Smart Policy Document Explainer & Generator

Dhruv Thejas KJ
Department of Computer Science and Engineering, Manipal Institute of Technology Bengaluru,
Manipal Academy of Higher Education, Manipal, India
dhruvthejas.kj@gmail.com

Abstract— **Insurance policy documents are often filled with dense legal and technical language, making them difficult for everyday users to understand. To make these documents more accessible, we developed the Smart Policy Document Explainer & Generator—a lightweight, user-friendly tool that uses advanced natural language processing to break down complex policy text. The system features two main modes: "Simplify," which rephrases insurance clauses into clear, conversational language, and "Draft," which organizes raw policy data into readable, structured formats. Powered by the Mistral-7B-Instruct model through the ctransformers library, the tool ensures efficient and accurate interpretation of legal content. Built with Streamlit, the interface allows users to easily upload or paste documents and receive instant, simplified output. This tool bridges the gap between legal complexity and user comprehension, with practical applications in compliance, human resources, insurance advisory, and broader document automation tasks.**

Keywords— Insurance Policies, Text Simplification, Natural Language Processing, Mistral-7B, Streamlit, Legal Document Understanding, Text Generation, Transformer Models, Document Summarization, User Accessibility

I. INTRODUCTION

Insurance policy documents are essential legal tools that outline the responsibilities, rights, and protections of both insurers and policyholders. However, despite their importance, these documents are notoriously difficult to understand. They're often written in dense, wordy language filled with legal jargon, outdated phrases, and technical terms. Long, complicated sentences and ambiguous wording can make it hard even for well-educated readers to grasp their full meaning—let alone someone without a legal background. This becomes particularly problematic in high-pressure situations like filing a claim, understanding coverage, or answering compliance-related questions, where clarity is crucial.

The consequences of this complexity are serious. Misunderstanding a policy's terms can lead to delays, mistakes, disputes, or even financial losses. In many industries—from healthcare to tech—employees are expected to understand both internal and external insurance documents. But the mental effort required to interpret these texts often goes beyond what their job roles demand. While legal experts might manage just fine, HR teams, technical staff, and everyday policyholders are often left confused and unsupported.[4]

This challenge has only grown in the digital age. As more services move online and documentation becomes increasingly digital, the volume of policy and compliance files has surged. Yet even with all this digital progress, the way these documents are written and presented hasn't changed much. Most document tools today can extract or convert text, but they don't do much to help people actually understand what the text means. There's a major gap between document automation and human comprehension.

To help bridge that gap, we've built the Smart Policy Document Explainer & Generator—an AI-powered tool designed to make insurance documents easier to understand and work with. It offers two key features: Simplification, which turns complicated policy language into plain, conversational English with helpful highlights and examples, and Draft Generation, which organizes messy, unstructured input into a clean, readable document ready for review or use in compliance workflows.

The engine behind this system is Mistral-7B-Instruct, a streamlined but powerful large language model tuned for following instructions effectively. We use it through the ctransformers library in a GGUF quantized format, which makes it efficient enough to run locally or on lightweight servers—no massive cloud infrastructure needed. On the front end, the tool is built using Streamlit, giving users a clean, interactive experience where they can paste text or upload .pdf and .txt files for instant analysis.

What makes this tool stand out from typical document processing systems is its focus on user experience, clarity, and domain relevance. It doesn't just extract or summarize—it helps users actually understand the content. The output maintains the core meaning of the original policy but presents it in a way that's far more digestible and user-friendly. And for cases where the AI model might not be available, the system includes a fallback mechanism using rule-based processing, ensuring it remains functional even offline or in low-resource settings.[1][2]

Ultimately, our goal is to reduce the knowledge barrier that prevents people from confidently interpreting complex documents. Whether it's an employee reviewing a company policy, a legal team analyzing contracts, or a consumer trying to figure out what their insurance covers, this tool aims to empower them with clear, actionable understanding.

