

# Kỹ Thuật Prompt: Một Cái Nhìn Chuyên Sâu

Chắt lọc tinh hoa từ các chuyên gia tại Anthropic

# Những Người Thợ Đằng Sau Con Chữ



**Alex**  
Developer Relations Lead

Từng là một kỹ sư prompt, kết nối giữa khía cạnh giải pháp và nghiên cứu.



**David Hershey**  
Customer Engagement

"Giúp khách hàng xây dựng hệ thống với các mô hình ngôn ngữ, tập trung vào các ứng dụng doanh nghiệp."



**Amanda Askell**  
Finetuning Team Lead

"Làm cho Claude trở nên trung thực và tốt bụng, kết nối nghiên cứu và tâm lý học mô hình."



**Zack Witten**  
Prompt Engineer

"Xây dựng các tài liệu giáo dục và công cụ nhằm nâng cao mặt bằng chung về kỹ thuật prompt trong xã hội."

# “Prompt Engineering” Không Chỉ Là Viết

## PROMPT

“Về bản chất, là **giao tiếp rõ ràng**. Giống như nói chuyện chuyện với một người.”

## ENGINEERING

“Đến từ quá trình **thử nghiệm và sai sót...** có một **nút khởi động** lại. Và còn là việc tích hợp các prompt vào một hệ thống lớn hơn.”

# Vòng Lặp Kỹ Thuật: Thủ Nghiệm, Lặp Lại, Hoàn Thiện

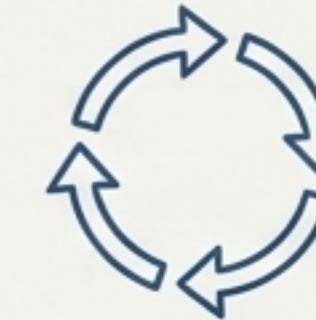


# Phẩm Chất Của Một Kỹ Sư Prompt Bậc Thầy



## Giao Tiếp Rõ Ràng

Khả năng mô tả các khái niệm một cách chính xác.



## Tinh Thần Lặp Lại

Sẵn sàng thử hàng trăm prompt trong 15 phút để tìm ra điểm tối ưu.



## Tư Duy Về Trường Hợp Biên

Luôn tự hỏi: “Điều gì sẽ xảy ra nếu đầu vào là một chuỗi rỗng? Hoặc không phải là một bộ dữ liệu?”



## Thấu Hiểu Hệ Thống

Suy nghĩ về nguồn dữ liệu, độ trễ và các đánh đổi khi tích hợp vào sản phẩm.

## Kỹ Năng Quan Trọng Nhất: Đọc Kỹ Đầu Ra Của Mô Hình

“**'Hãy nhìn vào dữ liệu của bạn'** là một câu nói kinh điển trong machine learning. Với prompting, điều tương đương là **'hãy nhìn vào đầu ra của mô hình'**. ”

- Zack Witten

# Phát Triển "Lý Thuyết Về Tâm Trí" Cho AI

## Hiểu Người Dùng

"Người dùng cuối thực sự sẽ gõ gì vào? Họ không dùng phím shift và mặc đầy lỗi chính tả. Họ nghĩ đó là **Google**."



## Hiểu Mô Hình

"Mô hình sẽ nhìn nhận chỉ dẫn của bạn như thế nào? Nó **thiếu bối cảnh** gì mà bạn đang mặc định là nó biết?"

# Phép Loại Suy Quyền Lực: Người Nhân Viên Thời Vụ

Hãy tưởng tượng bạn đang hướng dẫn một nhân viên thời vụ mới.

Họ rất thông minh và có năng lực, nhưng hoàn toàn không biết họ mình ng` tông biết gì về công ty hay nhiệm vụ cụ thể của bạn. Bạn sẽ giải thích cho họ như thế nào? Hãy bắt đầu prompt của bạn từ đó.



# Các Nguyên Tắc Vàng Của Prompting Hiệu Quả

**RÕ RÀNG:** Cung cấp cho mô hình toàn bộ bối cảnh của nhiệm vụ. Đừng bắt nó phải đoán.

**CHO LỐI THOÁT:** Hướng dẫn mô hình phải làm gì khi không chắc chắn (ví dụ: xuất ra thẻ <UNSURE>).

**YÊU CẦU PHẢN HỒI:** Hỏi thẳng mô hình: “Chỉ dẫn của tôi có điểm nào không rõ ràng không?”

**ĐỂ MÔ HÌNH TỰ GIÚP MÌNH:** “Bạn đã mắc lỗi này. Hãy viết lại chỉ dẫn của tôi để bạn không mắc lại lỗi đó nữa.”

# Vượt Ra Ngoài "Mánh Mẹo": Sự Tinh Tế Của Việc Gán Vai

## ✗ CÁCH CŨ (Kém hiệu quả)

Nói dối mô hình về một vai trò tương tự.

