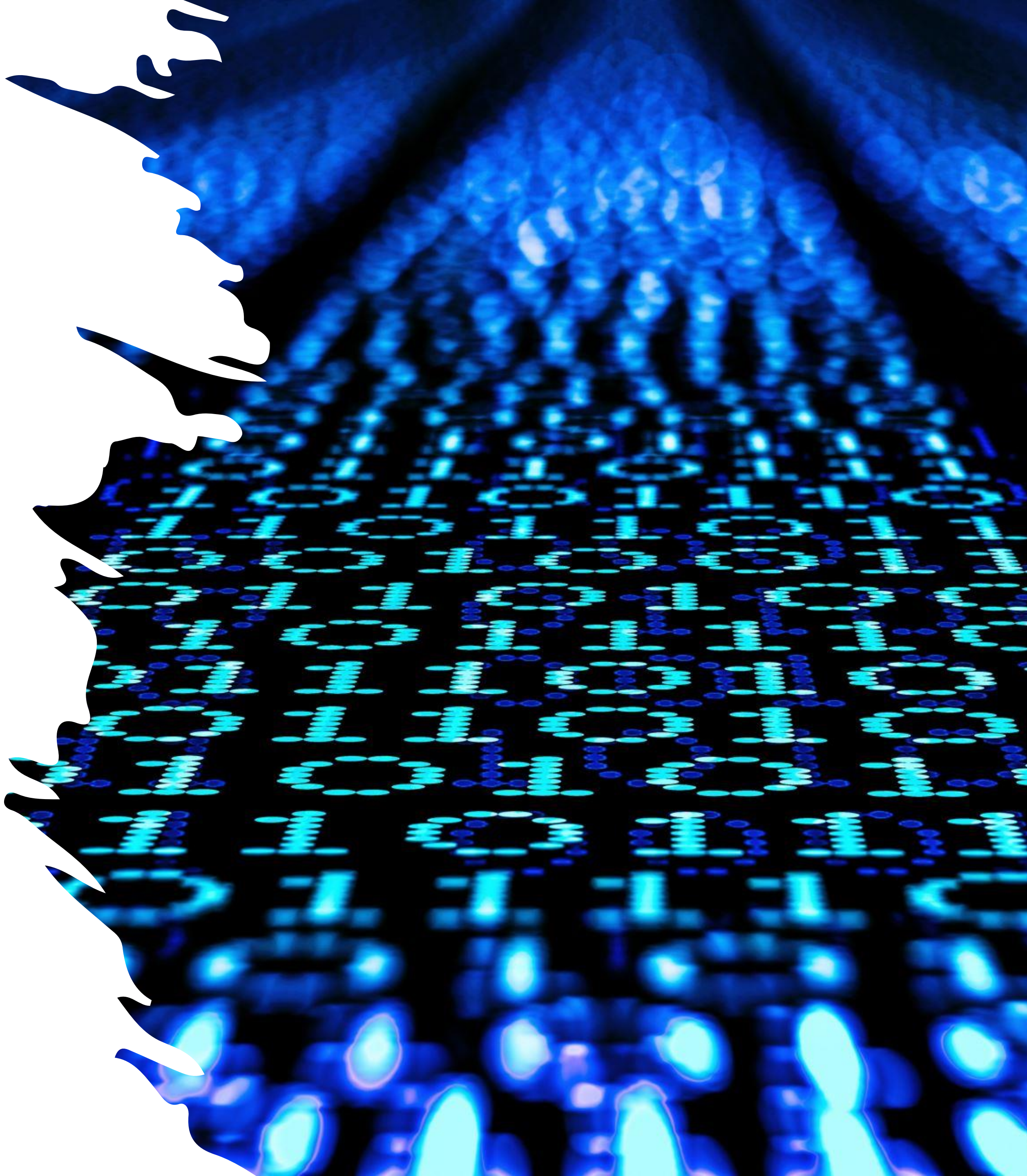
Toolbox for Agentic Workflows

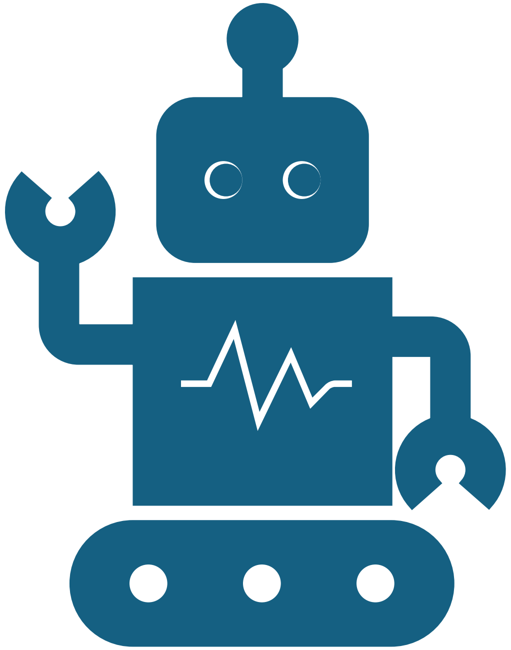
for Cybersecurity use

David Caswell



What are we talking about

- What is Agentic AI
- What are the key differences that make a system “Agentic”
- A couple of cybersecurity examples and use-cases
  - Multi-Document Agents
  - ReAct Agents
- Some observations design principles



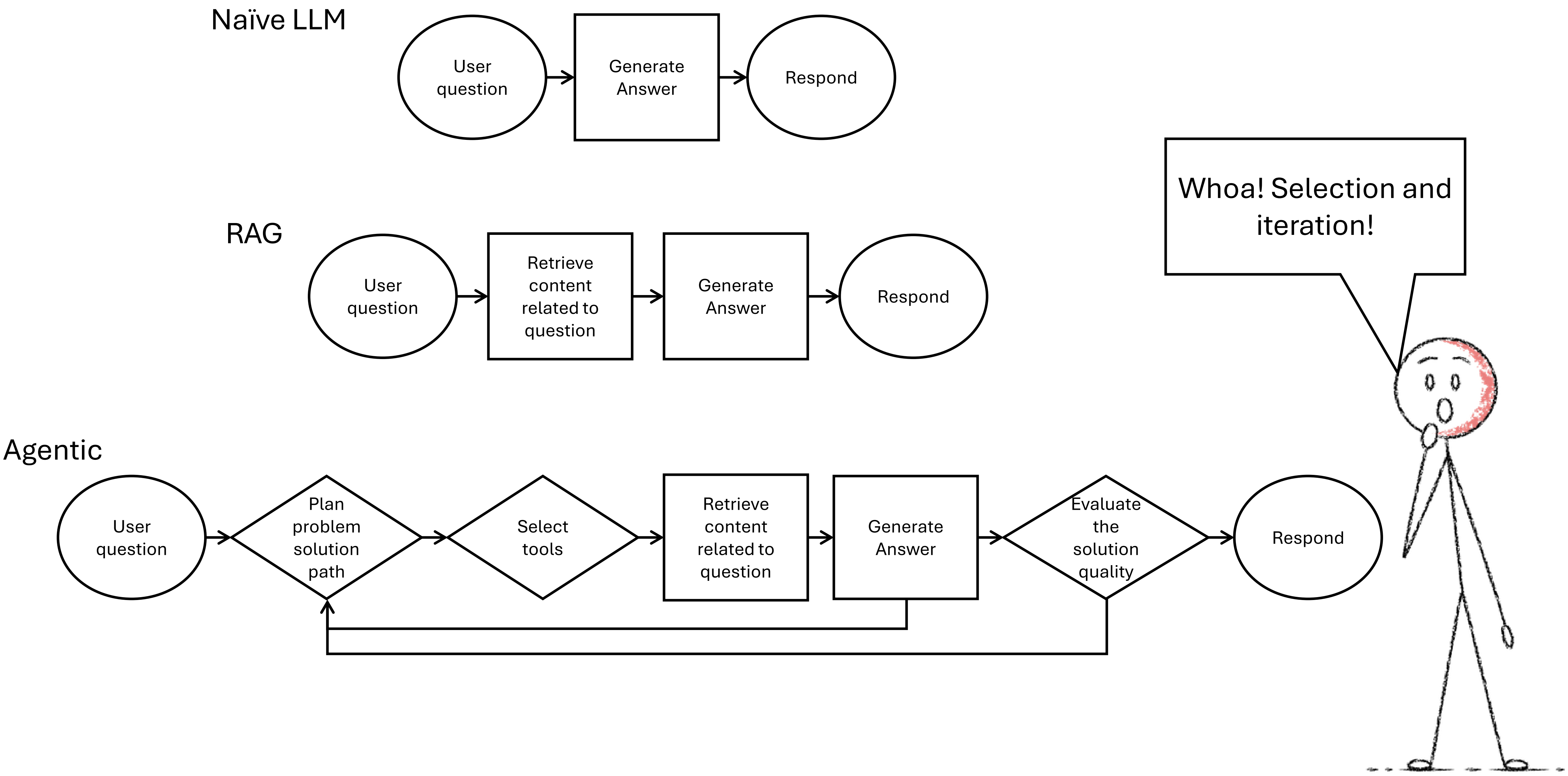
User: “What is Agentic AI”