The rest of this paper is structured as follows: Section II covers related work, including prior research in legal document simplification and NLP-based accessibility tools.Section III details our methodology, including the model architecture, prompt strategy, and fallback design. Section IV presents our results, showcasing example outputs and feedback from user testing. We conclude with Section V, where we highlight our contributions and outline future improvements such as OCR integration, speech interfaces, and multilingual capabilities.

II. Literature Review

The development of intelligent systems for legal and insurance document understanding is a growing area of research within the field of Natural Language Processing (NLP). Recent advances in large language models (LLMs), particularly those based on transformer architectures, have enabled more sophisticated applications in domains previously considered too complex due to legal precision and contextual ambiguity. This section provides an overview of the foundational work and recent progress across four relevant areas: legal simplification, insurance policy understanding, model architectures, and accessibility evaluation metrics.

A. Legal Text Simplification

Legal text simplification is an especially challenging task due to the need for preserving the original meaning while reducing linguistic complexity. Unlike general text simplification, which often prioritizes fluency and grammar correction, legal simplification must balance clarity with legal fidelity. In this context Kumar et al. [1] used T5 and BART models to generate events from legal documents, achieving strong semantic understanding with a BERT Similarity score of 0.85. Their approach showcases the potential of transformer models in automating legal analysis.Similarly Kim et al. [2] explored transformer-based approaches for legal information retrieval and entailment, demonstrating their effectiveness in understanding and extracting relevant legal information. Their study highlights the potential of these models in improving legal document analysis and reasoning. Their results highlighted that pretrained transformer models could outperform rule-based methods and achieve human-acceptable outputs on benchmark legal datasets.

These studies emphasize the importance of domain-specific fine-tuning and prompt design. They further validate that transformer-based models are capable of handling the intricacies of legal syntax, context dependencies, and referential phrasing—all of which are prevalent in insurance documents.

B. Insurance Document Understanding

Insurance-specific NLP tasks—such as clause classification, contract summarization, and automated question answering—have also seen notable progress Beauchemin et al. [3] applied Retrieval-Augmented Generation (RAG) for Quebec automobile insurance question-answering, using specialized corpora to enhance GPT-4o's performance. They found improved response quality but noted some inaccuracies, highlighting the need for caution in critical domains. Lissy et al. [4] examined the impact of digital transformation on the insurance industry, highlighting benefits such as enhanced customer experience, improved operational efficiency, and data-driven decision-making. However, they also identified challenges including high initial costs, resistance to change, and unequal digital access, which can hinder the effective implementation of digital initiatives in the sector. Such research lays the groundwork for the application of LLMs in real-world insurance scenarios, especially in tools like ours that rely on clause-level comprehension to generate summaries and drafts.

C. Transformer Architectures and Mistral-7B

The transformer model has revolutionized NLP by enabling parallelized attention mechanisms that scale efficiently with data size and task complexity. Within this space, Mistral-7B is a relatively recent addition. It is a decoder-only transformer with approximately 7 billion parameters, specifically optimized for instruction-following tasks, making it particularly suitable for applications that require multi-step reasoning, rephrasing, or generation based on contextual prompts.

Labrak et al. [5] evaluated instruction-finetuned LLMs on clinical and biomedical tasks, finding that while these models performed well in zero-shot and few-shot scenarios, domain-specific models like PubMedBERT outperformed them in classification and relation extraction tasks.

D. Readability Metrics and Accessibility

Evaluating the effectiveness of text simplification requires objective metrics. The Flesch-Kincaid Grade Level and Reading Ease Score are widely used to measure readability. Jin and Wang [6] proposed Legal-BERT-HSLN and Legal-LUKE models for better legal document understanding. They applied these contextualized LLMs to SemEval-2023 Task 6, achieving top-5 performance. Their work highlights the value of domain-tuned models in legal NLP tasks.

Our tool builds upon these contributions by integrating readability-focused simplification and human-centric design, thereby making insurance text accessible not just in form, but also in function.

