

# Báo Cáo So Sánh & Đề Xuất Giải Pháp OCR Tối Ưu Cho Hekate

---

## I. Mục tiêu & Bối cảnh

---

Trong bối cảnh Hekate triển khai AI Agent phục vụ khách hàng B2B và B2NGO, nhu cầu xử lý tài liệu hình ảnh bằng OCR trở nên cấp thiết. Giải pháp OCR cần:

- Hỗ trợ tiếng Anh & Pháp với độ chính xác cao.
  - Tốc độ xử lý tốt trên ảnh scan hoặc tài liệu có bố cục phức tạp.
  - Dễ tích hợp vào hệ thống AI hiện tại.
  - Bảo mật, triển khai linh hoạt cloud hoặc on-premise.
  - Chi phí thấp hoặc miễn phí.
- 

## II. Danh sách giải pháp OCR được phân tích

---

### 1. Tesseract OCR

---

- Open-source, do Google duy trì.
- - Hỗ trợ >100 ngôn ngữ, dễ tích hợp.
  - Độ chính xác thấp với ảnh kém chất lượng hoặc bố cục phức tạp.

### 2. Google Cloud Vision OCR

---

- Dịch vụ thương mại từ Google Cloud.
- - Độ chính xác rất cao, mạnh mẽ với ảnh phức tạp.
  - Tính phí theo API, yêu cầu kết nối mạng.

### 3. EasyOCR

---

- Open-source, dùng deep learning (PyTorch).
- - Tốc độ nhanh, hỗ trợ >80 ngôn ngữ.
  - Khó xử lý bảng biểu hoặc layout không chuẩn.

### 4. Kraken OCR

---

- Mã nguồn mở, mạnh với tài liệu cổ & viết tay.
- - Có thể huấn luyện lại mô hình.
  - Chậm, khó tích hợp, chưa tối ưu cho văn bản hiện đại.

### 5. PaddleOCR

---

- Framework deep learning từ Baidu.
- ◦ Rất chính xác, hỗ trợ detection + recognition, xuất JSON, chạy offline hoặc cloud.
- ◦ Cần cấu hình ban đầu nhưng tài liệu tốt.

### III. Tiêu chí đánh giá

Tiêu chí	Mô tả
Accuracy	Tỷ lệ nhận dạng đúng ký tự/từ/đoạn văn
Latency	Tốc độ xử lý trên ảnh trung bình
Multilingual	Số ngôn ngữ hỗ trợ, đặc biệt là tiếng Anh và Pháp
Ease of Integration	Khả năng tích hợp với Python/API, pipeline hiện có
Cost/License	Chi phí bản quyền, vận hành infra (cloud/on-premise)

### IV. Bảng so sánh tổng hợp

Công cụ	Accuracy	Latency	Multilingual	Dễ tích hợp	Chi phí sử dụng	Tổng điểm (1–5)
Tesseract	3	4	5	4	5	21
Google OCR	5	3	5	5	2	20
EasyOCR	4	5	4	4	5	22
Kraken	4	2	3	3	5	17
PaddleOCR	5	5	5	5	5	25

### V. Phân tích Trade-off

Cặp Trade-off	Nhận xét
Accuracy vs. Latency	Google OCR và PaddleOCR đều chính xác cao, nhưng PaddleOCR nhanh hơn.
Open-source vs. Thương mại	Google OCR mạnh nhưng tốn phí. PaddleOCR, EasyOCR miễn phí và mở rộng.
On-premise vs. Cloud	Google OCR chỉ chạy trên cloud. PaddleOCR linh hoạt hơn nhiều.

## VI. Giải pháp đề xuất

---

### Đề xuất sử dụng: PaddleOCR

#### Lý do chọn PaddleOCR:

---

- **Accuracy cao:** xấp xỉ Google Cloud Vision.
  - **Tốc độ nhanh:** phù hợp xử lý hàng loạt.
  - **Đa ngôn ngữ:** hỗ trợ >80 ngôn ngữ, trong đó có tiếng Pháp và Anh.
  - **Dễ tích hợp:** Python API, có module detection + recognition.
  - **Triển khai linh hoạt:** chạy tốt trên cả on-premise & cloud.
  - **Miễn phí hoàn toàn:** không phát sinh license phí hay chi phí API.
-