Analytics Vidhya

# Brief on Agentic RAG Systems


Instructor

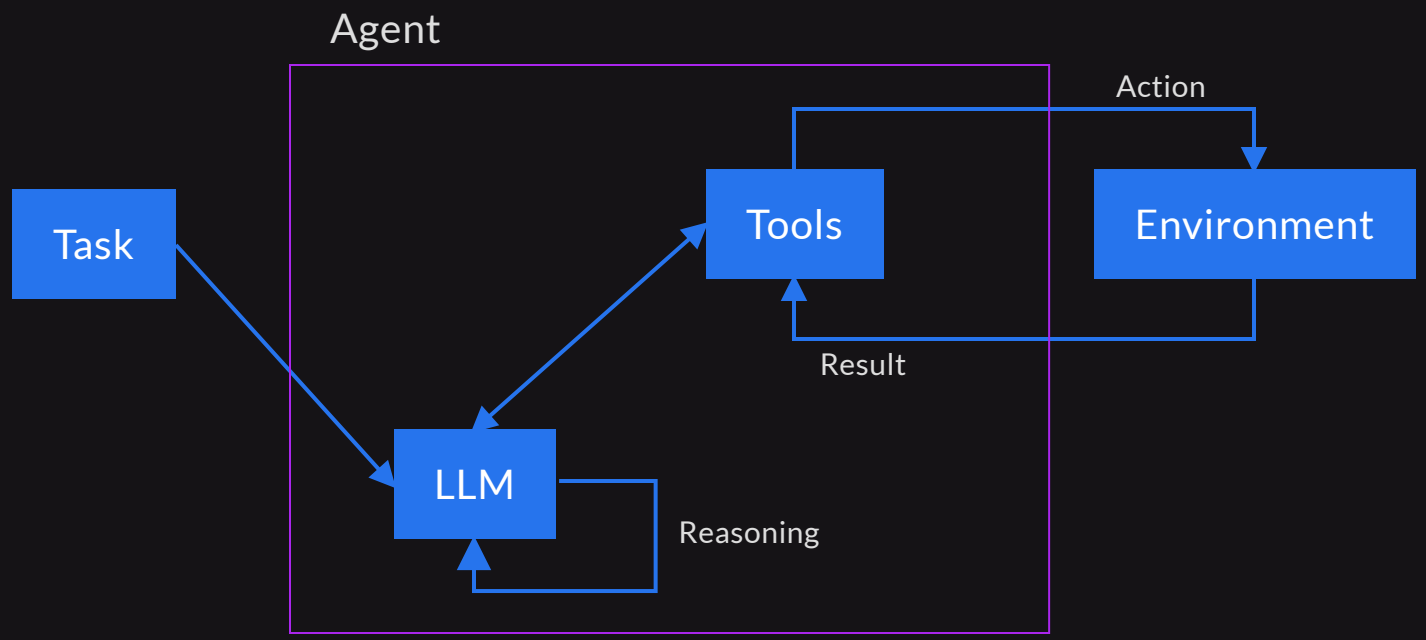
Dipanjan Sarkar

Head of Community & Principal AI Scientist at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

