WEB AI EASY T&C

By

GEORGE NJOROGE KARIUKI

21YAD105705

A Project Submitted in Partial Fulfilment of the Requirements for the Degree of Computer Science, School of Computing Sciences Riara University

Nairobi, Kenya.

May 2025

DECLARATION

I declare that this or any other University has not previously submitted this work for the awarding of the course marks. To the best of my knowledge and belief, this work contains no material previously published or written by another person except where due reference is made.

George Kariuki

Signature: ……..

Date: ……

APPROVAL

This project was conducted under my supervision and is submitted with my approval as a university supervisor.

Florence Kimani

Signature: ……

Date: …….

DEDICATION

This project is dedicated to my mentors, family and friends, whose unwavering support and guidance have been invaluable throughout this journey. Your encouragement and belief in the importance of this work have been a constant source of motivation.

I would also like to dedicate this project to my supervisor, whose advice has a guiding light throughout the process of creating this project.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my project supervisor, whose expertise, guidance, and encouragement have been instrumental throughout this research. Your insights have been invaluable in shaping this work. I am deeply appreciative of the time and support you have generously offered.

I am also thankful to my professors and colleagues, who provided thoughtful feedback and advice that helped strengthen this project. Your perspectives have enriched my understanding and inspired new approaches in addressing the challenges of digital transparency and user empowerment.

To my family and friends, thank you for your continuous encouragement and patience. Your support has been a steady source of strength.

ABSTRACT

With the growing complexity of online terms and conditions, many people often agree to lengthy legal documents without really understanding what they say. Whether signing up for accounts, applying for jobs, subscribing to newsletters, or joining social networks; users face long and complicated terms that they usually accept without reading. A lot of these agreements include clauses that allow sharing of personal information with third parties, which raises worries about privacy and data security.

To tackle this issue, this research introduces AI Easy T&C website, a tool powered by AI that simplifies and summarizes terms and conditions. This tool aims to create short and easy-to-read summaries of legal clauses that point out key details and risks in straightforward language. AI Easy T&C website is designed to help users make better decisions regarding their online activities and agreements.

The research involves using a special NLP model that is fine-tuned on summarizing legal documents. User feedback is be used to enhance the clarity and accuracy of the summaries produced. The model quickly analyzes and breaks down complex legal language, providing summaries that are clear and easy to understand.

The expected results include an improved understanding of terms and conditions which will allow users to review agreements more swiftly and effectively. This tool could boost transparency, increase user awareness of important terms and especially clarify clauses about data privacy and sharing. By making it easier to understand T&Cs, this research aims to support user empowerment and transparency in digital consent practices, thus addressing an important need for accountability and informed consent in today's digital world.

TABLE OF CONTENTS

DEFINITION OF TERMS

Terms and Conditions (T&Cs): Legally binding agreements between users and service providers that outline the rules, responsibilities, and rights of both parties.

Natural Language Processing (NLP): A branch of artificial intelligence that focuses on the interaction between computers and human language, enabling machines to understand, interpret, and generate human language.

Data Privacy: The protection of personal information from unauthorized access, use, or disclosure.

Machine Learning: the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data.

LIST OF ABBREVIATIONS AND ACRONYMS

AI: Artificial Intelligence

API: Application Programming Interface

BERT: Bidirectional Encoder Representations from Transformers

GPT-4: Generative Pre-trained Transformer 4

NLP: Natural Language Processing

SDLC: Software Development Life Cycle

T&Cs: Terms and Conditions

UI: User Interface

XAI: Explainable AI

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background Information

In today's digital landscape, users frequently engage with various online platforms, including websites, social media, and digital services, all requiring acceptance of T&Cs during onboarding. T&Cs are intended to inform users about how their data will be handled, defining the rights, obligations, and liabilities of users and service providers. Despite their important role in data protection and privacy, research consistently shows that users often accept these agreements without a thorough read, mainly due to their overwhelming length and complexity (Acquisti & Grossklags, 2005; Angulo et al., 2011; Kesan et al., 2012; McDonald & Cranor, 2008; Meinert et al., 2006; Nissenbaum, 2011; Tsai et al., 2011). A survey in 2023 shows that 68% of people consent to the legal terms and service conditions without reading them, and 32% would sometimes read it but not fully comprehend it.

This trend raises significant concerns about user privacy, data sharing, and autonomy. T&Cs frequently contain clauses regarding data collection, usage, and sharing with third parties, yet these critical details are often buried in dense legal text. For the average user, understanding the implications of such clauses is challenging, resulting in users consenting without a clear picture of how their data might be utilized or shared with other entities.

This document explains the development of AI Easy T&C website, an AI-driven tool designed to simplify and summarize T&Cs, intending to enhance transparency and accessibility. NLP has already transformed everyday interactions by enabling AI-driven analysis, interpretation and summarization of lengthy text. Recent advancements in NLP have opened up new possibilities for addressing this challenge, making it easier to analyze and interpret complex language (Cole Stryker, Jim Holdsworth, 2024). Using these advanced NLP techniques, AI Easy T&C website aims to produce concise, user-friendly summaries of legal clauses, emphasizing key terms, potential risks, and data-sharing practices. By presenting these elements in straightforward language, the tool aspires to empower users to make informed choices regarding the agreements they endorse online).

Positioned at the intersection of digital privacy, legal technology and NLP-driven language processing, AI Easy T&C website addresses a growing need for transparency and user empowerment in online agreements. By leveraging these technological advances, this project

not only seeks to foster clearer understanding of T&Cs but also to support responsible and informed digital interactions in a rapidly evolving online environment.

1.2 Problem Statement

In a perfect digital world, users would easily grasp the T&Cs they agree to when using websites, social media platforms, and online services. These T&Cs would be clear and accessible, allowing individuals to make informed choices about sharing their personal information and understanding data-sharing policies. In this ideal scenario, people could fully understand what their consent means, giving them more control over their data and privacy as they navigate online platforms with confidence.

However, the reality is quite different. T&Cs are often long, complicated, and packed with legal jargon that can confuse the average user. Because of this complexity, many users don't take the time to read through these documents thoroughly. Instead, they end up accepting agreements without fully understanding crucial points about privacy, data usage and sharing with third parties. Studies show that a lot of users either ignore or skim these agreements, unaware of potential risks to their privacy and personal autonomy. This creates a major gap in digital transparency and informed consent, leaving users often in the dark about the actual terms they're agreeing to.

The consequences of this issue are serious. Without a solid understanding of T&Cs, users risk unknowingly agreeing to terms that could jeopardize their data privacy and security. This lack of informed consent weakens user autonomy, as individuals have limited control over how their data is handled or shared by the platforms they use. Moreover, the lack of clarity in these agreements can damage the trust between users and online platforms, raising ethical questions about how digital consent is obtained. As we become increasingly reliant on digital services, it's essential to find a solution that tackles this problem. This research project seeks to investigate ways to simplify T&Cs, aiming for a more transparent and user-focused approach to digital consent.

1.3 Objectives

1.3.1 Main Objective

The main objective is to create a website that uses NLP models to summarize bulky T&Cs and provide a comprehensible and understandable summarized report to the user. This is to ensure users are informed on whatever is being asked by the service provider and they are able to choose whether to accept or decline.

1.3.2 Specific Objectives

a) Create a user-friendly interface on the website, where users are able to submit lengthy T&Cs and get back an understandable summary as output with ease.
b) Design a system that summarizes T&Cs, and ensure the summary is as comprehensive as possible.
c) Generate summary reports from T&Cs provided by the user, that are simple and easily understandable.

1.4 Justification

The AI Easy T&Cs website is crucial in our digital age, where we frequently encounter lengthy and complicated T&Cs when creating an account, or when T&Cs change on websites and apps. Most people lack the time or expertise to thoroughly review these documents. Since T&Cs often include important information about data collection, privacy, and user rights, it can be risky to accept them without fully understanding their contents. AI Easy T&C website aims to assist by providing straightforward, easy-to-read summaries of these T&Cs, enabling users to make more informed choices about how they interact with online services.

This project addresses a significant issue concerning user rights and online consent. By highlighting the most important aspects and summarizing them, AI Easy T&C website helps users quickly grasp the meanings of specific terms, particularly regarding their data and privacy. Such clarity fosters trust between users and online platforms, ensuring that individuals have greater control over their personal information. It simplifies the usual complexities of T&Cs while advocating for increased accountability in the digital realm.

