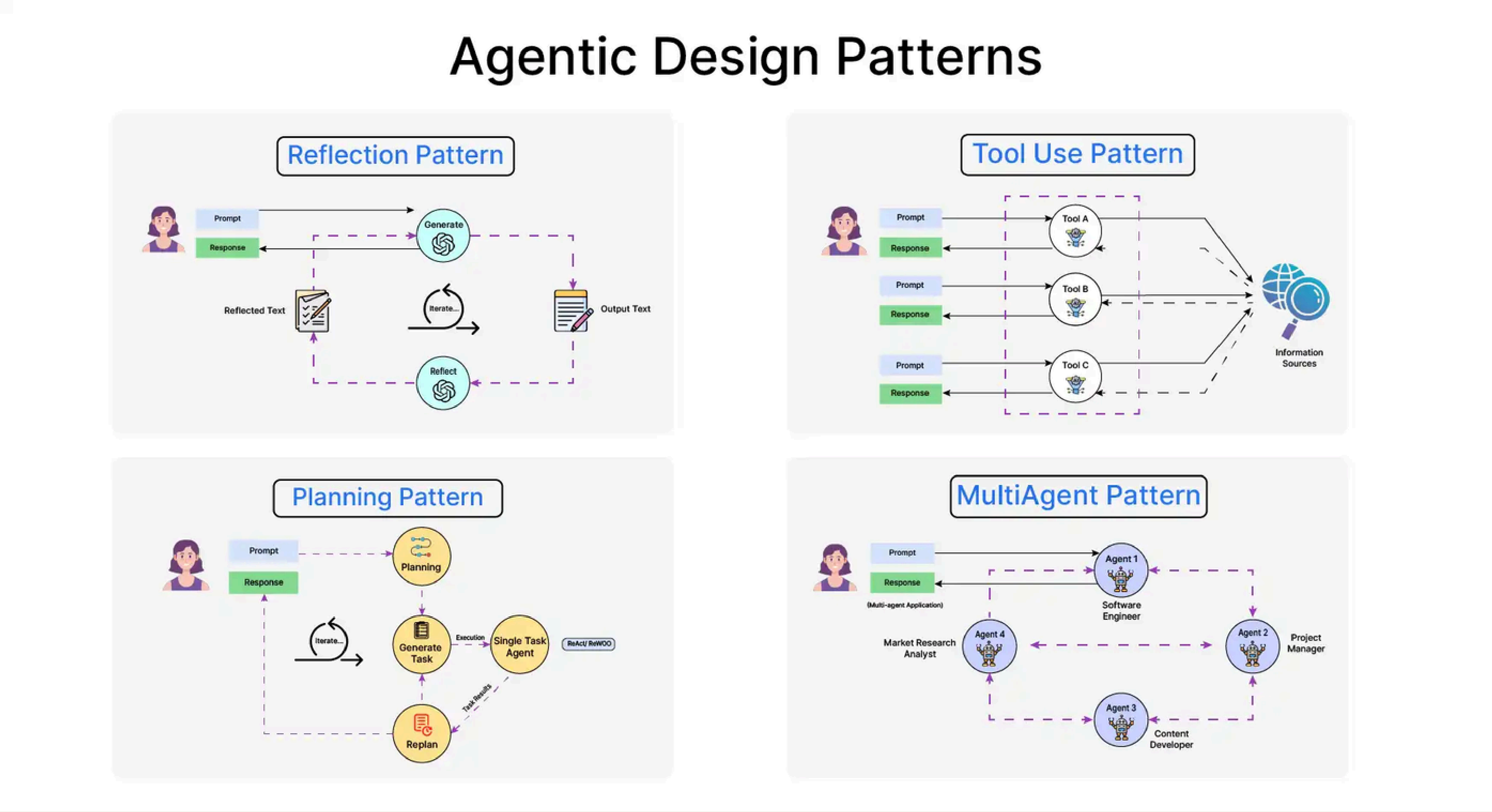
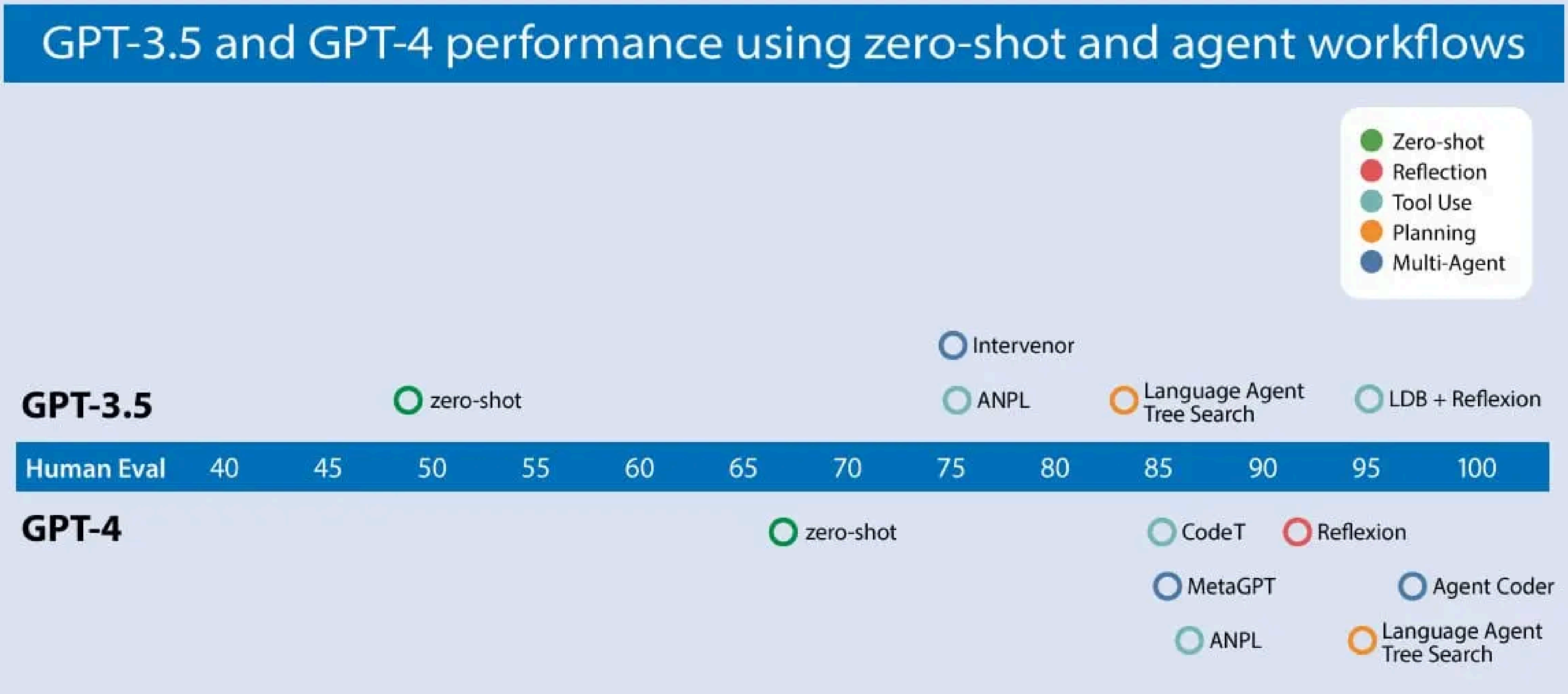
Top 4 Agentic AI Architecture Design Patterns



