

1. User Query Input

- **Interface:** Users interact with the system via a **Streamlit** web interface.
- **Input:** Users submit research-related queries through a text input field.

2. Query Processing and Agent Pipeline

- **Memory Check:** Upon receiving a query, the system first checks its memory cache to see if an answer to the query already exists.
- **Answer Retrieval:**
 - **Cache Hit:** If the answer is found in the cache, it is retrieved and presented to the user.
 - **Cache Miss:** If no cached answer exists, the system initiates the agent pipeline to generate a new response.

3. Agent Pipeline Execution

The agent pipeline comprises several stages:

- **Document Retrieval:** The system searches for relevant documents that pertain to the user's query.
- **Information Extraction:** Key information is extracted from the retrieved documents.
- **Answer Generation:** Using the extracted information, the system formulates a comprehensive answer to the user's query.
- **Answer Caching:** The generated answer is stored in the memory cache for future reference.

4. Response Delivery

- **Presentation:** The final answer is displayed to the user through the Streamlit interface.
- **Download Option:** Users have the option to download the generated answer in JSON format for offline use.