“Agentic AI refers to artificial intelligence systems that have some level of **autonomous behavior**, **goal-oriented actions**, and **adaptability**. Unlike traditional AI, which follows specific, pre-defined rules or algorithms, agentic AI exhibits qualities typically associated with ‘agency’ — **such as intentionality, responsiveness, and the ability to make decisions toward achieving a specific goal or objective.”**

-ChatGPT 4o

LLM vs Agentic

- Naïve LLM – chat, contextualized knowledge (RAG)
  - Uses pre-trained and/or static data
  - Not good at deterministic functions
  - Serially processes the query
- Agentic LLM
  - Allows non-linear evaluation
  - Leverages tools for deterministic parts
  - Enables a level of self-determination



Naïve LLM

User question

Generate Answer

Respond

RAG

User question

Retrieve content related to question

Generate Answer

Respond

Agentic

User question

Plan problem solution path

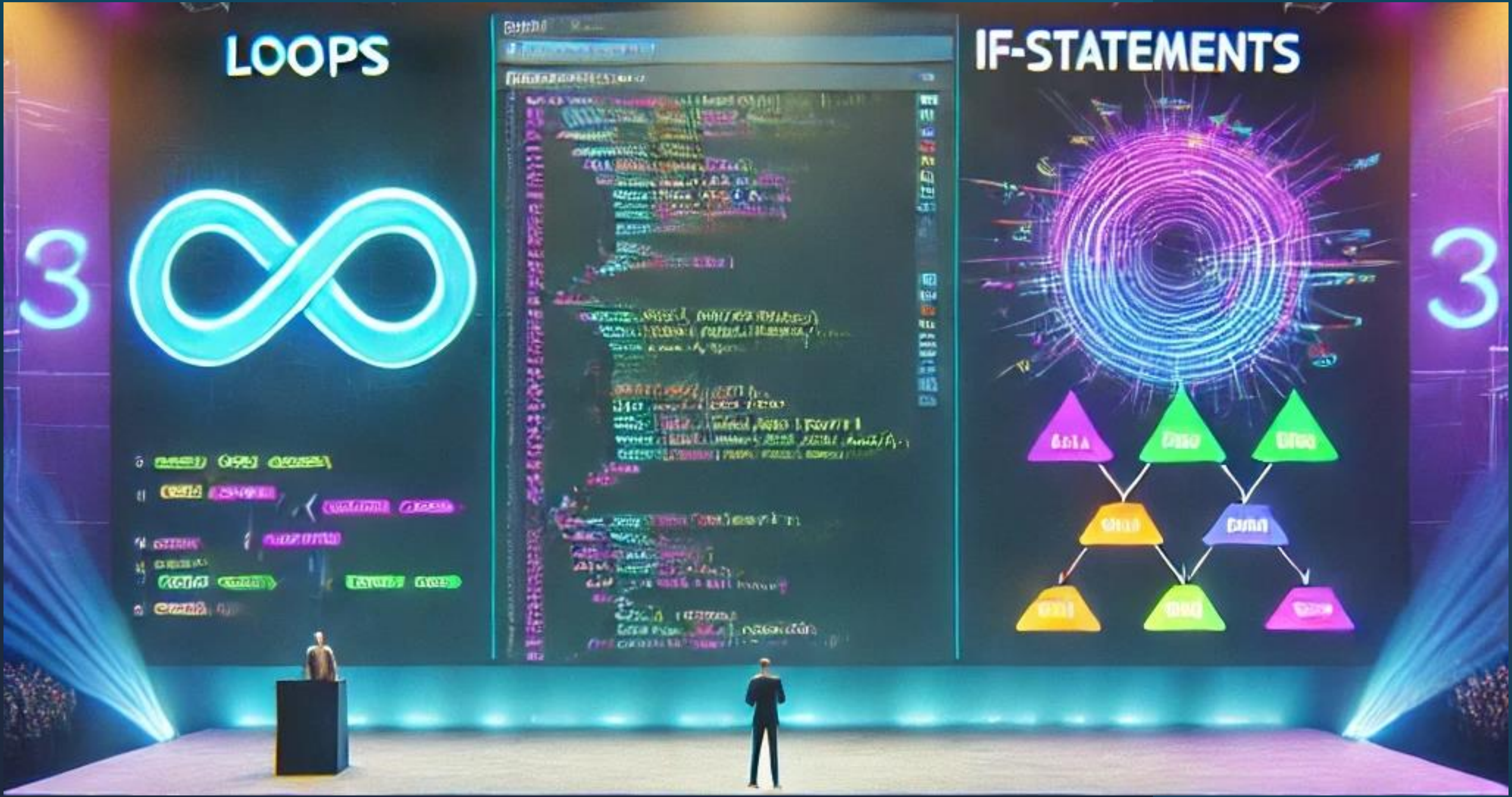
Select tools

