

# Kỹ Thuật Prompt Nâng Cao

Từ Nền Tảng đến Tiên Phong trong Tương Tác với AI

# Từ câu trả lời đơn giản đến giải pháp tinh vi...

Translate this.



Dịch câu này.

Làm thế nào để chúng ta khai phá toàn bộ  
tiềm năng của các Mô Hình Ngôn Ngữ Lớn?



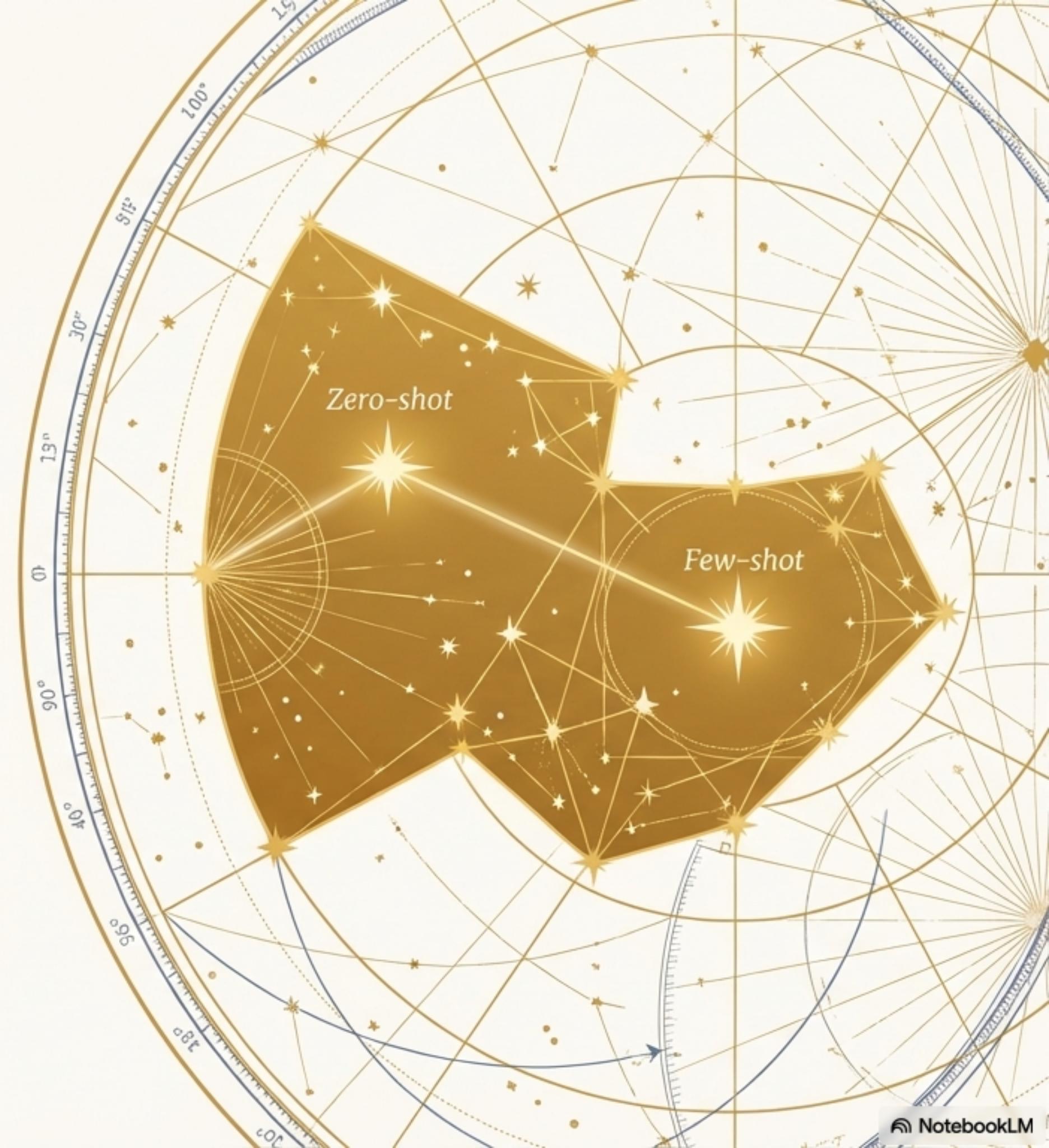
# Lộ Trình của Chúng Ta: Bản Đồ Thiên Hà Prompt

Mỗi kỹ thuật prompt là một vì sao. Khi kết hợp, chúng tạo thành những chòm sao mạnh mẽ, giúp giải quyết các vấn đề ngày phức tạp. Hãy cùng khám phá phồn bối chòm sao chính trong hành trình này.