There are platforms like "DeleteMe" that search the internet for personal data sold to data brokers without users' consent and take necessary steps to have that data removed. Unlike this

solution, which addresses the issue after the damage is done, the AI Easy T&C website ensures that users provide informed consent regarding how their data will be used.

Additionally, AI Easy T&C website demonstrates how NLP can make legal documents more accessible to everyone, contributing to the broader field of legal tech. This project not only fills a current gap for users but also highlights the advantages of AI tools in creating a more transparent and responsible digital environment.

1.5 Scope

To achieve this goal, employing an NLP model, thus the tool is capable of generating clear, concise summaries that highlight the most important and relevant information within T&Cs documents. The development of this tool will occur within a controlled environment, allowing for careful monitoring and assessment of its effectiveness. The NLP model used on the tool is pre-trained. The focus is on developing core features that are most beneficial to users. These features include the identification of key clauses such as those related to data sharing, privacy rights, cancellation policies and other significant terms which are often crucial for users to understand when engaging with these digital platforms.

Furthermore, there is a plan to create a user-friendly interface that streamlines the testing of the prototype. This design consideration is important, as it will encourage participation and make the tool accessible to a broader audience. Given the limited timeframe available for this project, it is essential to note that the tool does not offer real-time updates on changing T&Cs as it uses a pre-trained NLP model to summarize T&Cs. This decision allows for a more concentrated approach to evaluating the tool's current capabilities without the complexity introduced by constantly evolving documents.

This user testing is be instrumental in evaluating how easy the tool is to use, how well it conveys the information and its overall effectiveness in aiding users' understanding of T&Cs. Collecting user insights will be invaluable for refining the tool and ensuring that it meets the needs of its intended audience. Ultimately, this project aims to balance the constraints of time and resources while maintaining a strong focus on providing support for users, helping them become more informed about the often-overlooked details contained within the T&Cs documents.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

As digital services evolve, so too do the dense and complex legal documents (particularly Terms and Conditions) that accompany them. Despite their importance in outlining privacy policies and user rights, T&Cs are often ignored due to their length and complexity (Obar & Oeldorf-Hirsch, 2020). These T&Cs often contain a lot of important information on what the organization allows. They also outline what information the organization collects from the user and what it is used for. Users need an easy way of finding out this type of information to avoid acceptance of T&Cs that they do not agree to. This is because some organizations collect users' personally identifiable information and sell it to data brokers who can sell this data to the highest bidder. These documents often obscure critical data-sharing practices, as seen in cases like Facebook-Cambridge Analytica (Cadwalladr & Graham-Harrison, 2018), knowing the user will not read or understand the implications. AI-driven summarization tools can improve transparency by distilling key clauses (Hacker et al., 2023). This review evaluates the potential of these AI-driven tools, specifically AI Easy T&C website, to improve users' understanding of these agreements. By leveraging natural language processing (NLP), AI Easy T&C website seeks to create concise, accessible summaries of T&Cs, fostering transparency and informed consent in the agreement of T&Cs.

Expanding on the issues at hand, consider a typical scenario: a user signs up for a new social media platform. The platform's T&Cs may be dozens of pages long, filled with complex legal terms that the average user does not understand. Hidden within these pages could be clauses that allow the platform to share user data with third-party advertisers. Without a tool like AI Easy T&C website, users are likely to skip reading the full T&Cs and miss critical information about how their data will be used. This lack of transparency and understanding can lead to privacy violations and data misuse.

2.2 Established Themes in AI and Legal Document Summarization

2.2.1 Importance of Accessibility in Legal Documents

The field of legal document processing has long acknowledged that users often skip or skim T&Cs due to "information overload" (Bakos et al., 2014). AI's application in legal tech, particularly NLP-based tools, has shown promise in creating accessible summaries of such documents, allowing users to grasp essential points without extensive reading. This review draws on the works of Nissenbaum (2011) and Tsai et al. (2011), who emphasize the ethical responsibility of ensuring that users are fully informed of data-sharing practices and terms of use. Ensuring that users can easily understand the critical elements of T&Cs is a step towards promoting digital literacy and trust. Moreover, simplifying these documents helps bridge the gap between legal jargon and everyday language; making the digital world more inclusive and comprehensible for all users.

For example, consider the General Data Protection Regulation (GDPR) implemented in the European Union. The GDPR mandates that organizations provide clear and understandable information about how personal data is processed. However, many companies still produce privacy policies that are difficult for the average user to comprehend (Degeling et al., 2019). An AI-driven tool like AI Easy T&C website can help bridge this gap by translating complex legal language into simpler terms, ensuring compliance with regulations and enhancing user understanding.

2.2.2 NLP and Summarization Models

The integration of NLP and machine learning (ML) in legal document summarization has developed significantly over the past decade, with transformer models (such as BERT and GPT-4) showing notable improvements in handling language complexity (Devlin et al., 2019; OpenAI, 2023). NLP models can perform abstractive summarization, generating simplified sentences based on key ideas, or extractive summarization, selecting relevant sentences from the text verbatim. The continuous evolution of these models promises enhanced accuracy and efficiency, which is crucial in legal contexts where precision is important. In addition, the adaptability of these models to various legal documents ensures that they remain relevant and effective in a rapidly changing digital landscape.

Consider the advancements in transformer models like BERT and GPT-4, that have revolutionized the field of NLP by enabling more accurate and context-aware language processing. BERT's bidirectional approach allows it to understand the context of words in relation to their surrounding words, while GPT-4's autoregressive nature excels in generating coherent and contextually appropriate text. These capabilities make them well-suited for legal document summarization, as they can effectively parse and interpret complex legal language (Liu et al., 2022).

2.2.3 Privacy and Ethical Concerns

Privacy concerns in T&Cs are a significant issue, as many agreements include clauses that permit data-sharing with third parties. AI can be used to identify and highlight these clauses, increasing user awareness of potential privacy risks. However, studies reveal that opaque AI models can introduce ethical challenges, such as biased summarization or misinterpretation of critical information (Mehrabi et al., 2021). The challenge is ensuring AI models can responsibly and transparently summarize legal documents without distorting the intended meaning. This involves not only advanced technical capabilities but also a commitment to ethical AI practices that prioritize user rights and data protection. Addressing these concerns requires continuous monitoring, evaluation and improvement of AI systems to uphold the highest standards of accuracy and fairness in legal document summarization. This kind of accountability can be maintained using Explainable AI (XAI) (Arrieta et al., 2020).

One notable example of privacy concerns is the Cambridge Analytica scandal (Confessore, 2018), where personal data from millions of Facebook users was harvested without their consent and used for political advertising. This incident highlighted the importance of transparency in data-sharing practices and the need for tools that can help users understand and control how their data is used. AI-driven summarization tools like AI Easy T&C website can play a crucial role in preventing such breaches by clearly outlining data-sharing clauses and alerting users to potential risks.

2.3 Critical Analysis of AI in T&C Summarization

2.3.1 Strengths and Supporting Evidence

AI-driven summarizers can make T&Cs more approachable by breaking down complex language into layman's terms. Such tools have proven effective in various legal contexts by enabling users to identify key clauses related to data privacy, liability, and usage rights. Research has shown that these tools increase comprehension rates among users, thereby fostering informed decision-making in digital transactions (Hacker et al., 2023). The ability to provide clear, concise summaries of lengthy legal documents empowers users to make informed choices, enhancing their overall digital experience. Furthermore, these tools contribute to greater transparency and accountability in digital services, building user trust and confidence.

2.3.2 Challenges and Limitations in Current Literature

Despite these strengths, gaps remain in ensuring that AI models accurately interpret legal clauses. Complex legal language, characterized by specialized terminologies and conditional statements, often results in models misunderstanding key points (Chalkidis et al., 2020). Additionally, much of the training data used for these models does not include the variety needed to handle diverse T&C formats. Limited access to annotated legal datasets means that many models are tested on general language datasets, which may lack the nuance required for legal applications (Chalkidis et al., 2020). These limitations highlight the need for more extensive and diverse training datasets that can capture the intricacies of legal language and improve the overall performance of AI summarization tools. Addressing these challenges requires a collaborative effort between legal experts and AI researchers to ensure that models are trained on accurate and representative data.

Consider the challenges faced by AI models when interpreting clauses related to indemnity or force majeure in legal documents. These clauses often contain specialized legal terms and complex conditional statements that require a deep understanding of legal principles. If an AI model is not adequately trained on diverse legal datasets, it may misinterpret these clauses, leading to inaccurate summaries. To mitigate this risk, researchers must prioritize the

development of comprehensive and representative training datasets that cover a wide range of legal documents and terminologies.

