## **Overview**

This is an end-to-end, multi-agent, multi-modal AI system designed to automate and streamline the reporting, analysis, and resolution of real estate infrastructure issues. Built with scalability, professionalism, and user experience in mind, this assistant takes natural user input (via text, voice, or image), processes the information contextually, and delivers a structured report to both the user and relevant authorities. It features memory architecture, visual and language-based understanding, auto-documentation, and intelligent communication pipelines.

The system is composed of intelligent agents for dedicated tasks, a meta-controller for fallback and rerouting, and a modular backend compatible with microservices. It integrates cutting-edge AI models for language, vision, and speech processing, and incorporates automation pipelines for document generation and email dispatch.

## **Feature Summary (At a Glance)**

* **Multi-Agent System** with Meta-Agent fallback and routing
* **Multi-Modal Input**: Voice, Text, Image inputs with seamless switching
* **Dual Memory Architecture**:
  + Short-Term: Chat history for ongoing context
  + Long-Term: Case history logs for persistent memory
* **Emotion-Aware Dialogue**: Adjusts responses based on sentiment
* **Visual Labeling & Captioning**: Context-aware captioning + issue labeling with image overlap handling
* **Structured PDF Summary Generation** from conversations and media
* **Automated Authority Email Dispatch** with formal formatting
* **Location-Based Responses** with auto-detection of user region
* **Automated Contractor Suggestions** based on extracted location and issue type
* **Meta-Agent Control**: Automatic agent switching mid-conversation
* **Audio/Text Output** responses for accessibility
* **Security-Focused Troubleshooting Recommendations**
* **Proactive Alert System (Coming Soon)**: Scheduled follow-ups/reminders

## **Intelligent Components & Functional Breakdown**

### **1. Multi-Agent System**

* **Agent 1**: Domain-specific expert (real estate, infrastructure)
* **Agent 2**: General-purpose Legal assistant
* **Meta-Agent**: Detects query type, switches between agents mid-conversation or during task transitions

### **2. Input Modalities**

* **Text/Voice** → Handled by Agent 2 (transcribed via Whisper)
* **Image + Voice/Text** → Handled by Agent 1 (captioned + semantically processed)
* **Audio Output**: Optional speech responses generated for accessibility (Freezed in Beta)

### **3. Dual Memory Structure**

* **Short-Term Memory**: Maintains conversational context per session
* **Long-Term Memory**: Case log system to retrieve historical complaints/resolutions

## **🧾 Document Generation & Reporting**

### **Structured PDF Summary**

* Extracts issues, agent responses, technical recommendations, and resource links
* Clean format: no special characters or chatbot mentions
* Used for:
  + User-facing documentation
  + Authority-level dispatch

### **Issue Labeling & Visual Understanding**

* Context-preserving caption generation using overlapped-image segmentation
* Visual cues and inferred issue types from image metadata

## **Communication Pipeline**

* **Authority Report Emailing**:
  + Email sent with only structured content, Professional writeup, Profile based sendoff. (i.e - Regards User\_name, )
* **User Email Summary**:  
  Includes detailed document, issue discussion, recommendations
* **Proactive Alert System**:
  + Scheduled reminders (not yet active, in roadmap)

## **Contractor & Help Automation**

* **Contractor Suggestions**:
  + Based on geolocation and problem type
  + Contact info, relevant services, and reference links provided
* **Troubleshooting Recommendations**:
  + Includes security features (e.g., waterproofing for leaks)
  + Preventive tips, structural inspection guidance

## **Models & Technologies Used**

| **Component** | **Technology / Model Used** |
| --- | --- |
| Language Reasoning | Groq LLaMA3-8B |
| Image Captioning | Vision-Language Transformer (contextual, overlapping) |
| Voice Transcription | OpenAI Whisper |
| PDF Report Generation | FPDF |
| Email Automation | smtplib + secure SMTP (Gmail) |
| UI Frontend | Gradio |
| Agent Control Logic | Python-based Meta-Agent system |
| Memory Implementation | In-memory storage + external logging |
| Prompt Engineering | Dynamic template switching for all tasks |

## **📦 Unique Selling Points**

* Designed for both technical and non-technical users
* Fully modular: ready for containerization and microservice deployment
* Professional-level reporting with enterprise-ready automation
* Real-time decision switching between agents
* Easily extendable to different domains (e.g., healthcare, utilities, smart cities)
* Deployable as a standalone assistant or embedded module in larger platforms

## **📎 Project Links**

* **Loom Video Demo**: [Insert Loom Link Here]
* **GitHub Repository**: [Insert GitHub Repository Link Here]