Elevating the Use experience with AI – Product document Chatbots

Praveen Govindaraj

About me

- 📊 **Lead AI Engineer at Leading Telcom**
- Over 13.6 years of experience in Data science & analytics domain.
- 🚀 **Expertise Areas**:
- Customer Analytics, Anti-Money Laundering/Counter-Terrorism Financing
- (AML/CFT), Credit Risk Modelling, Data Warehousing & GenAI Platforms.
- 📉 **Achievements**:
- Reduced Customer Churn, Streamlined Processes, Enhanced Business performance
- Spearheading AI CoE
- 💡 **Proven Record**:

Delivering data-driven business improvements & GenAI platform.

Hobbies: Playing Piano 🎹 & Violin🎻.

💻 **Contributions**: Open Source Projects in GenAI.

✍️ **Interests**: Tech Blogging.

📚 Leisure: Reading Books.

Agenda

- Why product documentation is important for team or Company
- What are GenAI Agents
- Agents to enhance user experience on product documentation experience
- Solution and Demo
- RAG Agents Antipatterns

Why product documentation is important for team or Company

Enhances Customer Support and Satisfaction

Improves Efficiency and Speeds up Onboarding

Facilitates Knowledge Sharing and Consistency

What are GenAI Agents

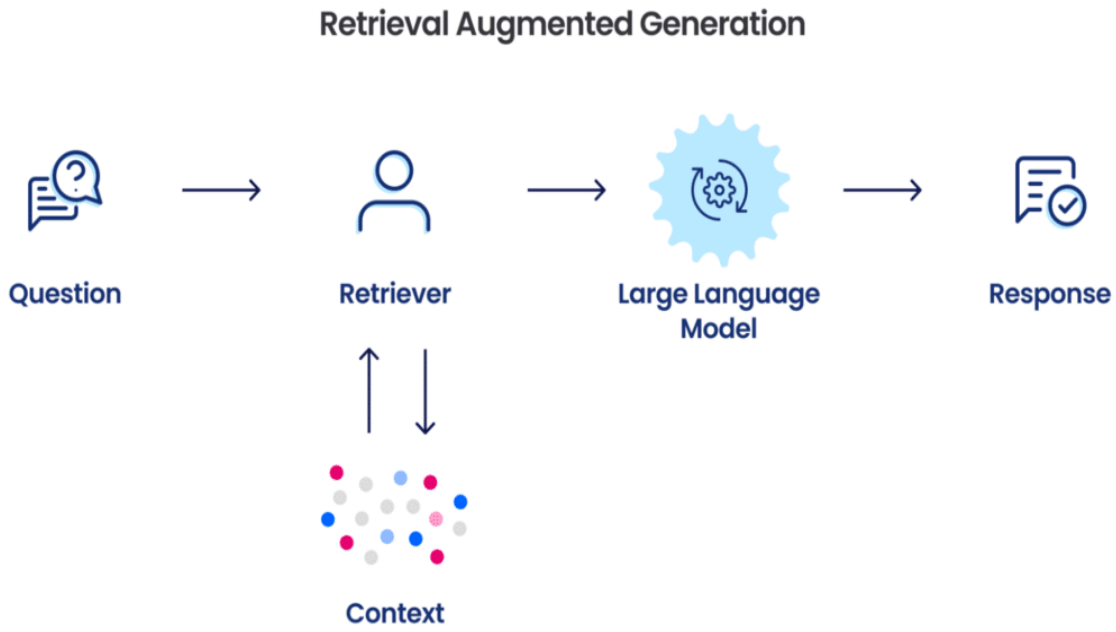
What is an LLM ?

An LLM, or Large Language Model, is like a super-smart writing assistant. It learns from a massive amount of text to understand and mimic human language. This lets it answer questions, write stories, and more, just by analysing the words you give it.

What is RAG ?