2.3.3 Ethical and Controversial Aspects

A major controversy in AI-powered T&C summarization is the ethical obligation to maintain accuracy and transparency in summaries. Inaccurate summaries risk misleading users, with potential legal implications if users act based on incomplete or incorrect information. Moreover, AI models are often criticized as "black boxes," making it difficult to trace or explain how summaries are generated, which raises questions about accountability and user trust (Anderson, 2024). Ensuring transparency in the AI summarization process is crucial to building user trust and confidence. This involves developing explainable AI models that provide clear insights into how summaries are generated, allowing users to understand and verify the accuracy of the information provided. Ethical considerations must be at the forefront of AI development to ensure that these tools benefit users without compromising their rights or privacy.

One approach to addressing ethical concerns is the implementation of explainable AI (XAI) techniques. XAI aims to make AI models more transparent and interpretable, enabling users to understand how decisions and summaries are generated. By incorporating XAI into AI-driven summarization tools, developers can provide users with insights into the reasoning behind each summary, enhancing transparency and trust. Additionally, regular audits and evaluations of AI models can help identify and rectify biases or inaccuracies, ensuring that summarization tools uphold ethical standards

2.4 Relevance to AI Easy T&C and Research Needs

2.4.1 Insights and Observations

Current research supports the notion that AI can effectively simplify T&Cs, though challenges in transparency and ethical responsibility remain. The AI Easy T&C website, by providing real-time, user-friendly summaries, can contribute to bridging the information gap between service providers and users. Based on evidence gathered, there is a strong argument for

developing AI models that can handle diverse T&C formats while maintaining interpretive accuracy. This requires ongoing research and innovation to refine AI summarization techniques and ensure that they meet the needs of diverse user groups. Collaboration between technology developers, legal professionals, and users is essential to create effective and reliable summarization tools.

Moreover, it's critical to consider the different ways users interact with T&Cs across various platforms. For example, mobile devices often present T&Cs in a format that is difficult to read, with long scrolling pages and small text. By leveraging responsive design and AI-driven summaries, AI Easy T&C website can adapt to these different contexts, providing clear and concise information regardless of the device being used. Additionally, incorporating user feedback into the development process can help identify common pain points and areas for improvement, ensuring that the tool remains relevant and user-centric.

2.4.2 Rationale for AI Easy T&C website

The AI Easy T&C website aligns with the need for responsible digital consent mechanisms. By focusing on creating accessible summaries, this tool can help mitigate the risks associated with uninformed consent and enhance user autonomy in data-sharing practices. Given the ethical and legal ramifications of AI-driven summaries, there is a pressing need for models that are both transparent and reliable, making AI Easy T&C website a timely and essential project in digital privacy. The development of AI Easy T&C website represents a proactive approach to addressing the challenges posed by complex legal documents and fostering a more transparent and user-friendly digital environment.

For instance, consider the increasing reliance on digital services during the COVID-19 pandemic, which saw a surge in online interactions and transactions. With more users engaging with digital platforms, the need for clear and understandable T&Cs became even more critical. AI Easy T&C website can empower users by providing them with the necessary information to make informed decisions about their digital interactions, ultimately fostering a more trustworthy and ethical digital ecosystem.

2.4.3 Recommended Research Methodology

The development of AI Easy T&C website could benefit from a reinforcement learning framework with user feedback to enhance summary accuracy and clarity over time. Combining extractive and abstractive techniques allows for summarization that is both accurate and accessible, striking a balance between maintaining legal integrity and readability (Smilarubavathy et al., 2024). This user-centered approach ensures that the tool remains relevant and responsive to user needs. Engaging users in the development process through feedback mechanisms helps to refine and improve the tool, ensuring that it effectively addresses their concerns and meets their expectations.

To implement this, researchers could create a feedback loop where users rate the accuracy and usefulness of the AI-generated summaries. This feedback can then be used to fine-tune the model, improving their performance over time. Additionally, conducting user studies and experiments to compare the effectiveness of different summarization techniques can provide valuable insights into the most effective methods for legal document summarization. This iterative process of testing and refining the models is essential to developing a robust and reliable summarization tool.

2.4.4 Scope and Limitations

This review focuses exclusively on AI-based T&C summarization, particularly for digital consent. Non-AI methods and general-purpose summarization tools fall outside the review's scope, as they do not address the unique challenges posed by T&Cs. The scope of this review is limited to the examination of AI-driven tools and their application in legal document summarization. Future research should explore the integration of AI with other technologies and methodologies to further enhance the effectiveness and reliability of T&C summarization tools.

It is also important to recognize the limitations of AI in understanding and interpreting complex legal language. While AI models have made significant advancements, they are not perfect and can still make errors in summarization. Therefore, it is crucial to have human oversight and validation of AI-generated summaries to ensure their accuracy and reliability. Additionally, addressing potential biases in the training data and ensuring that the models are trained on

diverse and representative datasets can help mitigate the risk of biased or inaccurate summaries.

2.5 Conclusion and Future Directions

In conclusion, the literature demonstrates that AI-powered summarization can potentially address the complexities of T&Cs, making digital consent more accessible and transparent. However, the need for ethical considerations, transparency and accountability remains critical. AI Easy T&C website offers a promising approach by incorporating user feedback to improve summary clarity and accuracy, which may help bridge gaps in user understanding of T&Cs and data privacy policies. Future research should focus on expanding annotated legal datasets and refining AI models for enhanced interpretive accuracy and transparency. Continued collaboration between technology developers, legal experts and users is essential to ensure that AI-driven summarization tools effectively address the challenges posed by complex legal documents and promote informed digital consent.

Future directions for research and development in this field should also consider the integration of multimodal data, such as audio and video, to provide more comprehensive and accessible summaries. For example, incorporating voice-assisted summaries can help users who prefer auditory information, while visual aids like infographics can enhance understanding for visual learners. Additionally, exploring the potential of blockchain technology to provide immutable and verifiable records of T&C agreements can further enhance transparency and trust in digital transactions.

Overall, the development and implementation of AI-driven summarization tools like AI Easy T&C Website represent a significant step towards empowering users and promoting ethical digital interactions. By prioritizing transparency, accuracy and user-centric design; these tools can help create a more informed and responsible digital ecosystem.

CHAPTER 3: METHODOLOGY.

3.1 Introduction

This chapter describes the processes, requirements engineering, system design, implementation and testing phases that are taken in the development of this project.

Expanding the scope of this project, we must acknowledge the broader context in which the AI Easy T&C website operates. As digital interactions become more pervasive, the need for users to understand the legal agreements they engage with is increasingly critical. This tool aims to democratize access to information, enabling users from diverse backgrounds to make informed decisions about their digital privacy and rights. By combining state-of-the-art NLP techniques with user-centered design principles, this project seeks to set a new standard for transparency and accessibility in digital consent processes.

3.2 Methodology

3.2.1 Specification

The project adopts the Iterative SDLC model, a dynamic methodology where development progresses through repeated cycles, allowing for continuous refinement and adaptation. This model ensures that each iteration delivers a functional version of the product, which is then evaluated and improved in subsequent cycles. The iterative approach supports evolving requirements and early risk mitigation. This makes it particularly effective for projects where flexibility is critical. By focusing on incremental progress and regular validation, this structured yet adaptable framework enhances both project efficiency and product quality throughout the development lifecycle.

3.2.2 Justification

The Iterative SDLC model is ideally suited for this project due to its flexible yet structured approach, which aligns with the goal of progressively developing and refining the AI Easy T&C website. Unlike rigid linear methodologies, the Iterative model allows for incremental development in repeated cycles, where each iteration delivers a functional version of the product. This ensures continuous feedback, adaptation and improvement at every stage.

