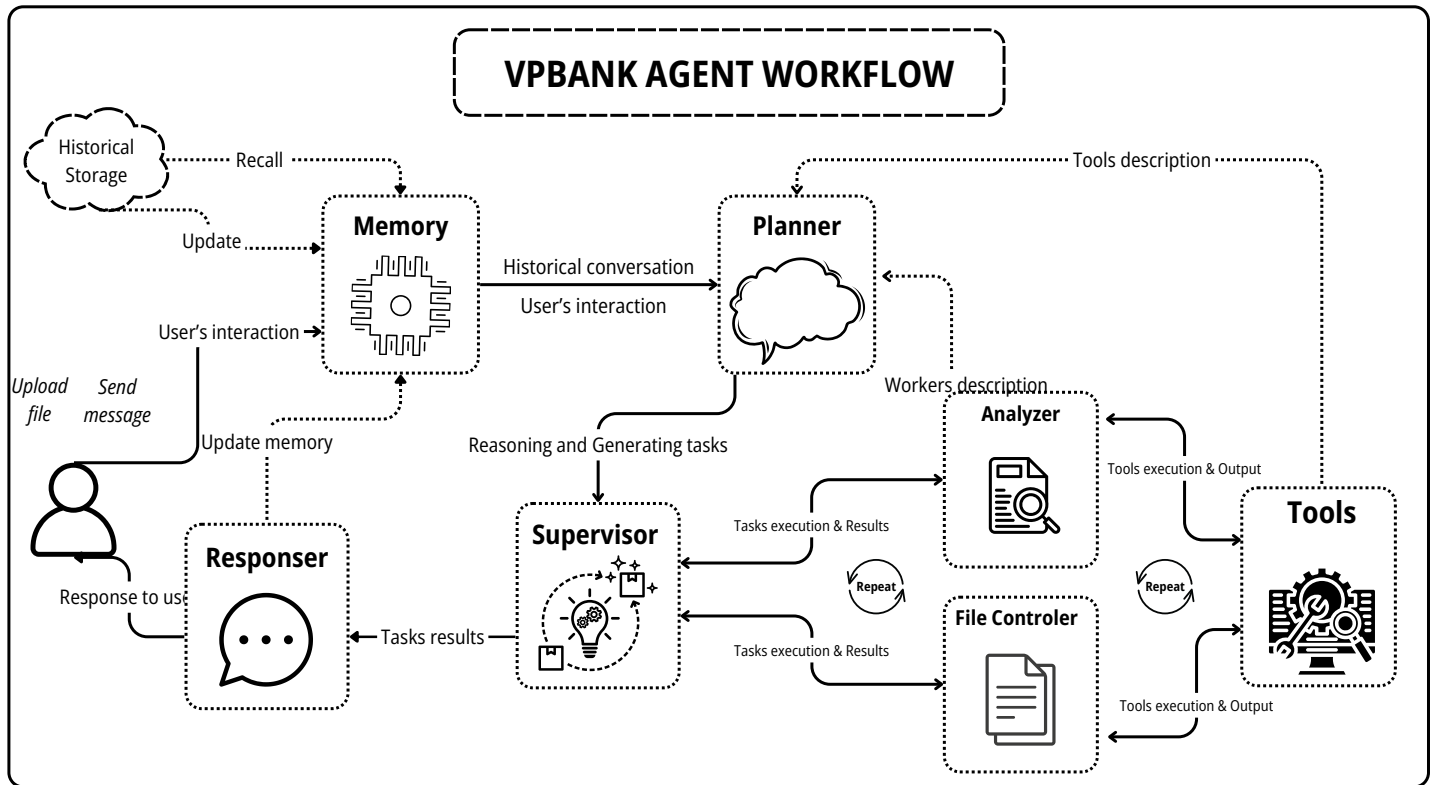


# Ứng dụng GenAI hỗ trợ phân tích báo cáo tài chính

## Tổng quan

Ứng dụng được thiết kế để tận dụng Generative AI (GenAI) và các khả năng AI tiên tiến như memory management, planning, tool utilization, and action execution hỗ trợ người dùng tương tác, phân tích chi tiết báo cáo tài chính, hoàn thành tờ trình tín dụng, ...



## Kiến trúc

### Các thành phần chính

#### 1. **Memory:**

- Lưu trữ lịch sử context và conversations.
- Cho phép agent ghi nhớ các hành động trước đó của người dùng.
- Lưu trữ LOCAL FILES dạng JSON theo user\_id, session\_id

#### 2. **Planning:**

- Reasoning: phân tích interaction của user, lên kế hoạch step-by-step.
- Kế hoạch bao gồm các hành động cần thực hiện, các tool cần sử dụng.

#### 3. **Tools:**

- Gồm các tools như:
  - Process file: Extract text content từ tài liệu gốc.
  - Write file: Lưu file đã xử lý vào hệ thống.
  - Read file: Đọc nội dung file đã xử lý.
  - Get files: Lấy danh sách các files người dùng đã upload.
  - ...

#### 4. **Workers**

- Gồm các workers thực hiện các tác vụ khác nhau:

- File controlr: Quản lý các file user đã upload (lưu trữ, xử lý, đọc, ...)
  - Analyzer: Chuyên gia phân tích tài liệu tài chính.
  - Thực hiện các tasks đã được planning bằng cách sử dụng các tools.
  - Tổng hợp kết quả các tasks đã thực hiện, đưa ra câu trả lời.
5. **Supervisor**
- Quản lý, phân chia các tasks cho workers
6. **Responder**
- Tổng hợp kết quả từ các workers để đưa ra câu trả lời.

## Workflows

1. **User Interaction:**
  - Người dùng tương tác với agent thông qua giao diện chat hoặc API.
  - Đầu vào được xử lý để xác định ý định và ngữ cảnh.
2. **Memory Recall:**
  - Agent truy xuất dữ liệu lịch sử conversation từ bộ nhớ.
3. **Planning:**
  - Agent tạo kế hoạch chi tiết cần thực hiện.
  - Kế hoạch được chia thành các bước step-by-step cụ thể.
4. **Supervisor:**
  - Dựa vào kế hoạch đã được tạo, Supervisor phân chia các tasks cần thực hiện cho từng workers.
  - Các workers có thể thực hiện các tasks tuần tự hoặc song song.
5. **Workers:**
  - Sử dụng các tools đã được định nghĩa để hoàn thành tasks được giao
  - Trả lại kết quả các task cho Supervisor
6. **Responder:**
  - Sau khi hoàn thành tất cả các tasks, Responder nhận các kết quả của Workers từ Supervisor
  - Phân tích, tổng hợp đưa ra câu trả lời cuối cùng.

## Cải thiện trong tương lai

- **Memory:**
  - Sử dụng cơ sở dữ liệu riêng biệt để lưu trữ conversations, user interactions.
  - Kết hợp vector database để lưu trữ long-term-conversation.
- **Planning:** Bổ sung Replanning agent để Observation hỗ trợ điều chỉnh các task phù hợp trong quá trình thực thi.
- **Workers:** Bổ sung thêm nhiều workers để thực hiện các task phức tạp hơn.
- **Tools:**
  - Thiết kế, bổ sung thêm nhiều tools giúp agent có thể hỗ trợ nhiều hơn như tạo biểu đồ, tạo file báo cáo, tờ trình, ...
  - Kết hợp nhiều sub agent hơn để thực thi nhiều task khác nhau.