# Chương I: Nền Tảng Prompt

Những viên gạch đầu tiên để xây dựng mọi tương tác hiệu quả với AI.  
Đây là điểm khởi đầu bắt buộc.



# Zero-shot Prompting



## Khái niệm Cốt lõi

Yêu cầu mô hình thực hiện một tác vụ mà không cung cấp bất kỳ ví dụ nào trước đó.

## Bối cảnh Áp dụng

Sử dụng khi tác vụ đơn giản, dễ hiểu và LLM đã được huấn luyện tốt về nó (ví dụ: tóm tắt, dịch thuật cơ bản, phân loại cảm xúc đơn giản).

## Cơ chế Hoạt động

1. Đưa ra chỉ thị (instruction) trực tiếp.
2. Cung cấp dữ liệu đầu vào (input).
3. Mô hình tự suy luận và tạo ra kết quả.

## Ví dụ Minh họa

### Prompt:

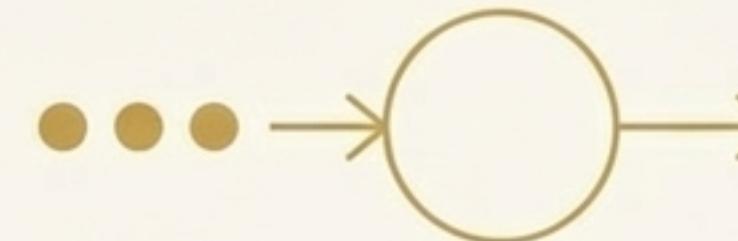
Phân loại cảm xúc của câu sau thành "Tích cực", "Tiêu cực" hoặc "Trung tính".

Câu: "Bộ phim này thật tuyệt vời, tôi rất thích nó!"

### Kết quả:

Tích cực

# Few-shot Prompting



## Khái niệm Cốt lõi

Cung cấp cho mô hình một vài ví dụ ("shots") về tác vụ cần thực hiện để định hướng và cải thiện chất lượng đầu ra.

## Bối cảnh Áp dụng

Khi tác vụ phức tạp hơn, cần định dạng đầu ra cụ thể, hoặc khi Zero-shot cho kết quả không nhất quán.

## Cơ chế Hoạt động

1. Cung cấp một vài cặp `[Input] -> [Output]` làm ví dụ.
2. Đưa ra input mới cần xử lý.
3. Mô hình học theo mẫu từ các ví dụ để tạo ra output mới.

## Ví dụ Minh họa

### Prompt:

Đây là một công cụ phân loại email.

### Ví dụ 1:

Email: "Giảm giá 50% tất cả sản phẩm!"

Phân loại: Quảng cáo

### Ví dụ 2:

Email: "Chào bạn, báo cáo quý 3 đã sẵn sàng."

Phân loại: Công việc

Email: "Bạn có muốn đi xem phim tối nay không?"

Phân loại: [Model prediction]

# Chương II: Khai Mở Năng Lực Lập Luận

Vượt qua câu trả lời tức thời. Hướng  
dẫn mô hình “suy nghĩ từng bước” để  
giải quyết các vấn đề logic và đa tầng.



# Chain-of-Thought (CoT)

O-O-O

## Khái niệm Cốt lõi

Hướng dẫn mô hình trình bày từng bước suy luận logic dẫn đến câu trả lời cuối cùng, thay vì chỉ đưa ra kết quả.

## Bối cảnh Áp dụng

Các bài toán số học, logic, suy luận nhân quả, hoặc bất kỳ tác vụ nào cần nhiều bước trung gian để đi đến kết luận đúng.

## Cơ chế Hoạt động

Thêm các cụm từ như “Hãy suy nghĩ từng bước một” (Let's think step by step) vào prompt để khuyến khích mô hình bộc lộ quá trình lập luận.

## Ví dụ Minh họa

Prompt:

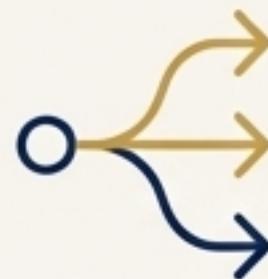
Q: Roger có 5 quả bóng tennis. Anh ấy mua thêm 2 hộp, mỗi hộp 3 quả. Hỏi bây giờ anh ấy có bao nhiêu quả? **Hãy suy nghĩ từng bước.**

A:

1. Roger ban đầu có 5 quả bóng.
2. Anh ấy mua thêm 2 hộp, mỗi hộp 3 quả, vậy số bóng mua thêm là  $2 * 3 = 6$  quả.
3. Tổng số bóng Roger có là  $5 + 6 = 11$  quả.