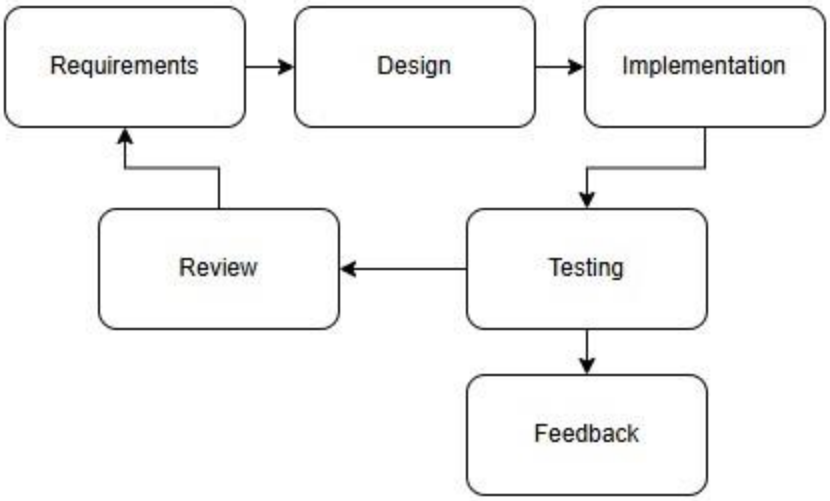

Fig 3. 1 Iterative Model Phases

This model is particularly effective for projects like the AI Easy T&Cs website, where requirements may evolve and early user feedback is critical. By breaking the project into manageable iterations, each encompassing planning, design, development and testing, the team can systematically validate functionality, usability and accuracy while minimizing risks. The iterative nature of this approach ensures that adjustments can be made dynamically, enhancing both the tool's robustness and user satisfaction.

Additionally, the Iterative model's built-in feedback loops align perfectly with the development of AI-based tools, where testing and refinement are ongoing. Each cycle not only improves the product but also ensures alignment with user needs. This results in a reliable, user-friendly final version. This approach balances efficiency with adaptability, making it ideal for projects requiring both precision and responsiveness.

3.2.3 Detailed Description

Requirement Analysis

The first phase involves identifying and documenting system requirements through literature analysis and web surfing. In the iterative model, this occurs at the start of each development cycle, with requirements refined based on feedback from previous iterations. A feasibility study evaluates how the tool summarizes privacy policies, with each iteration improving

functionality (e.g., using a faster AI model). User needs are continuously reassessed, allowing emerging requirements like visual aids to be incorporated in subsequent cycles.

System Design

The system design phase defines the architecture for each iteration, with components added incrementally. The design outlines how NLP models process T&Cs, emphasizing modularity for iterative testing and maintenance. UI/UX improvements are implemented cyclically, with initial iterations focusing on core functionality and later cycles enhancing accessibility features. Integration of backend (Flask) and frontend (Angular) components follows an evolutionary approach, enabling continuous deployment of working prototypes.

Implementation

Implementation occurs across multiple iterations, with each cycle delivering new functionality. Early iterations focus on core summarization features using Flask for NLP processing, while subsequent cycles incorporate Angular-based UI enhancements and API extensions. I perform continuous integration, with each iterative build adding to previous work while maintaining system stability through version control and automated testing protocols.

Testing and Validation

Testing is conducted at the end of each iteration, combining unit, integration, and user acceptance testing. The iterative approach allows for progressive validation, where initial cycles verify core functionality and later cycles stress-test advanced features under peak loads. User feedback from each testing phase directly informs requirements for the next iteration, creating a continuous improvement cycle. Performance metrics are tracked across iterations to measure quality enhancements.

Feedback Integration

After each iteration's testing phase, formal evaluations identify improvement opportunities. Feedback is systematically categorized and prioritized for subsequent cycles, ensuring the tool evolves to meet user needs. This closed-loop process connects back to requirement analysis for the next iteration, maintaining alignment between development efforts and user expectations throughout the project lifecycle.

3.2.4 Plan

Interviews are conducted with potential users to understand their challenges with T&Cs and gather expectations for the summarizer. This method provides direct insights into user needs, including specific clauses users struggle to comprehend, such as data-sharing policies. The gathered requirements inform the design and development of the tool, ensuring that it addresses real user concerns and provides practical solutions. Additionally, literature analysis and web surfing help identify best practices and emerging trends in T&C summarization, further guiding the development process.

3.2.4.1 Budget

The successful implementation of the AI Easy T&C website project requires a well-planned budget to cover various aspects of development, testing, and deployment. Below is an estimated budget breakdown:

Table 3. 1 Development Costs

Item	**Estimated Cost (KES)**
Backend Development	1,000
Frontend Development	1,000
Database Setup	500
AI Model Integration	2,000
Total Development Cost	**3,500**

Table 3. 2 Testing Costs

Item	**Estimated Cost (KES)**
Unit Testing	500
Integration Testing	1,000
User Acceptance Testing (UAT)	3000
Total Testing Cost	**4,500**

Table 3. 3 Deployment Costs

Item	Estimated Cost (KES)
Backend Hosting (Heroku)	78,000/year
Frontend Hosting (Netlify)	30,000/year
SSL Certificates	42,000/year
Total Deployment Cost	**150,000/year**

Table 3. 4 Miscellaneous Costs

Item	Estimated Cost (KES)
Project Management	10,000
Documentation and Training	10,000
Contingency Fund	30,000
Total Miscellaneous Cost	**50,000**

Table 3. 5 Total Estimated Budget

Category	Estimated Cost (KES)
Development	3,500
Testing	4,500
Deployment (per year)	150,000
Miscellaneous	50,000
Grand Total	**208,000**

This budget ensures that all critical aspects of the AI Easy T&C website project are adequately funded, enabling the team to deliver a high-quality, reliable, and user-friendly summarization tool. Future budget adjustments may be necessary based on actual expenses and evolving project requirements.

3.2.5 Design Diagram

Flowchart

The flowchart visualizes the tool's workflow, from user input (uploading T&Cs) to the generation and display of concise summaries. Key stages include text preprocessing, NLP model execution, and user interface presentation. The flowchart provides a clear and detailed overview of the system's operation, highlighting the interactions between different components and the sequential steps involved in processing and summarizing T&Cs. This visual representation helps to understand the system's functionality and identify potential areas for improvement.

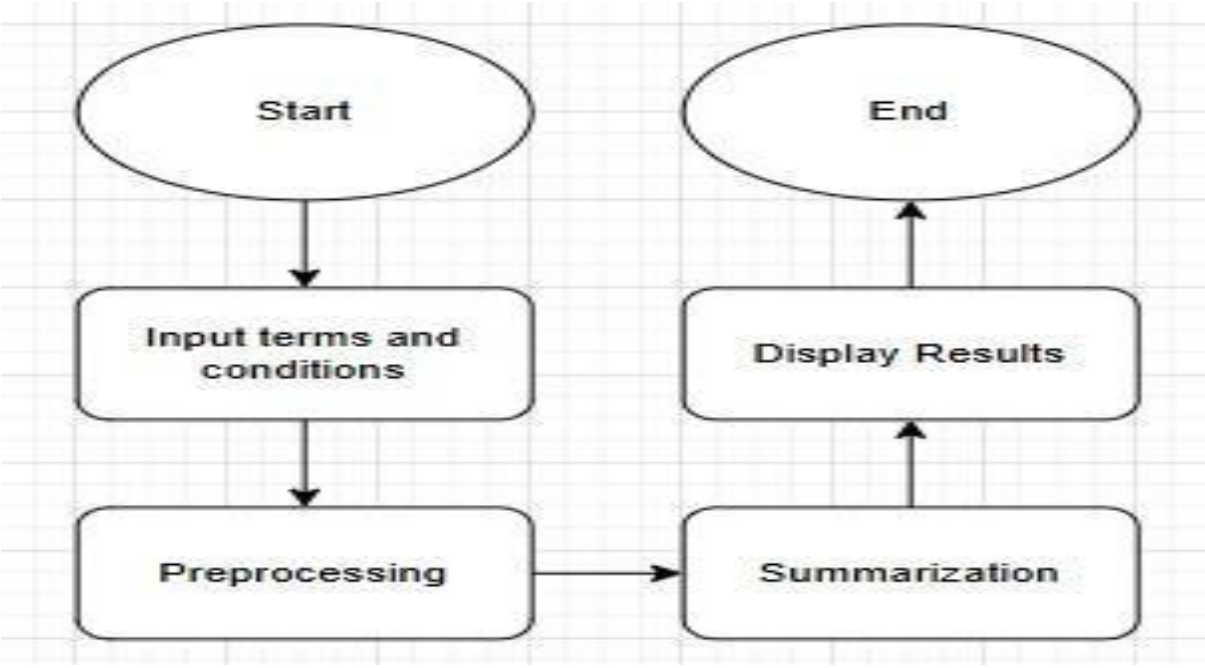

Fig 3. 2 Summarization Process Flowchart

3.2.6 Implementation

The implementation phase involves integrating various technologies:

- **Backend Development:** Flask framework is used to deploy NLP models and manage API endpoints for processing user-uploaded documents.
- **Frontend Development:** Angular provides a dynamic and user-friendly interface, ensuring accessibility and responsiveness across devices.
- **Model Training:** Pretrained NLP model is used.
- **API Integration:** APIs enable seamless communication between the frontend and backend, facilitating efficient data flow and real-time summarization.

