



AI TECH SNAPSHOT

# RAG & LLMs

JULY 2025



Unlike search engines, LLMs generate direct answers without links, which can create an impression of deeper understanding. LLMs do not comprehend content; they predict the most statistically probable words based on their training data. They do not “know” the information—they generate likely responses that fit the prompt.

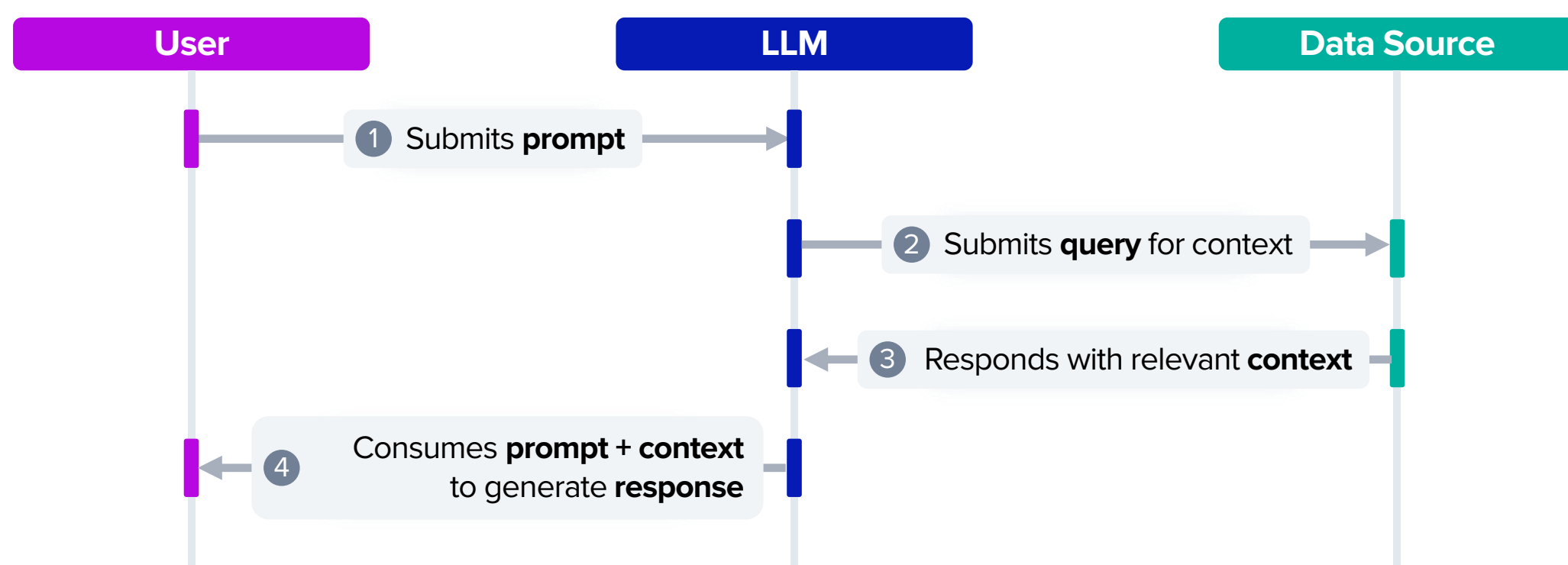
Additionally, LLMs are limited to the knowledge within their training data and cannot access real-time or comprehensive information. Retrieval Augmented Generation (RAG) addresses this limitation by integrating LLMs with external, current data sources.





# What is RAG?

RAG allows LLMs to draw on knowledge sources, such as company documents, databases, or regulatory content, that fall outside their original training data. By referencing this information in real time, LLMs can generate more accurate, relevant, and context-aware responses.





# Why RAG Matters for Business Leaders

LLMs generate responses based solely on information contained within their training data. Each user query is addressed using this fixed dataset, which may be limited or outdated.

RAG enhances LLM capabilities by enabling them to access external, up-to-date knowledge bases beyond their original training scope. While not infallible, RAG makes responses more accurate and contextually relevant.

Beyond internal data, RAG can harness internet-based information for real-time market intelligence, customer sentiment analysis, regulatory updates, supply chain risk monitoring, and talent insights, amplifying AI's business impact.





# Getting Started with RAG

Organisations can unlock RAG's potential through Custom GPTs; tailored LLM instances enhanced with external data sources. This enables specialised responses grounded in specific databases or documents.

Key use cases include:



## **Analytics & Intelligence**

Combine internal reports and market data for richer insights



## **Executive Briefings**

Automate strategy summaries from live data feeds.



## **Customer & Partner Support**

Deliver instant, precise answers using internal knowledge.



## **Compliance & Risk**

Query regulatory documents to mitigate risks.