Đáp án cuối cùng là 11.

# Từ Một Chuỗi Suy Luận đến Nhiều Con Đường Tư Duy



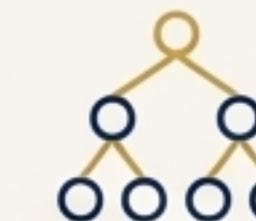
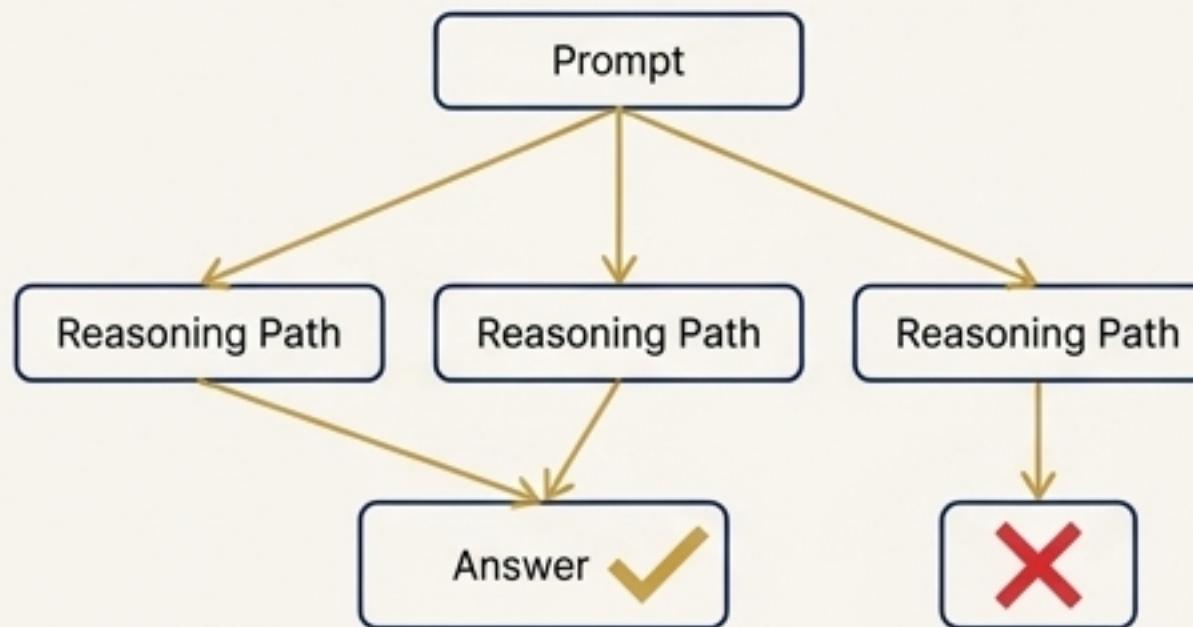
## Self-Consistency

### Khái niệm

Chạy cùng một prompt CoT nhiều lần để tạo ra nhiều chuỗi suy luận khác nhau, sau đó chọn câu trả lời xuất hiện 1 khác nhau, sau đó chọn câu trả lời xuất hiện nhiều nhất (majority vote).

### Mục đích

Tăng độ tin cậy và chính xác bằng cách giảm thiểu ảnh hưởng của một chuỗi suy luận sai lầm.