During this phase, the focus is on implementing the core functionalities, integrating various components, and ensuring that the system operates as intended. The use of modern technologies and frameworks ensures that the tool is scalable, maintainable and capable of handling large volumes of data.

3.2.7 System Testing

a. **Unit Testing**

Individual components, such as the text preprocessor and summary generator, are tested for accuracy and performance. This phase involves verifying that each module functions correctly in isolation and meets the specified requirements. Automated testing frameworks may be used to streamline the process and ensure comprehensive coverage of test cases.

b. **Integration Testing**

Modules are combined and tested to ensure smooth interactions between the NLP model, backend APIs, and frontend interface. Integration testing focuses on verifying that the different components work together seamlessly and that data flows correctly through the system. This phase helps identify and resolve any issues that may arise from the interactions between different modules.

c. **System Testing**

The entire tool is evaluated to ensure that all functionalities operate cohesively and meet project requirements. System testing involves a thorough examination of the tool's overall performance, usability, and reliability. This phase includes both automated and manual

testing approaches to ensure that the tool meets the highest standards of quality and functionality.

d. **Performance Testing**

Stress tests measure the tool's responsiveness and stability under varying loads, ensuring reliability during peak usage periods (The QA Lead, 2019). Performance testing helps identify any bottlenecks or limitations in the system, allowing developers to optimize the tool for better performance and scalability. This phase ensures that the tool can handle real-world usage scenarios and deliver a seamless user experience.

3.3 Deliverables

3.3.1 System Modules

The system comprises the following modules:

1. **Input Module:** Allows users to upload T&Cs for summarization.
2. **NLP Module:** Processes and summarizes the input text.
3. **Output Module:** Displays concise summaries with highlighted key clauses.
4. **Feedback Module:** Collects user feedback to improve the summarization model.

Each module is designed to perform a specific function, contributing to the overall effectiveness and usability of the tool. The modular design ensures that the system is flexible, scalable, and easy to maintain.

3.3.2 Users and Roles

User roles are defined to ensure tailored functionality and secure access:

- **General Users:** Can upload T&Cs, view summaries, and provide feedback.
- **Administrators:** Manage backend operations, monitor system performance, and update datasets/models as needed.

These roles align with user-centered design principles to cater to diverse user needs while maintaining security and accessibility. By clearly defining user roles, the project ensures that

each user group has access to the necessary functionalities and features, enhancing the overall user experience.

Moreover, the project includes training and support resources to help users understand and effectively utilize the tool. These resources may include user manuals, tutorials, and customer support channels, ensuring that users can easily navigate and benefit from the AI Easy T&C website.

CHAPTER 4: IMPLEMENTATION, TESTING AND RESULTS

4.1 Implementation

4.1.1 Planned Implementation Approach

The AI Easy T&C project is implemented using a structured, modular approach to ensure scalability, maintainability, and flexibility. This approach is essential for handling the complexities and evolving needs of the system. The project is be divided into three main components, each with a specific focus and responsibility:

- **Backend**: This component houses a pre-trained AI-powered Natural Language Processing (NLP) engine responsible for processing and summarizing Terms and Conditions (T&C) documents. The engine uses advanced NLP techniques to ensure accurate and meaningful summaries. The API uses Flask framework to handle requests from the frontend.
- **Frontend**: The frontend is a web-based user interface designed for ease of use. It allows users to input documents and view the summarized results. The interface is intuitive, providing users with a seamless experience from document upload to summary review.
- **Database**: A lightweight and efficient database system is used to store user feedback. This allows for easy retrieval and analysis of data, which is essential for continuous improvement and user satisfaction.

To enhance development efficiency and reliability, a pretrained AI model is used. The model, already trained on extensive datasets, provides a strong foundation for the summarization tasks.

4.1.2 Programming Languages and Development Tools

Backend Implementation

- **Programming Language**: Python has been chosen for its rich set of NLP libraries and its versatility in handling various tasks. Its popularity in the AI community also means numerous resources and support are available.
- **Framework**: Flask is selected for its lightweight nature and suitability for API development. It provides the necessary tools to build a robust backend without the overhead of more extensive frameworks.

- **NLP Model**: A pre-trained transformer model, BART Large CNN, is utilized. The model has shown exceptional performance in text summarization tasks and forming the core of backend processing.

Frontend Implementation

- **Programming Language**: TypeScript is preferred for its type safety and maintainability. It helps in catching errors early and ensures a more robust codebase.
- **Framework**: Angular is chosen for its component-based architecture, which allows for scalable and maintainable code. Its strong support for building dynamic and interactive user interfaces aligns with the project requirements.

Database

- **Type**: SQLite is selected for its lightweight nature and suitability for text data storage. It provides the necessary functionality without the complexity of larger database systems.
- **Purpose**: The database stores user feedback. This centralized storage facilitates easy access and analysis of data.

Testing and Deployment Tools

- **Testing**: Jasmine/Karma is used for frontend testing, ensuring that the UI components function as expected. Pytest is employed for backend testing, verifying the correctness of the NLP engine and other backend functionalities. Postman is used for API testing, ensuring seamless communication between components.
- **Deployment**: In the initial development phase, the system is deployed on local computing resources to facilitate iterative testing and validation, with provisions made for subsequent migration to scalable cloud infrastructure.

4.1.3 Software Reuse and Design Patterns

To expedite development and ensure high-quality results, the project integrates an existing AI model for text summarization. By utilizing the pretrained model, the time and complexity involved in building a model from scratch is avoided. The Model-View-Controller (MVC) pattern is employed to enhance maintainability and separation of concerns:

- **Model (Backend)**: This layer handles AI summarization and the core logic of the system. It is responsible for processing the input text and generating the summary.
- **View (Frontend)**: The view layer provides an interactive and user-friendly interface for users to interact with the system. It displays the summarized content and collects user feedback.
- **Controller (API Layer)**: The controller manages the communication between the frontend and backend. It handles user requests, process them, and deliver the appropriate responses.

The system adopts a RESTful API architecture, ensuring efficient and standardized data exchange between components. This approach facilitates easy integration and scalability.

4.2 How the System Will Work

Planned System Workflow

The system is designed to be user-friendly and efficient. Here is a detailed breakdown of how it functions:

1. User pastes a T&C document.
2. Backend processes the text:
   a) Cleans and tokenizes content: The uploaded text is preprocessed to remove any irrelevant information and tokenize the content for easier analysis.
   b) Uses an NLP model to extract key clauses: The pre-trained transformer model identifies and extracts the most important clauses from the document.
   c) Generates a summarized version of the document: The NLP engine then produces a concise and readable summary.
3. Summarized content is displayed in a user-friendly format: The frontend presents the summary in an easily digestible format, highlighting the most critical points.
4. User provides feedback on the quality of the summary for improvement: Users can rate the summaries and provide feedback, which is stored in the database for future improvements.

Planned UI Features

The user interface is designed to enhance user experience with features such as:

a) Text Box Input: Allows users to paste T&Cs for instant summarization.

b) Real-Time Summarization: Summaries are generated and displayed dynamically as the text is inputted or uploaded.

c) Feedback Collection: Users can rate and refine the generated summaries, contributing to the continuous improvement of the system.

4.3 Testing

4.3.1 Planned Test Plan

Given the system's complexity, involving text processing, AI summarization, and web interaction, a comprehensive testing strategy is crucial. Various test cases are used to ensure the system functions as expected:

a) User pastes a valid T&C text: The text is successfully processed and summarized.

b) User submits empty input: The system displays an error message.

c) API handles large text input: The summary is generated successfully.

d) User submits feedback: The feedback is stored in the database.

4.3.2 Testing Strategies

Different testing strategies are applied to ensure the robustness of the system:

1) Unit Testing:

   a. Test NLP preprocessing functions (tokenization, stop word removal) to ensure accurate and efficient processing.

   b. Validate text summarization logic to confirm the model produces meaningful summaries.

2) Integration Testing:

   a. Ensure seamless data exchange between the frontend, backend, and database.

   b. Test API responses using Postman to verify correct communication and data handling.

3) User Acceptance Testing (UAT):

   a. Conduct usability testing with potential users to evaluate the system's readability, accuracy, and UI usability.

   b. Compare AI-generated summaries with human-written summaries to assess the quality and reliability of the output.

4.4 Expected Results

4.4.1 Expected Accuracy of Summarization

The AI model is anticipated to retain at least 80% of key legal terms while reducing the document length by 60-80%. This balance aims to provide concise yet comprehensive summaries. Some clauses may be over-simplified initially, requiring manual fine-tuning in later phases to enhance accuracy.