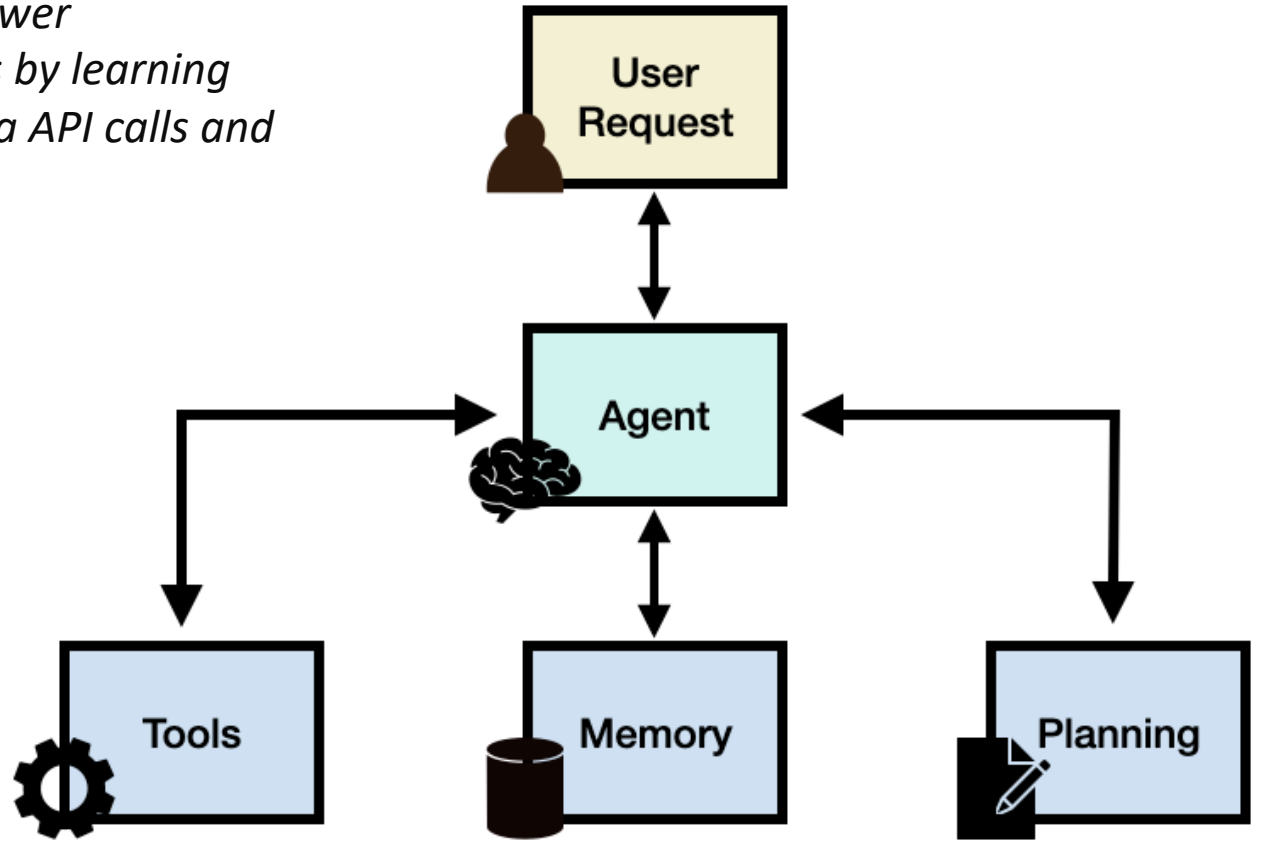
RAG, or Retrieval-Augmented Generation, is like a smart assistant that first finds relevant info from a big database, then uses it to craft detailed answers.