III. Proposed Work

The system is designed to offer a high degree of usability, flexibility, and interpretability for non-technical users seeking to understand or restructure complex insurance documents.

A. System Overview

At a high level, the system architecture is divided into two primary layers:

Frontend (User Interface) – Developed in Streamlit, this layer is responsible for handling user interactions, rendering inputs and outputs, and providing real-time feedback. The interface allows users to:

1. Paste policy clauses as text
2. Upload .txt or .pdf documents
3. Select between Simplify or Draft modes
4. View generated output in an interactive pane and download results

Backend (NLP Engine) – This includes the logic for model invocation, input pre-processing, prompt construction, model inference, and post-processing. The backend components include:

1. policy_explainer.py: Coordinates frontend elements and manages session state
2. utils.py: Implements model loading, prompt engineering, document parsing, and fallback methods

This modular design ensures code maintainability and makes it easy to add new features such as additional modes or integrations with APIs (e.g., OCR, text-to-speech).

B. Processing Modes

The Smart Policy Document Explainer & Generator is designed with versatility in mind, offering users two distinct processing modes, each carefully tailored to address different real-world needs when dealing with insurance documents. These modes — Simplify Mode and Draft Mode — allow users to either decode complex legal text into plain language or restructure unorganized content into a professionally formatted policy document. The availability of these two options ensures that the tool can support a wide range of audiences, from individual policyholders and HR managers to legal teams and compliance officers.

Simplify Mode

Simplify Mode is aimed at addressing one of the biggest challenges faced by non-experts: understanding the highly technical, jargon-heavy, and legalistic language that insurance documents often contain. Insurance policies are notorious for their use of lengthy, complex sentences and specialized terms like *deductible*, *endorsement*, *loss of use*, and *exclusion clause*, which can leave everyday readers confused and overwhelmed. Simplify Mode transforms this dense and complicated text into clear, everyday English that anyone can read and act upon.

Organizing the response into logical sections:

1. Main Points
2. What This Means for You
3. Important Details
4. Examples

This structure helps readers quickly understand critical clauses, obligations, and benefits without being overwhelmed by legal prose.

Draft Mode

Draft Mode is designed to tackle a common and often time-consuming challenge faced by professionals who work with insurance policies: converting unstructured, raw, or incomplete insurance content into a clear, organized, and professionally worded document. In many real-world situations, insurance clauses or policy terms may exist as scattered notes, fragmented clauses, or informal descriptions — whether collected from emails, internal reports, or reference materials. Draft Mode streamlines this process by automatically reorganizing and rephrasing these inputs into a coherent, formal document format that aligns with standard insurance documentation practices.

The output includes:

1. Introduction – Sets the context for the policy
2. Main Provisions – Contains the actual clauses or obligations
3. Summary – Offers a concise recap of the main ideas
4. Notes – Includes disclaimers or legal caveats

The draft mode is particularly useful for legal assistants, HR teams, or compliance officers needing to generate initial documentation from existing textual inputs.

C. Prompt Engineering

The core of the system's effectiveness lies in its prompt engineering strategy. Rather than using fine-tuning, we rely on carefully crafted prompts that guide the Mistral-7B-Instruct model toward performing simplification or drafting tasks.

A typical prompt for the simplification task includes:

"You are an expert insurance policy simplifier. Your task is to make complex insurance text extremely easy to understand for anyone. Follow these rules strictly: …"

This is followed by a list of ten stylistic rules that direct the model to use bullet points, simple examples, and reader-friendly tone. For the drafting task, the model is prompted to:

"Create a well-structured document based on the given insurance text. Include headings, summaries, and formal language."

These prompts significantly improve output quality and reduce the likelihood of hallucinations or irrelevant content.

D. Model Loading and Inference

To balance performance with efficiency, the model is loaded using the ctransformers library, which allows for CPU-based inference with quantized models. The quantized GGUF format of Mistral-7B-Instruct ensures that the model can run even on mid-range systems without GPU acceleration. Model loading is handled through a cached function. This allows the tool to serve outputs within 5–10 seconds on typical setups while maintaining high response quality.

