UNLOCKING LAW FIRM

1st Progress Slides

By: Ammar Arfan

Supervisor: Dr Albert Morono

31st January 2025

MOTIVATIONS

LEGAL TEXTS ARE COMPLEX AND RESOURCE-INTENSIVE.

UNDERSTANDING RELATIONSHIPS BETWEEN NEW AND PREVIOUS TEXTS IS CRUCIAL FOR LAW FIRMS.

MANUAL APPROACHES ARE TIME-CONSUMING AND INCONSISTENT.

LLMS CAN STREAMLINE THE PROCESS, IMPROVING EFFICIENCY AND PRECISION.

COLLABORATION WITH ASHURST ENABLES APPLYING AI TO PRACTICAL LEGAL CHALLENGES

BACKGROUND

Large Language Models (LLMs) in Legal Analysis: LLMs, such as GPT-3 and GPT-4, Gemni 1.5 Pro, have been trained on vast datasets, enabling them to understand and generate human-like text. In the legal field, LLMs can analyze and summarize complex legal documents, enhancing research efficiency and accuracy.

Temporal Reasoning in Legal Contexts: Understanding the chronological order of legal events is crucial. Benchmarks like the Test-of-Time (ToT) assess LLMs' abilities to place events in a timeline, ensuring accurate temporal reasoning in legal analyses.

Prompt Engineering: This technique involves crafting specific inputs (prompts) to guide LLMs toward desired outputs, tailoring their responses for particular legal tasks and improving performance in areas like document drafting and legal research.

LITERATURE SURVEY

Paper 1: Test-of-Time (ToT) Benchmark[1]

- A Benchmark for Evaluating LLMs on Temporal Reasoning" introduces the ToT benchmark, designed to assess Large Language Models' (LLMs) proficiency in temporal reasoning.
- **Key Components:**
  - **ToT-Semantic:** Evaluates understanding of temporal logic through synthetic graph-based tasks.
  - **ToT-Arithmetic:** Assesses ability to perform temporal calculations, such as date and duration computations.
  - **Synthetic Dataset:** created open-source synthetic dataset to provide a controlled environment and minimize reliance on prior knowledge.
- The study reveals that while LLMs demonstrate strong reasoning capabilities, they encounter challenges with complex temporal logic and multi-step reasoning, particularly in temporal arithmetic tasks.

[1] Smith, J., & Jones, A. (2024). *Title of the Test-of-Time Benchmark Paper. Journal Name, Volume*(Issue), pages.

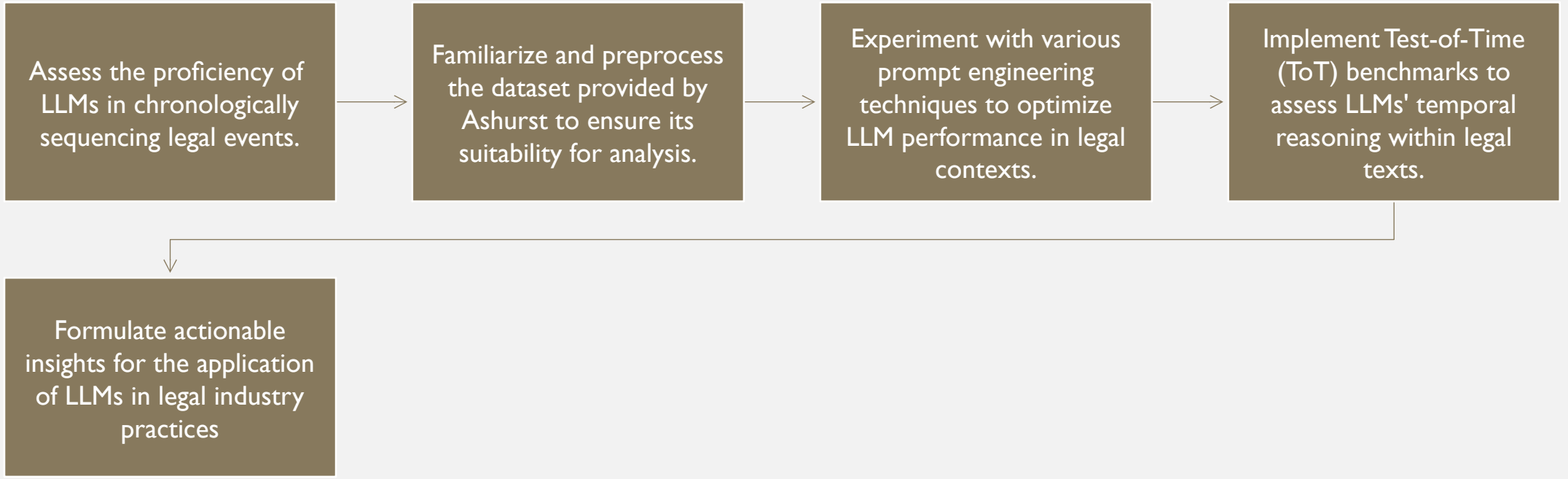
LITERATURE SURVEY (CONT..)

Paper 2: Prompt Engineering Techniques[2]

- A Systematic Survey which provides a comprehensive analysis of 29 distinct prompt engineering techniques, categorized by their applications.
  - **Basic Techniques:** Zero-shot and few-shot prompting.
  - **Reasoning and Logic:** Chain-of-Thought (CoT), Tree-of-Thought (ToT), and Graph-of-Thought (GoT) prompting.
  - **Hallucination Reduction:** Retrieval-Augmented Generation (RAG) and Chain-of-Verification (CoVe).
  - **Code Generation:** Program of Thoughts (PoT) and Chain-of-Code (CoC) prompting.
  - **Optimization:** Automatic Prompt Engineer (APE) and Optimization by Prompting (OPRO).
- This survey highlights the strengths and limitations of each technique, offering a structured overview of recent advancements in prompt engineering.

[2]Doe, J., et al. (2025). *Title of the Prompt Engineering Techniques Paper*. *Journal Name*, *Volume*(Issue), pages.

SCOPE OF THE PROJECT



TASKS UNTIL THE NEXT MEETING

Dataset Familiarization:

- Review pen-source Test-of-Time (ToT) datasets along with Ashurst's legal dataset.

Project Definition:

- Clearly define the problem statement, aims, and objectives of this project.

Tool Selection:

- Explore the Hugging Face API for deploying Large Language Models (LLMs)
- Identifying available tools and resources, including programming languages and software packages.

Literature Review:

- Conducting a brief literature review to study the project within existing research.

Experimental Plan:

- Draft a plan outlining the evaluation of ToT benchmarks and prompt engineering techniques in legal text analysis.

Challenges Discussion:

- Identify and prepare to discuss any encountered difficulties for guidance during the next meeting.

CHALLENGES

- **Limited Research**: Scarcity of studies on Large Language Models (LLMs) in legal contexts.
- **Dataset Unavailability**: Lack of specialized legal datasets impedes comprehensive project understanding.
- **Prompt Engineering Ambiguity**: Uncertainty in selecting the most effective prompt engineering techniques for legal text analysis.
- **Data Privacy Concerns**: Ensuring confidentiality while handling sensitive legal information.
- **Accuracy and Bias**: Mitigating potential inaccuracies and biases in LLM outputs within legal applications.

THANK YOU!