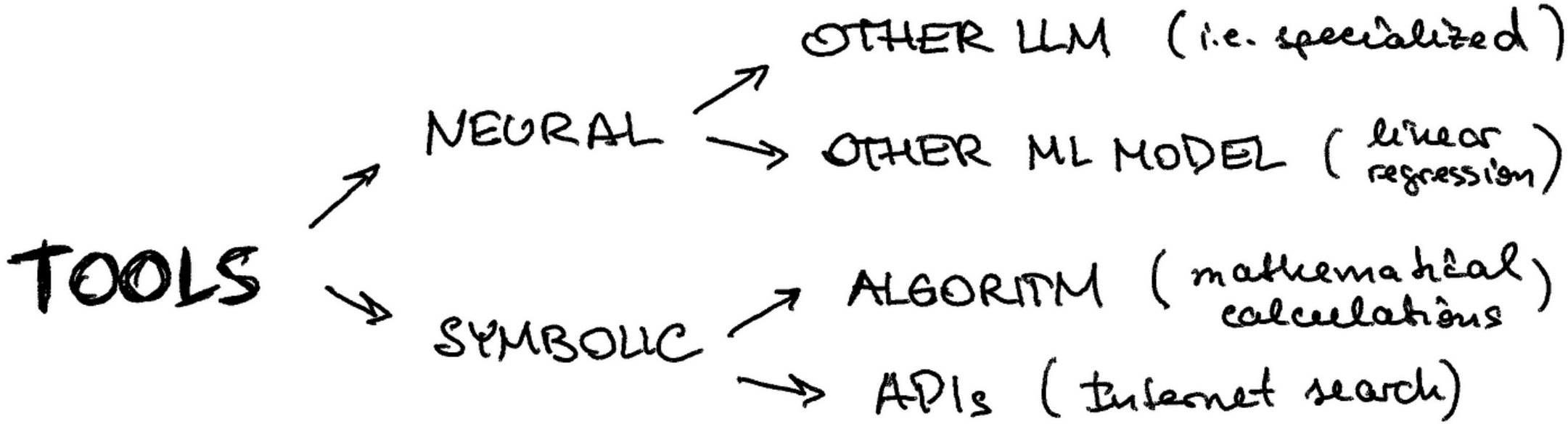
It combines quick research with creative writing, providing informed and engaging responses.