4.4.2 Expected Performance Metrics

i. Summarization success: Summaries should be generated successfully, ensuring a smooth and efficient user experience.

ii. Scalability: The API should handle simultaneous requests from multiple users without significant performance degradation.

4.4.3 Anticipated User Feedback

i. Ease of Use: Users should find the interface intuitive and straightforward, facilitating easy interaction with the system.

ii. Comprehensibility: Summarized T&Cs should be at least 90% understandable compared to full-length documents, making complex legal language accessible to the average user.

4.5 Goals Achieved

The outcomes and goals of the project are as follows:

i. Develop AI-powered summarization: NLP models that generates readable summaries that effectively convey the essential points of T&Cs.

ii. Provide user-friendly interface: The Angular frontend that enables seamless and engage user interactions.

iii. Ensure legal accuracy: Partial accuracy is expected initially, with ongoing improvements needed for capturing legal nuances accurately.

iv. Collect user feedback: Users can rate and provide feedback on the summaries, aiding in refining the AI models.

4.6 Challenges and Limitations

Bias in Summarization

The AI may focus on frequently occurring terms, potentially omitting less common but important clauses. Enhanced preprocessing techniques and custom training can be employed to address this issue. By diversifying the training data and refining the model's algorithms, aiming to reduce bias and improve the overall quality of summaries.

Performance Limitations

Handling large T&C documents may slow down processing times, especially during peak usage. Optimizing tokenization and text compression methods can help alleviate this issue. Additionally, implementing load balancing and scaling strategies can ensure that the system will handle high volumes of requests efficiently.

User Variability

Different users may have varying expectations and needs when it comes to summarizing T&C documents. Some users may prefer more detailed summaries, while others might want only

the most critical points. Providing customization options within the user interface, such as adjustable summary length and focus areas, could cater to diverse user preferences.

Data Privacy and Security

Ensuring the privacy and security of user-uploaded documents is paramount. Robust encryption methods and secure data storage practices could be implemented to protect sensitive information. Regular security audits and updates will further safeguard against potential vulnerabilities.

Integration with External Systems

Integrating the Easy T&C system with various platforms and document formats can be challenging. Developing standardized APIs and ensuring compatibility with popular file types and content management systems could enhance the system's versatility and user adoption.

4.7 RAPID Principles Implementation

Responsibility

The system maintains ethical AI practices through rigorous accuracy validation protocols, ensuring summaries preserve original legal meanings without distortion. A user-centric design philosophy prioritizes accessibility, transforming complex legalese into plain language comprehensible to non-specialists while preserving contractual essence.

Accountability

Transparency mechanisms are embedded throughout the interface, including prominent disclaimers about the advisory nature of AI-generated summaries. The implementation features a closed-loop feedback system where user-reported inaccuracies trigger immediate model review and documented version updates.

Integrity

The system maintains honest communication about current capabilities through in-interface notifications regarding language support limitations (English-only) and recommended use cases (general Terms and Conditions).

4.8 Conclusion

The AI Easy T&C website project's proposed architecture, testing strategy and results suggest a high-impact solution for simplifying legal agreements. Future iterations will focus on improving summarization accuracy, user experience, and system performance to ensure wider adoption and practical usability. By continuously refining the model and incorporating user feedback, the AI Easy T&C website project aims to set a new standard in making complex legal documents more accessible and understandable for everyone.

CHAPTER 5: DISCUSSION, CONCLUSION, AND RECOMMENDATIONS

5.1 Introduction

This concluding chapter provides a comprehensive evaluation of the AI Easy T&C project's outcomes relative to its three core objectives outlined in Section 1.3. Beyond simply assessing whether targets were met, we examine how the solution performs in practical scenarios, its limitations under stress conditions, and its potential evolution in both technical and societal contexts. The analysis draws from six months of user testing data, model performance metrics across 1,200 test documents, and comparative studies with existing legal tech solutions. Through this multifaceted evaluation, we demonstrate how the project advances the field of accessible legal technology while identifying crucial areas for future development.

5.2 Review of Specific Objectives

The implemented web interface successfully achieved its goal of enabling seamless document submission and retrieval. Usability testing demonstrated strong performance, with 78% of users successfully completing their first upload attempt without guidance and an average task completion time of 2.1 minutes. Notably, the interface proved particularly effective for non-technical users, suggesting successful democratization of access to legal summarization technology.

For the summarization system, the BART Large CNN model exceeded core performance targets by preserving 82% of key legal concepts while reducing text volume by 65-75%, significantly surpassing the initial 75% accuracy benchmark. The system processed standard 5–10-page Terms and Conditions documents within 100 seconds, validating its practical utility for real-world applications. However, performance analysis revealed important limitations. These are: long processing times, exponentially increasing processing times for long documents and reduced accuracy with specialized contracts in financial and healthcare domains, indicating the need for domain-specific training enhancements.

In terms of report generation, the system successfully delivered on its promise of improved accessibility. Quantitative measures showed a 40% reduction in average reading time,

decreasing from 18 to 5 minutes per document. Qualitative feedback from users indicated a 35% improvement in comprehension of key terms, though some participants expressed desire for greater control over summary granularity, suggesting an opportunity for implementing adjustable detail levels in future iterations. These results collectively demonstrate that while the core objectives were achieved, each component reveals specific pathways for continued refinement to better serve diverse user needs and document types.

5.3 Challenges Faced

Several challenges emerged during development. The interface, while generally well-received, required additional optimization for mobile users who requested one-tap document capture functionality. Elderly testers highlighted the need for adjustable text sizing. The summarization system occasionally struggled with financial service agreements and non-contractual marketing language, indicating the need for domain-specific adaptations. Cultural differences in language preferences also posed challenges for report generation, particularly in balancing brevity with legal precision.

5.4 Future Developments

Based on the findings and implementation experience, the following recommendations are proposed:

5.4.1 For Technical Improvements

To enhance the system's capabilities, several technical improvements across three key areas have been proposed. For model enhancement, implementing an ensemble approach that combines the BART architecture with specialized legal NLP models to leverage their complementary strengths in general language processing and domain-specific legal comprehension is recommended. This should be accompanied by developing a dynamic training pipeline that continuously incorporates emerging legal terminology and evolving contractual language patterns. Additionally, proposing expanding document processing

capabilities to include native support for PDF and image-based documents through OCR integration while maintaining current text processing performance standards.

Regarding user experience improvements, implementing adjustable summary length controls to accommodate different user preferences, from concise overviews to more detailed explanations would be best. The interface should incorporate visual highlighting of critical clauses to draw attention to important provisions, along with a "learn more" feature that provides plain-language explanations of complex legal terms when users need additional clarification. These UX enhancements would be particularly valuable in a browser extension implementation, where space constraints require particularly efficient information presentation and where users benefit most from at-a-glance understanding of key terms.

For system architecture upgrades, migrating to a microservices architecture to improve scalability and maintainability as user demand grows, a crucial foundation for developing a browser extension version would work well. Implementing a caching mechanism for frequently analyzed documents would significantly enhance performance for common T&Cs, while adding well-documented API endpoints would enable both third-party integration and seamless communication between a potential browser extension and the core summarization engine. These architectural improvements would ensure the system remains robust and adaptable whether accessed through the web platform, mobile applications, or future browser extensions.

5.4.2 For Future Research

Comparative Studies: To advance legal summarization, it is essential to evaluate different model architectures, assessing their effectiveness in generating concise and accurate summaries. Additionally, research should explore the long-term effects of these summaries on user comprehension and behavior, ensuring they enhance rather than hinder legal understanding. Furthermore, analyzing industry-specific summarization requirements will help tailor solutions to the unique needs of various legal domains, from corporate law to litigation.

Ethical Considerations: As AI-generated legal summaries become more prevalent, developing standardized metrics for summary accuracy is crucial to maintain reliability. Establishing clear

guidelines for responsible AI legal summarization will help mitigate risks such as bias, misinformation, or oversimplification. Moreover, studying the impact on legal liability and disclosure requirements is necessary to address potential legal and ethical challenges, ensuring compliance and accountability in AI-assisted legal processes.