Retrieve content related to question

Generate Answer

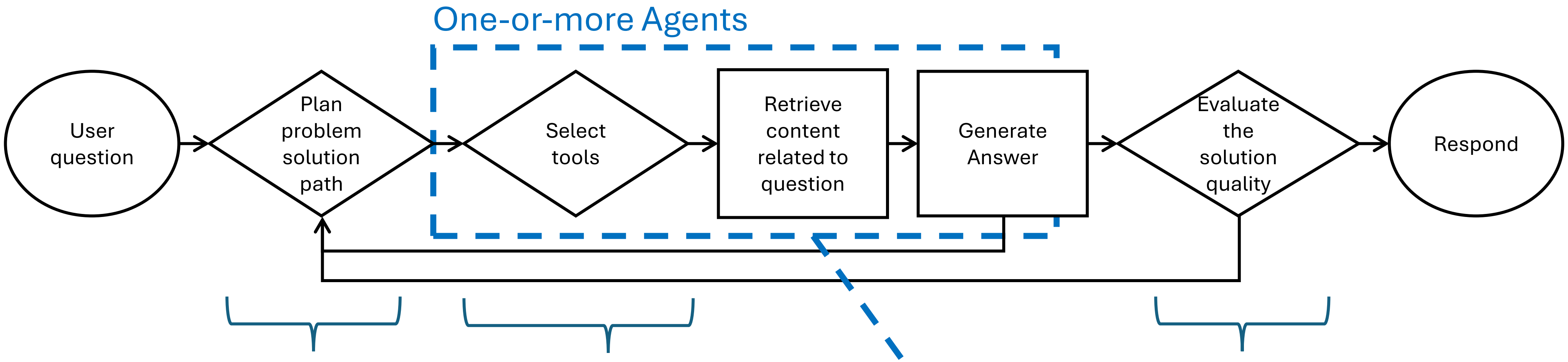
Evaluate the solution quality

Respond



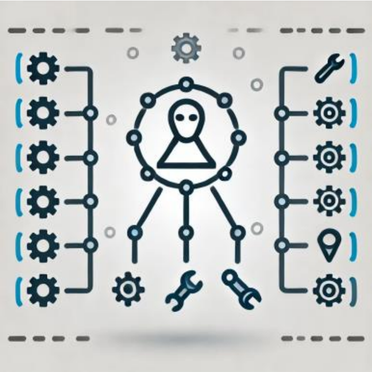
Dall-e

Four capabilities that make a workflow agentic

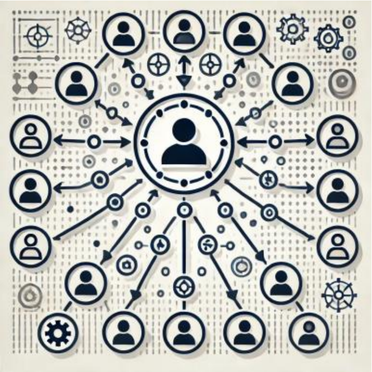

One-or-more Agents
User question
Plan problem solution path
Select tools
Retrieve content related to question
Generate Answer
Evaluate the solution quality
Respond

Planning

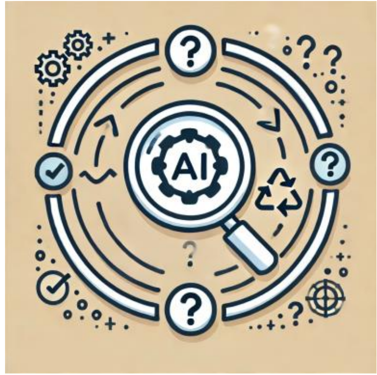
Tool Use



Multi-Agent



Reflection

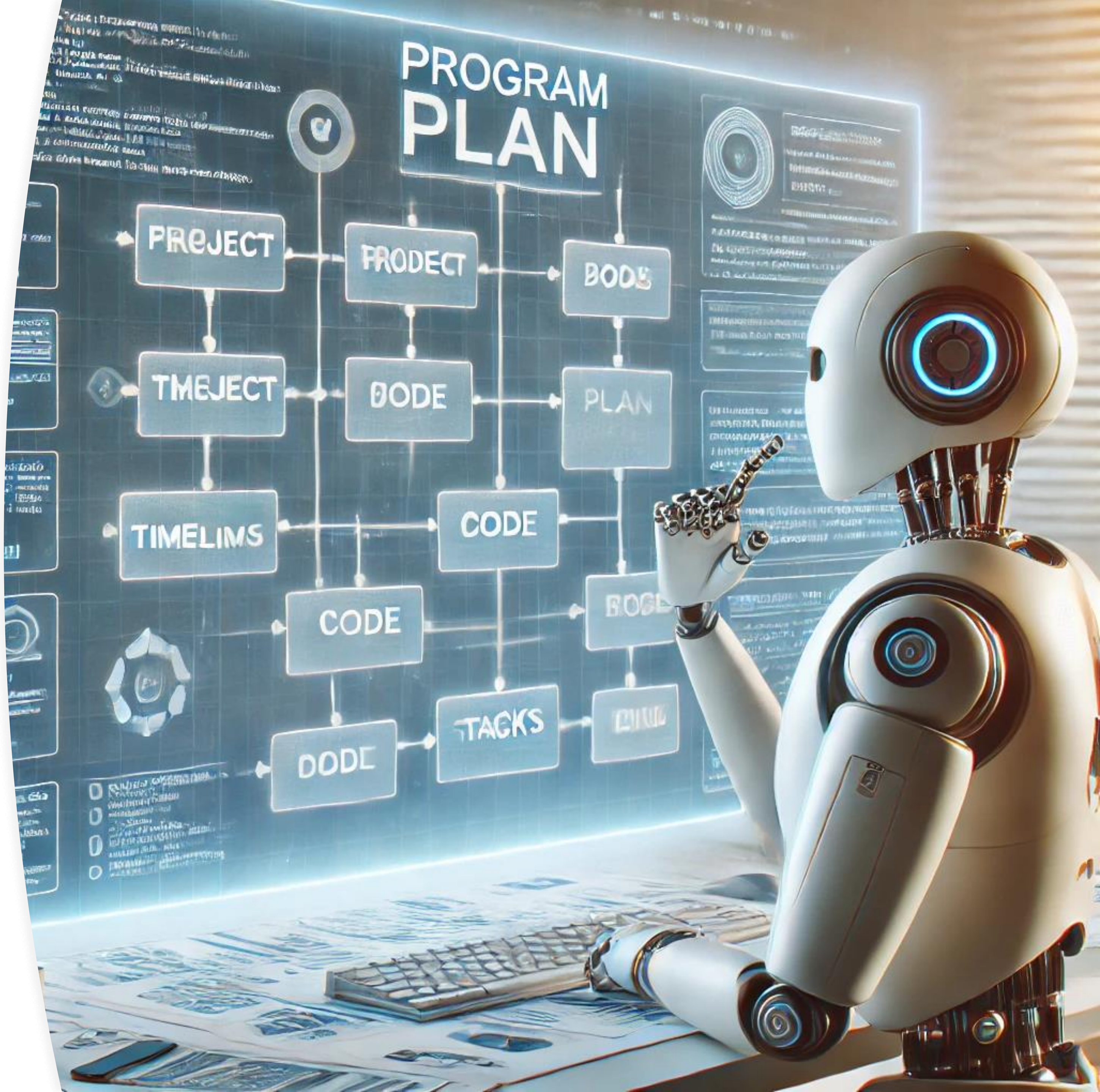


Planning

Determining steps that are required to solve the problem

Some Approaches

- Chain of Thought - decomposes problems into intermediate steps and solves
- ReAct (Reasoning and Acting) – structured decomposition of inputs to steps and function calls [blog, LlamaIndex]
- ReWOO (Reasoning with Open Ontology) – brings in info from various knowledge domains
- Language Agent Tree Search (LATS) - planning/acting/reasoning within Monte Carlo Tree Search