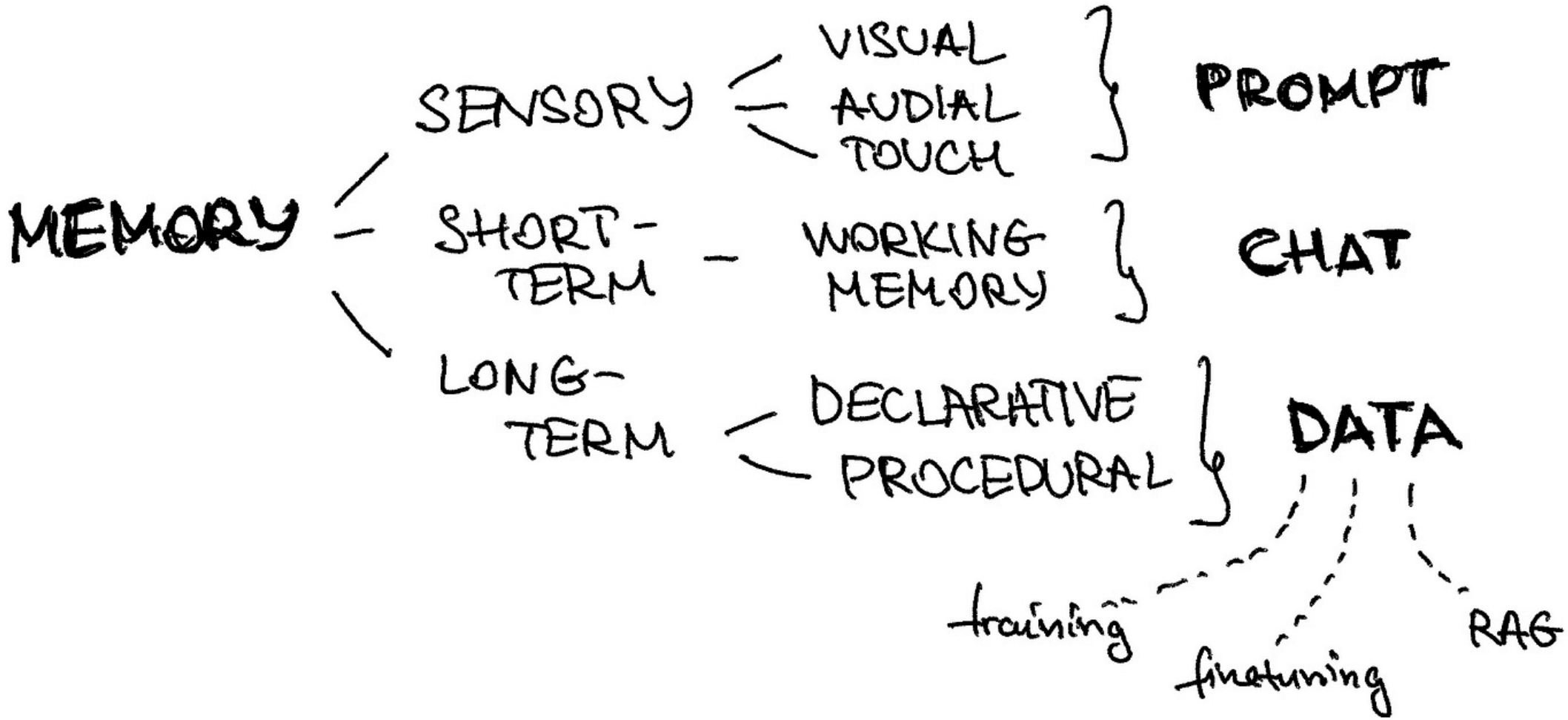


What is LLM an Agent ?

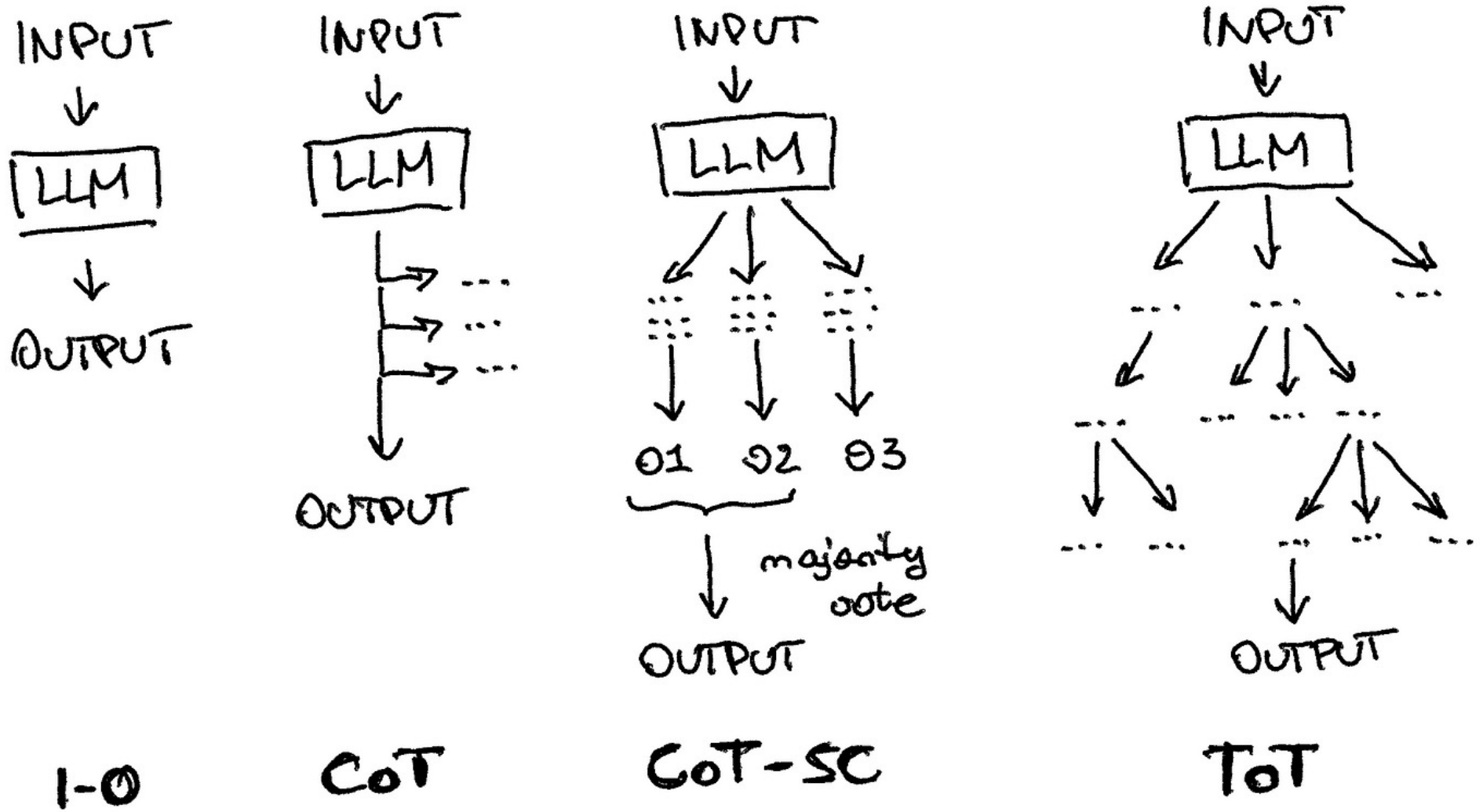
Very smart digital assistants that can write text, answer questions, and even create stories or solve problems by learning from vast amounts of information on the internet via API calls and work well with other software tools