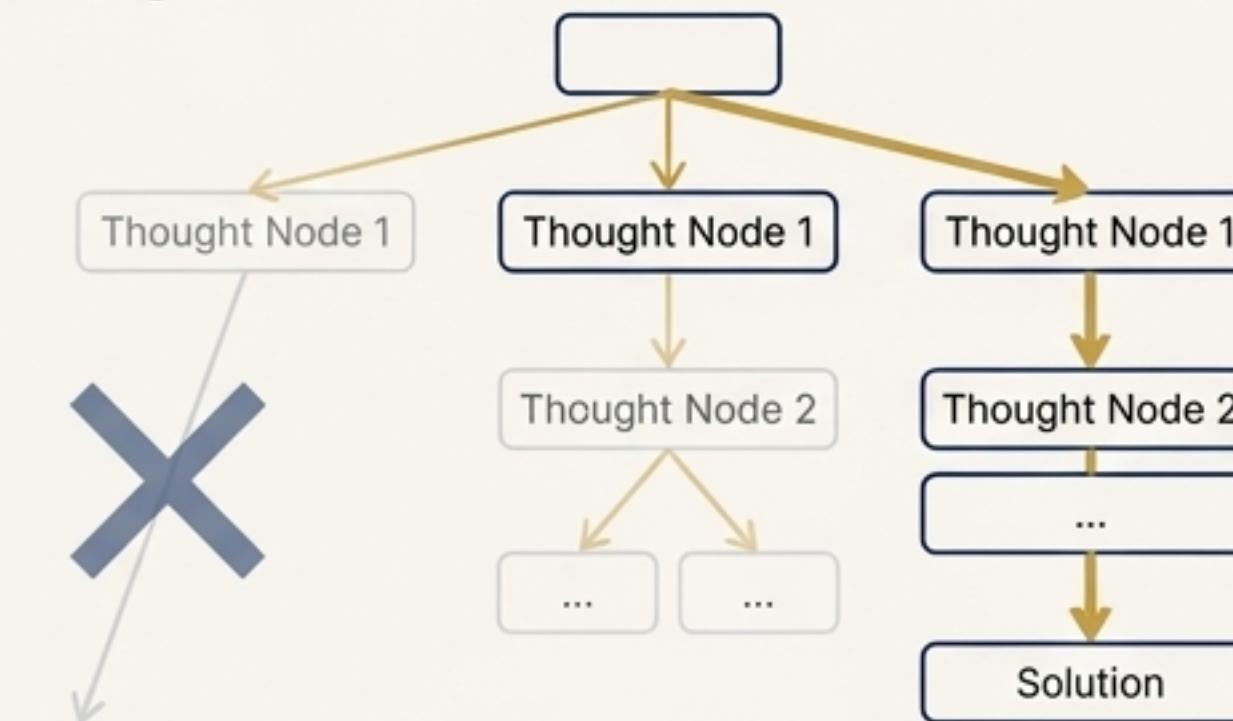
## Tree of Thoughts (ToT)

### Khái niệm

Cho phép mô hình khám phá nhiều nhánh suy luận khác nhau một cách có hệ thống, tự đánh giá và lựa chọn hướng đi hứa hẹn nhất tại mỗi bước.

