

How JP Morgan Built An AI Agent for Investment Research with LangGraph | LangChain Interrupt



LangChain
134K subscribers

Subscribe

2.1K



Share

Clip

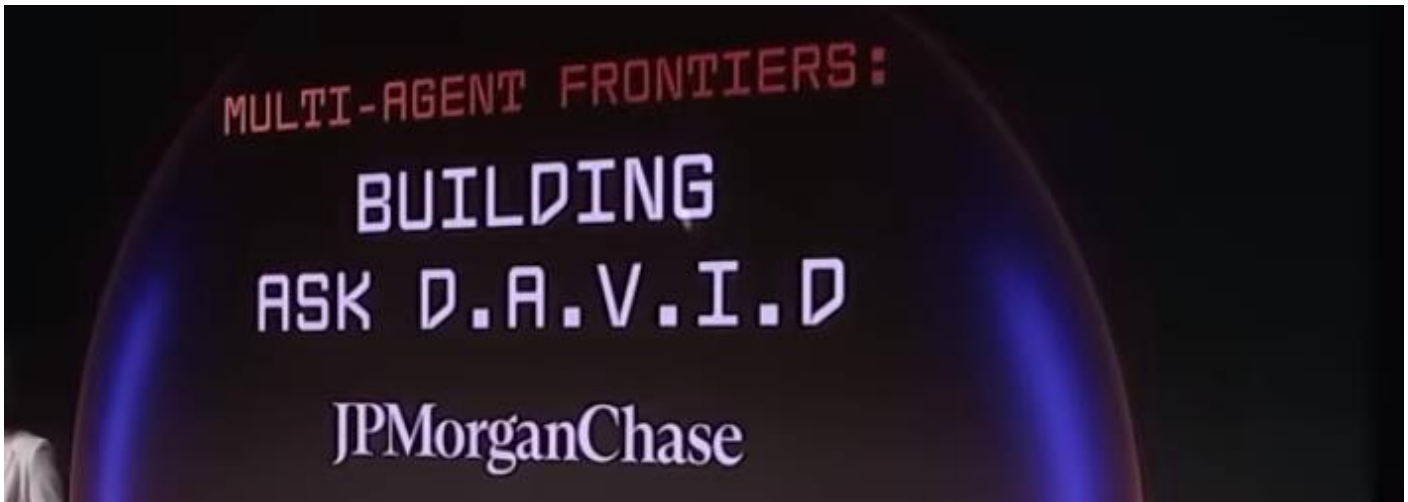
Save



96,241 views May 29, 2025 #LangChain #LangGraph #MultiAgent

Watch all of our recorded sessions from Interrupt here: <https://interrupt.langchain.com/video...>

David Odomirok and Zheng Xue from JP Morgan Chase Private Bank reveal how they built "Ask David" - a sophisticated multi-agent AI system designed to automate investment research for thousands of financial products. With billions of dollars in assets at stake, this isn't just another chatbot - it's an enterprise system built with human oversight for high-stakes financial decisions.



Today

JPMorgan

Organization



→Private Bank specializing in a diverse range of investment products tailored for our clients.

→**Investment Research Team** curates and manages a comprehensive list of investment opportunities.

Challenge



The team often receives numerous inquiries about investment products, requiring individuals to manually research and compile data to provide comprehensive responses.

→ Manual and time-consuming research
→ Lack of product insights
→ Scalability constraints

Goal



Automating investment research to deliver fast and precise results.

