

GIỚI THIỆU VỀ AI AGENT



★ AI Agent là gì?

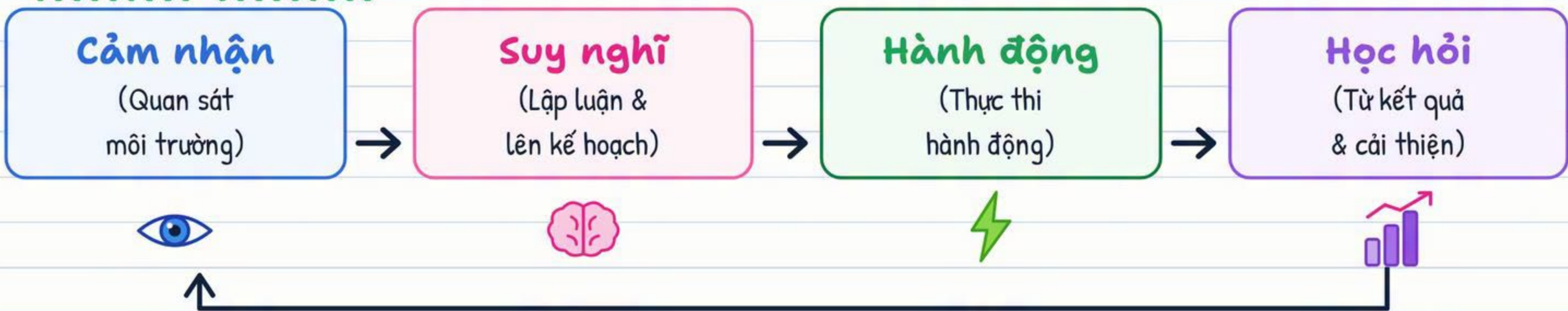


AI Agent là một hệ thống **tự chủ** có thể **cảm nhận** môi trường xung quanh, đưa ra **quyết định**, thực hiện **hành động** và **học hỏi** từ kết quả để **đạt được những mục tiêu cụ thể**.

★ Agent khác gì AI thường / Chatbot?

AI thường / Chatbot	AI Agent
- Chỉ phản hồi khi được hỏi - Không có bộ nhớ (hoặc rất hạn chế) - Không thể hành động ngoài thực tế - Cần con người ra lệnh từng bước - Ví dụ: ChatGPT trả lời câu hỏi	- Chủ động làm việc hướng tới mục tiêu - Ghi nhớ ngữ cảnh & học hỏi - Có thể hành động thông qua tools - Có thể làm việc độc lập - Ví dụ: Đặt vé máy bay, gửi email, tìm kiếm & tóm tắt tài liệu

★ Ý tưởng cốt lõi



★ Ví dụ thực tế

Bạn yêu cầu một AI travel agent:



"Lên kế hoạch đi Đà Lạt 3 ngày, ngân sách dưới 5 triệu."

Agent tự tìm chuyến bay, khách sạn, thời tiết, lập lịch trình, đặt chỗ rồi gửi lại kế hoạch cho bạn - hoàn toàn tự động.



Tóm lại, AI Agent là những hệ thống **hướng mục tiêu, tự chủ**, có thể **suy nghĩ, hành động** và **học hỏi**.



AI AGENT HOẠT ĐỘNG NHƯ THỂ NÀO?

★ AI Agent hoạt động theo một **vòng lặp liên tục** để đạt mục tiêu bằng cách tương tác với **môi trường** xung quanh.



★ Điểm mấu chốt



- ⇒ Vòng lặp này (Quan sát → Suy nghĩ → Hành động → Học hỏi) chính là **trái tim** của một AI Agent.
- ⇒ Agent liên tục **cải thiện** nhờ học từ **phản hồi**.



Hãy hình dung AI Agent giống như một con người - nó **quan sát**, **suy nghĩ**, **quyết định**, **hành động** và **học** từ trải nghiệm.

★ Ví dụ

Bạn nhờ AI Agent: **"Tìm giúp tôi chiếc laptop tốt nhất dưới 20 triệu."**



- **Quan sát** : Đọc yêu cầu của bạn.
- **Suy nghĩ** : Hiểu nhu cầu, so sánh cấu hình & giá.
- **Hành động** : Tìm trên các website, lọc kết quả.
- **Phản hồi** : Thu thập dữ liệu (giá, đánh giá, nhận xét).
- **Lặp lại** : Tinh chỉnh và đưa ra gợi ý tốt nhất.