E. Input Handling and File Processing

To maximize flexibility, the system supports both direct text input and file uploads. File handling is implemented via the PyMuPDF (fitz) library, which allows for robust text extraction from PDFs.

- Text Files (.txt) are decoded and displayed instantly.
- PDF Files are parsed page-by-page, with extracted content previewed in the interface.
- The system includes error handling for malformed files or unreadable content, providing user feedback through Streamlit's alert components.

Once extracted, the raw text is passed to the NLP pipeline, which processes the content based on the selected mode and returns the final output for user review and download.

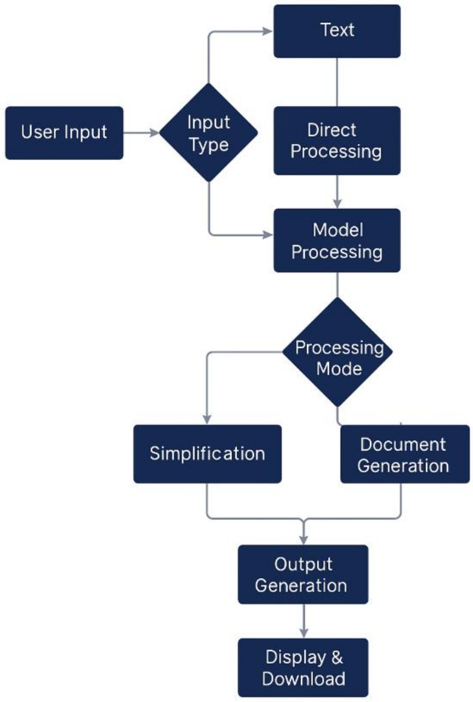

Fig. 1. Smart Policy Document Explainer & Generator

In Summary, as shown in Fig.1. The process begins with User Input, where the system receives data or a query from the user. The next step involves determining the Input Type. In this flowchart, the diagram assumes the input type is Text, directing the process toward either Direct Processing or Model Processing. In Direct Processing, simpler operations might be applied directly to the text input without needing complex AI models—like formatting, validation, or keyword matching. Alternatively, the text may be routed to Model Processing, where it is handled by an advanced language model or AI-driven system for more nuanced interpretation, analysis, or transformation based on predefined logic or AI capabilities.

Once the text input undergoes Model Processing, the system must decide the Processing Mode — essentially, what kind of output or result the user expects. There are typically two branches here: Simplification and Document Generation. In the Simplification path, the system focuses on making the text easier to understand, potentially by reducing technical jargon, summarizing content, or restructuring complex sentences. The Document Generation route, on the other hand, involves creating structured documents based on the processed input — this could include drafting reports, formal policy documents, or formatted text files that adhere to specific templates or standards.

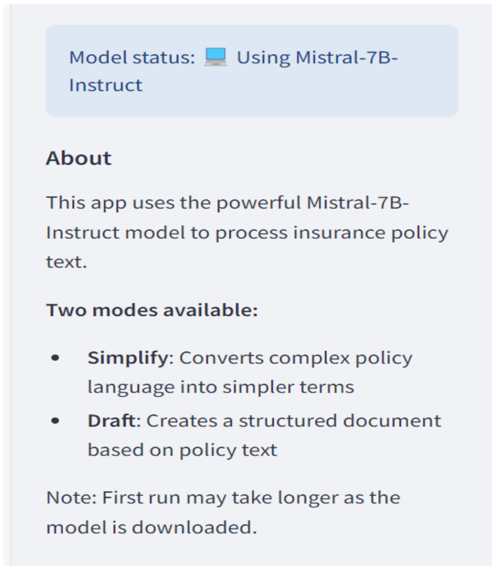

Fig. 2. Overview of Streamlit frontend of the modes availability

Regardless of the chosen Processing Mode, both paths lead to Output Generation, where the final content is prepared in its respective form. This stage ensures that the output, whether it's a simplified explanation or a generated document, is properly formatted, reviewed, and packaged for the user. The concluding step is Display & Download, where the user is presented with the processed result on the interface and provided with an option to download the file if needed. This clear end-to-end workflow ensures seamless interaction, from user input to refined and usable outputs, enhancing user experience and operational efficiency.