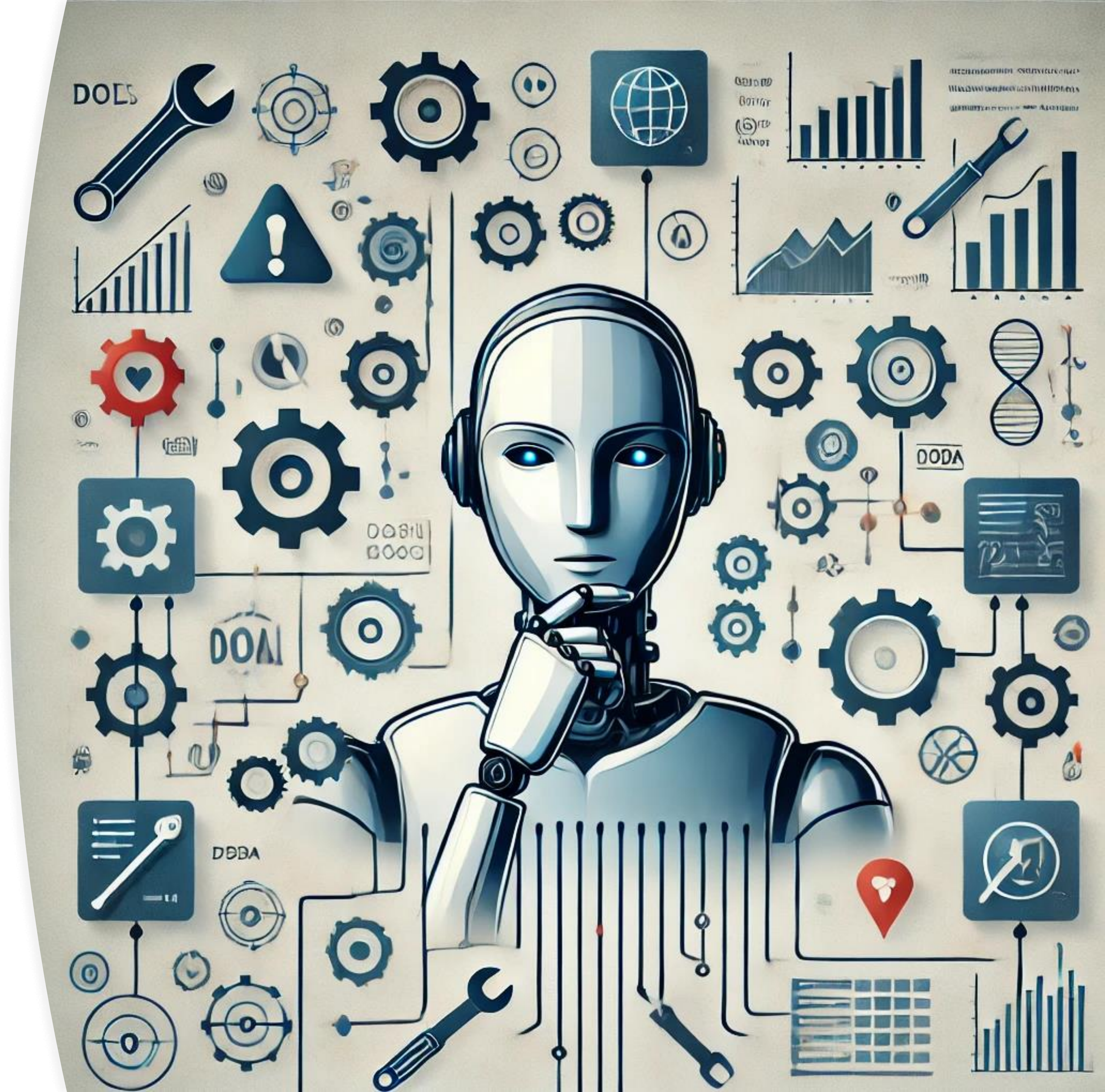


Tool Use

- Agent understands a variety of different tool options
- Determines what tools are applicable to the question
- Calls the tools and incorporates the response

<u>Some Approaches</u>

- Function Calling Agents
- Chain-of-Abstraction - dynamic function generation as part of the prompting process [llamaIndex]



Multi-Agent

Use of multiple agents for expert services or for feedback

Some Approaches

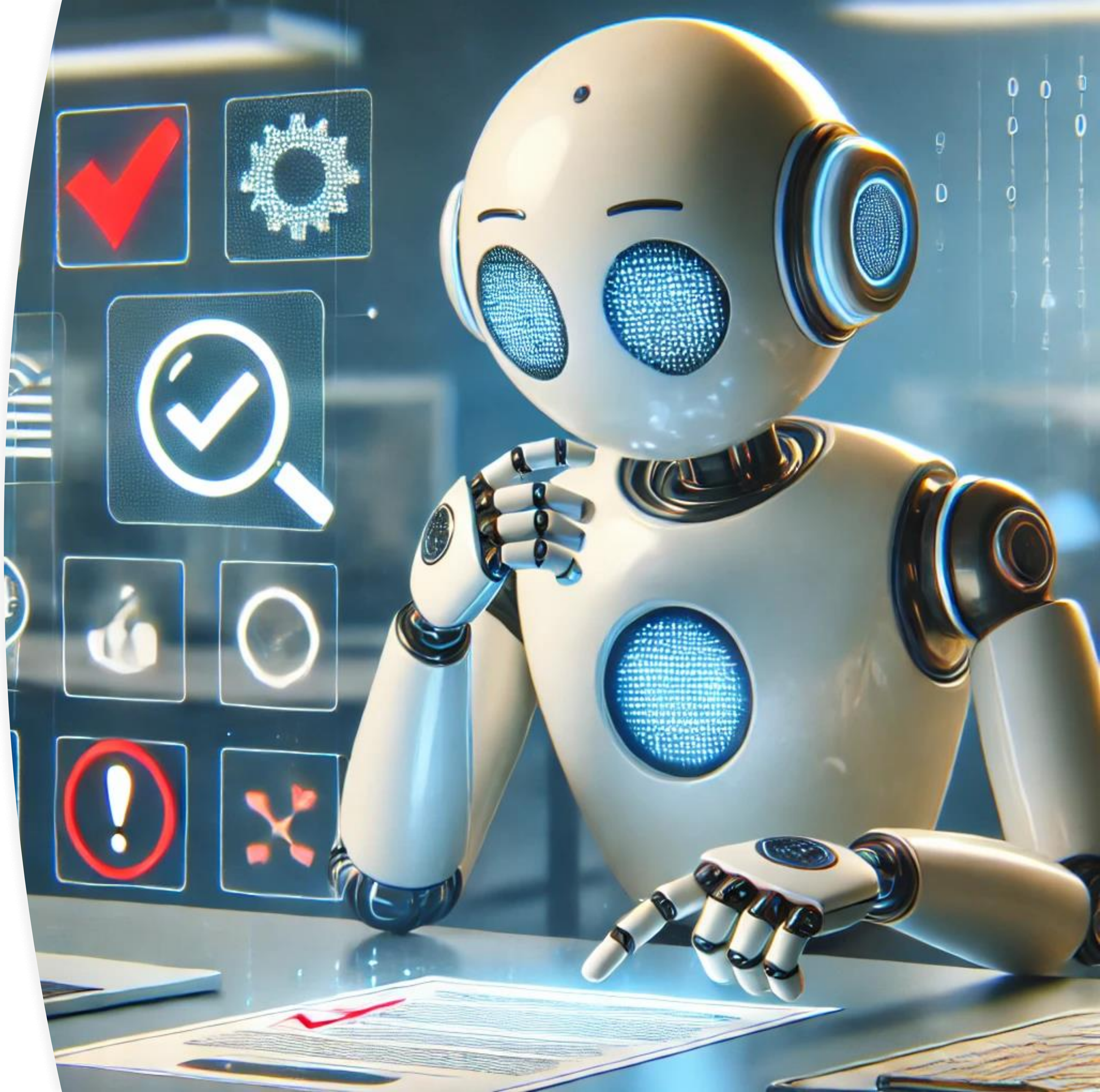
- Collaborative [AutoGen, langgraph]
- Adversarial
- Supervised
- Hierarchical – [LLMCompiler, LlamaIndex, Expert System Delegation, LLM Orchestration]

