

# Add AI Features & Logic

Team Name: Team Invictus      Chatbot Name: Beacon AI

**Implementation Note:** The entire chatbot system was hard-coded end-to-end, including intent handling, conversation branching, validation rules, and routing logic, to demonstrate full control over chatbot behavior.

Beacon AI is an anonymous, AI-powered chatbot designed to help citizens safely report corruption. Existing complaint portals are often complex, unsafe, and opaque, discouraging citizens from reporting incidents. Beacon AI addresses this gap by offering a simple, conversational, and privacy-first reporting experience.

The chatbot focuses on structured data capture, emotional sensitivity, and lawful escalation. It is designed to guide users step by step, without judgment, while ensuring anonymity and trust throughout the process.

## 1. INTENT & ENTITY DESIGN

The intent design of Beacon AI is centered on reducing user fear while ensuring accurate and complete information capture. Instead of form-heavy interfaces, the chatbot uses conversational prompts to gradually gather details.

### Intent 1: Report Corruption

Triggered when a user wants to report bribery, misuse of authority, or unethical conduct. Sample training phrases include: *I want to report corruption, An officer asked for a bribe, Someone demanded extra money.*

### Intent 2: Seek / Provide Incident Details

Activated when users provide contextual information such as location, department, timing, or supporting evidence. Sample phrases include: *It happened yesterday in Ludhiana, Passport office asked for Rs. 500, Municipal office delay.*

### Intent 3: Track Complaint

Allows users to track complaint progress using an anonymous case ID. Sample phrases include: *Track my complaint, Check my case status, What happened to my report?*

#### Entities Defined:

- Location** (city, state, or office name) extracted using NLP entity recognition;
- Authority / Department** (passport office, municipality, RTO) identified via contextual classification;
- Evidence Type** (photo, receipt, document) detected from user input or uploads.

## 2. AI / NLP FEATURES IMPLEMENTED

Beacon AI integrates multiple AI and NLP capabilities to ensure accurate understanding, emotional sensitivity, and reliable information processing across all conversations.

| AI / NLP Feature              | How It Improves the Chatbot  |
|-------------------------------|--|
| Intent Recognition            | Identifies whether the user wants to report corruption, provide details, or track a case |
| Entity Extraction             | Automatically extracts location, department, and incident type from free-text input      |
| Sentiment Analysis            | Detects fear or anxiety and adjusts responses to be calm, supportive, and reassuring     |
| FAQ Integration               | Answers common questions related to anonymity, safety, and reporting process             |
| Conditional Logic & Branching | Dynamically adapts follow-up questions based on use responses and comfort level          |

These features ensure that Beacon AI remains context-aware, trustworthy, and accessible, even for users unfamiliar with legal or bureaucratic procedures.

## 3. ADVANCED ENHANCEMENTS

### PII Redaction (Auto-Anonymization):

Personal identifiers such as names, phone numbers, and addresses are automatically removed before any report is stored or routed, ensuring user safety and privacy.

### Credibility Scoring:

Each report is assigned a credibility score based on completeness, consistency, and availability of supporting evidence, helping prioritize genuine cases.

### Pattern-Based Corruption Detection:

Using clustering and pattern analysis, Beacon AI identifies repeated mentions of the same offices or departments across multiple anonymous reports to surface systemic corruption trends.

### Context-Aware Follow-up Questioning:

The chatbot dynamically adjusts follow-up questions based on sentiment, completeness of details, and presence of evidence, improving data quality without overwhelming the user.

### Retrieval-Augmented Generation (RAG):

Beacon AI references verified legal and policy sources such as the Prevention of Corruption Act and whistleblower protection frameworks to ensure factual, non-hallucinatory responses.

### GPT-based NLP Engine:

A GPT-based language model enables natural conversation, contextual follow-ups, and safe handling of sensitive user inputs.

## 4. TESTING AI FEATURES

Beacon AI was tested using multiple real-world inspired queries to validate intent detection, entity extraction, sentiment handling, and routing accuracy.

| Test Query                      | Detected Intent / Entity | Result  |
|---------------------------------|--------------------------|---------|
| I want to report a bribe        | Report Corruption        | SUCCESS |
| Passport office asked for money | Report + Department      | SUCCESS |
| Yesterday in Ludhiana           | Location Entity          | SUCCESS |
| I am scared to complain         | Sentiment Detected       | SUCCESS |
| I have a photo receipt          | Evidence Entity          | SUCCESS |
| Track my complaint              | Track Complaint          | SUCCESS |
| Municipal office corruption     | Report Corruption        | SUCCESS |
| Will my name be revealed?       | FAQ Intent               | SUCCESS |
| Where will this be sent?        | FAQ Intent               | SUCCESS |
| Check case ID 4582              | Track Complaint          | SUCCESS |

**Refinements Made:** Based on testing outcomes, emotional responses were improved, follow-up questions were simplified, and department detection accuracy was enhanced.

## CONCLUSION

Beacon AI demonstrates the responsible application of AI to empower citizens, reduce fear around corruption reporting, and transform isolated complaints into structured, verifiable, and actionable insights while maintaining privacy and trust.

[Back to Safety](#)



#### SECURE CONNECTION

ENCRYPTED

SESSION: 6564fca...

Hello. I am Beacon AI, your compassionate and anonymous assistant. I'm here to help you report corruption safely. Please know that your identity is fully protected, and I will guide you through this process one step at a time. How can I assist you today?

6:08 PM

I was charged higher than the MRP

6:29 PM

I'm so sorry to hear that you were overcharged. That can be really frustrating and unfair. Could you walk me through what happened from start to finish?

11:59 AM

Type your message securely...

END-TO-END ENCRYPTED & ANONYMOUS



# Beacon AI

The government-grade, anonymous corruption reporting system.  
Powered by advanced AI to protect your identity and ensure justice.



#### 100% Anonymous

No personal data collected. Your identity remains hidden forever.



#### End-to-End Encrypted

Military-grade encryption ensures your evidence is secure.



#### AI Sanitization

Automated PII redaction removes traces before human review.

[Start Secure Report](#) →

[Check Status](#)

