

# Tổng Kết – Kỹ thuật tạo lệnh & Thực Hành

Chúng ta đã đi một chặng đường dài – từ người dùng ChatGPT thông thường, đến **hiểu rõ cách AI tư duy** và **tự điều chỉnh đầu ra theo mục tiêu**. Điều này không chỉ là kỹ thuật, mà còn là một **tư duy sáng tạo và hệ thống**.

## 1. Tokenization – AI "Đọc" Văn Bản Như Thế Nào?

- **Token = đơn vị nhỏ nhất** mà AI xử lý được (thường là từ, cụm từ, thậm chí là một phần từ).
- Ví dụ: "strawberry" có thể bị chia thành "straw" + "berry" hoặc thậm chí "str" + "aw" + "berry" tùy mô hình.
- Điều này ảnh hưởng:
  - Độ dài tối đa của prompt.
  - Cách AI hiểu và phản hồi câu hỏi.

Bài học:

- Luôn **đoán trước** token hóa có thể xảy ra → viết prompt rõ ràng, ngắn gọn.
- Dùng [tiktoken viewer](#) để kiểm tra nếu cần.

## 2. Tối Ưu Prompt – Càng Ngắn Gọn, Càng Hiệu Quả

- Prompt càng dài → AI càng bị "chia trí".
- Hãy **"nói chuyện với AI như nói chuyện với người thông minh nhưng mất tập trung"**.
- Dùng kỹ thuật như:
  - Bullet points
  - Hạn chế từ không cần thiết
  - Gợi ý phong cách mong muốn

## 3. System Message – Định Hình "Tính Cách" Của AI

- Đây là **nền tảng cho hành vi của AI**: lịch sự, nghiêm túc, hài hước, sáng tạo...
- Ví dụ:
  - "You are a legal advisor for small businesses in Vietnam."
  - "You're a witty poet who only responds in rhymes."

Mở rộng:

- LM Studio cho phép thiết lập system message cố định ngay trong phần cấu hình.

## 4. Các Tham Số Mô Hình - Đặt Đúng, Hiệu Quả Tăng Gấp Đôi

Tham số	Ý nghĩa chính
temperature	Độ ngẫu nhiên (0.1 = chính xác, 1.2 = sáng tạo)
top_k	Chọn trong K từ khả thi nhất
top_p	Chọn từ theo xác suất cộng dồn P
repetition_penalty	Phạt nếu AI lặp lại từ/ngữ đã dùng
presence/frequency penalty	Phạt nếu AI đề cập lại từ đã có

## Mẹo:

- Dự án đòi hỏi sự **ổn định**: temperature thấp, penalty cao
- Dự án đòi hỏi sự **sáng tạo**: temperature cao, penalty thấp

## 5. Ứng Dụng Trong LM Studio - Thực Hành Là Mastery

Chúng ta đã:

- Thiết lập system messages
- Tùy chỉnh các tham số: temperature, top\_k, repetition\_penalty, v.v.
- Thử nghiệm nhiều mô hình khác nhau với **GGUF** (Mistral, Zephyr, Phi-2...)
- Tạo trò chơi: tung xúc xắc, oẳn tù tì, chọn dâu tây...

## 6. Vượt Xa ChatGPT - Trở Thành Prompt Engineer Thực Thụ

Bạn giờ đã có thể:

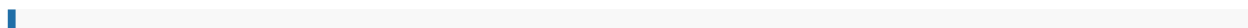
- Tự xây dựng **prompt hệ thống**
- **Điều chỉnh đầu ra** theo mục tiêu
- Tích hợp vào ứng dụng thực tế như chatbot, trợ lý ảo, công cụ sáng tạo

## 7. Sửa Lỗi & Tối Ưu Prompt

Prompt tốt = **hiểu mô hình + thực hành lặp đi lặp lại**. Không có prompt hoàn hảo ngay từ đầu.  
Bạn cần:

- Thử - Sai - Sửa
- Ghi chép thay đổi
- Kiểm thử đa dạng dữ liệu đầu vào

**Hãy nhớ:**



☐ **Mastery = Practice x Curiosity**

- Bạn không còn chỉ là người dùng ChatGPT nữa.
- Bạn đã hiểu AI từ trong ra ngoài.
- Bạn **có thể điều khiển, sáng tạo, và xây dựng tương lai với AI.**

**Tác giả: Đỗ Ngọc Tú**

**Công Ty Phần Mềm VHTSoft**

Phiên bản #1

Được tạo 24 tháng 4 2025 15:20:46 bởi Đỗ Ngọc Tú

Được cập nhật 26 tháng 4 2025 02:30:05 bởi Đỗ Ngọc Tú