TOOLS
NEURAL
OTHER LLM (i.e. specialized)
OTHER ML MODEL (linear regression)
SYMBOLIC
ALGORITM (mathematical calculations)
APIs (Internet search)


MEMORY
SENSORY
VISUAL
AUDIAL
TOUCH
PROMPT
SHORT-TERM
WORKING MEMORY
CHAT
LONG-TERM
DECLARATIVE
PROCEDURAL
DATA
training
finetuning
RAG

Planning


INPUT
LLM
OUTPUT
INPUT
LLM
OUTPUT
INPUT
LLM
O1
O2
O3
majority vote
OUTPUT
INPUT
LLM
OUTPUT
I-O
CoT
CoT-SC
ToT

CHALLENGES OF SILOED IMPLEMENTATION

Road to robust LLM Agentic systems

1

Unenhanced LLM Workflows

- Lack of fact fullness led to hallucination.
- Bias

SOLUTION 01

LLM with relevant context information helps to mitigate the hallucination.

2

RAG workflows

- Good at sharing the relevant answer to question.
- Lacks recent information and connectivity. information.

SOLUTION 02

Equipped with tools and connectivity to internet and intranet to get the relevant information via API.

3

LLM Agent

- Complex system integration
- Quite hard to validate accuracy due to complex integration