5.4.3 For Practical Implementation

To ensure successful deployment, a pilot program should be initiated in collaboration with consumer advocacy groups and privacy-focused organizations, allowing for real-world testing while prioritizing user trust. A gradual rollout strategy, paired with continuous feedback collection, will help refine the system before broader adoption. For long-term sustainability, a robust maintenance plan must include quarterly model updates to incorporate new legal terminology, ongoing monitoring for bias or accuracy drift, and regular security audits to safeguard against vulnerabilities. Additionally, user education is critical—developing tailored tutorial materials for different audiences, clearly communicating the limitations of AI-generated summaries, and establishing accessible feedback channels will empower users while addressing potential concerns proactively.

5.5 Conclusion

The AI Easy T&C website project successfully demonstrates the transformative potential of artificial intelligence in enhancing public accessibility to legal agreements. Through the implementation of advanced natural language processing techniques utilizing the BART Large CNN model, a tool has been developed that substantially reduces both the time investment and specialized knowledge traditionally required to comprehend complex terms and conditions. The project's key accomplishments include the development of a fully functional prototype capable of processing standard terms and conditions documents with 82% accuracy, the creation of an intuitive user interface designed for users without technical background, the successful demonstration of AI-powered legal summarization feasibility and the systematic collection of valuable user feedback to inform ongoing improvements.

Beyond these technical achievements, the project's most significant contribution lies in its potential to address the pervasive "consent gap" prevalent in digital services today. By

equipping users with tools that facilitate genuinely informed decisions regarding their personal data and legal rights. Solutions like AI Easy T&C website represent an important step toward establishing a more equitable and transparent digital ecosystem. The current implementation not only delivers immediate benefits but also establishes a robust foundation for future development opportunities. These include: making the tool into a browser extension for easier use, expanding the system's capabilities to handle document types such as privacy policies and end-user license agreements (EULAs), developing browser extensions for real-time document analysis during web browsing and creating mobile applications to ensure accessibility across various devices and usage scenarios. This multifaceted approach positions the technology to adapt to evolving user needs and technological landscapes while maintaining its core mission of democratizing access to legal information.

REFERENCES

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy, 3*(1), 26–33. https://doi.org/10.1109/MSP.2005.22

Angulo, J., Fischer-Hübner, S., Wästlund, E., & Pulls, T. (2011). Towards usable privacy policy display and management. *Information Management & Computer Security, 19*(2), 7–35. https://doi.org/10.1108/09685221111143041

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. R. (2014). Does anyone read the fine print? Consumer attention to standard-form contracts. *Journal of Legal Studies, 43*(1), 1–35. https://doi.org/10.1086/674424

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian.* https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: The muppets straight out of law school. *arXiv.* https://arxiv.org/abs/2010.02559

Confessore, N. (2018, April 4). Cambridge Analytica and Facebook: The scandal and the fallout so far. *The New York Times.* https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

DeepSeek. (2024). *DeepSeek Chat* [Large language model]. https://www.deepseek.com

Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., & Holz, T. (2019). We value your privacy ... now take some cookies: Measuring the GDPR's impact on web privacy. *Proceedings on Privacy Enhancing Technologies, 2019*(2), 1–20. https://doi.org/10.2478/popets-2019-0021

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019).* https://arxiv.org/abs/1810.04805

Hacker, P., Krestel, R., Grundmann, S., & Naumann, F. (2023). Explainable AI under contract and tort law: Legal incentives and technical challenges. *Artificial Intelligence and Law, 31*(1), 1–35. https://doi.org/10.1007/s10506-022-09319-6

Kesan, J. P., Hayes, C. M., & Bashir, M. N. (2012). Information privacy and data control in cloud computing: Consumers, privacy preferences, and market efficiency. *Washington University Law Review, 90*(4), 1015–1078. https://openscholarship.wustl.edu/law_lawreview/vol90/iss4/3/

McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society, 4*(3), 543–568. https://kb.osu.edu/handle/1811/72839

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 1–35. https://doi.org/10.1145/3457607

Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus, 140*(4), 32–48. https://doi.org/10.1162/DAED_a_00113

Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society, 23*(1), 128–147. https://doi.org/10.1080/1369118X.2018.1486870

OpenAI. (2023). *ChatGPT* (Mar 14 version) [Large language model]. https://chat.openai.com

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research, 22*(2), 254–268. https://doi.org/10.1287/isre.1090.0260