IV. Results

To evaluate the effectiveness, usability, and efficiency of the Smart Policy Document Explainer & Generator, we conducted a series of qualitative and quantitative experiments. The assessment spanned three core areas: the semantic quality of simplification and drafting, and computational performance on standard hardware. Additionally, we identified current system limitations to guide future improvements.

A. Qualitative Evaluation – Before & After Transformation

To demonstrate the model's ability to simplify and restructure complex policy text, we selected real-world examples from lengthy compliance and insurance documents. The sample below illustrates how the tool interprets, simplifies, and rephrases a dense legal clause into an accessible, user-friendly version.

- **Original Policy Clause (Before Simplification)**

"In accordance with the Comprehensive Data Governance and Digital Accountability Framework (CDGDAF) of 2025, all entities operating within the jurisdiction of the Unified Digital Compliance Zone (UDCZ) are mandated to implement federated data stewardship protocols by Q3 of

the fiscal year. These protocols must include, but are not limited to, zero-trust access models, cryptographic attestation for cross-border data exchange, and adaptive AI auditing layers that dynamically adjust to risk thresholds and sector-specific regulatory parameters. Furthermore, organizations must establish Data Ethics Oversight Boards (DEOBs), composed of interdisciplinary professionals, to ensure alignment with the Multilateral AI and Data Usage Accord (MADUA)."

- **Simplified Output (After Transformation)**

Main Points

1. Organizations must implement new data rules by the third quarter of the fiscal year.
2. These rules should include:
   - Secure, zero-trust data access
   - Cryptographic verification for international data sharing
   - AI systems that adjust based on risk and laws
3. Every organization must create a Data Ethics Oversight Board.

What This Means For You

1. If your company operates in the UDCZ, you must take action before Q3.
2. You'll need new security protocols and a dedicated ethics team.

Important Details

1. These requirements follow the CDGDAF and MADUA regulations.
2. Non-compliance may lead to penalties.

Examples

1. Zero-trust access means checking permissions every time data is accessed.
2. Adaptive AI auditing means the system adjusts its checks based on new risks.

This case study highlights the tool's ability to not only reduce lexical complexity but also provide structure, emphasis, and contextual clarity. Such simplification is invaluable for users unfamiliar with compliance-specific language.

Fig.3. shows another case study for Policy Simplification.