Dipanjan (DJ)

Why AI Agents?



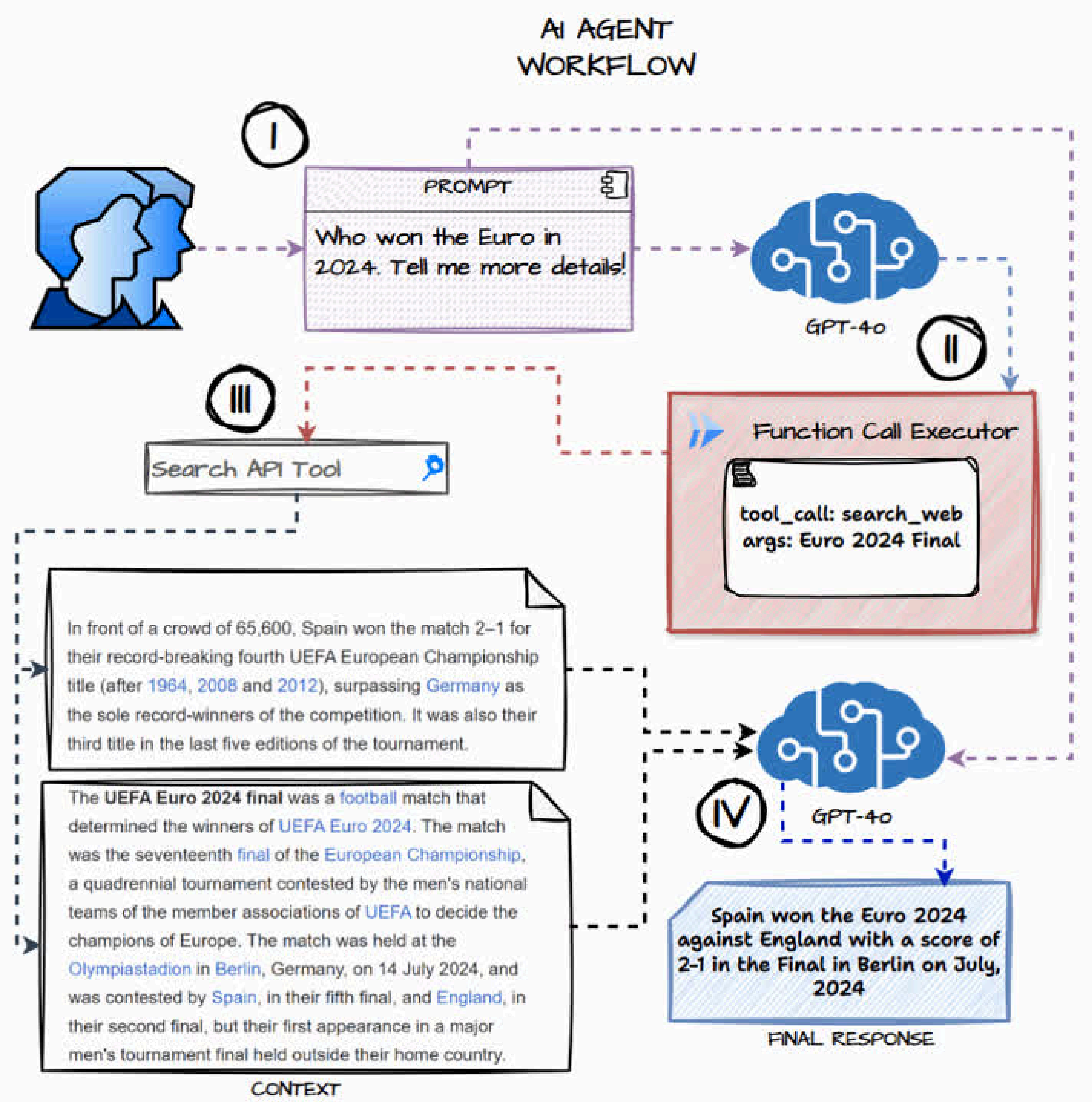
Performance of GPT-3.5 and GPT-4 (zero-shot) on HumanEval, along with algorithms that use agent workflows on top of GPT-3.5 or GPT-4. Thanks to Joaquin Dominguez and John Santerre for help with this analysis.

- **LLMs have increased performance using Agentic workflows**

- **GPT-3.5 powered Agentic AI Systems achieved up to 95.1% on HumanEval coding benchmark**

- **Easy to connect LLMs, tools and prompts along with external data to build simple and complex agentic workflows**

How AI Agents work?



Who won the Euro in 2024. Tell me more details!