Reflection

Agent evaluates the output to determine if it needs additional work

Some Approaches

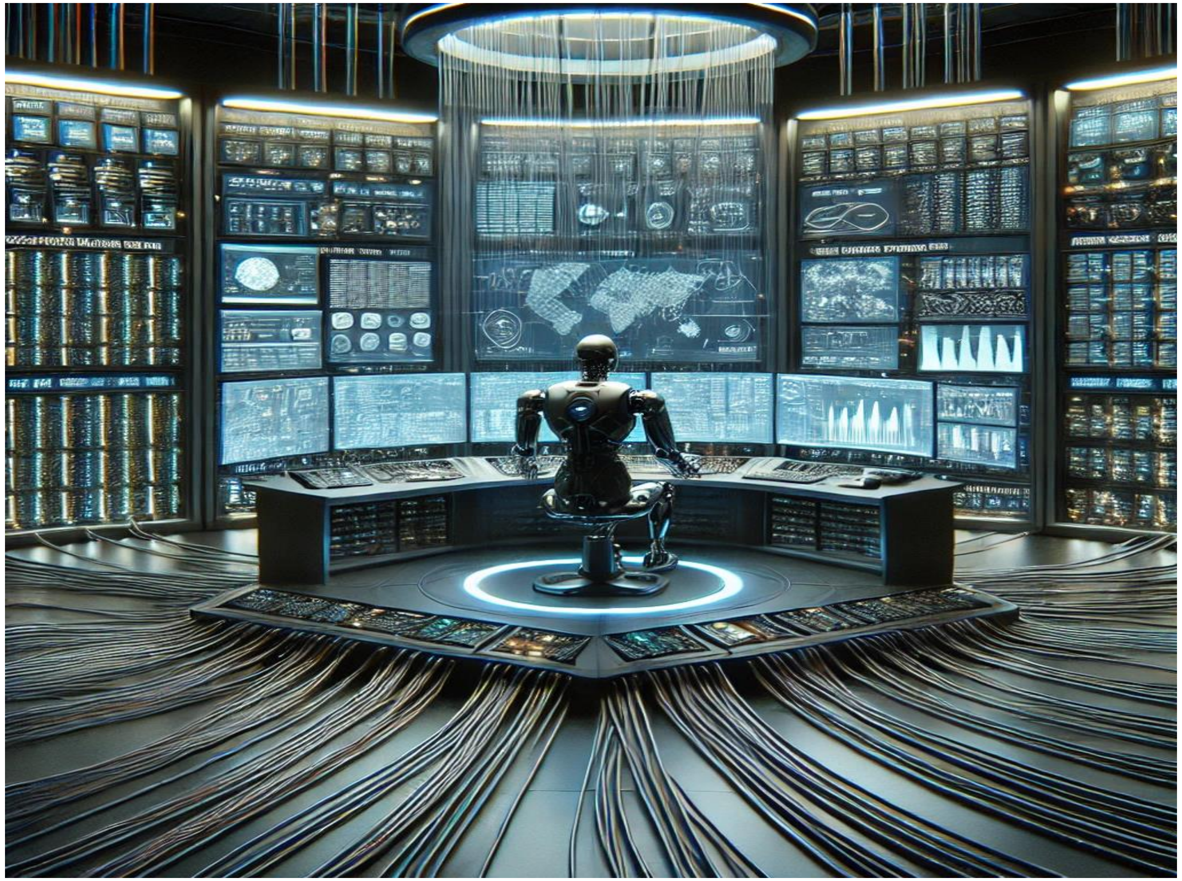
- Self-RAG: evaluate and refine its own output through a self-critique loop [original, original-github, langgraph]
- Self-Refine
- Reflexion
- CRITIC