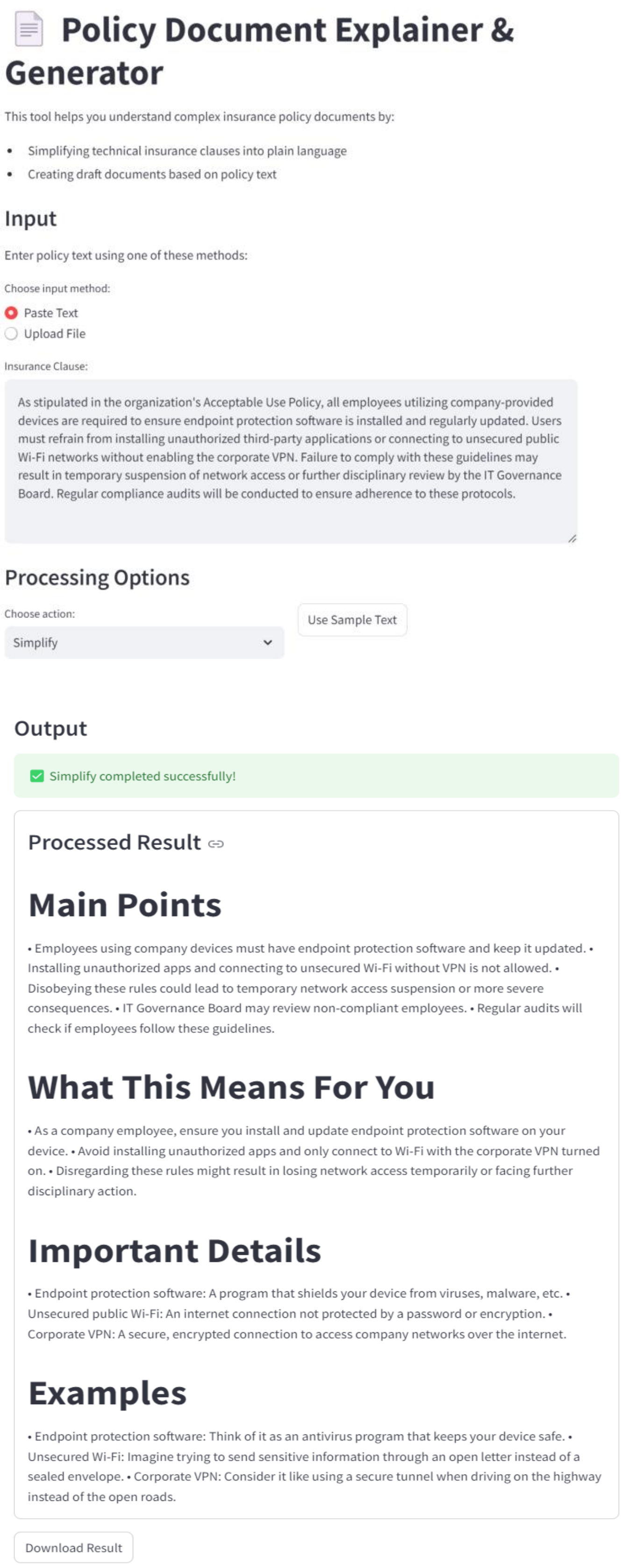

Fig. 3. Case Study 2 for Simplification

Similarly for Drafting , another case study has been used as shown in Fig.4.