SOLUTION 03

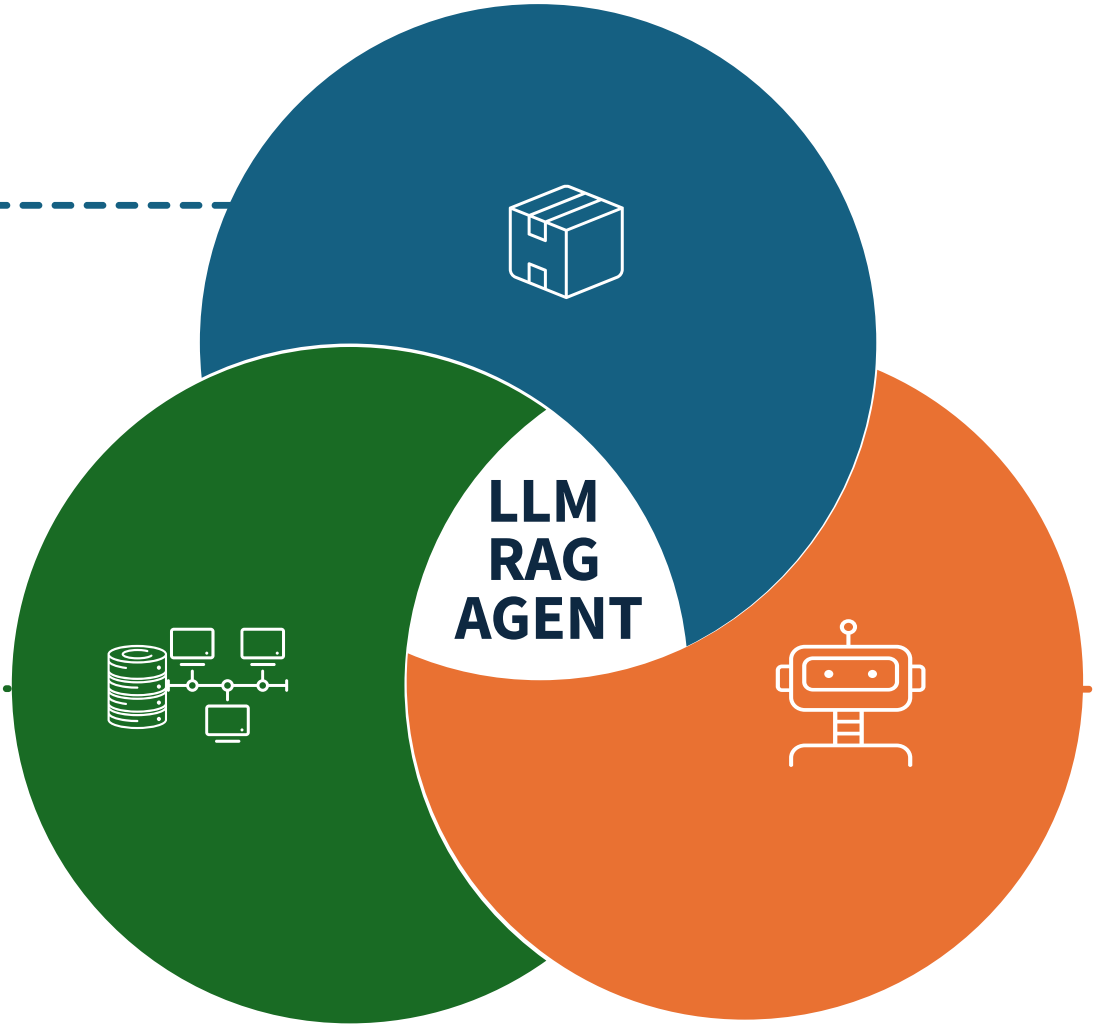
Ensure individual components are validated and tested as whole system.

Confluence of LLM , RAG and Agent

Collaborative Intelligence LLM,RAG and Agent Integration

LLM

Large Language Models excel at generating relevant responses to prompts, possessing the capability to reason and reflect.