★ Vì sao điều này quan trọng?



- ✓ Giúp agent **tự chủ** và **hướng mục tiêu**.
- ✓ Cho phép xử lý các **tác vụ phức tạp** theo từng bước.
- ✓ Giúp agent **thích nghi** và **tiến bộ** theo thời gian.

CÁC THÀNH PHẦN CỐT LÕI CỦA AI AGENT



★ Một AI Agent không chỉ là một mô hình. Nó gồm nhiều **thành phần** phối hợp chặt chẽ với nhau.



★ Nói đơn giản



Memory cung cấp ngữ cảnh, **LLM** suy nghĩ, **Planning** quyết định, **Tools** giúp hành động, **Environment** đưa vào dữ liệu, **Knowledge** làm agent thông minh hơn.

★ Ví dụ

Một agent đặt tour dùng **memory** (các chuyến đi trước), **tools** (API chuyến bay), **knowledge** (cẩm nang du lịch), **planning** (các bước lịch trình), **environment** (website, người dùng) để thực hiện **actions** (đặt chỗ, xác nhận, thông báo).



☆ **Tất cả các thành phần phối hợp với nhau tạo nên một agent mạnh mẽ!**

CÁC LOẠI AI AGENT



★ AI Agent có thể được thiết kế theo nhiều cách khác nhau, tùy vào cách chúng đưa ra **quyết định** và đạt được **mục tiêu**.

1. Simple Reflex Agent 	Ra quyết định chỉ dựa trên đầu vào hiện tại hoặc tình huống trước mắt, theo các quy tắc điều kiện - hành động.	Ví dụ: Máy điều nhiệt - nếu nhiệt độ > 30°C thì bật điều hoà.
2. Model-Based Reflex Agent 	Duy trì một mô hình nội bộ (bộ nhớ) về thế giới để xử lý quan sát thiếu sót và dự đoán kết quả.	Ví dụ: Robot hút bụi ghi nhớ bản đồ ngôi nhà của bạn.
3. Goal-Based Agent 	Hành động để đạt được mục tiêu cụ thể . Nó lập kế hoạch và chọn các bước dẫn tới mục tiêu.	Ví dụ: Dẫn đường GPS - đưa bạn tới đúng điểm đến.
4. Utility-Based Agent 	Chọn hành động tối đa hoá "độ hữu ích" hay mức hài lòng, chứ không chỉ dừng ở việc đạt mục tiêu.	Ví dụ: App giao đồ ăn - chọn tuyến nhanh nhất & trải nghiệm tốt nhất.
5. Learning Agent 	Cải thiện hiệu năng theo thời gian nhờ học từ kinh nghiệm và phản hồi nhận được.	Ví dụ: Hệ thống gợi ý ngày càng chuẩn hơn nhờ phản hồi của bạn.

★ So sánh nhanh

Simple Reflex	Model-Based	Goal-Based	Utility-Based	Learning Agent
				
Không có bộ nhớ Tuân theo quy tắc Chỉ biết hiện tại	Có bộ nhớ Hiểu tình huống tốt hơn	Biết rõ mục tiêu Lập kế hoạch Đạt mục tiêu	Chọn phương án tốt nhất Tối ưu lợi ích Không chỉ mục tiêu	Học từ dữ liệu Tiến bộ theo thời gian Ngày càng thông minh

★ Điều cần nhớ



Càng đi từ **Simple Reflex** → **Model-Based** → **Goal-Based** → **Utility-Based** → **Learning Agent**, agent càng thông minh, linh hoạt và đủ sức xử lý sự phức tạp của thực tế.