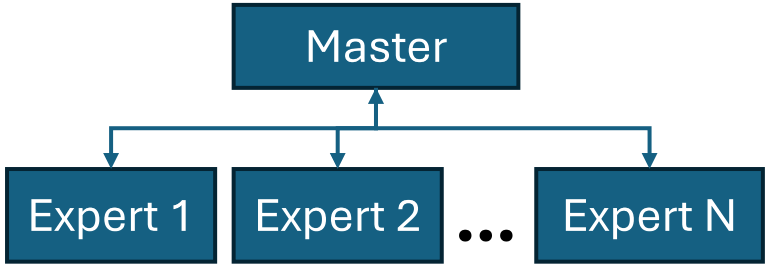
Example Use-Cases

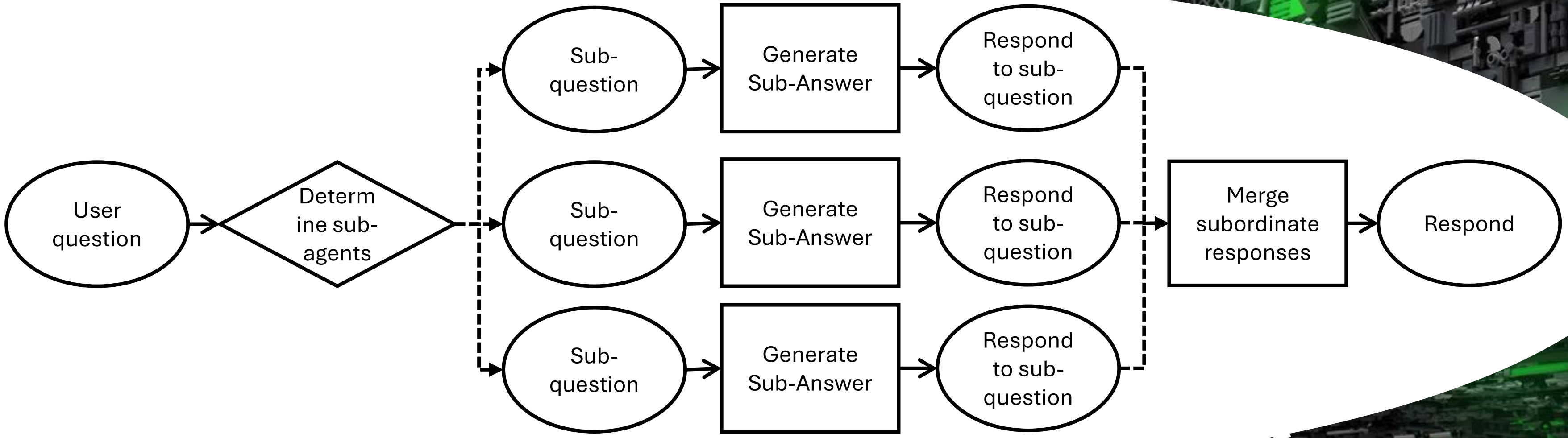
ISSO Standards Service

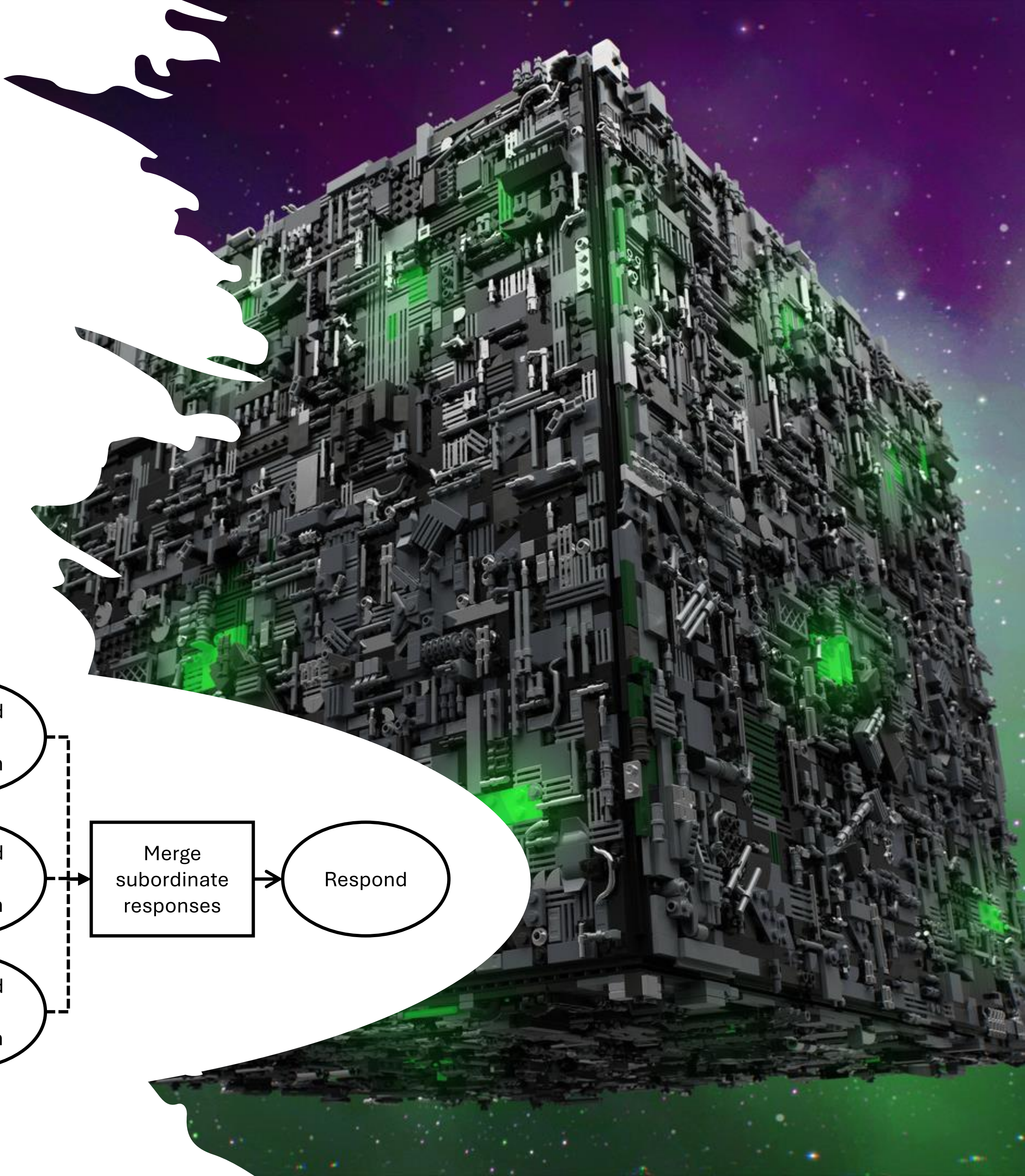
ISSO Standards Service



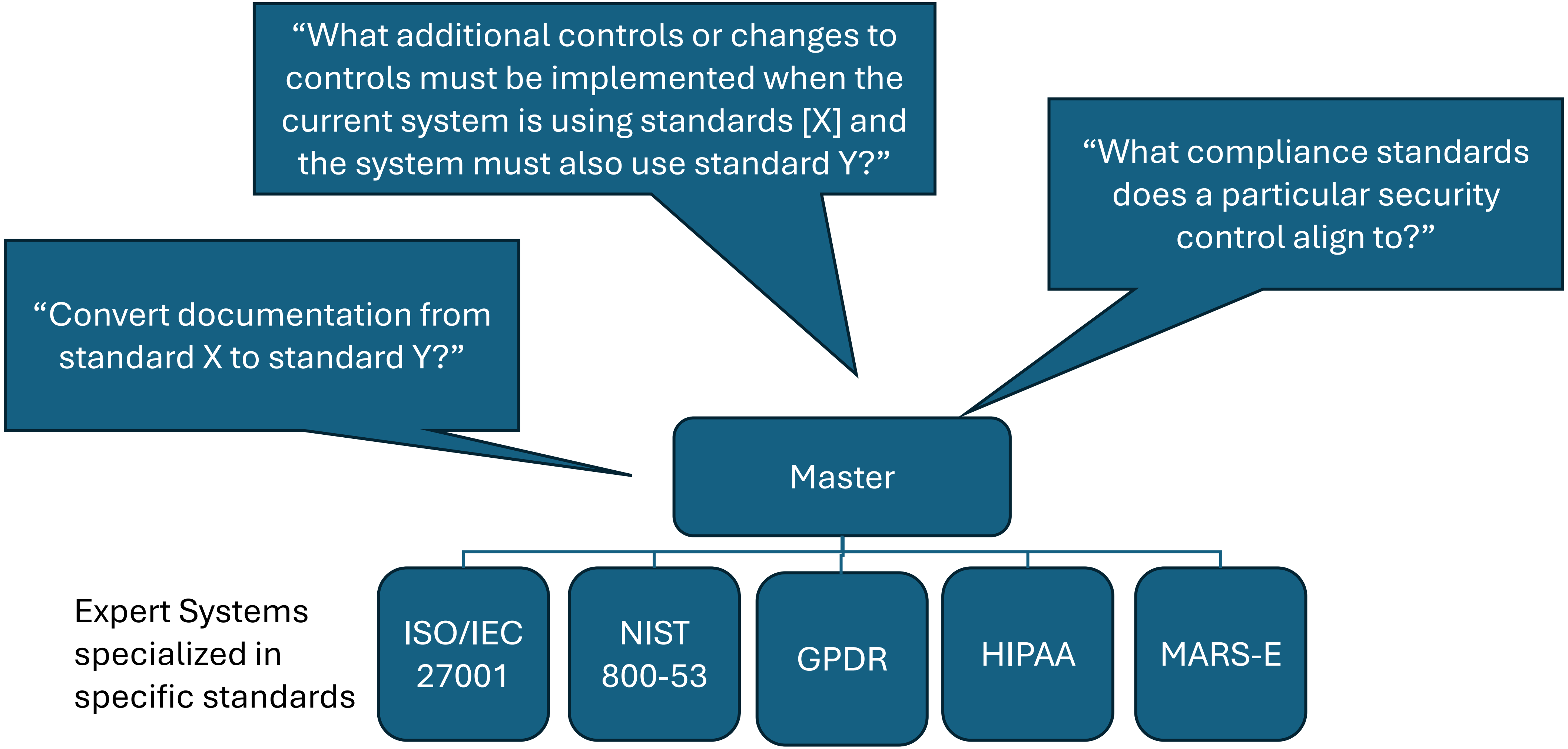
Multi-Agent System
(Hierarchical Assimilator)







Example: Info-Sec Standards References



"What additional controls or changes to controls must be implemented when the current system is using standards [X] and the system must also use standard Y?"

