GenAI Hackathon T$O - 2025 Submission

Team : Meme Machine
- Tanav Kolar
- Shreyansh Kumar Nayak

Problem Statement: Can AI be the Bridge Between Citizens and Government ?

Boudhi: AI-powered Government Scheme Finder & Screener

---

1. Problem Statement

Access to Government Welfare Schemes in India: A Fragmented Reality

India has over **10,000+ government welfare schemes** distributed across central, state, and local governments. These span sectors such as agriculture, healthcare, education, employment, housing, and more. However, despite the abundance of schemes, their **awareness, discoverability, and accessibility remain extremely low**—especially among rural populations, marginalized communities, and the digitally illiterate.

Key Challenges

1. **Fragmented Information Sources**

   ○ Scheme details are spread across different departmental websites with varying formats, outdated portals, and broken links.

2. **Low Awareness & Digital Literacy**

   ○ Target beneficiaries often do not know which schemes exist, let alone which they are eligible for. Many users cannot navigate websites or understand legal/technical jargon.

3. **Language & Accessibility Barriers**

   ○ Government portals are primarily in English or Hindi, making them inaccessible to large parts of the population who speak regional languages.

4. **Lack of Personalization**

   ○ No existing platform offers personalized scheme discovery based on a user's life situation (e.g., widow, farmer, student from SC/ST).

5. **Redundant Human Dependency**

   ○ Citizens often rely on intermediaries, cyber cafes, or field officers to apply for schemes, which introduces delays, misinformation, and sometimes corruption.

---

2. Our Approach and Solution

Vision

To build an **AI-powered, multilingual, accessible, and personalized scheme discovery and screening assistant** that empowers every citizen to **understand, discover, and apply** for schemes relevant to them — using just **natural language**.

Solution Architecture: Boudhi Chatbot

◆ **Core Components**

1. **Semantic Scheme Retrieval (Vector Search)**

   ○ Scraped and structured 1000 government schemes from myscheme.gov.in.
   ○ Embedded the full context of each scheme (name, benefits, eligibility, application process, etc.) using `intfloat/e5-large` model.

   ○ Indexed all schemes into **Pinecone** — enabling **semantic similarity search** on user queries.

2. **Multimodal User Interface**

   ○ **Voice input** support via custom ASR module.

- ○ **Text input** with auto-translation for multilingual access.

- ○ UI powered by **Streamlit** for fast deployment.

3. **Scheme Screener**

- ○ A guided eligibility check tool that asks the user relevant questions.

- ○ Filters out schemes using pre-coded eligibility rules (age, caste, gender, income, etc.).

4. **Text-to-Speech (TTS)**

- ○ Converts results into voice in the user's native language using gTTS.

5. **LLM-Powered Answers**

- ○ Groq-hosted LLMs (Mistral/LLaMA) take top-matching schemes and generate human-like responses to user queries.

Data Pipeline Overview

- ● **Step 1**: Scrape schemes from India.gov.in and state portals.

- ● **Step 2**: Normalize the data into structured JSON — with fields like eligibility, benefits, documents, etc.

- ● **Step 3**: Embed scheme paragraphs using `e5-large`.

- ● **Step 4**: Upsert scheme vectors + metadata into Pinecone.

- ● **Step 5**: During inference, embed query → retrieve top N schemes → pass to LLM.

Example Use Case

User: "I am a widow with two daughters. Are there any educational scholarships or pension schemes I can apply for?"

- Boudhi transcribes voice → translates to English (if needed)

- Embeds query and performs vector similarity search

- Retrieves top schemes like *Indira Gandhi National Widow Pension Scheme* or *Post-Matric Scholarship for Girls*

- Uses LLM to summarize and present the best options

- Optionally screens user further to confirm eligibility

---

3. Future Improvements

Knowledge Graph Integration

- Build a **Neo4j-based graph database** for schemes, categories, beneficiaries, and relationships.

- Enables better linking between central/state schemes, benefit dependencies, and duplicate overlaps.

End-to-End Application Flow

- Allow users to **apply for schemes directly through the chatbot**.

- Integrate with **DigiLocker**, **Aadhaar**, and **eSign** for document uploads and verification.

Regional Dialect and Offline Support

- Train **custom ASR models for dialects** (e.g., Bundeli, Maithili).

- Support **offline mode** via lightweight models and local embeddings for rural kiosks.

Integration with Citizen Service Hubs

- Enable Boudhi in **Common Service Centres (CSCs)**, **Panchayat Bhawans**, and **Jan Suvidha Kendras**.

Complaint and Grievance Redressal

- Integrate CPGRAMS or state grievance portals.

- Help users lodge complaints regarding non-disbursement or application delays.

Impact Potential

- Rural citizens discover schemes independently, reducing dependency on intermediaries.

- Students, widows, farmers, and elderly can find financial aid easily.

- Multilingual and voice-first design reaches the underserved and low-literacy population.

- Personalized eligibility screening prevents information overload.

👥 Team

- **Tanav Kolar**
- **Shreyansh Kumar Nayak**