

Case Study – Ứng dụng RAG vào Sách Dạy Nấu Ăn

Chúng ta vừa thiết lập nền tảng cho những gì bạn sẽ đạt được trong phần này. Và bây giờ, đã đến lúc **đi sâu vào một Case Study cụ thể** – một ví dụ thực tế, giúp bạn **áp dụng RAG và OpenAI vào thế giới dữ liệu phi cấu trúc**.

Tại sao lại là Sách Dạy Nấu Ăn?

Bạn có thể thắc mắc: “Tại sao lại chọn sách dạy nấu ăn? Nghe có vẻ đơn giản quá!”

Thật ra, **đây là một ví dụ cực kỳ thông minh và thiết thực**. Hãy thử tưởng tượng:

- Sách dạy nấu ăn thường là **tập hợp khổng lồ các công thức, mẹo nấu ăn và kiến thức ẩm thực**.
- Tuy nhiên, khi bạn cần tìm **một công thức phù hợp với nguyên liệu bạn có**, hoặc **thay thế một thành phần vì dị ứng** – việc tìm kiếm trong sách là rất khó khăn.

Đây chính là **bài toán lý tưởng cho RAG**: Làm thế nào để AI có thể **hiểu, tìm kiếm và trả lời các câu hỏi ngữ cảnh** dựa trên dữ liệu rối rắm này?

Case Study – Từng Bước Giải Quyết

Chúng ta sẽ đi qua quy trình xử lý cụ thể như sau:

Bước 1: Chuyển Đổi PDF Thành Hình Ảnh

Tài liệu ban đầu là một file PDF – chứa các trang sách với công thức nấu ăn, mẹo vặt và hình ảnh món ăn. Tuy nhiên, **GPT không đọc được PDF tốt**.

Vì vậy, chúng ta sẽ:

- Dùng công cụ như `pdf2image` để **chuyển các trang PDF thành ảnh**.
- Ví dụ: Trang PDF chứa công thức "Bánh mì bơ tỏi" sẽ được chuyển thành `page_1.jpg`.

Bước 2: Dùng GPT cho OCR và Hiểu Nội Dung

Khác với OCR truyền thống (như Tesseract chỉ nhận diện ký tự), chúng ta sẽ:

- Dùng GPT để "đọc hiểu" nội dung từ hình ảnh, không chỉ trích xuất văn bản.
- GPT có thể hiểu bố cục, cấu trúc và nội dung như:
 - Tên món ăn
 - Danh sách nguyên liệu
 - Các bước nấu
 - Gợi ý hoặc chú thích

Ví dụ thực tế: Với ảnh trang công thức, GPT có thể trả về kết quả như:

```
{  
  "title": "Bánh Mì Bơ Tỏi",  
  "ingredients": ["Bánh mì baguette", "Bơ", "Tỏi băm", "Mùi tây"],  
  "steps": ["Làm nóng lò", "Trộn bơ với tỏi", "Phết lên bánh mì", "Nướng 10 phút"],  
  "tags": ["ăn nhẹ", "chay", "phù hợp cho bữa tối"]  
}
```

Bước 3: Làm Sạch và Cấu Trúc Dữ Liệu

Không phải dữ liệu nào cũng cần giữ lại. Chúng ta sẽ:

- Loại bỏ các đoạn quảng cáo, nội dung không liên quan
- Chuẩn hoá cấu trúc dữ liệu để AI có thể tìm kiếm tốt hơn

Ví dụ: Một trang sách có thể chứa cả lời tác giả, nhưng chúng ta chỉ giữ phần công thức.

Bước 4: Tạo Embeddings cho Truy Xuất

Dữ liệu sau khi đã sạch sẽ và được cấu trúc, sẽ được chuyển thành **embeddings** – đại diện dạng số cho AI:

- Chúng ta sẽ dùng OpenAI Embedding API hoặc `text-embedding-3-small` để tạo vector.
- Các vector này sẽ được lưu vào hệ thống truy xuất như **FAISS** hoặc **Weaviate**.

Điều này giúp chúng ta tìm kiếm nội dung nhanh chóng bằng ngữ nghĩa, không chỉ từ khoá.

Bước 5: Xây dựng RAG & Trả Lời Câu Hỏi

Đây là phần "ma thuật":

Bạn sẽ xây dựng một hệ thống có thể **trả lời các câu hỏi phức tạp**, ví dụ như:

- "Tôi chỉ có bột mì, trứng và sữa, tôi có thể nấu món gì?"
- "Tôi bị dị ứng hạt, có thể thay thế hạnh nhân bằng gì trong món bánh này?"
- "Hãy gợi ý một món tráng miệng cho tiệc sinh nhật 10 người."

Với RAG, hệ thống sẽ:

- Truy xuất** đoạn văn phù hợp từ sách nấu ăn (retrieval)
- Dùng GPT để trả lời có ngữ cảnh, tự nhiên và hữu ích** (generation)

Ứng Dụng Thực Tế Rộng Hơn

Mặc dù chúng ta đang dùng sách nấu ăn để thực hành, bạn hoàn toàn có thể áp dụng quy trình này vào các lĩnh vực khác:

- Nghiên cứu khoa học (PDF nghiên cứu)
- Tài liệu pháp lý (hợp đồng, luật)
- Review sản phẩm khách hàng
- Báo cáo tài chính

Tóm tắt các bước:

Bước	Mô tả
1	Chuyển PDF sang ảnh
2	Dùng GPT để trích xuất và hiểu nội dung (OCR nâng cao)
3	Làm sạch, cấu trúc và gắn thẻ dữ liệu
4	Tạo embeddings và lưu vào cơ sở tìm kiếm
5	Truy xuất và tạo phản hồi bằng RAG

Tác giả: **Đỗ Ngọc Tú**
Công Ty Phần Mềm **VHTSoft**

Phiên bản #1

Được tạo 5 tháng 5 2025 14:19:21 bởi Đỗ Ngọc Tú

Được cập nhật 6 tháng 5 2025 09:52:34 bởi Đỗ Ngọc Tú