"What compliance standards does a particular security control align to?"

"Convert documentation from standard X to standard Y?"

Master

Expert Systems specialized in specific standards

ISO/IEC 27001

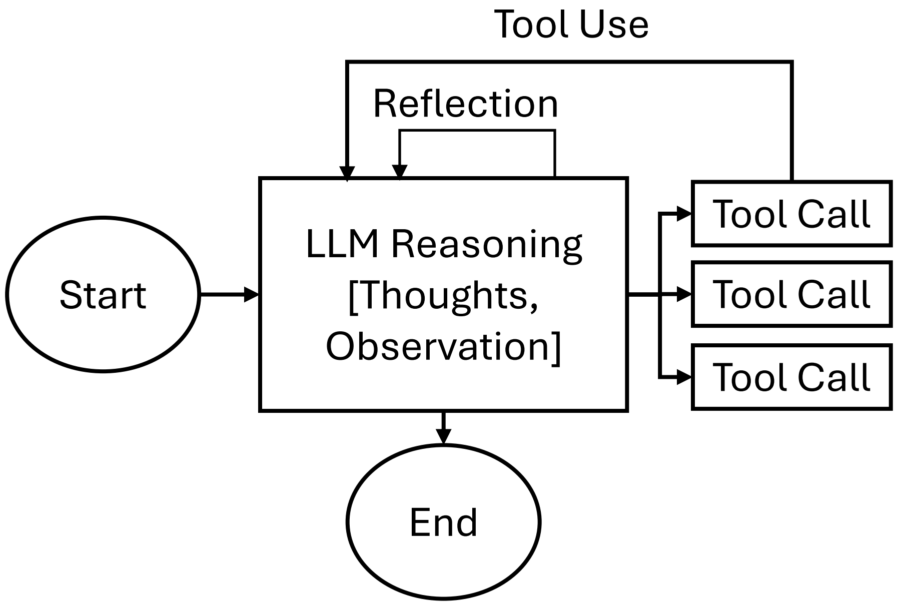
NIST 800-53

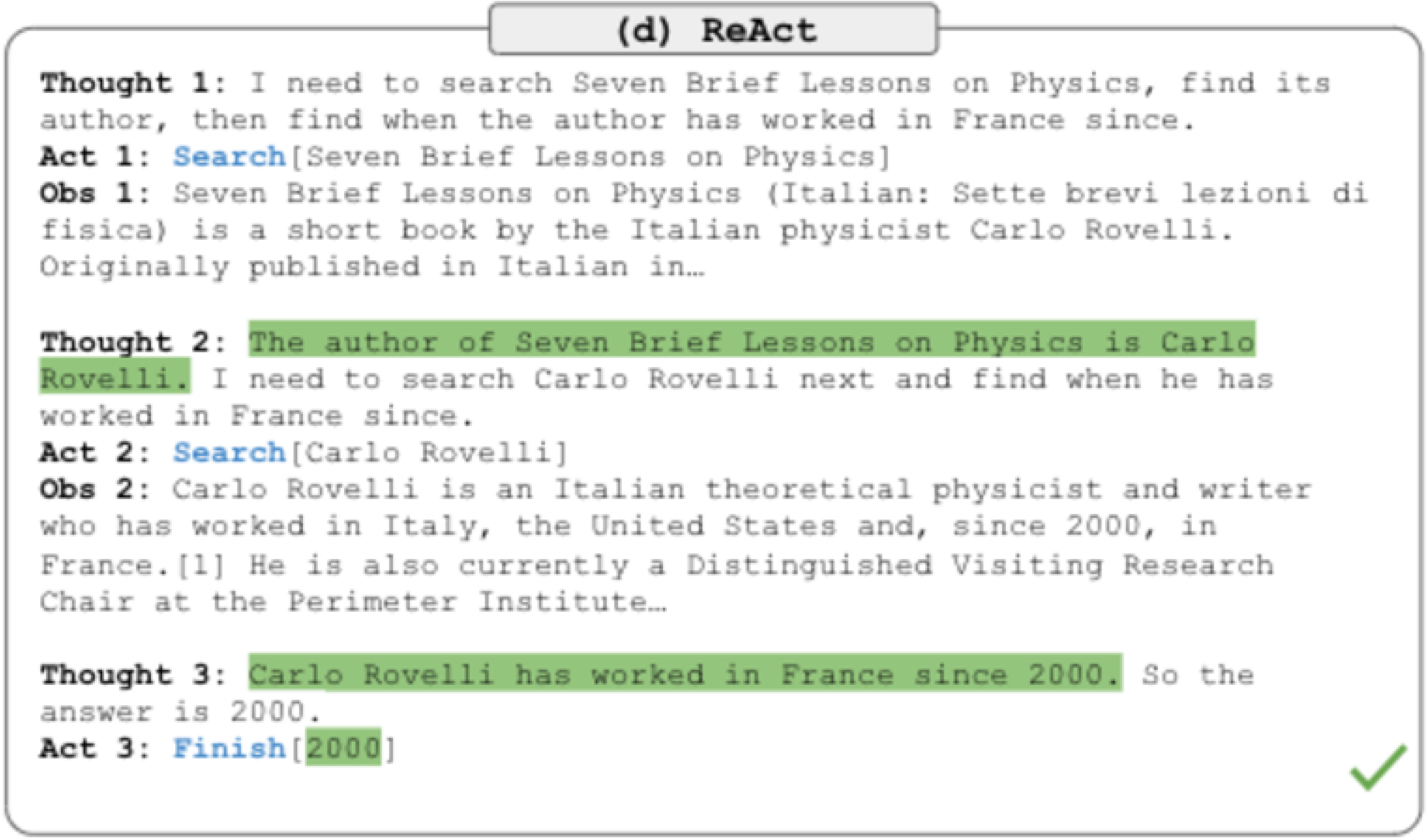
GPDR

HIPAA

MARS-E

Reasoning/Acting (ReAct)





(d) ReAct

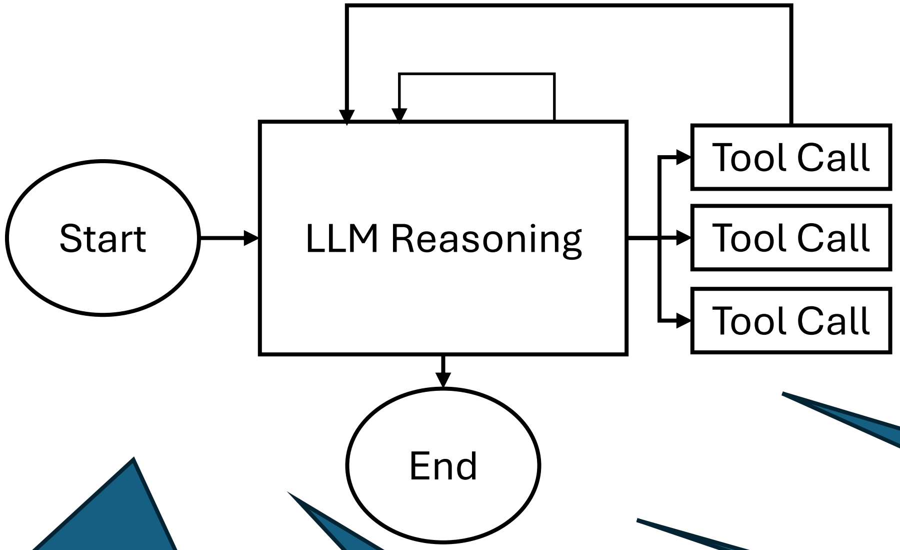
Thought 1: I need to search Seven Brief Lessons on Physics, find its author, then find when the author has worked in France since.
Act 1: Search[Seven Brief Lessons on Physics]
Obs 1: Seven Brief Lessons on Physics (Italian: Sette brevi lezioni di fisica) is a short book by the Italian physicist Carlo Rovelli. Originally published in Italian in…