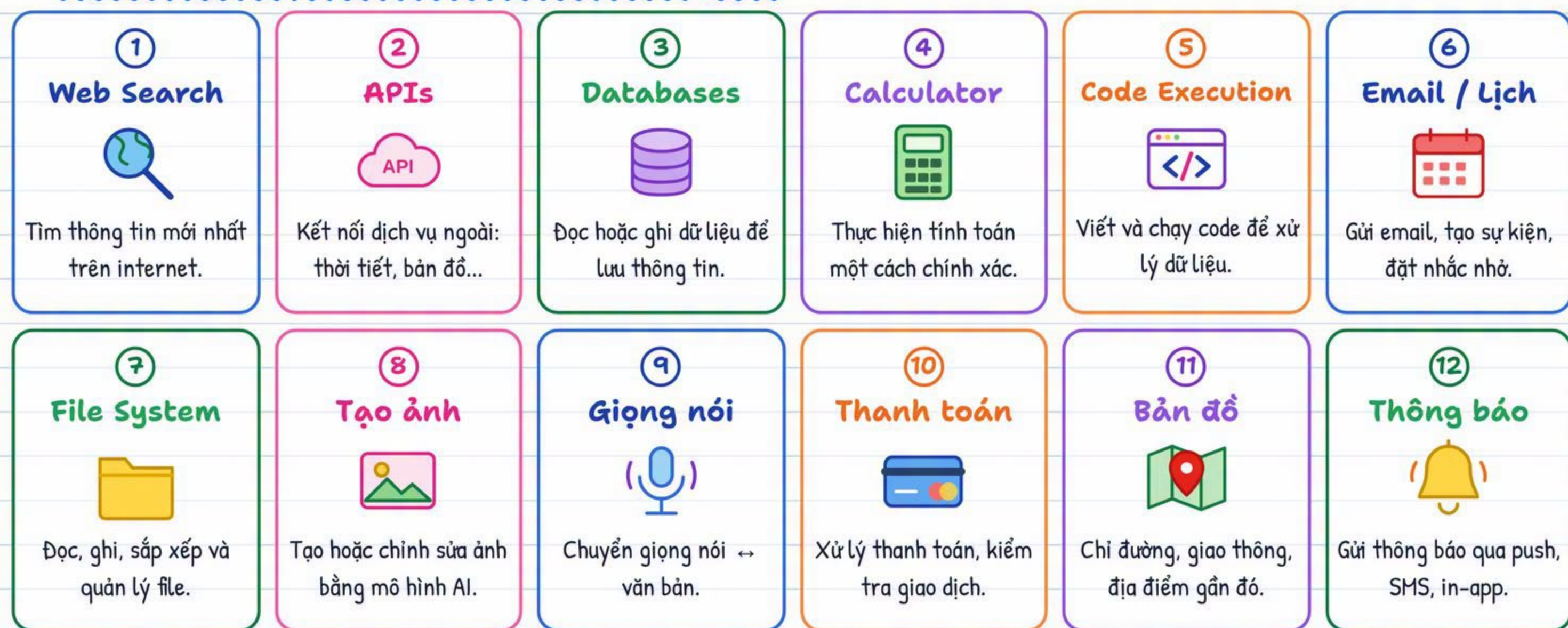


TOOLS & FUNCTION CALLING



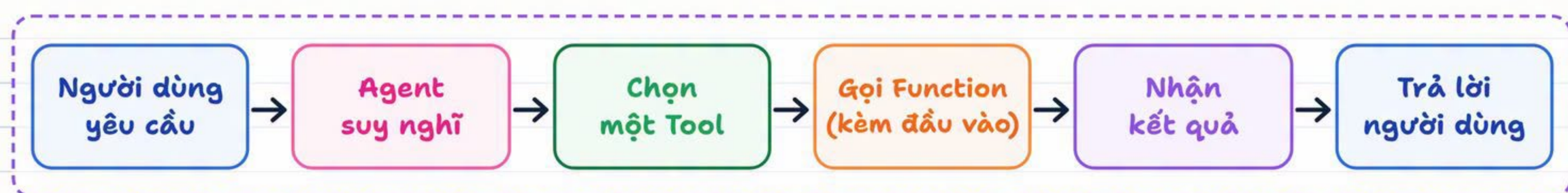
- ★ Bản thân **LLM** chỉ có thể trò chuyện và sinh văn bản, nhưng **không làm được gì** trong thế giới thực. **Tools** chính là thứ mang lại cho AI Agent khả năng **hành động**.

★ Các Tools phổ biến cho AI Agent



★ Function Calling là gì?

Function calling (hay tool calling) là khi mô hình quyết định dùng một **tool** bằng cách gửi đúng **đầu vào (arguments)** cho tool đó.



★ Vì sao Tools lại quan trọng?

- Cho phép agent **tương tác** với thế giới thực.
- Mở rộng **đáng kể** những gì agent có thể làm.
- Giúp agent **chính xác**, **hữu ích** và **đáng tin cậy** hơn.
- Không có tools, agent chỉ dừng ở mức **"nói chuyện"**.



BỘ NHỚ TRONG AI AGENT



★ Vì sao Memory lại quan trọng?