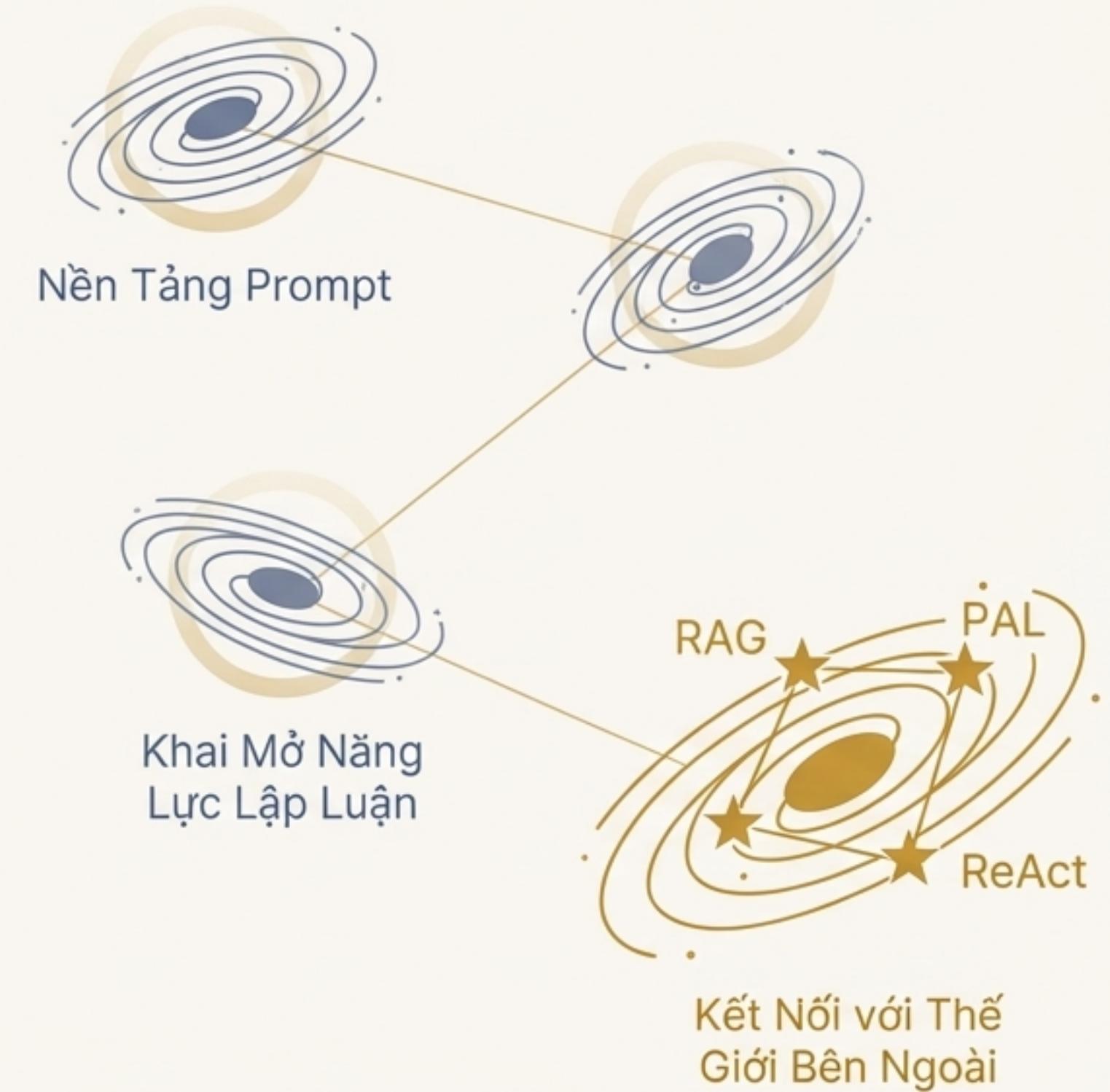
### Mục đích

Giải quyết các vấn đề phức tạp đòi hỏi sự tìm kiếm, khám phá hoặc lên kế hoạch chiến lược, nơi một con đường duy nhất có thể không đủ.

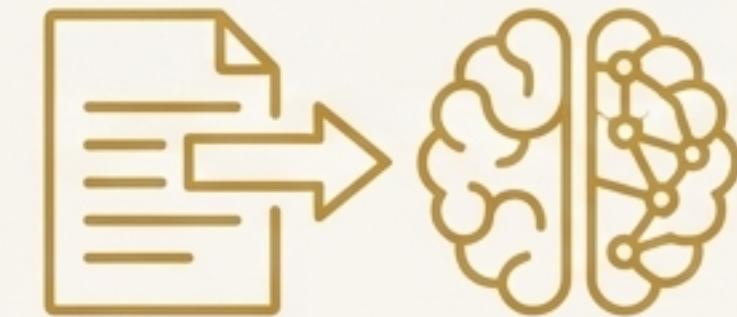


# Chương III: Kết Nối với Thế Giới Bên Ngoài

Phá vỡ giới hạn của dữ liệu huấn luyện. Trang bị cho LLM khả năng truy xuất thông tin mới và sử dụng các công cụ bên ngoài.



# Retrieval Augmented Generation (RAG)



**Khái niệm Cốt lõi:** Tăng cường kiến thức cho LLM bằng cách truy xuất thông tin liên quan từ một kho dữ liệu bên ngoài (ví dụ: cơ sở dữ liệu, tài liệu nội bộ) trước khi tạo ra câu trả lời.

**Bối cảnh Áp dụng:** Trả lời câu hỏi dựa trên thông tin mới nhất, giảm ảo giác (hallucination), xây dựng chatbot dựa trên kiến thức chuyên biệt (ví dụ: tài liệu pháp lý, y tế).

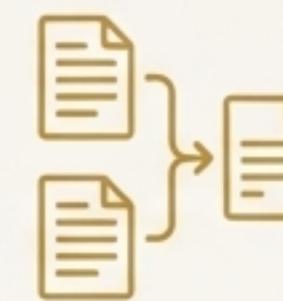
## Cơ chế Hoạt động



User Query



1. Retrieve: Tìm kiếm tài liệu liên quan.



2. Augment: Bổ sung context vào prompt.



3. Generate: LLM tạo câu trả lời.

# Vượt Ngoài Ngôn Ngữ: Dạy LLM Hành Động và Sử Dụng Công Cụ

## Program-Aided Language Models (PAL)



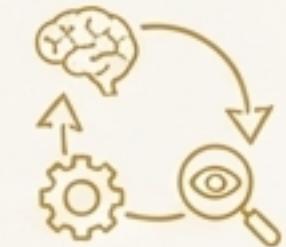
- Khái niệm:** Hướng dẫn LLM viết và thực thi mã (ví dụ: Python) để giải quyết các vấn đề đòi hỏi tính toán chính xác, thay vì tự tính toán.
- Mục đích:** Tận dụng sự chính xác của máy tính cho các tác vụ logic, toán học, nơi LLM có thể mắc lỗi.