Tomorrow

J.P.Morgan



Domain Knowledge

Structured Data →
RDS, Data mesh



Unstructured Data →
Emails, Meeting notes, Zoom

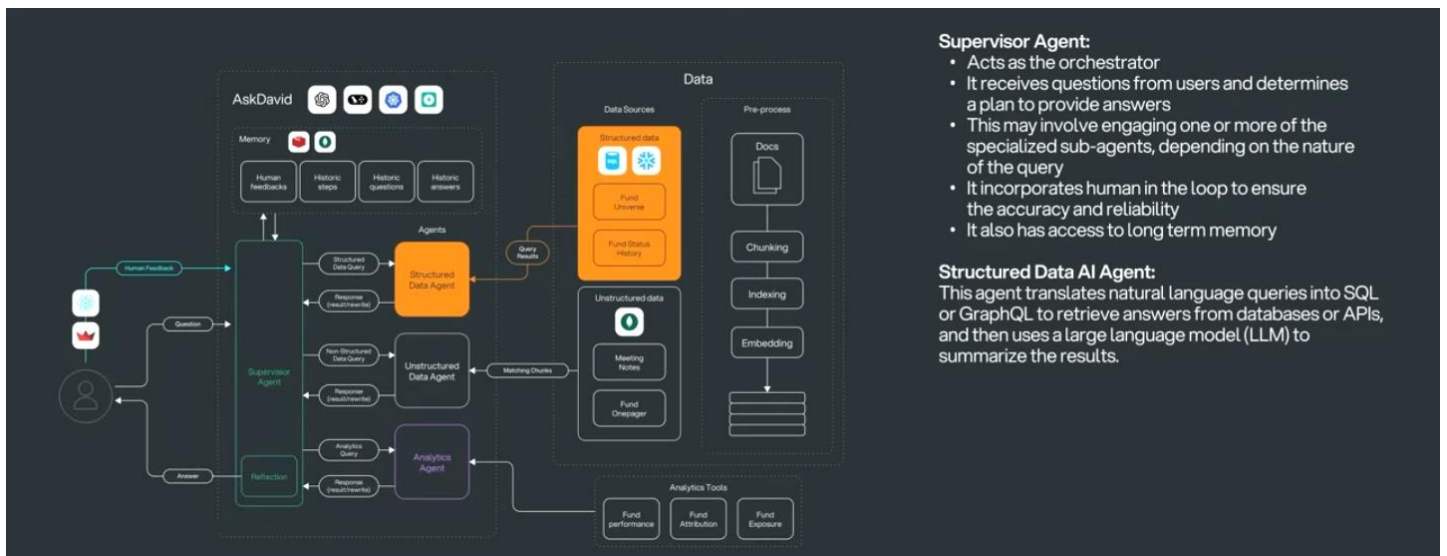
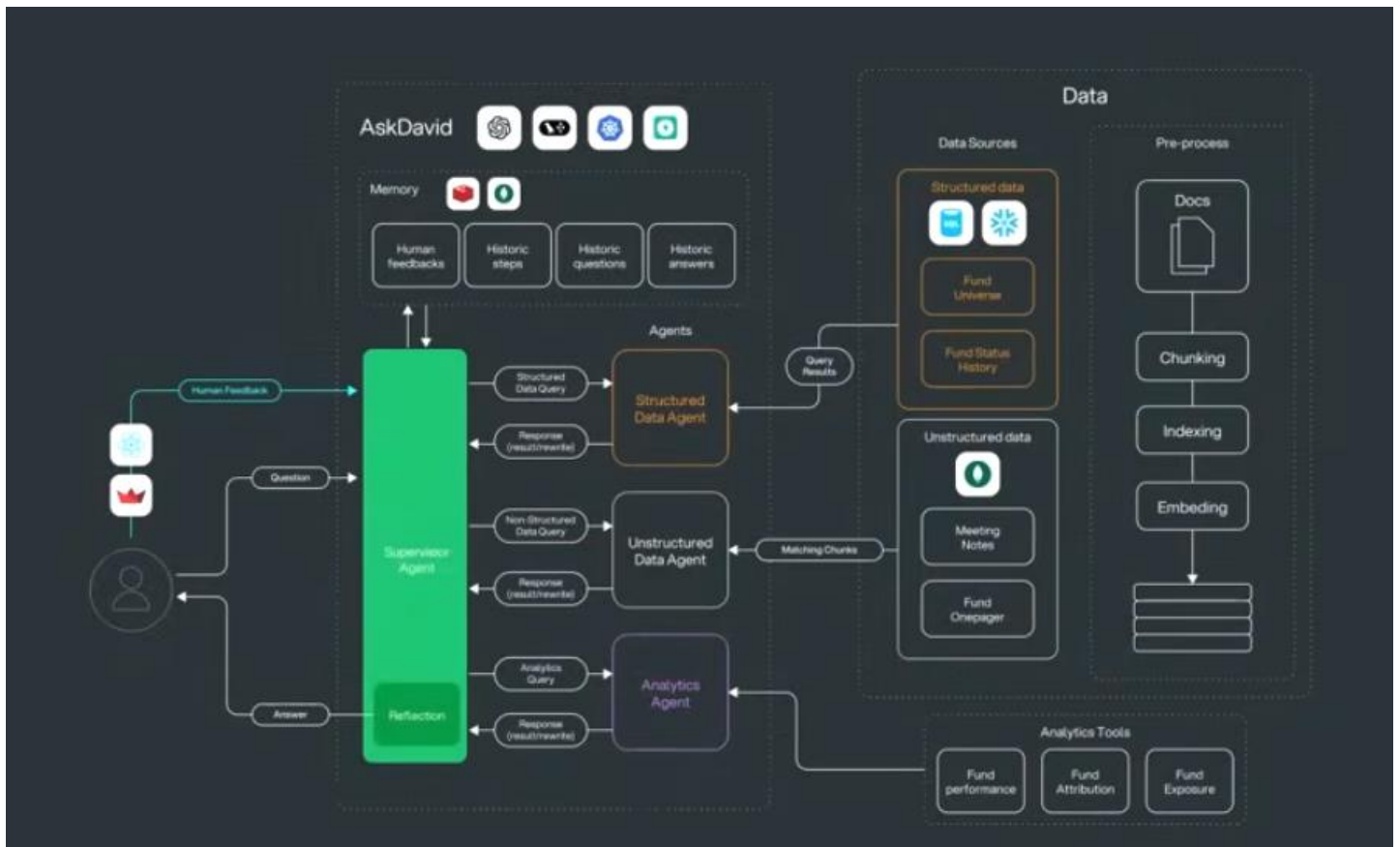


