An IoT-enabled Smart Office via Autonomous LLM Agents

Abstract

Modern smart office environments demand adaptive and intelligent systems to seamlessly manage an expanding network of interconnected IoT devices and user interactions. However, traditional rule-based automation frameworks lack the contextual understanding and flexibility to operate in dynamic, multi-device office ecosystems. This is further complicated by ambiguous device references, heterogeneous APIs, and inflexible automation routines, all of which limit the scalability and usability of current solutions.

To address these limitations, we propose an **IoT-enabled smart office framework powered by autonomous Large Language Model (LLM) agents**. This system leverages multimodal inputs—including natural language and images—to enable intuitive user interaction and dynamic device control. A suite of specialized tools is integrated into the agent, each addressing a specific challenge in the smart office context:

- **Device Disambiguation Tool** identifies the correct device when user commands are vague or multiple devices match the description (e.g., "Turn on the big screen").
- **API Tool** retrieves real-time data from web services, APIs, or office-specific knowledge bases such as schedules, weather, or stock updates.
- **Code Interpreter Tool** performs logic execution, calculation, and state reasoning based on system context or user queries.
- **Code Generation Tool** dynamically creates executable code to automate tasks or establish personalized office routines (e.g., auto-launching a daily stand-up meeting).
- **Condition Monitoring Tool** allows the agent to observe specific environmental or device states (e.g., "Notify me when the projector is off"), supporting event-triggered actions.
- **Device Interaction Tool** handles device communication with internal subtools such as:
    - **Device Planner**, which breaks user intent into actionable device commands.
    - **Device Command Executor**, which performs low-level device control actions like switching on a display or opening a presentation.
    - **Device State Tracker**, which monitors device status to support real-time responsiveness.

These tools work collaboratively to resolve vague user instructions, plan multi-step actions, interact with both virtual and physical systems, and generate automation logic on the fly. By autonomously adapting to user intent and environmental context, our framework promises enhanced usability, flexibility, and efficiency in real-world smart office scenarios.

System Model

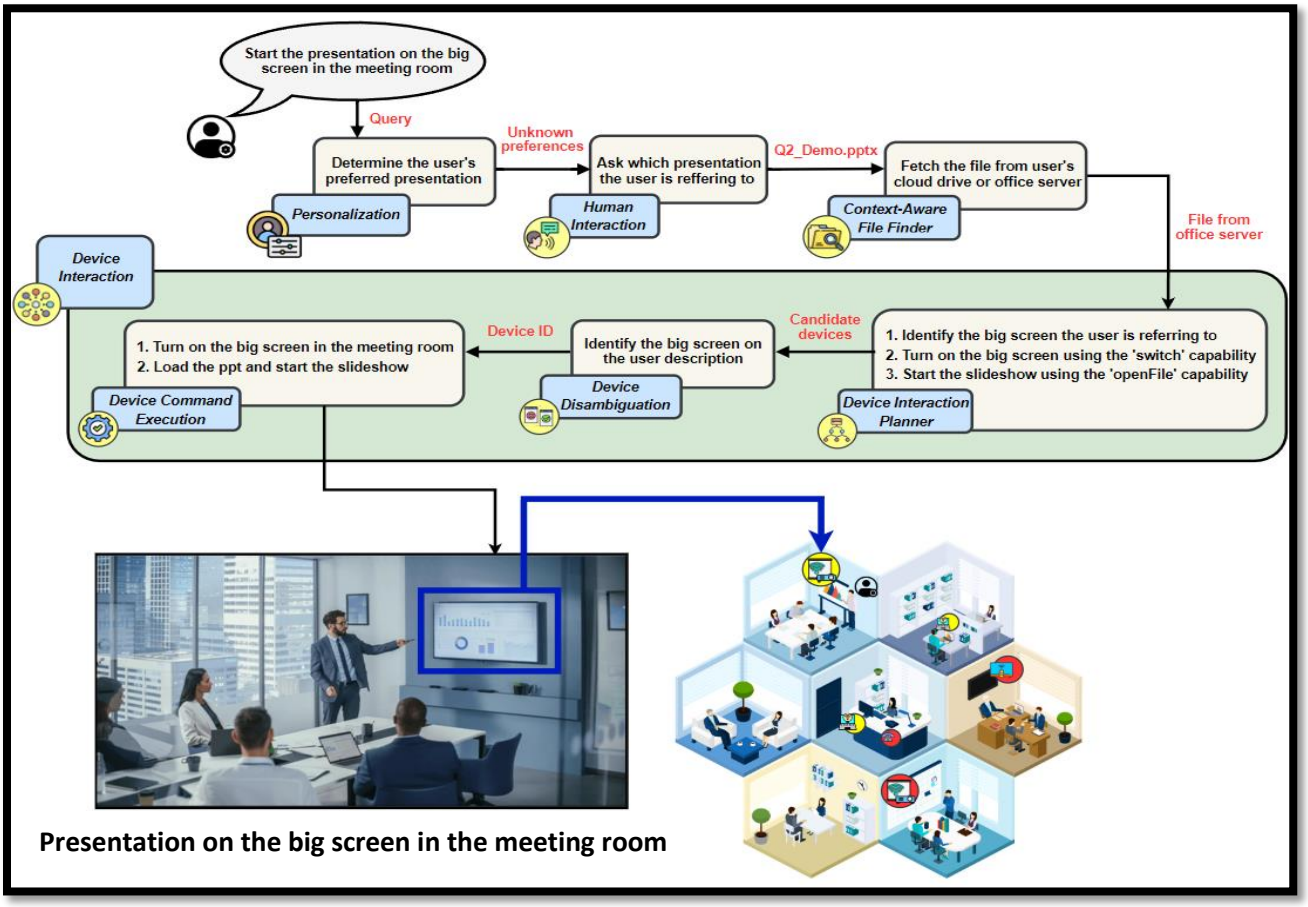

Figure 1: Demonstration of An IoT-Enabled Smart Office via Autonomous LLM Agents

The proposed system operates through a dynamically constructed sequence of LLM-driven prompts, where each user request initiates a structured reasoning loop. This process allows the agent to iteratively decide the next best action, verify the success of each step, interact with external tools or the user when clarification is needed, and determine when the task has been completed or requires adaptation.