### Program-Aided Language Models (PAL)

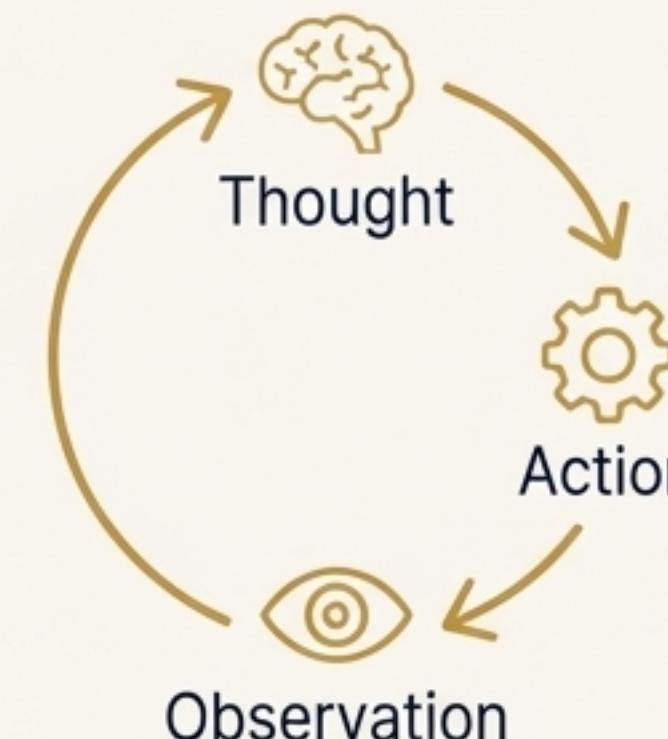
Thay vì tự tính ` $(7*8)+9$ `, LLM sẽ tạo ra code:

```
print((7*8)+9)
```

