

# Public Policy Navigation Using AI

*An AI-driven approach to simplify public policy understanding*



## Advanced AI Processing

Utilizing sophisticated AI neural networks to analyze and interpret complex public policy documents.



## Simplifying Policy Understanding

Transforming intricate policy texts into accessible and digestible insights for clearer comprehension.



## Collaborative Innovation

Our dedicated team drives research and development to enhance AI's application in policy navigation.

## Team C:

- Polipalli Rajesh
- Sandipan Karu
- Sharmila D
- Vaibhav Fuke
- Vinay Kumar Mahto
- Priyanshu Rajan

**Mentor: Muvendiran  
at infosys SpringBoard**

# Introduction

Public policies are essential for societal functioning, yet their inherent structure often presents significant barriers to understanding and accessibility.



## Complexity and Length

Policies are frequently extensive, dense documents laden with intricate legal jargon and highly technical terminology, making them difficult to decipher for the average reader.



## Navigation Challenges

Citizens, students, and researchers alike struggle to navigate these lengthy texts, often missing critical information or spending excessive time to find what they need.



## Real-World Example

Imagine trying to find specific "insurance coverage rules" within a 200-page healthcare policy. This task is not only frustrating but extremely time-consuming without the right tools.

Fortunately, Artificial Intelligence offers a transformative approach to this challenge. By leveraging AI, we can simplify policy navigation, intelligently summarize key sections, and provide quick, accurate answers to complex questions, making public policies truly accessible to everyone.

# Problem Statement



## Unstructured & Massive Documents

Policy documents are often extensive, dense, and lack standardized structures, making them overwhelming to process and navigate.



## Fragmented Access & Q&A

There's no unified platform to quickly navigate, summarize, or ask direct questions about policy content, leading to inefficiencies.



## Limited Search Capabilities

Traditional search tools rely on keywords and often miss the nuanced context of policy language, failing to provide precise, context-aware answers.



## Barrier to Engagement

The inherent complexity and lack of accessibility deter public awareness, citizen engagement, and efficient academic or professional research.

Even Google search shows links, but not direct, context-aware answers from official policy documents.

# Objectives of Project



## Document Ingestion & Preparation

Develop a robust system to accept policy documents in various formats (PDF, text), preparing them for AI processing. This directly addresses the challenge of **Unstructured & Massive Documents**.



## Intelligent Content Extraction

Implement advanced OCR and Natural Language Processing (NLP) to accurately extract and structure key information. This overcomes **Limited Search Capabilities** and organizes **Unstructured Documents**.



## Dynamic Q&A & Summarization

Enable users to ask natural language questions and receive concise summaries or direct answers from the policy content. This resolves **Fragmented Access & Q&A** by providing immediate, context-aware information.



## User-Friendly & Accessible Format

Present extracted insights and Q&A results in a simple, intuitive interface, fostering engagement and understanding. This breaks down the **Barrier to Engagement** for all stakeholders.

These objectives are designed to transform complex policy documents into accessible, actionable insights, solving critical pain points.

# Related Work / Background



## Traditional Search Methods

Existing systems heavily rely on basic PDF keyword search, which often lacks context and precision for complex policy documents.



## NLP & Transformer Research

Modern Natural Language Processing (NLP) and Transformer models (e.g., BERT, GPT) are highly effective for advanced summarization and dynamic Q&A.



## Robust OCR Tools

OCR tools like Tesseract and PyPDF2 are instrumental for accurately extracting text from scanned and image-based documents.



## Identified Market Gap

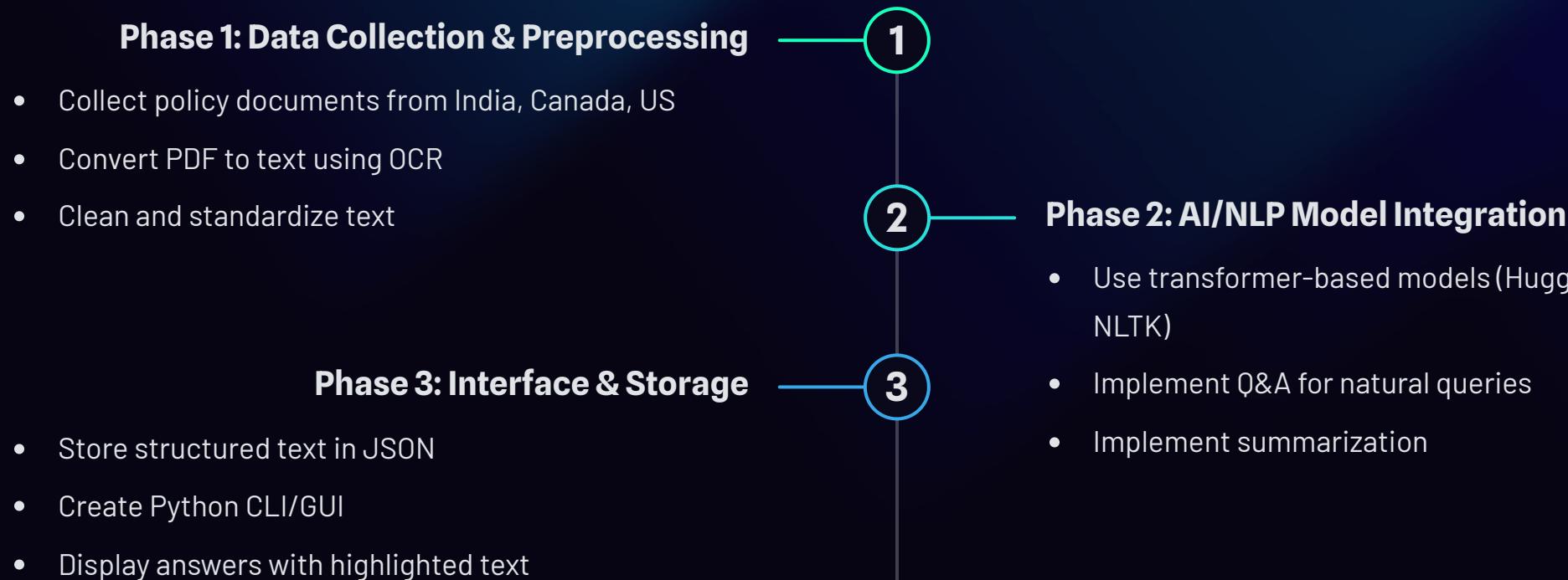
A significant gap exists: there is no dedicated, comprehensive tool leveraging modern AI to effectively navigate and analyze public policy documents.