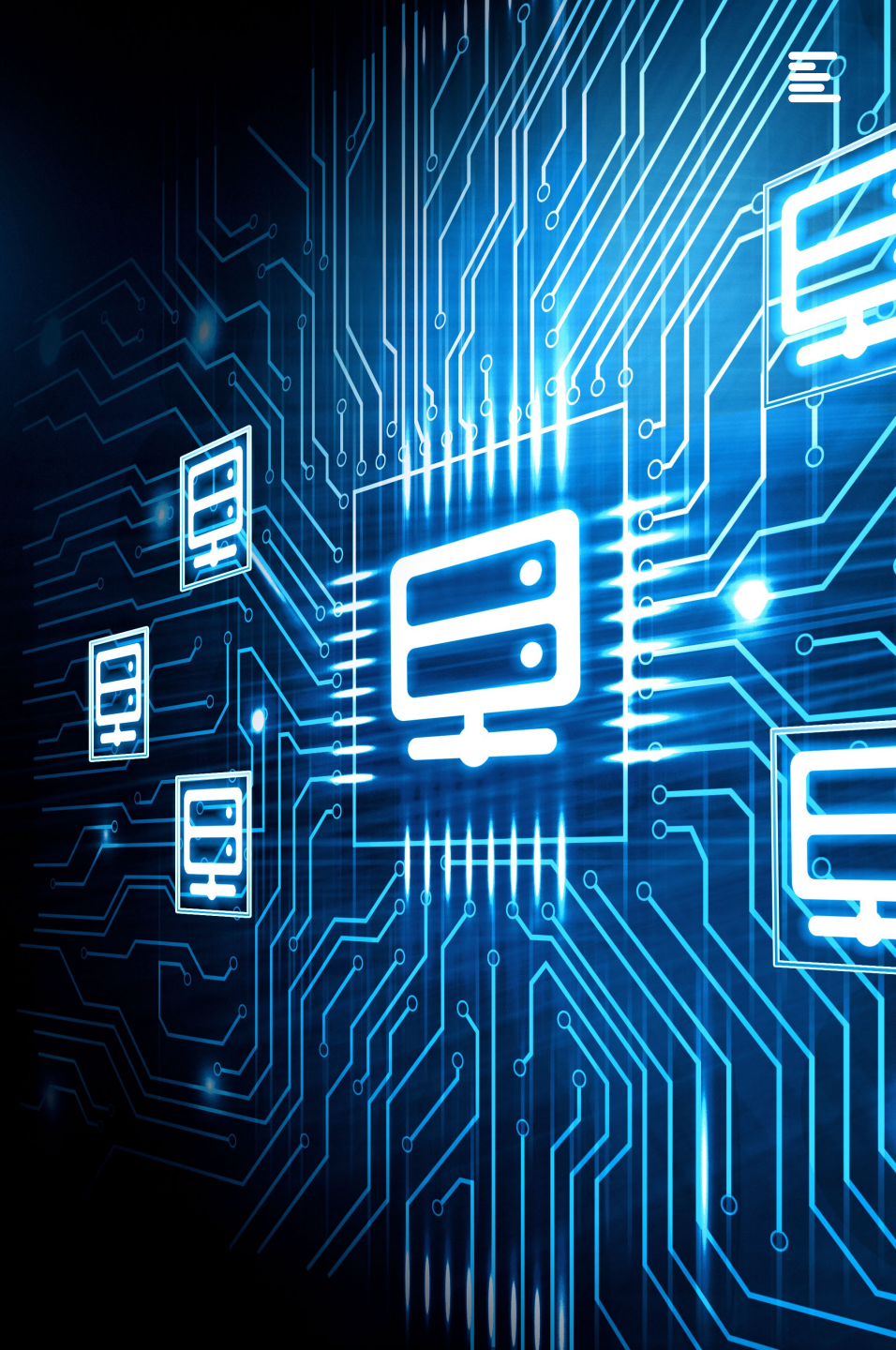
## **Training & Onboarding**

Streamline new hire familiarisation with company policies.

## Ecosystem Opinion

RAG enhances LLMs but has inherent limitations. Its effectiveness depends on the quality and organisation of external data; poorly maintained sources can lead to inaccurate outputs. Additionally, LLMs do not inherently manage privacy or security concerns, so measures such as role-based access controls and compliance audits are necessary to protect sensitive information.

For organisations, adopting RAG involves balancing innovation with governance. Effective implementation requires integrating RAG's capabilities with structured data management and security practices to support reliable, compliant, and efficient AI-driven outcomes.







# For more Ecosystem Insights, visit



[info@ecosystem.io](mailto:info@ecosystem.io)



[www.ecosystem.io](http://www.ecosystem.io)