## ReAct (Reason + Act)

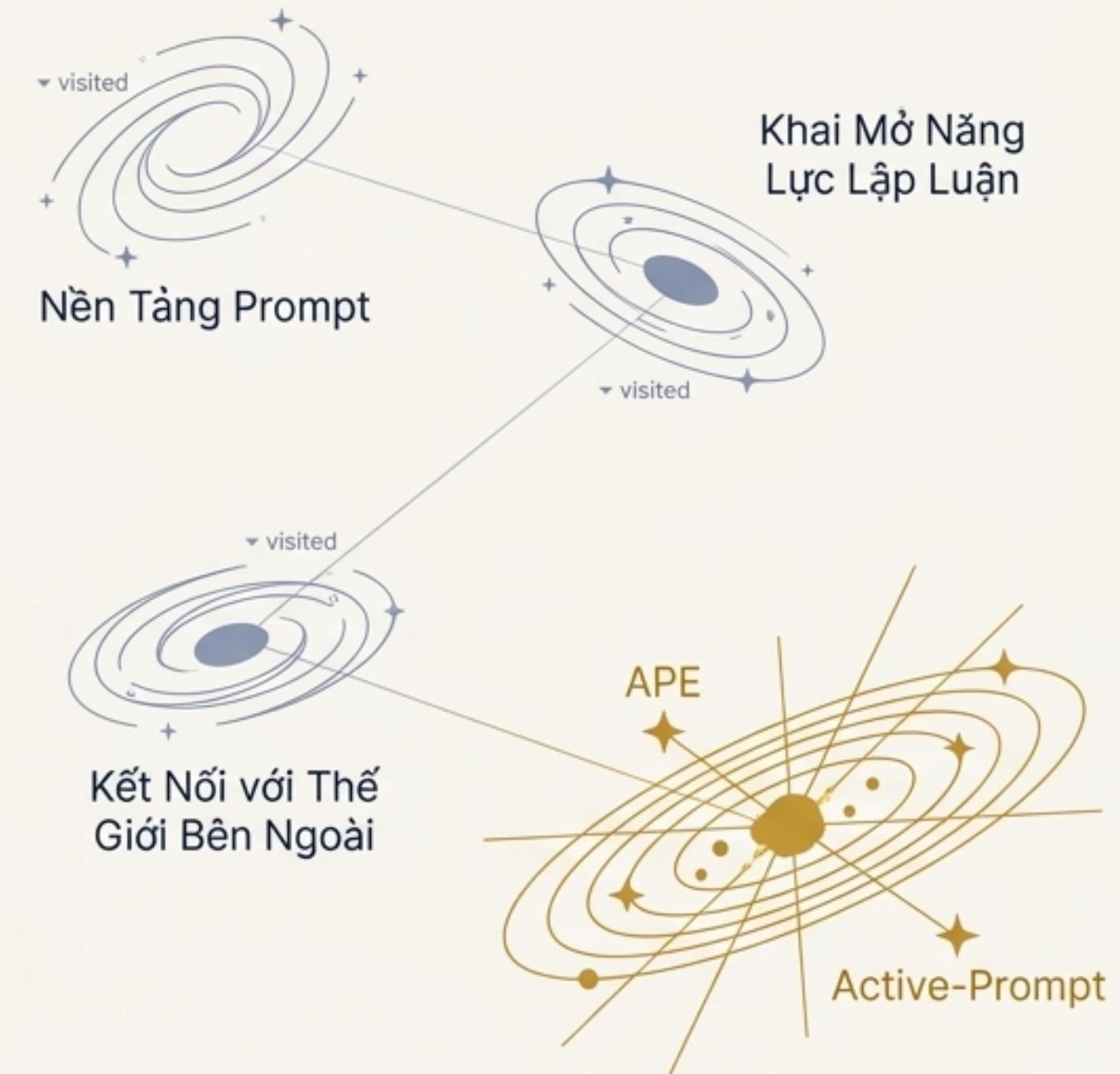


- Khái niệm:** Kết hợp xen kẽ các bước 'Suy luận' (Reason) và 'Hành động' (Act) để giải quyết các tác vụ đòi hỏi tương tác với môi trường bên ngoài (ví dụ: tìm kiếm web, sử dụng API).
- Mục đích:** Cho phép LLM tạo ra một kế hoạch, thực hiện một hành động, quan sát kết quả và điều chỉnh kế hoạch cho đến khi hoàn thành mục tiêu.



# Chương IV: Tiên Phong Tự Động Hóa

Khi AI tự mình tạo ra những prompt tối ưu. Khám phá các kỹ thuật meta giúp tự động hóa và tinh chỉnh quá trình prompt engineering.

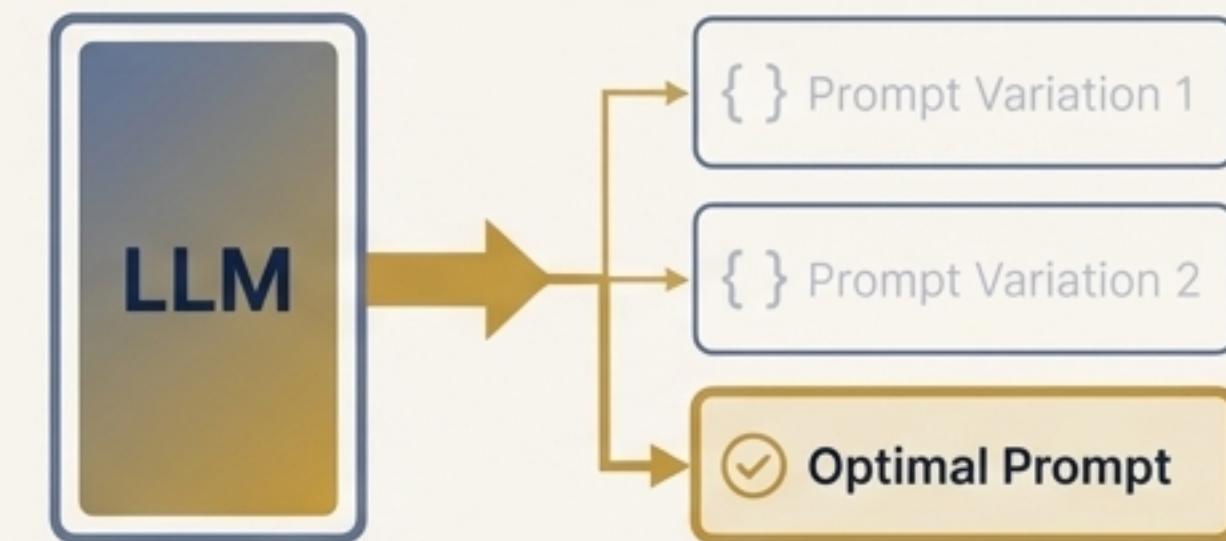
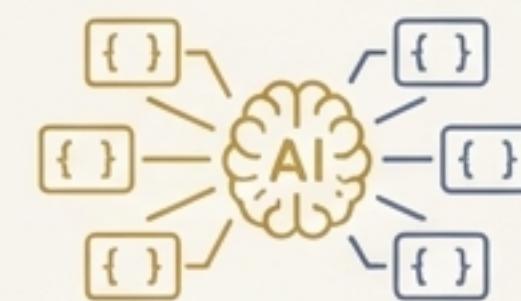


