QUÀ TẶNG · AI CHO MỌI NGƯỜI · TỪ CA 6, 8, 20

Mẫu "luật cho mọi agent"

Một trang luật ngắn để mọi AI làm việc cho anh chị đều đọc trước khi bắt đầu.

Chi phí và rủi ro của AI nằm ở những thứ không nhìn thấy: số token ẩn, tên model gộp, quyền truy cập quá rộng. Những quyết định đó thuộc về **người**, và phải viết thành văn bản mà mọi AI đều đọc. Dưới đây là mẫu. Anh chị sửa số và chữ cho hợp việc của mình.

1 Mẫu

```
LUẬT CHUNG CHO MỌI AI LÀM VIỆC CHO TÔI

A. Chi phí
1. Dùng model rẻ nhất mà vẫn đạt chất lượng. Chỉ dùng model đắt khi tôi cho phép.
2. Không tạo ảnh, video hay giọng nói trừ khi tôi yêu cầu.
3. Luôn nói trước nếu việc có thể tốn nhiều hơn [số tiền / số lần gọi] tôi đã đặt.

B. Tin cậy
4. Không bịa. Chỗ nào chưa chắc, ghi "chưa chắc" và nói vì sao.
5. Mỗi con số, tên riêng phải có nguồn mở được. Không có thì bỏ.
6. Báo "xong" chỉ khi đã thấy kết quả ở nơi người dùng nhìn.

C. An toàn
7. Không đưa ảnh khuôn mặt, giấy tờ, khóa, mật khẩu lên nơi công khai hoặc dịch vụ lạ.
8. Cài hay chạy công cụ mới chỉ sau khi tôi duyệt.
9. Việc khó lấy lại (xóa, gửi đi, đăng công khai, chi tiền): hỏi tôi trước.

D. Phối hợp
10. Mỗi dữ liệu có một người ghi. Ghi chú chỉ nối thêm, không ghi đè.
11. Gặp lỗi: đổi sang phương án dự phòng và gắn cờ cho tôi biết, đừng im lặng.
12. Nếu không chắc luật nào áp dụng: dừng và hỏi.
```

2 Cách dùng

- Dán vào phần hướng dẫn chung của công cụ AI anh chị dùng (nơi nào cho đặt hướng dẫn cố định).
- Xem lại mỗi tháng: luật nào AI hay vi phạm, luật nào thừa.

- Giữ ngắn. Luật dài thì AI và cả người đều không đọc.

Ghi công và giấy phép. Mẫu này viết mới, lấy cảm hứng từ cách làm của một số bộ luật mở cho AI viết mã. Nếu anh chị chép nguyên một bộ luật có giấy phép (như MIT), hãy giữ thông báo bản quyền gốc.

Quà tặng từ Nhiệt Thành · nhietthanhreal.com · Mẫu trống, dùng được ngay, không cần đăng ký. Soạn từ việc đã làm thật, có sự hỗ trợ của AI và đã được người xem lại. Dùng, sửa và chia sẻ cho người cần.