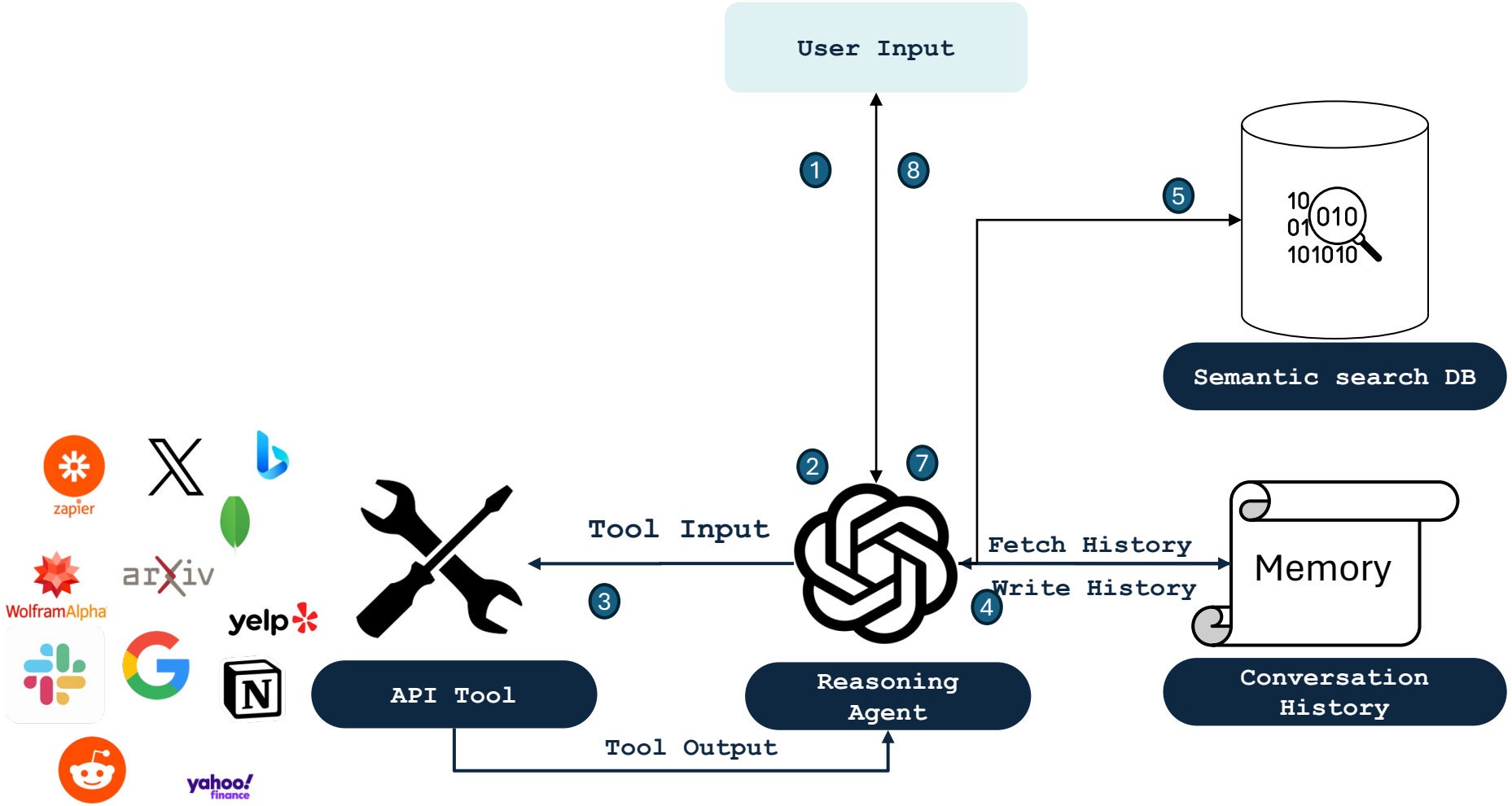
RAG

Enhance the performance of LLM by providing richer, more pertinent contextual information, thereby reducing the occurrence of inaccurate or fabricated responses.

AGENT

It possesses tools, memory, and the ability to plan, enabling it to coordinate, plan, and execute tasks effectively.