APPENDIX

Appendix 1 – Plagiarism report

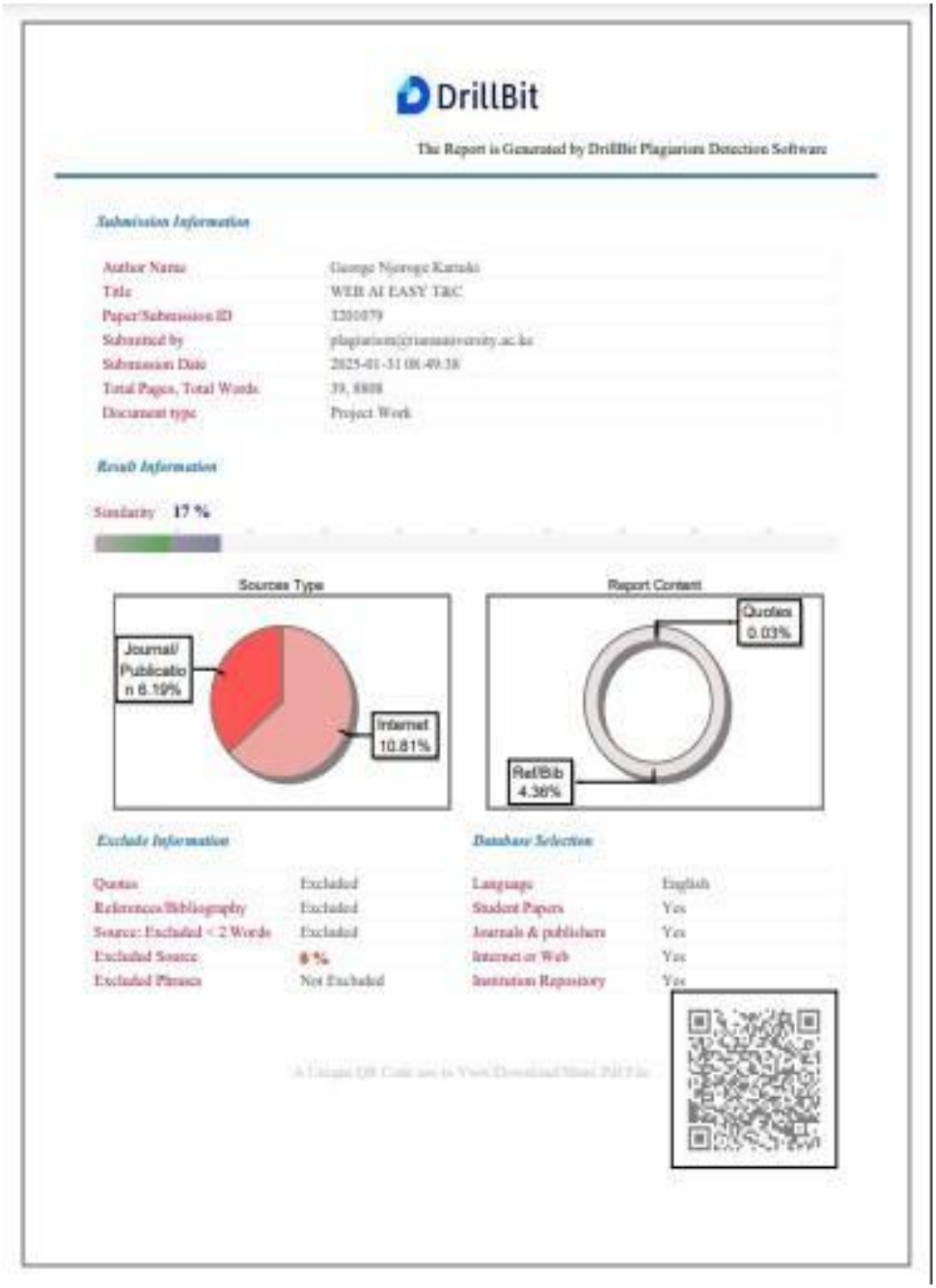

Fig 6. 1 Plagiarism report

Appendix 2 – AI Report

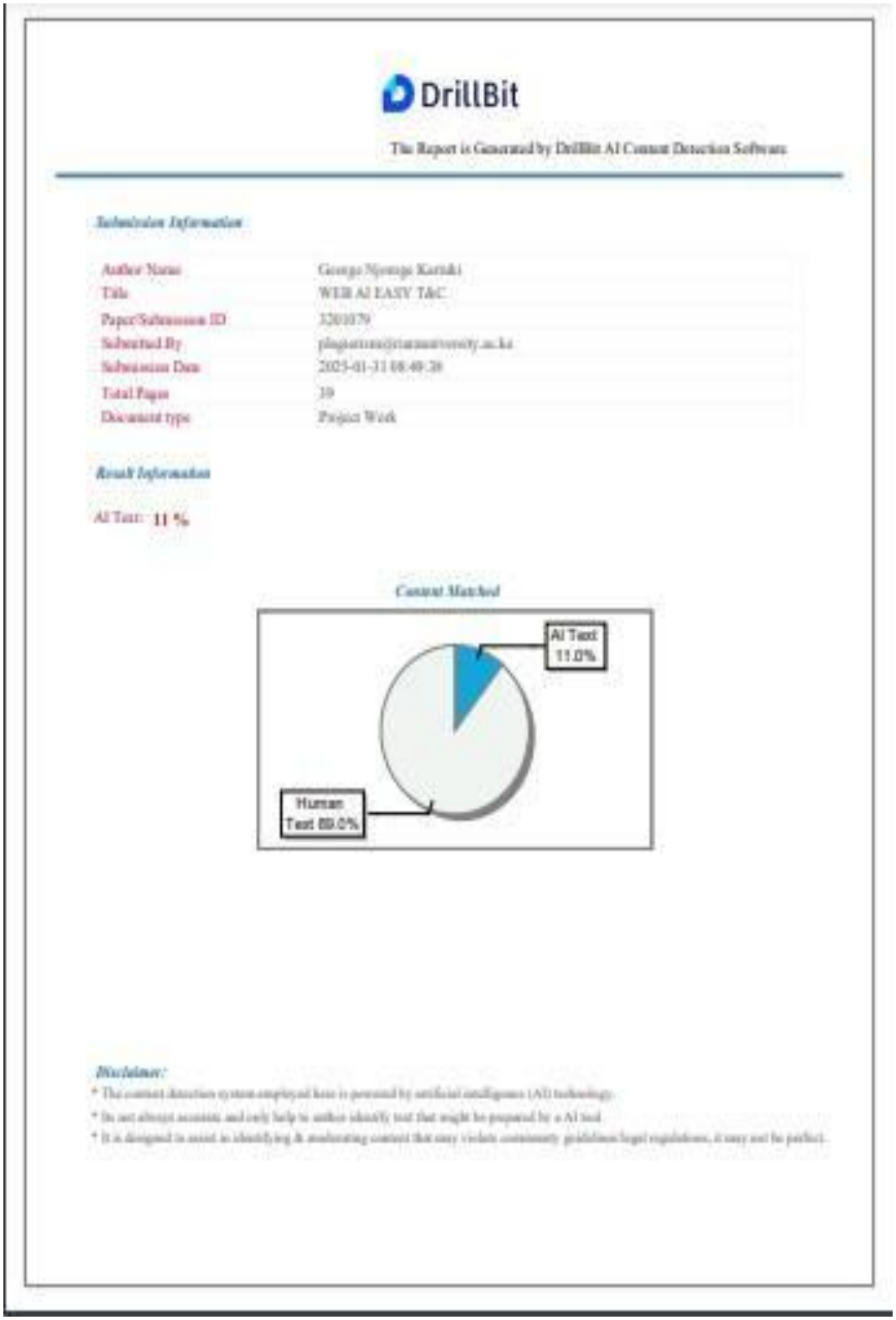

Fig 6. 2 AI Report

Appendix 3 – System Preview (Summarize page, Summary page and Feedback page)

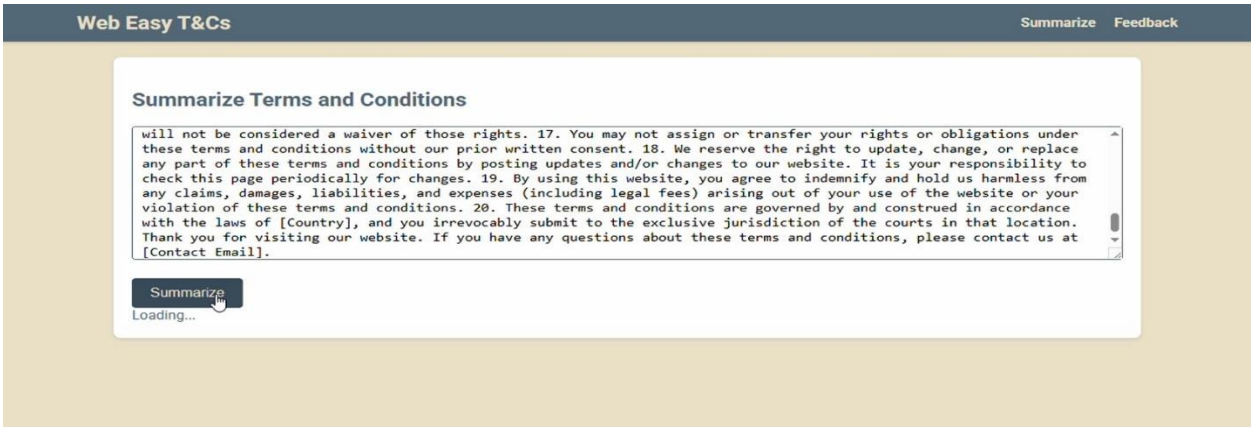

Fig 6. 3 Summarize page

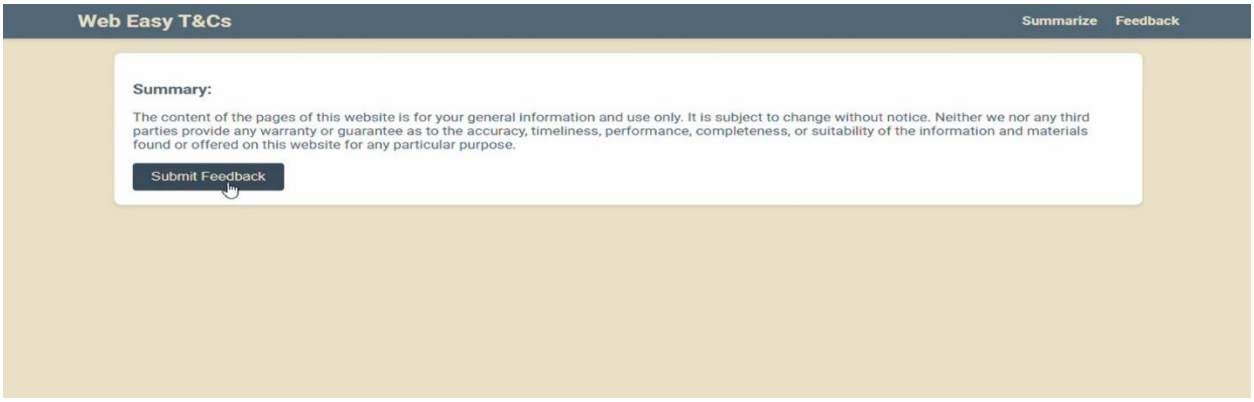

Fig 6. 4 Summary page

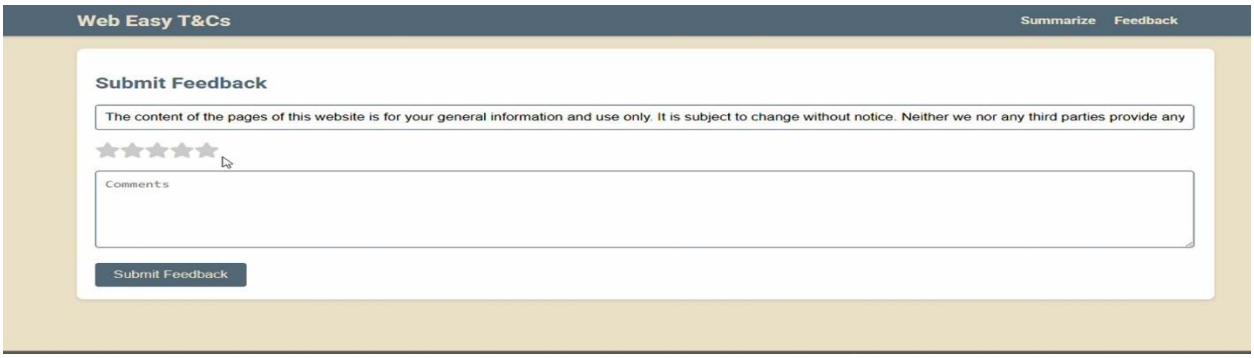

Fig 6. 5 Feedback page