Proprietary Models and Analytics →
APIs, Libraries



The Approach

J.P.Morgan

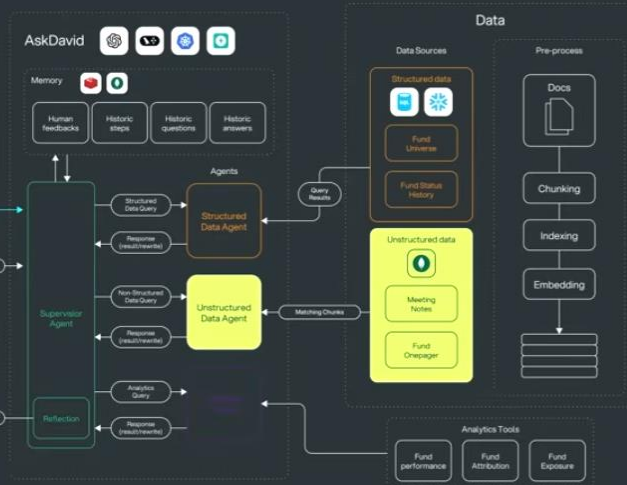


Supervisor Agent:

- Acts as the orchestrator
- It receives questions from users and determines a plan to provide answers
- This may involve engaging one or more of the specialized sub-agents, depending on the nature of the query
- It incorporates human in the loop to ensure the accuracy and reliability
- It also has access to long term memory

Structured Data AI Agent:

This agent translates natural language queries into SQL or GraphQL to retrieve answers from databases or APIs, and then uses a large language model (LLM) to summarize the results.