tool_call: search_web
args: Euro 2024 Final

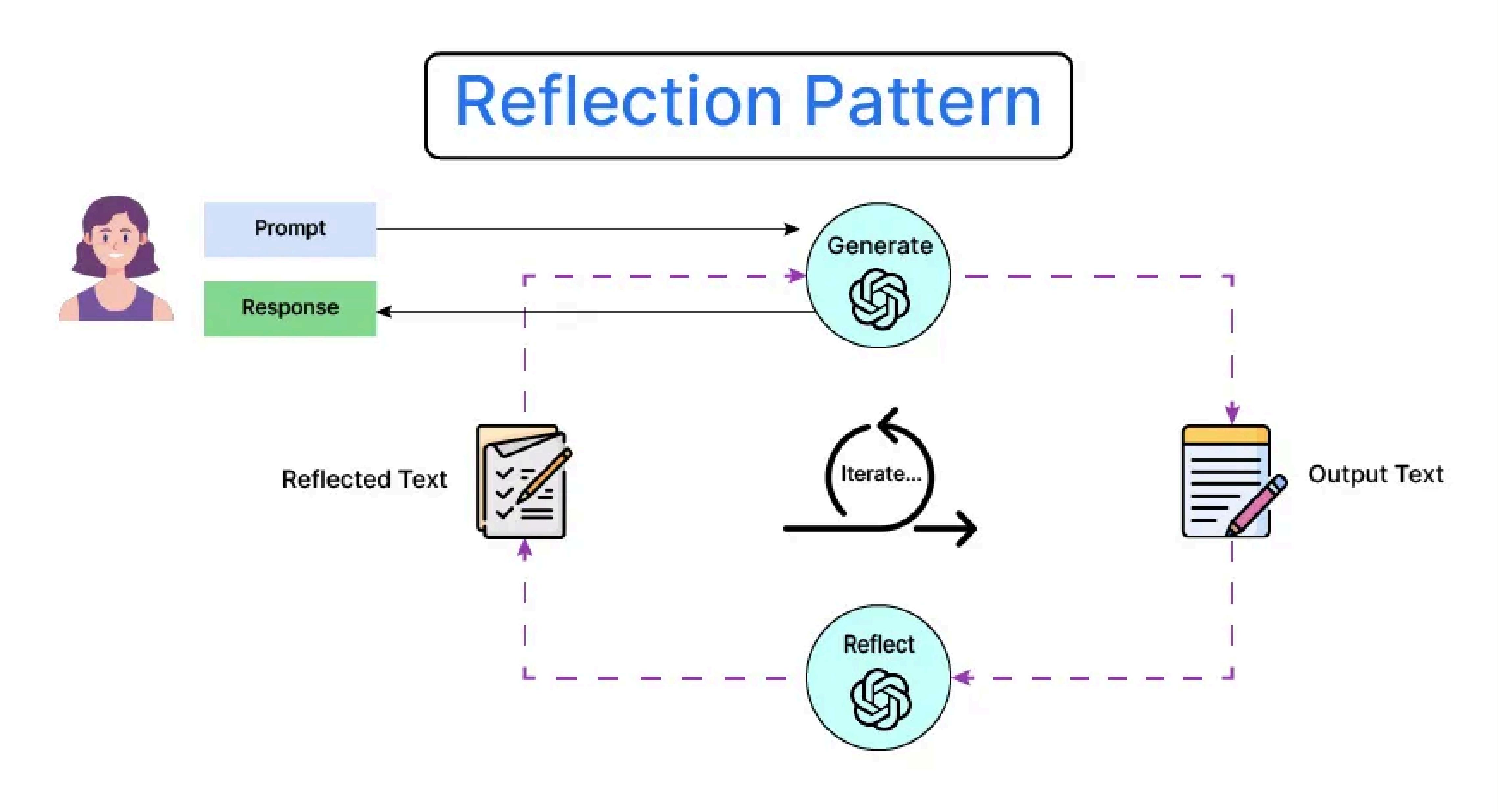
In front of a crowd of 65,600, Spain won the match 2–1 for their record-breaking fourth UEFA European Championship title (after 1964, 2008 and 2012), surpassing Germany as the sole record-winners of the competition. It was also their third title in the last five editions of the tournament.

The **UEFA Euro 2024 final** was a football match that determined the winners of UEFA Euro 2024. The match was the seventeenth final of the European Championship, a quadrennial tournament contested by the men's national teams of the member associations of UEFA to decide the champions of Europe. The match was held at the Olympiastadion in Berlin, Germany, on 14 July 2024, and was contested by Spain, in their fifth final, and England, in their second final, but their first appearance in a major men's tournament final held outside their home country.

Spain won the Euro 2024 against England with a score of 2-1 in the Final in Berlin on July, 2024

- **I. Start with the initial instruction prompt**
- **II. LLM processes the prompt and decides what tool to call based on the available tools (function calling)**
- **III. The specific tool and tool arguments are then executed by the system to get some new context information - in this case search results**
- **IV. This new information is passed along with the initial prompt to the LLM to give a response or call more tools**