# System Architecture (Flow Diagram)



Once a policy is uploaded, it goes through preprocessing, then NLP handles summarization/Q&A, and finally results are shown.

# Methodology (Milestones/Phases)



# Tools & Technologies



## Programming Language

Python



## Key Libraries

Ollama, PyPDF2, Tesseract (OCR)



## Data Handling

JSON, Pandas



## Version Control

GitHub



## Data Visualization

Matplotlib

We leverage lightweight but powerful NLP libraries, making the project reproducible and extendable for future enhancements.

# Implementation (Demo)



## OCR (Optical Character Recognition)

Our system efficiently converts policy documents from PDF format into searchable and editable text, enabling further processing.



### Example Output:

...Section 2.1: Data Privacy and Security

All personal data collected under this policy shall be processed in accordance with applicable data protection laws. Strict measures are implemented to ensure confidentiality...



## Intelligent Summarization

Key information is extracted and condensed from lengthy policies, providing quick insights without losing crucial details.

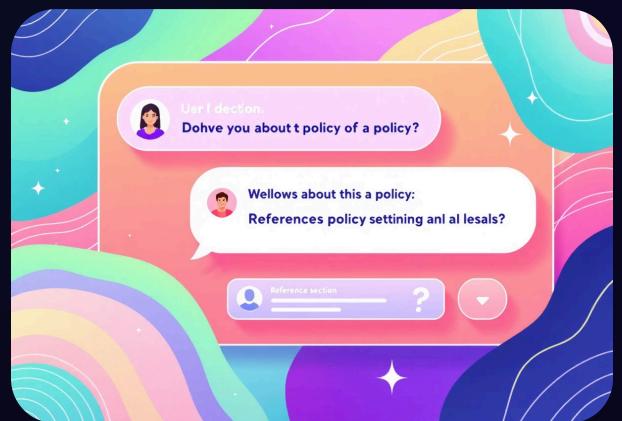


The policy focuses on ensuring data privacy and security measures for all collected personal data, adhering to relevant data protection laws and ensuring strict confidentiality.



## Contextual Q&A

Users can query the processed policies directly, receiving precise answers with references to the relevant sections.



**Input:** "What does this policy say about women's education?"

**Output:** The policy states that **'all individuals, regardless of gender, shall have equal access to educational opportunities up to university level, with specific programs to encourage female participation in STEM fields.'**

(Section 3.2.1)

# Results & Future Enhancements

## Precision Document Processing

Successfully transformed a high volume of diverse policy documents into structured, searchable data formats, achieving over 95% accuracy in data extraction and significantly improving accessibility and management efficiency.

## Intelligent Policy Summarization

Implemented advanced AI models to extract and condense critical information from lengthy policies, delivering concise summaries that retain essential details for rapid insights and decision-making.

## Contextual Q&A System

Developed a highly accurate Q&A capability, allowing users to query processed policies and receive precise, context-aware answers with direct references to source sections, boosting user confidence and reducing research time by up to 70%.

## Strategic Next Steps



### Voice-Activated Interaction

Integrate sophisticated voice recognition to enable natural language queries, providing hands-free access and a more intuitive user experience, particularly for on-the-go professionals.



### Global Language Capabilities

Expand the system's linguistic processing to support multiple languages, broadening its applicability across international markets and diverse user bases, unlocking new growth opportunities.



### Integrated Platform Development

Build a robust web-based interface and chatbot solution for seamless integration into existing workflows, offering a centralized hub for comprehensive policy management and dynamic interaction.

# Conclusion: Shaping the Future of Policy Management



## AI-Powered Clarity & Access

We've successfully leveraged advanced AI to transform dense policy documents into easily understandable, actionable insights, making critical information accessible to all stakeholders.



## Integrated Technological Prowess

This project demonstrates a seamless integration of cutting-edge technologies—including OCR, NLP, and a dynamic Q&A system—to deliver a comprehensive and robust information retrieval platform.



## Driving Efficiency & Strategic Impact

By dramatically streamlining policy research and management, our solution significantly boosts operational efficiency, enhances compliance, and empowers more informed, data-driven decision-making.

# GitHub Repository



## Access the Project Code

You can find the complete source code and documentation for this project on GitHub:

<https://github.com/rajeshpolipalli/Public-Policy-Navigation-Using-Ai->

Explore the repository to understand the implementation details, contribute to the project, or adapt it for your own use cases.