Supervisor Agent:

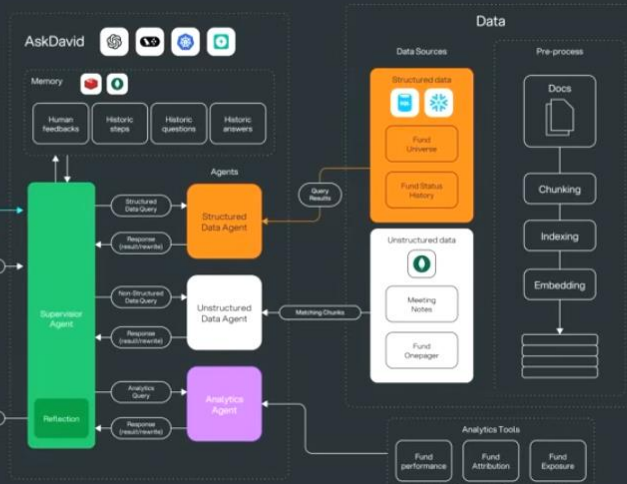
- Acts as the orchestrator
- It receives questions from users and determines a plan to provide answers
- This may involve engaging one or more of the specialized sub-agents, depending on the nature of the query
- It incorporates human in the loop to ensure the accuracy and reliability
- It also has access to long term memory

Structured Data AI Agent:

This agent translates natural language queries into SQL or GraphQL to retrieve answers from databases or APIs, and then uses a large language model (LLM) to summarize the results.

Unstructured Data AI Agent:

This RAG agent utilizes a vector database and hierarchical index retrieval to store documents and search for relevant information chunks, followed by summarization using an LLM.



Supervisor Agent:

- Acts as the orchestrator
- It receives questions from users and determines a plan to provide answers
- This may involve engaging one or more of the specialized sub-agents, depending on the nature of the query
- It incorporates human in the loop to ensure the accuracy and reliability
- It also has access to long term memory

Structured Data AI Agent:

This agent translates natural language queries into SQL or GraphQL to retrieve answers from databases or APIs, and then uses a large language model (LLM) to summarize the results.

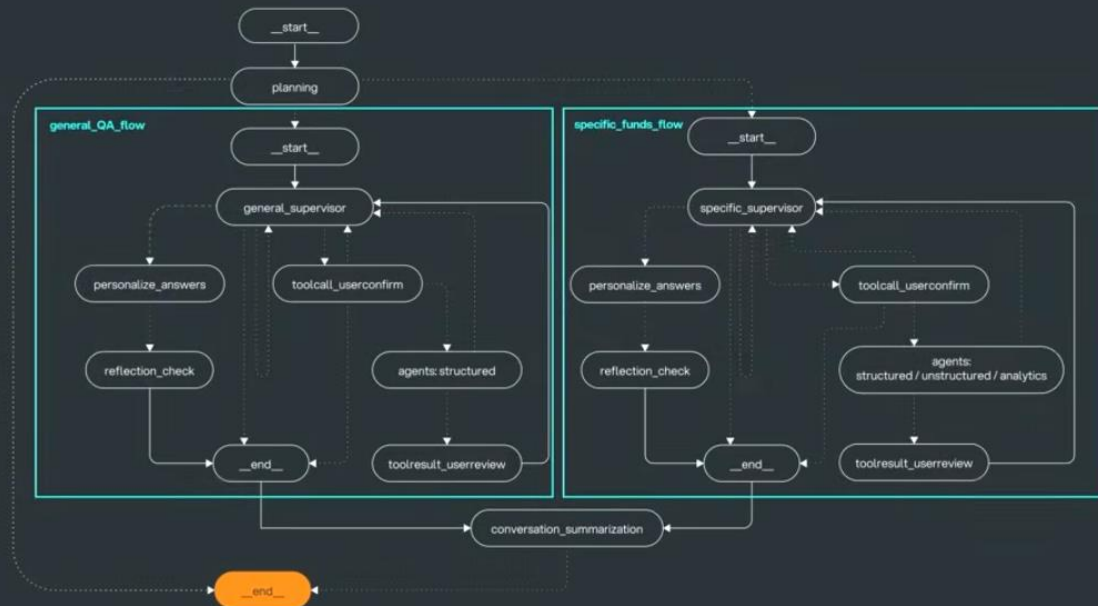
Unstructured Data AI Agent:

This RAG agent utilizes a vector database and hierarchical index retrieval to store documents and search for relevant information chunks, followed by summarization using an LLM.