Pattern 01 - Reflection



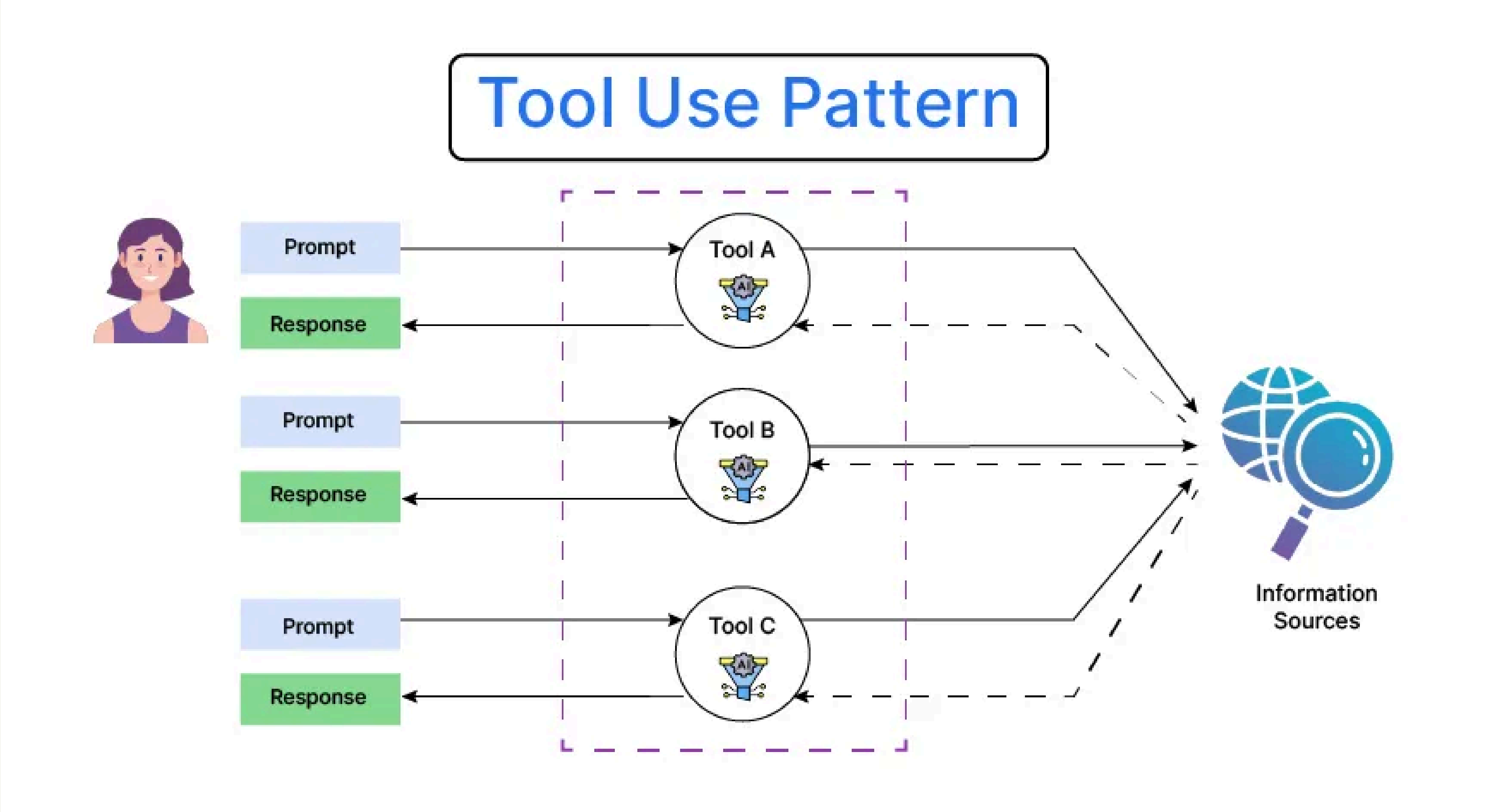
- Reflection Pattern focuses on improving the AI agent's ability to evaluate and refine its own outputs

- This self-critique loop of generating and reflecting is not limited to a single iteration

- System can repeat the reflection process as many times as necessary to achieve a refined result