- Không có bộ nhớ, agent **quên sạch** mọi thứ sau mỗi bước.
- Memory giúp agent hiểu **ngữ cảnh**, **cá nhân hoá** phản hồi và ra quyết định tốt hơn theo thời gian.
- Nó cho phép agent **học** từ những trải nghiệm đã qua.



Memory

=

Agent thông minh hơn

★ Các loại bộ nhớ

1. Short-Term Memory (STM) - Ngắn hạn

- Lưu thông tin trong **cuộc hội thoại** hoặc tác vụ hiện tại.
- Dung lượng hạn chế.
- Dùng để theo dõi các tin nhắn hoặc các bước gần đây.
- Ví dụ: Lịch sử chat trong một cuộc trò chuyện.

2. Long-Term Memory (LTM) - Dài hạn

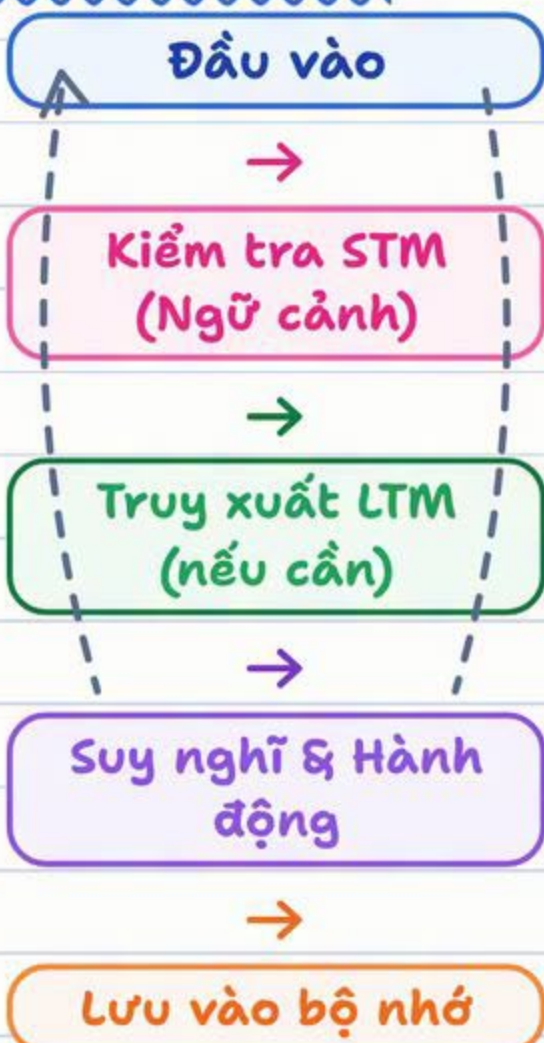
- Lưu thông tin trong **thời gian dài**.
- Dùng được xuyên suốt nhiều phiên làm việc.
- Thường lưu ở database ngoài hoặc vector store.
- Ví dụ: Sở thích người dùng, đơn hàng cũ, knowledge base.

★ STM và LTM khác nhau ra sao?

Tiêu chí	Short-Term Memory	Long-Term Memory
Mục đích	Xử lý ngữ cảnh hiện tại	Lưu tri thức để dùng về sau
Thời gian	Tạm thời (vài phút / một hội thoại)	Lâu dài (ngày, tháng, năm)
Dung lượng	Hạn chế	Rất lớn (mở rộng được)
Lưu trữ	Trong bộ nhớ (model / session)	Bên ngoài (DB, Vector Store, ổ đĩa)
Ví dụ	Tin nhắn chat, các bước vừa thực hiện	Hồ sơ người dùng, tài liệu, sở thích

★ Memory hoạt động ra sao?

- 1 Agent nhận đầu vào.
- 2 Kiểm tra bộ nhớ ngắn hạn để lấy ngữ cảnh.
- 3 Nếu cần, truy xuất thông tin từ bộ nhớ dài hạn.
- 4 Dùng bộ nhớ để suy nghĩ và hành động.
- 5 Thông tin mới được lưu ngược lại vào bộ nhớ.



★ LTM được lưu ở đâu?

- Cơ sở dữ liệu (SQL / NoSQL)
- Vector Database (Pinecone, Weaviate, Chroma, FAISS...)
- File / Tài liệu
- Lưu trữ đám mây (Cloud Storage)

★ Memory giúp Agent tốt lên thế nào?