Analytics Agent:

ReAct agent with access to analytics APIs as tools.

Graph: Workflow



Example: Why was a fund terminated?

J.P.Morgan

Planning

- Determines the intention
- Invokes the corresponding subgraph

Subgraph - Specific Funds Flow

- delegates to doc_search_agent

Doc Search Agent

- Conducts mongoDB search to retrieve relevant information

Supervisor

- Personalization based on the role of human
- Reflect to ensure the accuracy and relevancy of answers
- Summarize conversation, update memory, and return answer

```
planning 0.57s
AzureChatO... gpt-4o-2024-... 0.55s
PydanticToolsParser 0.00s
specific_funds_flow 22.76s
specific_supervisor 0.51s
AzureChat... gpt-4o-2024-... 0.49s
finish_supervisor 0.00s
toolcall_userconfirm 0.01s
route_after_human 0.00s
agents *Structured *Unst... 14.76s
doc_search_agent 14.44s
MongoDBCommentarySea... 0.13s
AzureCha... gpt-4o-20... 14.30s
finish_agents 0.32s
specific_supervisor 1.80s
AzureChat... gpt-4o-2024-... 1.47s
finish_supervisor 0.31s
personalize_answers 5.67s
AzureChat... gpt-4o-2024-... 5.65s
PydanticToolsParser 0.00s
reflection.check 0.00s
conversation_summarization 1.57s
```

HUMAN
Why was fund put on Terminated?

AI

doc_search_agent call_I1ihfcik7PbuYguH3u8aM4

query: reason for fund being put on terminated

YAML

TOOL

call_I1ihfcik7PbuYguH3u8aM4

was terminated due to concerns about its performance.

Despite the fund performing well and outperforming in each of the past four quarters, the primary concern was that the rebound in performance was not as strong as expected. There were no changes to the people, philosophy, or process of the fund, but the committee decided to terminate it based on the underwhelming rebound in performance [1][a131e/c4-8796-411c-8601-07faa9e510c5].