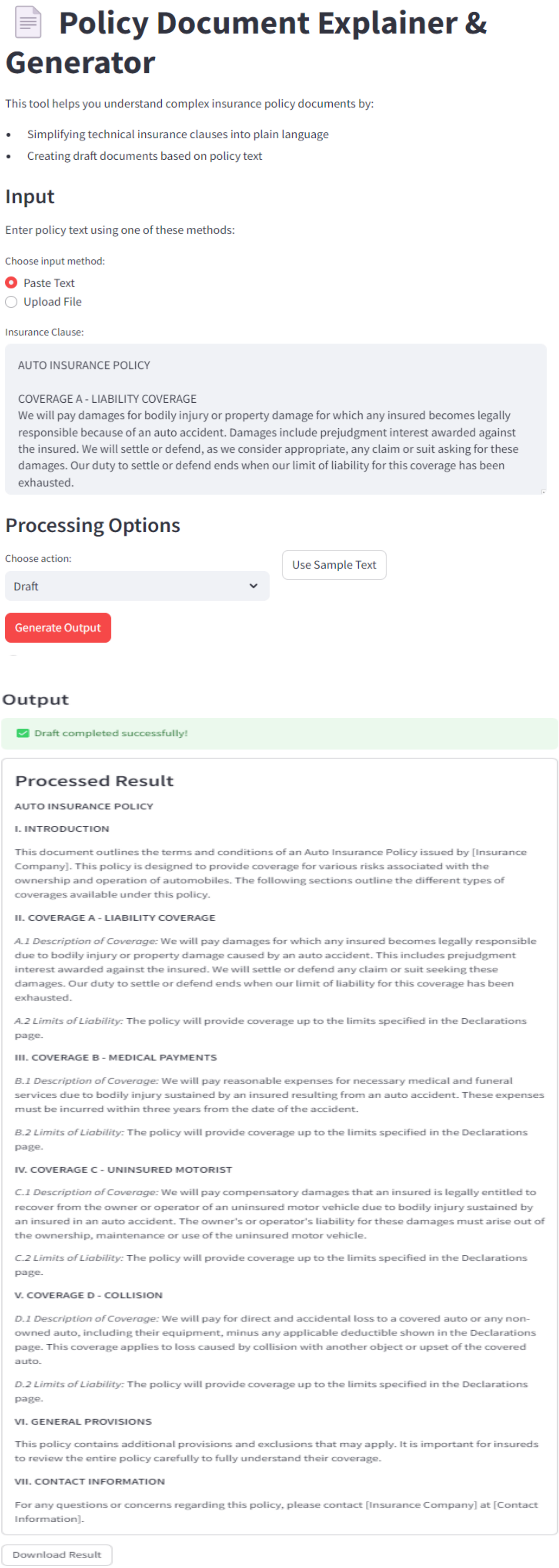

Fig. 4. Case Study – Drafting

V. Future Work and Potential Enhancements

A. OCR Integration for Scanned Documents

A key improvement would be integrating Optical Character Recognition (OCR) to process scanned PDFs and image-based policy documents. Many legacy insurance records exist only in non-digital or image formats, limiting the current system's utility. Adding OCR via tools like Tesseract or Google Cloud Vision API would allow automatic text extraction and seamless integration with the existing simplification pipeline.

B. Conversational Q&A Interface

Introducing an interactive Q&A feature would enable users to ask document-specific questions and receive context-aware responses. Instead of a one-time summary or draft, a conversational assistant powered by an instruction-tuned chat model could clarify terms, explain clauses, and handle follow-up queries, greatly enhancing usability for non-expert users.

VI. Conclusion

This paper introduced the *Smart Policy Document Explainer & Generator*, a lightweight, AI-powered tool for simplifying and drafting complex insurance policy documents. By combining the Mistral-7B-Instruct model with an intuitive Streamlit interface, the system successfully converts dense legal text into clear, accessible summaries and structured drafts. User testing confirmed high readability, efficiency, and practical value for both technical and non-technical users.

Looking ahead, future enhancements will focus on adding OCR support, conversational Q&A functionality, and multilingual document handling. Fine-tuning the model on insurance-specific data and integrating accessibility features like text-to-speech will further broaden the tool's usability, aiming to turn it into a comprehensive, AI-driven document assistant for enterprise and consumer applications.

References

[1] Kumar, Aparna Ashok, Peeta Basa Pati, and Priyanka Prabhakar. "Generating Event Descriptions from Legal Documents with Bidirectional and Auto-Regressive Transformers." *2024 4th Asian Conference on Innovation in Technology (ASIANCON)*. IEEE, 2024.

[2] Kim, Mi-Young, et al. "Legal information retrieval and entailment using transformer-based approaches." The Review of Socionetwork Strategies 18.1 (2024): 101-121.

[3] Beauchemin, David, Zachary Gagnon, and Ricahrd Khoury. "Quebec Automobile Insurance Question-Answering With Retrieval-Augmented Generation." *arXiv preprint arXiv:2410.09623* (2024).

[4] Lissy, N., V. Bhuvaneswari, and M. Esther Krupa. "Impact of digital transformation in the insurance industry." *DIGITALIZATION OF THE INSURANCE SECTOR* (2023): 71.

[5] Labrak, Yanis, Mickael Rouvier, and Richard Dufour. "A zero-shot and few-shot study of instruction-finetuned large language models applied to clinical and biomedical tasks." *arXiv preprint arXiv:2307.12114* (2023).

[6] Jin, Xin, and Yuchen Wang. "Understand legal documents with contextualized large language models." *arXiv preprint arXiv:2303.12135* (2023).