# Tự Động Tối Ưu Hóa: Các Kỹ Thuật Meta-Prompting

## Automatic Prompt Engineer (APE)

**Khái niệm:** Sử dụng một LLM để tự động tạo và chọn lọc các chỉ thị (instructions) tốt nhất cho một tác vụ cụ thể, dựa trên một vài ví dụ về kết quả mong muốn.

**Mục đích:** Tự động hóa quá trình tìm kiếm prompt hiệu quả nhất, thay vì con người phải thử nghiệm thủ công.

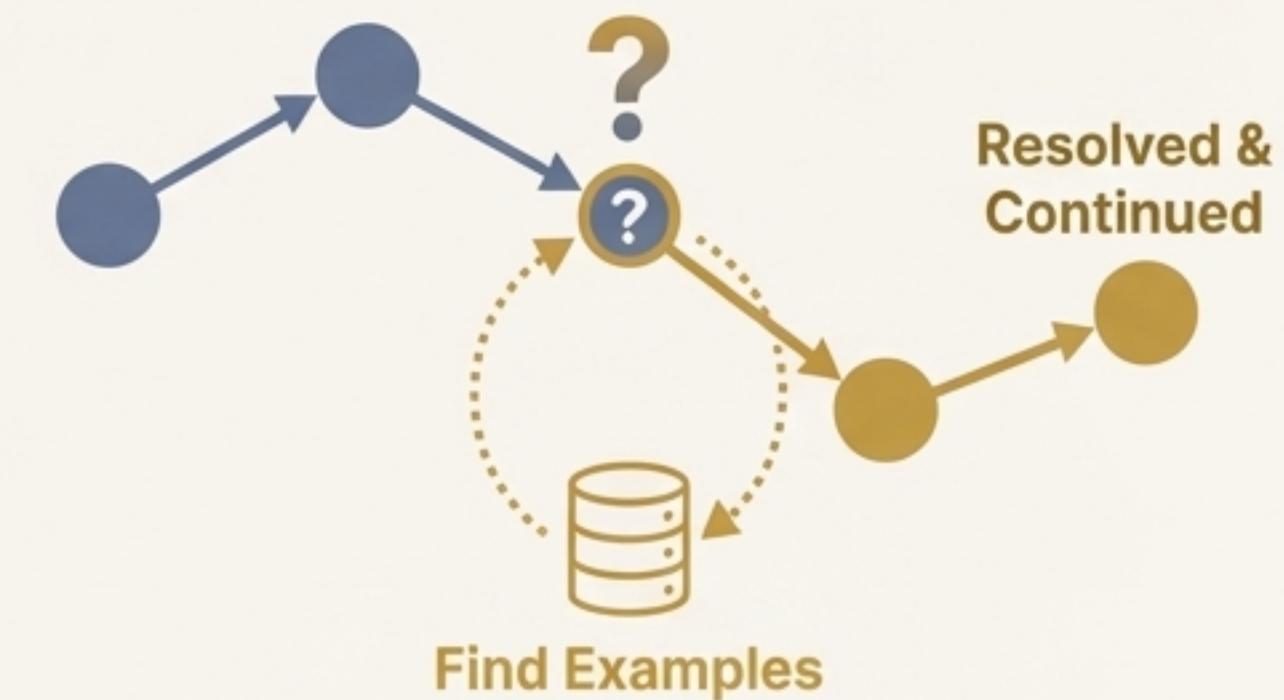


## Active-Prompt



**Khái niệm:** Mở rộng CoT bằng cách chủ động xác định những câu hỏi không chắc chắn nhất (uncertain questions) và tìm kiếm ví dụ (demonstrations) phù hợp để trả lời chúng.

**Mục đích:** Tăng cường khả năng suy luận cho các vấn đề phức tạp bằng cách 'học' một cách có chủ đích từ các ví dụ quan trọng nhất.



# Bộ Công Cụ Prompt Engineering của Bạn



Từ những chỉ thị đơn giản đến các hệ thống lập luận và hành động tự động, giờ đây bạn đã có bản đồ để khai phá toàn bộ sức mạnh của các mô hình ngôn ngữ lớn.