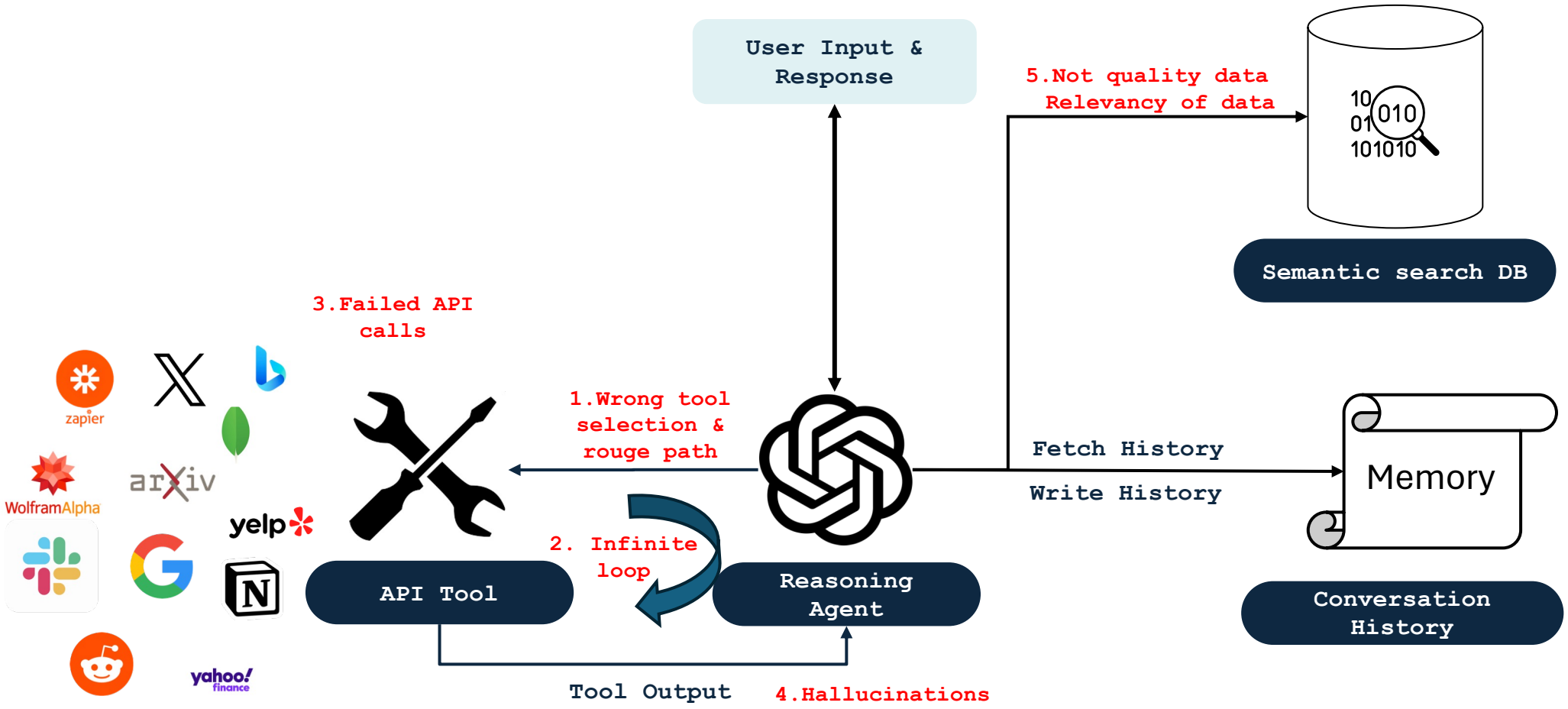
Typical LLM Data Agent



Demo

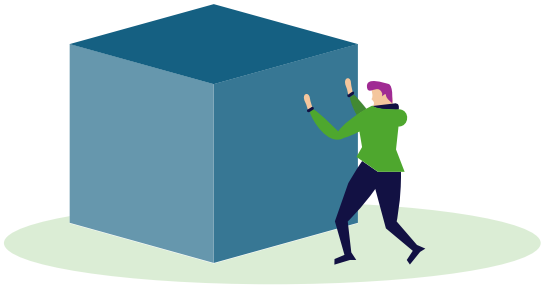
LLM Data Agents Antipatterns

LLM Data Agents – Anti patterns



LLM Data Agents – Anti patterns (1/2)

Better architecture for sustainable outcomes



Anti patten # 01

Wrong tool selection and rouge path

Solution 01

The selection of tools and API agents is rigorously verified and standardized.

Anti patten # 02

Agent is repeatedly using tools such as internet search without exit criteria

Solution 02

Ensure that the exit criteria for tool searches are clearly defined, including a specified iteration limit.

Anti patten # 03

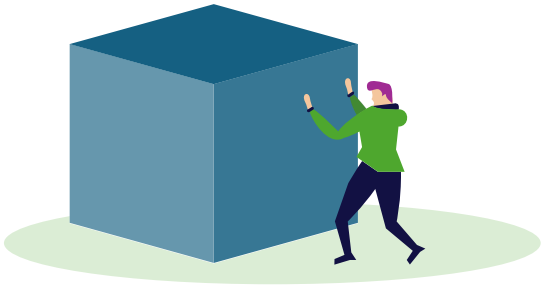
Failed API Calls

Solution 03

Implement retry mechanisms and manage exceptions appropriately within the workflow.

LLM Data Agents – Anti patterns (2/2)

Better architecture for sustainable outcomes



Anti patten # 04

Occurrence of inaccurate or fabricated responses from LLM. i.e. Hallucination

Solution 04

Align LLM with specific goals based on business needs. i.e. Finetuning

Anti patten # 05

Not good quality of data and relevancy of data

Solution 05

Ensure the data within the vector store is of high quality and regularly updated in a timely manner.

Take away

- **RAG ARCHITECTURE**
- **LLM RAG AGENTS**
- **HOW LLM RAG AGENTS HELPS TO IMPROVE USE EXPERIENCE ON PRODUCT DOCUMENTS**
- **AGENT EVALUATION**

Thank you

Presentation materials