

— A COMPLETE GUIDE TO —

AI AGENTS

Concepts, Architecture, Frameworks
and the Future of Intelligent Systems



A COMPLETE GUIDE TO

AI AGENTS

Concepts, Architecture, Frameworks
and the Future of Intelligent Systems



01

Understanding the Limitations of LLMs



02

LLM Model Data vs Vector Database



03

External Sources of Knowledge in LLMs



04

RAG Overview



05

Types of Memory in LLMs



06

Key Elements of an AI Agent



07

MCP Overview



08

AI Agent Frameworks



09

Enterprise AI Controls



10

The Complete AI Ecosystem



Samir Solanki



linkedin.com/in/samir38

UNDERSTANDING THE LIMITATIONS OF LLMs

Exploring the boundaries, challenges, and real-world realities of Large Language Models.





KNOWLEDGE FRESHNESS

LLMs have a knowledge cutoff – they may not know events, updates, or information that happened after their training.



CHALLENGE

LLMs are trained on data available up to a certain cutoff date. Any information, events, regulations, product updates, or market changes that occur after that date are unknown to the model.

IMPACT ON APPLICATIONS



Outdated answers and recommendations



Incorrect or misleading information



Missed recent developments and opportunities



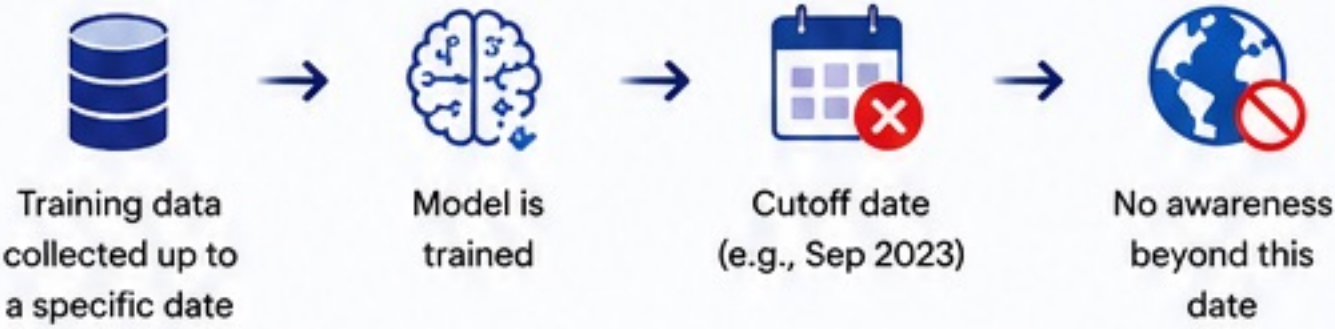
Poor user trust and experience



EXPLANATION

An LLM only “knows” what it was trained on. It does not have real-time awareness of the world. Without external information, it cannot provide accurate responses about anything beyond its knowledge cutoff.

WHY IT HAPPENS



i The model cannot browse the internet or fetch new data on its own.



PRACTICAL EXAMPLE

Scenario: An LLM with a knowledge cutoff of September 2023 is asked about the latest iPhone model released in 2024.



What is the latest iPhone model?



As of my knowledge cutoff in September 2023, the latest iPhone model is **iPhone 15**.



The model is unaware that iPhone 16 was released in September 2024.

MORE REAL-WORLD EXAMPLES



Stock Prices

Asking for today’s stock price – model may return outdated or incorrect data.

AAPL Price (May 17, 2024)
\$189.98



Government Policies / Regulations

New policies or tax rules introduced after the cutoff won’t be reflected.

New Tax Rule 2024

NEW



Product Launches / Features

New product releases, features, or updates won’t be known.



iPhone 16
Released Sep 2024

NEW



News & Events

Recent events, elections, or announcements will be missing.

Elections 2024 Results
Updated Nov 2024

NEW



SOLUTION

Provide the model with up-to-date information when needed.

REAL-TIME INFORMATION ACCESS

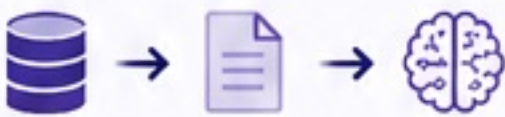
Integrate with search engines, APIs, or live data sources to fetch latest information on demand.



Use when: Real-time facts, prices, news, events.

RETRIEVAL-AUGMENTED GENERATION (RAG)

Retrieve relevant, up-to-date documents from your knowledge base and provide them as context to the LLM.



Use when: Internal docs, knowledge base, domain specific data.

USER-SUPPLIED CONTEXT

Allow users to provide the latest information in the prompt (e.g., news article, report, data file).



Use when: User has the latest data or context.

Example Context
News: Company X launched Product Y on May 10, 2024.

PERIODIC DATA UPDATES

Continuously update your internal databases, knowledge bases, and embeddings.



Use when: You control the data pipeline and refresh cycle.



KEY TAKEAWAY

LLMs are powerful, but their knowledge has a timestamp. To keep answers accurate and relevant, connect them to fresh data.