Thought 2: The author of Seven Brief Lessons on Physics is Carlo Rovelli. I need to search Carlo Rovelli next and find when he has worked in France since.
Act 2: Search[Carlo Rovelli]
Obs 2: Carlo Rovelli is an Italian theoretical physicist and writer who has worked in Italy, the United States and, since 2000, in France.[1] He is also currently a Distinguished Visiting Research Chair at the Perimeter Institute…

Thought 3: Carlo Rovelli has worked in France since 2000. So the answer is 2000.
Act 3: Finish[2000]

✓

Example: Security Operations Center Agent



- Security service relations
- Attack path relationships

- SIEM
- Security Device
- CMDB

- Threat Intelligence
- IoC Libraries
- Threat Libraries

"What is the most at-risk point of access for the set of services being monitored?"

"Are there any malicious activities consistent with the behavior of a particular APT?"

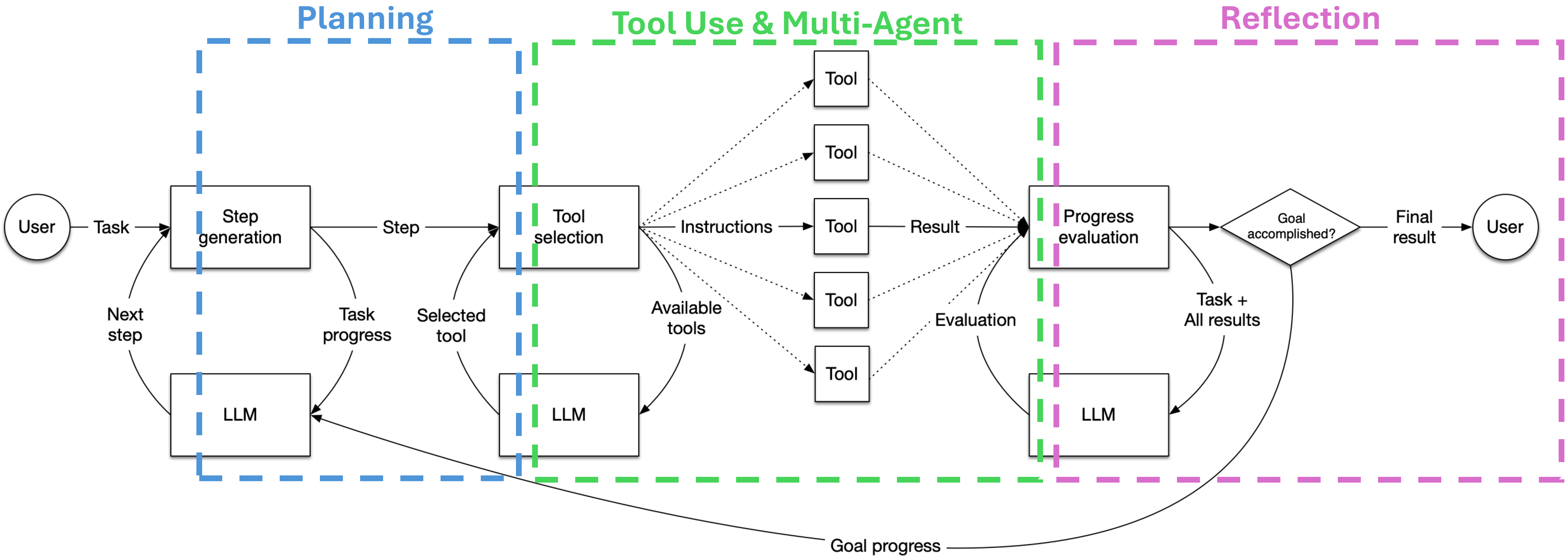
"What are likely next steps for an attack assuming a threat actor made access at a particular point?"

"What are IOCs that we should prioritize?"

Key Design Principles

- Extract any deterministic functions wherever possible, even if only quasi-deterministic such as evaluation schemas
- Don’t assume the system “knows” anything- provide critical knowledge where possible (content relationships, function/tool effects, etc)
- Assume the LLM is one opinion, get lots of opinions and try to merge when feasible particularly in the planning stages
- When possible, focus each LLM call on as small and compartmentalizable question as you can craft it
- Determine the minimal level of key feedback points for human-in-the-loop
- Build in traceability at the onset

Putting it all together



Frameworks:

LangGraph

crewai

LangChain

LlamaIndex

Diagram Source: LlamaIndex

Additional Resources

Function Considerations

- Math Functions
- Weather
- Stock Values
- News
- Web
- Commercial Info
- Cyber Threat Intel
- Chat Feeds
- Sentiment Analyzers
- Classification Services
- Image Parsers

Query Improvement Patterns

- HyDE (Hypothetical Document Embedding) – create a hypothetical document from a larger document and use for downstream embedding
- Query-Rewriting – try to rewrite the query to help retrieval
- Step-back Prompting - generalizes the query to help with retrieval
- Human-in-the-loop interaction – states that allow for human intervention in the workflow

Data Acquisition Sources

Static Data Store (SQL, NoSQL)	Query history, static data that can be accessed generally via a bespoke Tool
Deterministic Functions (Tools, APIs)	Conduct deterministic calculate and return result (math, static data analysis,…)
Vector Database	Search for content by semantic similarity based on embedding function
Graph Database	Search for relationship of information or across shared attributes
Dynamic Functions (Coming Soon)	use a subordinate agent to dynamically build a query to retrieve data from more complex data stores

Other References

- Four AI Agent Strategies That Improve GPT-4 and GPT-3.5 Performance
- Choosing Between LLM Agent Frameworks | by Aparna Dhinakaran | Sep, 2024 | Towards Data Science
- Agentic Workflows in 2024: The ultimate guide, Vellum.ai
- Qineng Wang[†], Zihao Wang[†], Ying Su, Hanghang Tong, and Yangqiu Song, Rethinking the Bounds of LLM Reasoning: Are Multi-Agent Discussions the Key?, 2402.18272v1
- [2406.14550] GraphReader: Building Graph-based Agent to Enhance Long-Context Abilities of Large Language Models
- Understanding Agentic Concepts in LLM Workflows | by Pankaj | Oct, 2024 | Medium
- langgraph/docs/docs/concepts/agentic_concepts.md at main · langchain-ai/langgraph
- Advanced RAG 06: Exploring Query Rewriting | by Florian June | Medium
- langgraph/docs/docs/concepts/agentic_concepts.md at main · langchain-ai/langgraph
- https://github.com/langchain-ai/langgraph/blob/main/docs/docs/concepts/agentic_concepts.md
- What is Agentic AI Tool Use Pattern? - Analytics Vidhya
- Choosing Between LLM Agent Frameworks | by Aparna Dhinakaran | Sep, 2024 | Towards Data Science
- GraphRAG: Unlocking LLM discovery on narrative private data - Microsoft Research
- Knowledge Graph vs. Vector RAG: Benchmarking, Optimization Levers, and a Financial Analysis Example