YAML

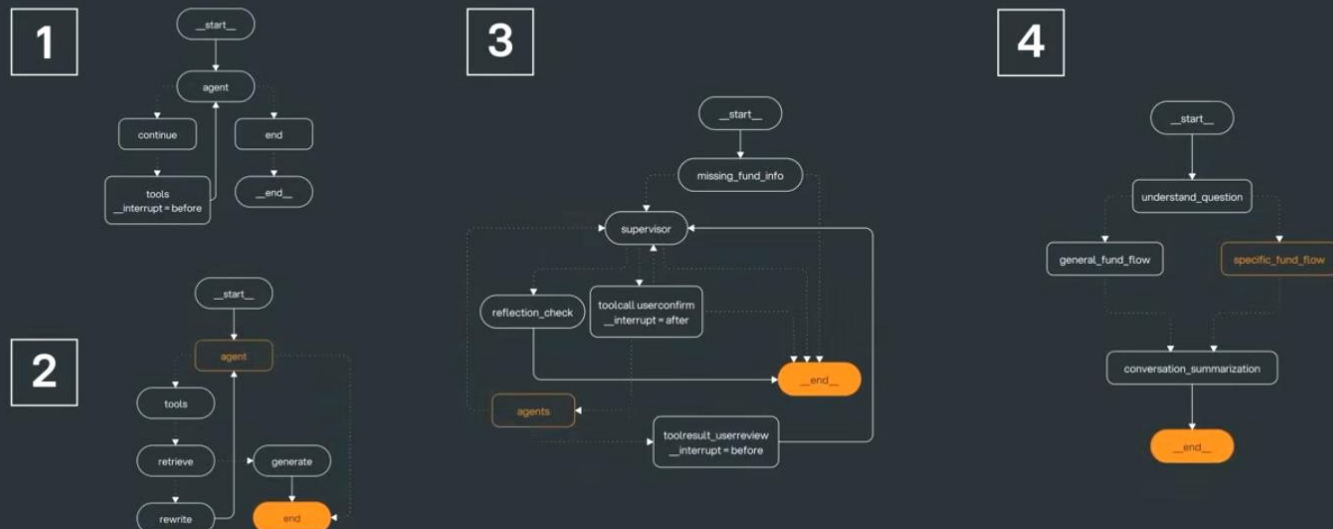
RAW

AI

was terminated due to concerns about its performance. Despite the fund performing well and outperforming in each of the past four quarters, the primary concern was that the rebound in performance was not as strong as expected. There were no changes to the people, philosophy, or process of the fund, but the committee decided to

Lessons Learned

Lesson 1: Start Simple and Refactor Often



Lesson 2: Evaluation Driven Development

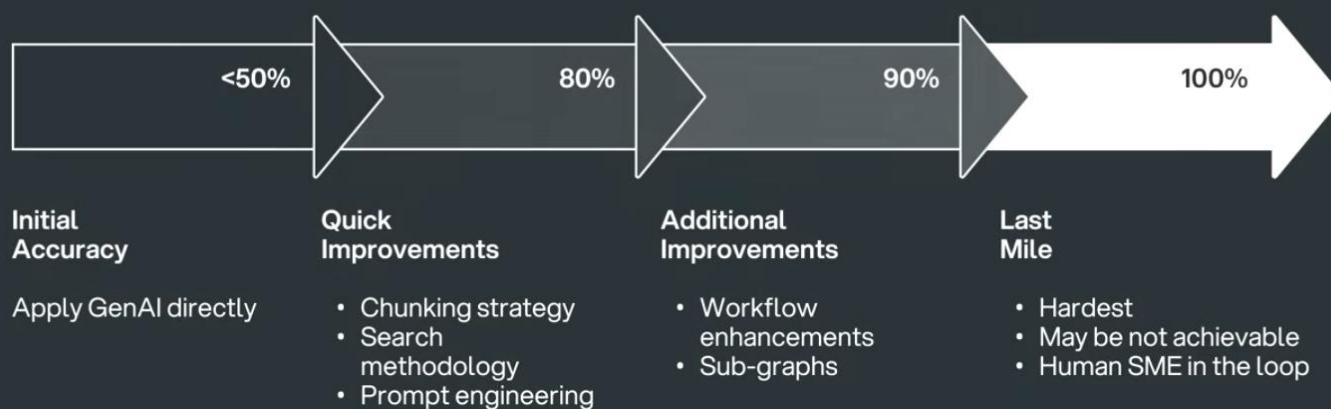
→ GenAI has much shorter dev phase than traditional AI projects, which enables rapid prototyping and iteration

→ Domain Confidence requires extensive evaluation and refinement

- Independent Evaluation of Subagents: Assess subagents and the whole flow separately.
- Diverse Metrics for Agent Types: Tailor metrics to the agent design.
- Evaluation with/without Ground Truth: Not all metrics need ground truth; valuable insights can be gained without it.
- LLM as Judge + Human SME Review: Use automated evaluations to scale and reduce SME workload.



Lesson 3: Human SME in the Loop



Iterate fast,
evaluate early,
include Human SME
in the loop

Thank you!

jane.xue@jpmchase.com
david.j.odomirok@jpmorgan.com

J.P.Morgan

May 14th, 2025

@ Interrupt,
The AI Agent Conference
by Langchain