- ✓ Ghi nhớ sở thích của người dùng.
- ✓ Đưa ra câu trả lời nhất quán, đúng trọng tâm.
- ✓ Học hỏi từ các tương tác trước đó.
- ✓ Cải thiện cá nhân hoá và trải nghiệm người dùng.
- ✓ Xử lý được tác vụ phức tạp, nhiều lượt.



Bộ nhớ tốt biến AI Agent từ một con bot **hay quên** thành một **đồng đội đáng tin cậy**.



LẬP KẾ HOẠCH & LẬP LUẬN TRONG AI AGENT



★ Planning là gì?

Planning là khả năng của agent trong việc chia một **mục tiêu lớn** thành các **bước nhỏ hơn** và quyết định **trình tự hành động** tốt nhất.



★ Agent lập luận & lập kế hoạch thế nào?



★ Ví dụ: Nghiên cứu một chủ đề



★ Các kỹ thuật lập luận

- **Chain of Thought (CoT):** Suy nghĩ theo từng bước một.
- **Self-Ask:** Agent tự đặt câu hỏi phụ để giải bài toán phức tạp.
- **ReAct (Reason + Act):** Lập luận về hành động kế tiếp rồi mới làm.
- **Reflection:** Nhìn lại kết quả, học từ sai lầm và cải thiện.

★ Điểm cần nhớ

- ✓ Planning giúp agent xử lý tác vụ phức tạp, nhiều bước.
- ✓ Reasoning giúp agent ra quyết định thông minh ở mỗi bước.
- ✓ Vòng lặp Quan sát → Học hỏi → Lập lại kế hoạch giúp agent linh hoạt.
- ✓ Planning tốt + reasoning tốt = agent đáng tin cậy.



Ghi nhớ: AI Agent không chỉ là **biết**, mà là **suy nghĩ**, **lập kế hoạch**, **hành động** và **học hỏi** liên tục.



VÒNG LẶP AGENT - VẬN HÀNH RA SAO



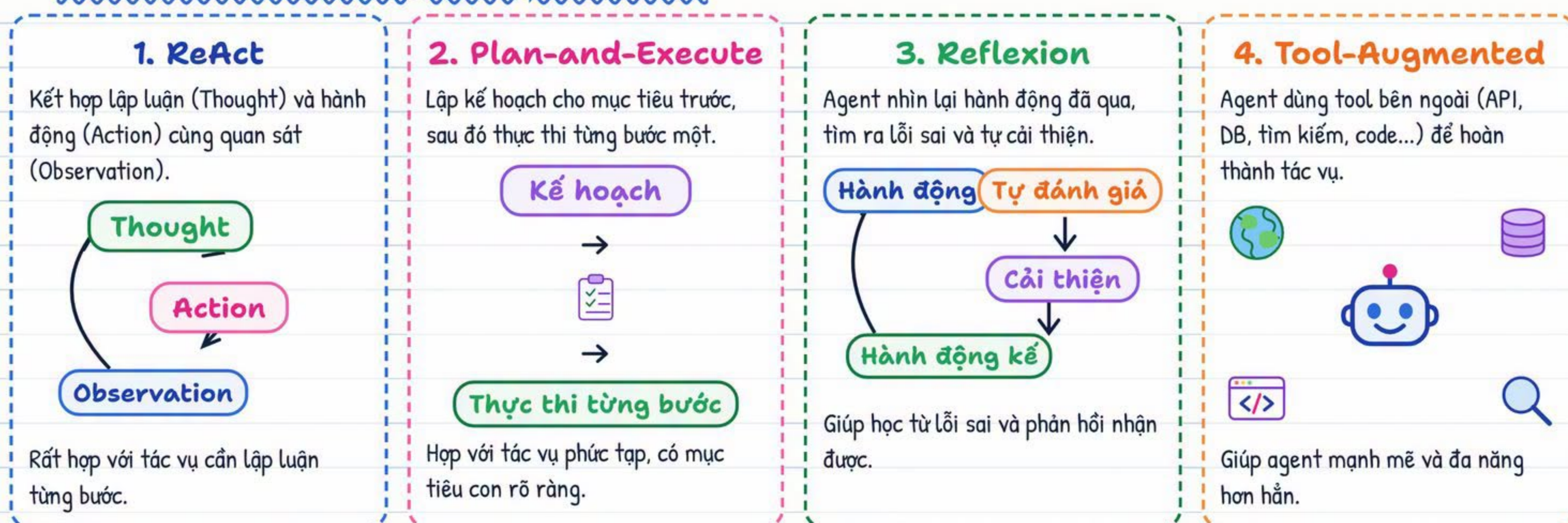
★ Agent Loop là gì?