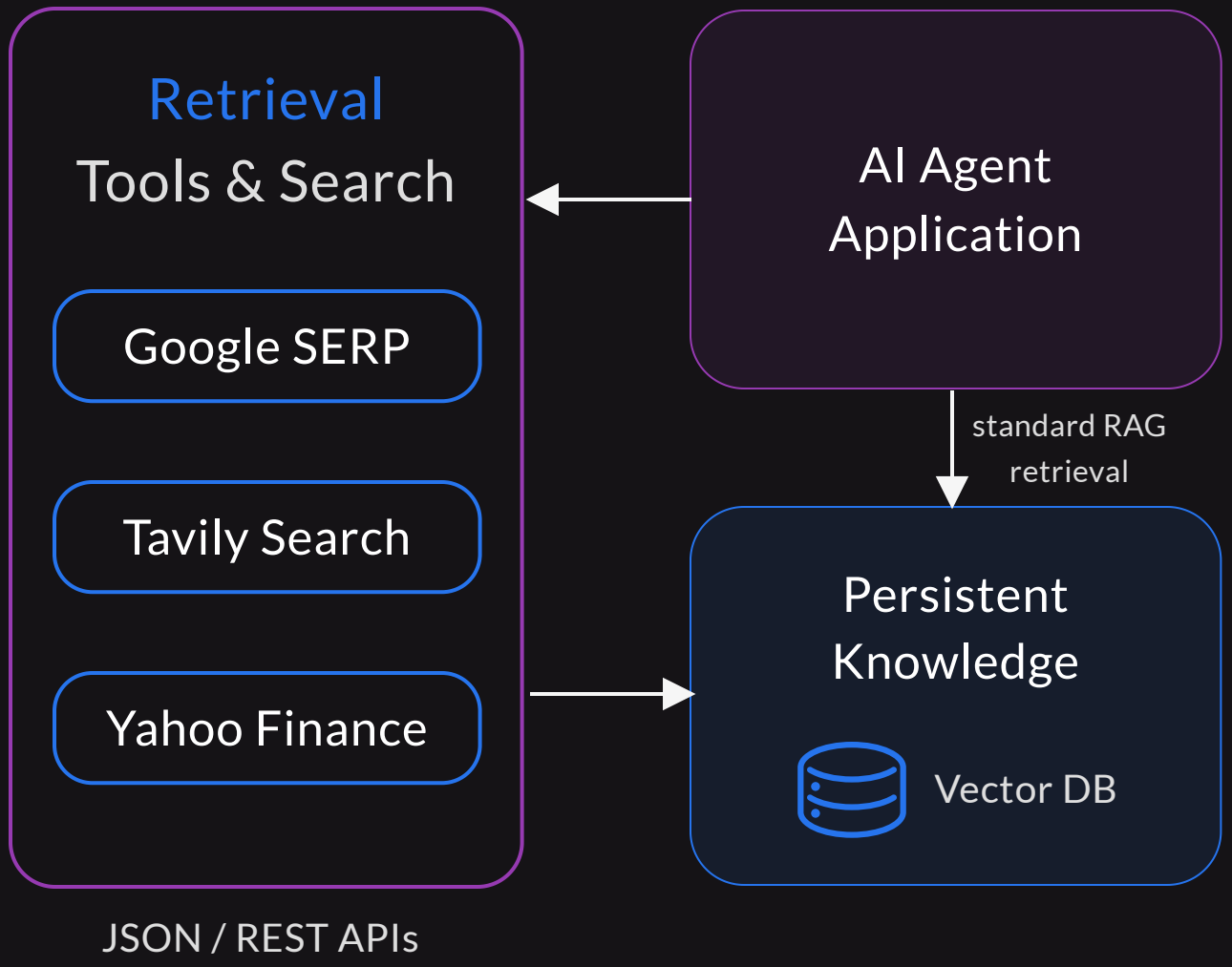
Published Author

# AI Agents



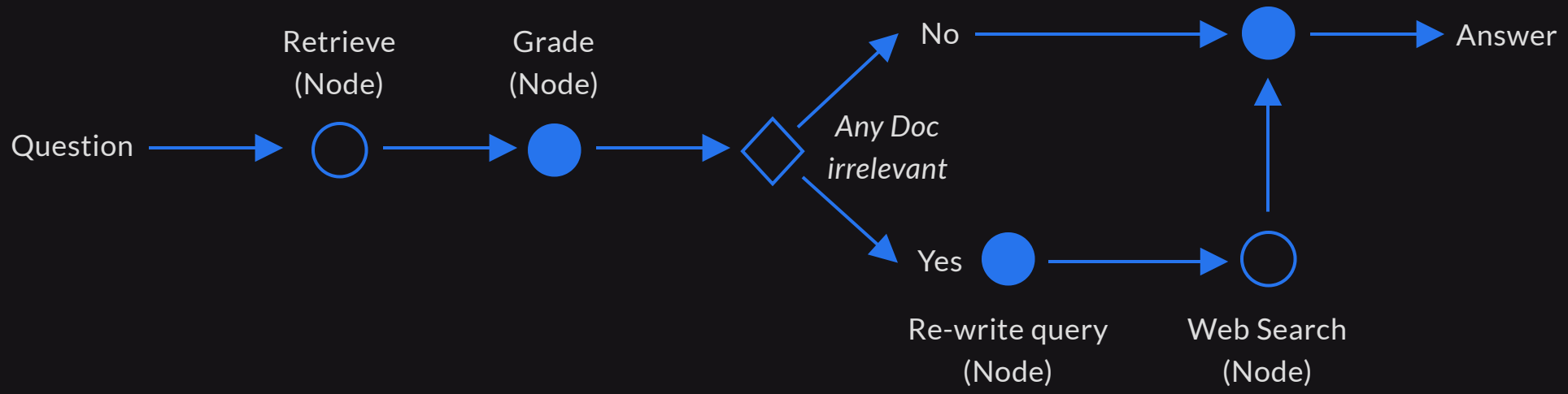

- Technically an AI Agent is a combination of LLMs, prompts and tools
- Flow starts with a user query or task
- The LLM usually reasons about what to do next in the cycle like Chain of Thought
- The LLM might call one or more tools to get relevant information from external sources
- The above steps might happen multiple times till the LLM has enough information to give a response

# Agentic RAG



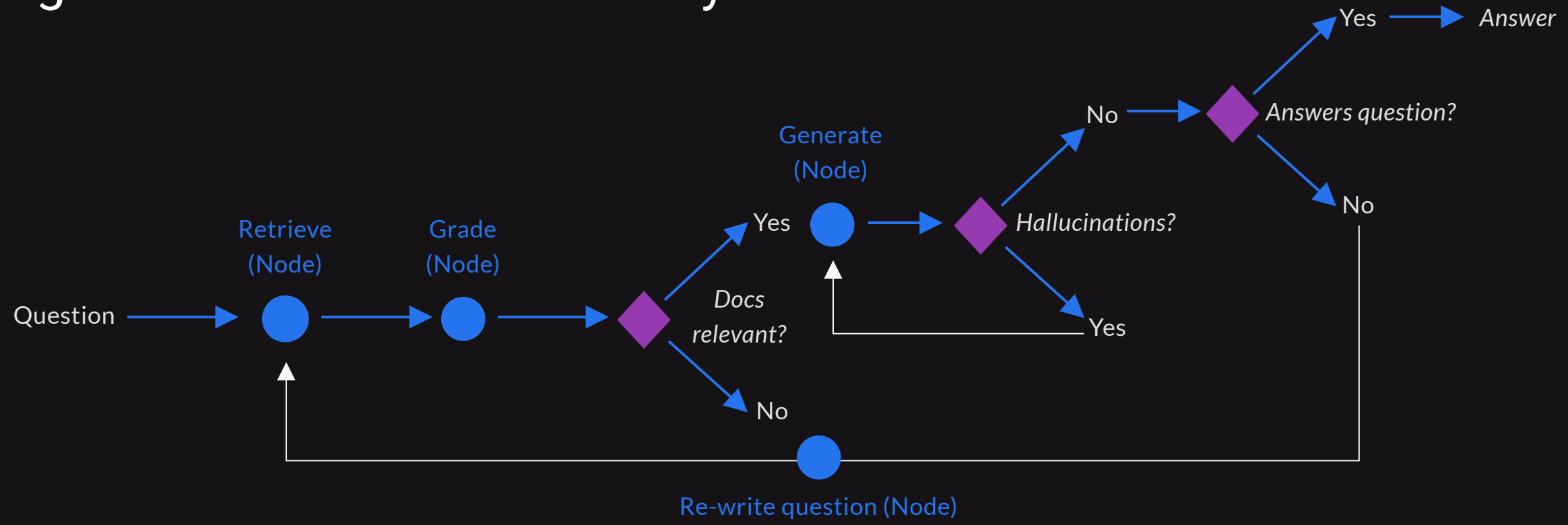

- Agentic RAG is a combination of AI Agents and RAG Systems
- Leverages retrieval and search tools to access live real-time data besides the vector database
- Can be extended to add in multiple levels of complex flows to validate retrieval, response generation and check for hallucinations
- Examples include Agentic Corrective RAG, Self-Reflective RAG and more

# Agentic Corrective RAG System




- Use Agentic flows and build a graph-based network
- Utilize a powerful language model to assess if the retrieved context from the vector database is sufficient to answer the query.
- If it's applicable, follow the standard RAG flow; otherwise, use web search tools to obtain live contextual information to answer the query.

# Agentic Self-Reflective RAG System




- Use standard vector database retrieval for context based on query
- Leverages agentic reflection pattern to use an LLM to reflect on the context and check for relevancy
- Also checks for hallucinations and if the question is answered using the same pattern to make the system more accurate

Analytics Vidhya

# Thank You

Analytics Vidhya