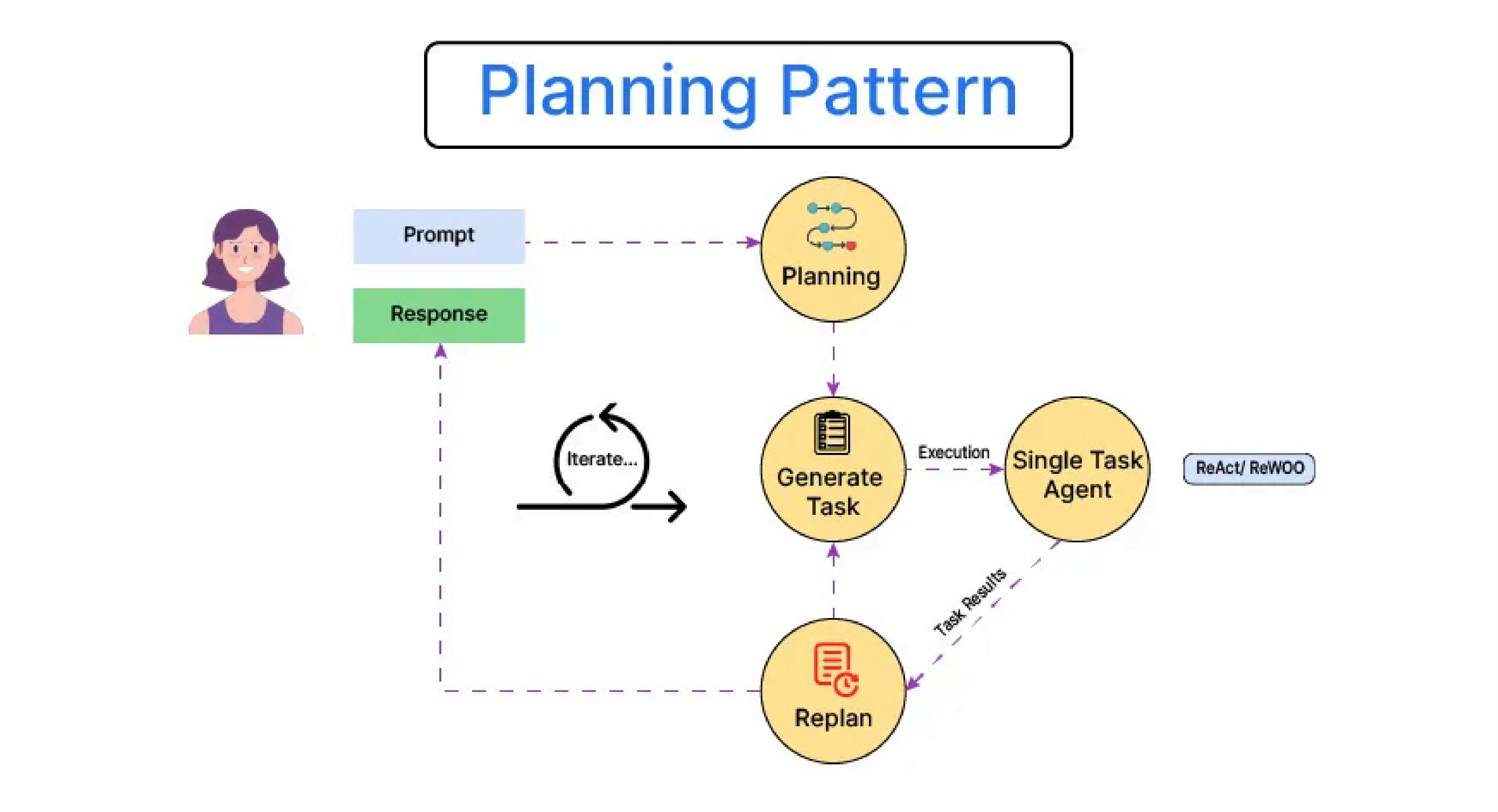
- Self-Reflection RAG is a popular Agentic RAG system using this pattern

Pattern 02 - Tool Use



- Tool Use Pattern significantly broadens an LLM's capability by allowing it to interact with external tools and APIs
- Agentic AI systems using this pattern can access databases, search the web, or even execute complex functions via programming languages like Python
- Very useful to augment RAG systems with capabilties to answer questions based on real-time searches
- Most useful pattern to automate diverse tasks using a set of tools