Agent loop là chu trình liên tục trong đó agent **cảm nhận** môi trường, **suy nghĩ**, **hành động** và **học** từ chính kết quả của mình.

Cảm nhận → Suy nghĩ →
Hành động →
Quan sát → Học hỏi →
Lặp lại ☆



★ Các kiến trúc Agent phổ biến



★ Những điểm cần nhớ

- Agent hoạt động theo vòng lặp - **Cảm nhận**, **Suy nghĩ**, **Hành động**, **Quan sát**, **Học hỏi**.
- Vòng lặp tiếp diễn đến khi đạt **mục tiêu** hoặc hoàn thành tác vụ.
- Mỗi bài toán và môi trường sẽ dùng **kiến trúc** khác nhau.
- Vòng lặp tốt + **phản hồi** = agent thông minh và đáng tin cậy hơn!



Một agent loop **vững chắc** chính là **bí quyết** đằng sau mọi AI Agent thành công.



HỌC HỎI & CẢI THIẾN TRONG AI AGENT



★ Vì sao học hỏi lại quan trọng?

- Không học hỏi, agent mãi giậm chân tại chỗ và lặp lại **sai lầm**.
- Học hỏi giúp agent **thích nghi** với tình huống mới và **cải thiện** hiệu năng.
- Nó khiến agent **thông minh**, **hiệu quả** và **đáng tin cậy** hơn theo thời gian.



Học hôm nay,
Giỏi ngày mai!

★ Các kiểu học của AI Agent



★ Quy trình học tổng quát của Agent



★ Agent tiến bộ nhờ đâu?

- ① **Nhiều dữ liệu hơn** → càng nhiều trải nghiệm, agent học càng tốt.
- ② **Thuật toán tốt hơn** → học nhanh hơn và chính xác hơn.
- ③ **Phản hồi & phần thưởng** → giúp ra quyết định tốt hơn.
- ④ **Huấn luyện liên tục** → giữ cho agent luôn cập nhật.
- ⑤ **Thích nghi** → agent điều chỉnh theo môi trường và tác vụ mới.

★ Thách thức khi học

- **Dữ liệu nhiễu** / thiếu sót có thể khiến agent học sai.
- **Overfitting** – agent học quá khớp dữ liệu huấn luyện rồi thất bại ngoài thực tế.
- Cần nhiều **tài nguyên tính toán** & thời gian để huấn luyện.
- Thiết kế **phần thưởng** trong reinforcement learning rất khó.
- **Thiên lệch** trong dữ liệu tạo ra agent thiếu công bằng.



Học hỏi chính là chìa khoá biến một agent **tốt** thành một agent **xuất sắc** và **thông minh**.



XÂY DỰNG AI AGENT - THÀNH PHẦN CHÍNH



★ Các thành phần chính là gì?

Mọi AI Agent đều được tạo nên từ những **khối xây dựng** thiết yếu, hoạt động ăn khớp với nhau.

Công thức cho một Agent xuất sắc

Perception + Memory + Reasoning
+ Action + Tools
= Agent thông minh ❤️



★ Các thành phần phối hợp ra sao?



★ Tools mà Agent hay dùng

Web Search	→ Tìm kiếm trên internet
API Calls	→ Lấy dữ liệu thời gian thực
Code Interpreter	→ Chạy code, tính toán
File System	→ Đọc / ghi file
Database	→ Lưu và truy xuất dữ liệu
Browser Automation	→ Tự động thao tác web
Communication	→ Gửi email, tin nhắn

★ Ví dụ trong thực tế

- Trợ lý chat** (ví dụ ChatGPT)
 - Hiểu câu hỏi, suy nghĩ và trả lời.
- Trợ lý cá nhân** (Siri, Alexa)
 - Đặt nhắc nhở, trả lời, điều khiển thiết bị.
- Xe tự hành**
 - Cảm nhận môi trường, suy luận và hành động.
- Bot giao dịch**
 - Phân tích thị trường, quyết định và đặt lệnh.
- Hệ thống gợi ý**
 - Học sở thích người dùng và đề xuất nội dung.



Điều cốt lõi



AI Agent không chỉ là một mô hình – đó là một hệ thống biết **cảm nhận**, **suy nghĩ**, **hành động** và **học hỏi**, dùng **tools** để đạt mục tiêu.

