





Next-gen intelligent agent support using **Agentic Al**



LEMEN

- Ton That Huu Luan
- Nguyen Vinh Gia Bao
- Le Viet Bao
- To Bao Minh Hoang
- Nguyen Huu Quoc Huy

ACADEMIC SUPERVISOR

■ Dr. Ginel Dorleon

INDUSTRY SUPERVISORS

- Mr. Phi Huynh Ngoc
- Mr. Thien Nguyen Le Quynh

Project Background

The central aim of this project is to develop an enhanced iteration of BonBon by incorporating intelligent conversation-routing mechanisms that transfer interactions to human specialists when the AI system reaches its operational limits. This enhancement seeks to address persistent shortcomings in contemporary AI-driven support systems, particularly their inability to manage highly complex queries and their restricted access to real-time information.

To achieve this, the upgraded system adopts a multi-agent architecture designed to facilitate comprehensive query resolution and seamless collaboration with human agents. Through the principles of multi-agent systems (MAS), the framework fosters intelligent communication and coordination between agents, enabling AI algorithms to effectively process data, generate insights, and support informed decision-making. The architecture is underpinned by tools hosted on a dedicated MCP server, ensuring advanced interoperability and streamlined integration among diverse AI agents.

Key Features

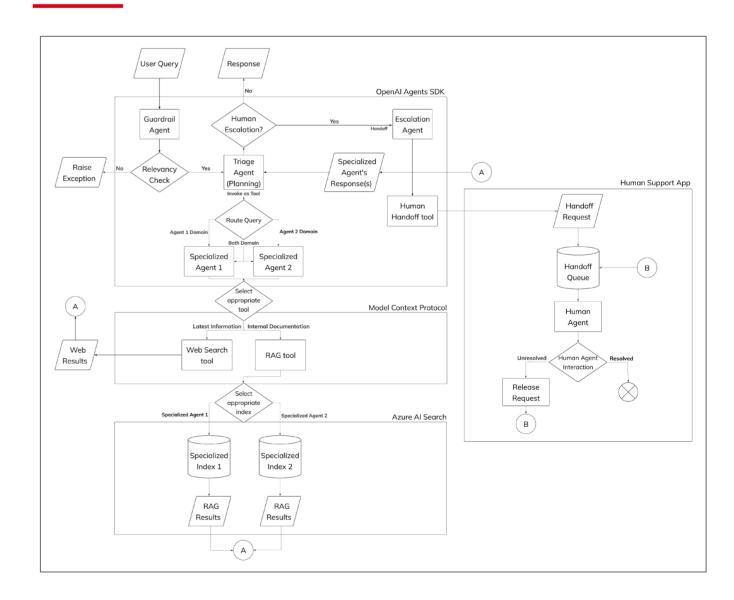
- Knowledge-based responses: The specialized agents retrieve relevant information in their domain based on the user's query and generate the response. The triage agent uses the specialized agents' response to reply to the user.
- Query with uploaded documents/images: User can send a query along with their documents/images, and the AI agent responds in accordance with the information extracted from the attached files.
- External tools for AI agents: The specialized agents use tools from an external MCP server, such as web searching, knowledge-based searching, and a loan calculator, to expand their capability in delivering detailed and accurate responses to the user.
- Escalation to human support: If the Triage determines that the specialized agent's response is insufficient, it transfers the conversation to the appropriate department, where a human agent takes control.

Motivation

While chatbots perform well with routine queries, they often struggle with complex or atypical cases. Without effective escalation, users risk frustrating interaction loops, an issue reflected in findings that 63% of users abandon a service after a single poor chatbot experience.

The project's vision is beyond financial services toward a scalable, cross-sector architecture for agentic Al. Lessons from the financial domain informed the development of a more versatile framework that supports autonomous, adaptive, and modular agents capable of managing complex workflows and responding dynamically to evolving standards and user needs.

Al Flow Diagram



System Architecture