Pattern 03 - Planning



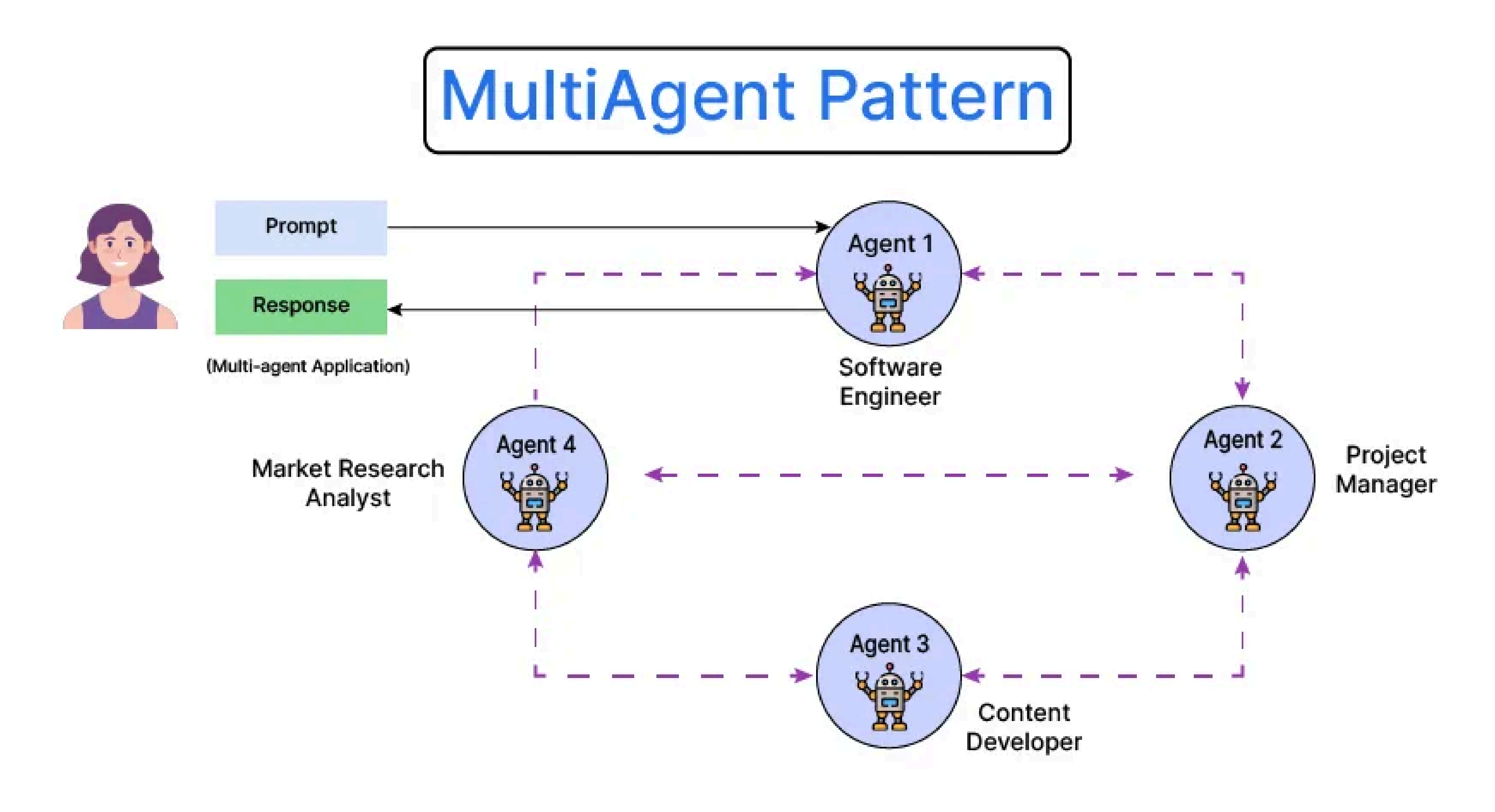
- **Planning Pattern enables an LLM to break down large, complicated tasks into smaller, more manageable components**

- **Planning equips an agent with the ability to react to requests and strategically structure the steps before execution**

- **Create a roadmap of subtasks, determining the most efficient path to completion**

- **ReAct (Reasoning and Acting) and ReWOO (Reasoning With Open Ontology) further extend this approach by integrating decision-making and contextual reasoning into the planning process**

