legal chatbot assistant for query-based explanations.
- Integration with document management systems for enterprise use.

References

- ❑ Hugging Face Transformers Library.
- ❑ NLTK & SpaCy documentation.
- ❑ Legal Text Simplification Research Papers.
- ❑ Flask & Bootstrap Official Documentation.
- ❑ Infosys Springboard Learning Resources.

Thank you

Project: ClauseEase | Contract Language Simplifier
Infosys SpringBoard Virtual Internship 6.0

ClauseEase

Transforming Legal Documents through Intelligent Automation

- ✓ AI-Driven Contract Understanding
- ✓ Multi-Format Uploads (PDF, DOCX, TXT)
- ✓ Simplify Complex Legal Language
- ✓ Clear Contracts. Confident Decisions