Ví dụ: “Bạn là một giáo viên...”

## ✓ CÁCH MỚI (Hiệu quả hơn)

- **Sử dụng phép ẩn dụ**

“Hãy dùng thang điểm của một bài tập ở trường trung học để **đánh giá biểu đồ** này.” (Chỉ rõ quy mô phân tích mong muốn).

- **Nói sự thật**

“Hãy nói sự thật về nhiệm vụ. Ví dụ: ‘Tôi là một nhà nghiên cứu AI đang xây dựng một bộ dữ liệu đánh giá.’ Mô hình đủ thông minh để hiểu.”

# Kỹ Thuật Phải Thích Nghi Với Bối Cảnh

## DOANH NGHIỆP (Enterprise)

### Mục tiêu:

Độ tin cậy & Nhất quán.  
Độ tin cậy & Nhất quán.

### Kỹ thuật:

Nhiều ví dụ cụ thể,  
định dạng nghiêm ngặt.

## NGHIÊN CỨU (Research)

### Mục tiêu:

Sự đa dạng & Thăm dò  
năng lực.

### Kỹ thuật:

Các ví dụ mang tính  
minh họa (không phải  
theo nghĩa đen), tránh  
“mớm lời” cho mô hình.

## TRÒ CHUYỆN (Chat / Personal)

### Mục tiêu:

Một câu trả lời đúng  
duy nhất cho một lần.

### Kỹ thuật:

Đối thoại lặp lại, có  
con người giám sát  
(human-in-the-loop).

# Con Đường Tới Sự Tinh Thông

## Đọc & Phân Tích

*"Mỗi khi tôi thấy một prompt tốt... tôi sẽ đọc kỹ, cố gắng phân tích nó đang làm gì và tại sao."*

## Đẩy Giới Hạn

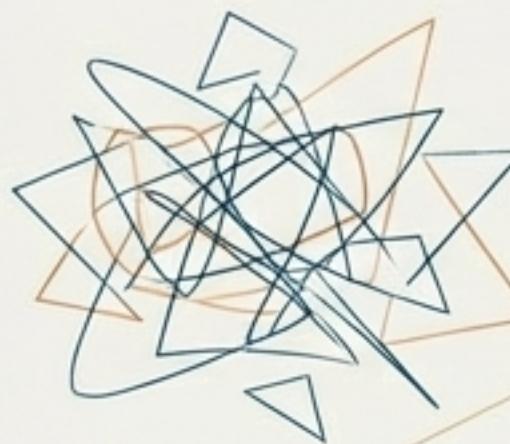
*"Cố gắng bắt mô hình làm điều gì đó bạn nghĩ nó không thể làm. Đó là những lúc tôi học được nhiều nhất."*

## Chấp Nhận “Cày Cuốc”

Kỹ năng thực sự đến từ sự lặp lại chuyên sâu và đầy hứng thú. Những người giỏi nhất là những người thực sự thích quá trình này.

# Từ “Mánh Mẹo” Đến Tính Năng Cốt Lõi

Mỗi khi chúng ta tìm ra một ‘mánh’ prompting hiệu quả, mục tiêu tiếp theo là huấn luyện khả năng đó vào thẳng mô hình. Vì lý do đó, các kỹ thuật tốt nhất luôn có vòng đời ngắn.



**MÁNH MẸO**



**TÍNH NĂNG CỐT LÕI**

Ví dụ: Kỹ thuật "**chain-of-thought**". Trước đây cần phải yêu cầu, giờ đây mô hình tự động thực hiện khi nhận diện được bài toán.

# Sự Đảo Ngược Vĩ Đại: Từ Hướng Dẫn Sang Hợp Tác

“Prompting sẽ chuyển từ việc bạn hướng dẫn một nhân viên thời vụ, sang việc bạn tham vấn một chuyên gia. Mô hình sẽ bắt đầu ‘phỏng vấn’ bạn để hiểu rõ điều bạn thực sự muốn.”



Hướng Dẫn



Hợp Tác



# Mục Tiêu Tối Thượng: Ngoại Hiện Hóa Tư Duy Của Bạn

Cốt lõi của prompting bậc thầy là khả năng **phân tích những ý tưởng phức tạp** trong đầu bạn và làm cho chúng trở nên rõ ràng, dễ hiểu đối với một **thực thể thông minh khác**. Đó là nghệ thuật làm cho **tư duy của bạn** trở nên **hữu hình**.