Pattern 04 - Multi-Agent

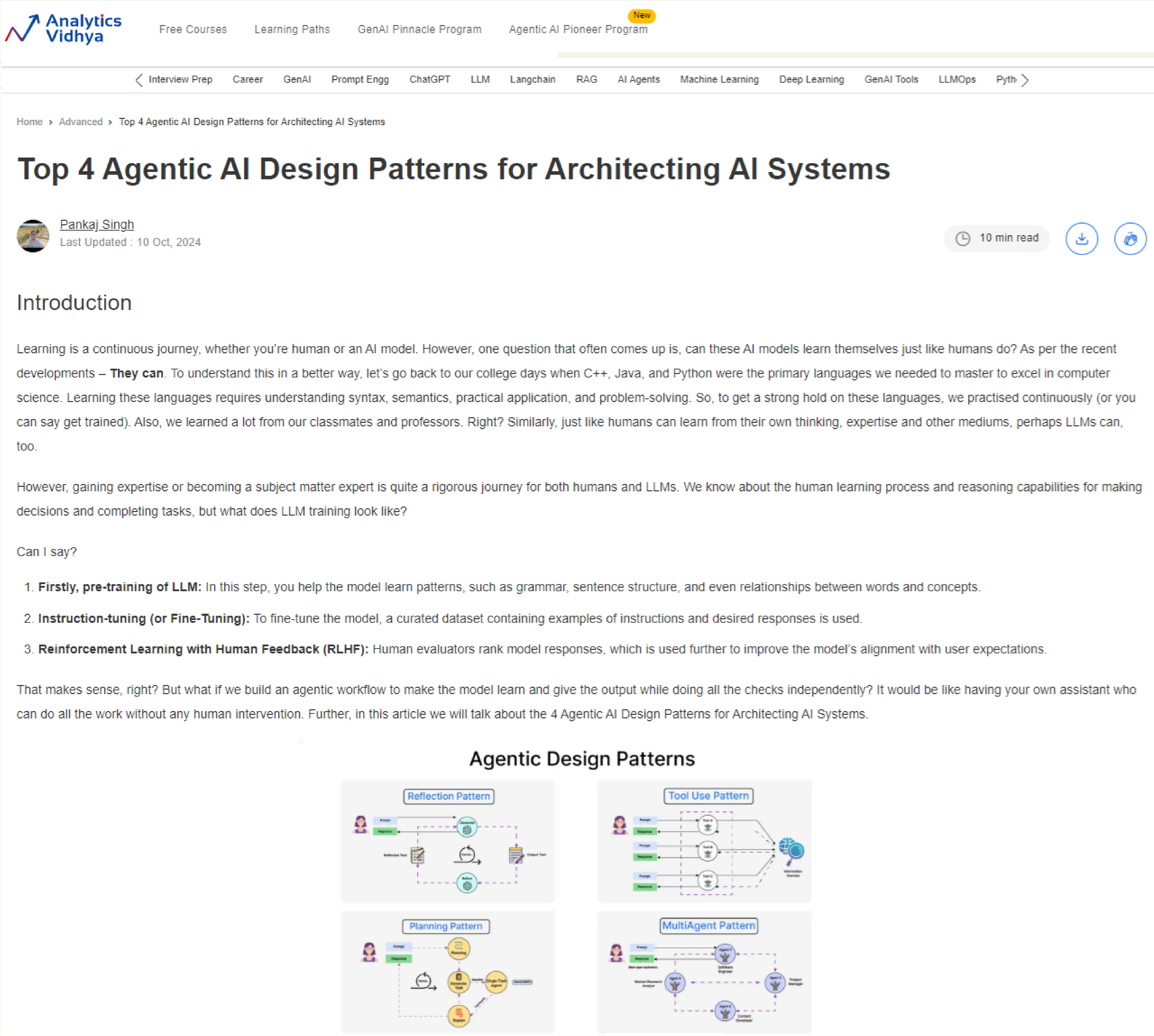


- **Multi-Agent Pattern builds upon the concept of delegation, akin to project management in human teams**

- **Involves assigning different agents (which are instances of an LLM with specific roles and tools) to handle various subtasks**

- **Several types of multi-agent system patterns:**
  - **Collaborative Agents:** Multiple agents work together on different parts of a task, sharing progress and building toward a unified result
  - **Supervised Agents:** A central supervisor agent manages other agents, coordinating their activities and verifying results to ensure quality
  - **Hierarchical Teams:** A structured system where higher-level agents oversee lower-level agents

Detailed Article



Top 4 Agentic AI Design Patterns for Architecting AI Systems

Pankaj Singh
Last Updated : 10 Oct, 2024

Introduction

Learning is a continuous journey, whether you're human or an AI model. However, one question that often comes up is, can these AI models learn themselves just like humans do? As per the recent developments – **They can**. To understand this in a better way, let's go back to our college days when C++, Java, and Python were the primary languages we needed to master to excel in computer science. Learning these languages requires understanding syntax, semantics, practical application, and problem-solving. So, to get a strong hold on these languages, we practised continuously (or you can say get trained). Also, we learned a lot from our classmates and professors. Right? Similarly, just like humans can learn from their own thinking, expertise and other mediums, perhaps LLMs can, too.

However, gaining expertise or becoming a subject matter expert is quite a rigorous journey for both humans and LLMs. We know about the human learning process and reasoning capabilities for making decisions and completing tasks, but what does LLM training look like?

Can I say?

1. **Firstly, pre-training of LLM:** In this step, you help the model learn patterns, such as grammar, sentence structure, and even relationships between words and concepts.
2. **Instruction-tuning (or Fine-Tuning):** To fine-tune the model, a curated dataset containing examples of instructions and desired responses is used.
3. **Reinforcement Learning with Human Feedback (RLHF):** Human evaluators rank model responses, which is used further to improve the model's alignment with user expectations.

That makes sense, right? But what if we build an agentic workflow to make the model learn and give the output while doing all the checks independently? It would be like having your own assistant who can do all the work without any human intervention. Further, in this article we will talk about the 4 Agentic AI Design Patterns for Architecting AI Systems.

CHECK OUT THE DETAILED ARTICLE HERE