**Open AI and Azure - Hybrid Architecture**

Kevin 20250407

Let’s design a **hybrid AI OCR architecture** that combines the strengths of **Azure OCR** for structured field extraction and **OpenAI Vision (GPT-4V)** for semantic reasoning, summaries, and edge cases like tax documents or financial reports.

**🧠 Hybrid OCR Architecture — High-Level Overview**

**🧩 Components:**

1. **Document Ingestion**
   * **Upload Center: supports file types (.pdf, .jpg, .png, etc.)**
   * **Classifier (based on keywords or AI): detects document type (invoice, bank statement, etc.)**
2. **Routing Logic**
   * **Rule-based or ML-based dispatcher:**
     + **Azure OCR** for invoices, receipts, statements, bills, etc.
     + **OpenAI Vision** for reasoning, summaries, validation, and free-form documents.
3. **Azure OCR Module**
   * **Uses:**
     + **Prebuilt Models**: Invoices, Receipts, Business Cards
     + **Custom Trained Models**: T4, T2, CRA, tax letters
     + Outputs: JSON with labeled fields (vendor, total, line items)
4. **OpenAI Vision Module**
   * **Uses:**
     + Prompt templates based on doc type
     + Extracts missing fields, summarizes key info, performs anomaly checks
   * **Useful for:**
     + Reasoning: “Is this payment overdue?”
     + Summaries: “Who is the sender and what's the total due?”
     + Labeling: T4 → “Line 14: Employment Income = $45,000”
5. **Post-processing & Validation**
   * **Merges Azure + OpenAI outputs**
   * **Adds confidence scoring, highlights low-confidence fields**
   * **Human-in-the-loop (optional): reviewers approve flagged fields**
6. **Output to Excel / Database**
   * **Tabular export of line items**
   * **Summary table with 7 fields (vendor, date, subtotal, tax, total, payment method, doc type)**

**🗂️ Document Routing Strategy**

| **Doc Type** | **Route to Azure** | **Route to GPT-4V** | **Notes** |
| --- | --- | --- | --- |
| Invoices | ✅ | ✅ (only if validation needed) | Azure extracts fields, GPT checks due/past-due |
| Receipts | ✅ | ❌ | Azure handles POS and restaurant receipts |
| Credit Card Statements | ✅ | ✅ (for description explanation) | Azure for table, GPT for unusual charges |
| T4 / T2 / Tax Forms | ✅ (custom model) | ✅ (labeling & reasoning) | Combine both |
| Financial Statements | ⚠️ (limited) | ✅ (summarization, checks) | GPT can extract key metrics |
| Hydro/Utility Bills | ✅ | ✅ (for summaries) | Azure extracts, GPT adds insight |
| Property Tax | ✅ | ✅ | Azure extracts balance due, GPT adds payment instructions |
| General Letters | ❌ | ✅ | GPT excels at summarizing letter intent |

**⚙️ Tech Stack Example**

| **Layer** | **Tools** |
| --- | --- |
| **Frontend** | React + Upload Centre (Browse / Email Forward / API Link) |
| **Backend** | FastAPI or Node.js |
| **OCR Processor** | Azure Document Intelligence SDK + GPT-4 Vision |
| **Document Classifier** | Azure Form Recognizer / Custom AI model |
| **Storage** | Azure Blob or AWS S3 |
| **Database** | PostgreSQL or MongoDB |
| **Excel Export** | Pandas + OpenPyXL |
| **Queue/Batch** | Celery or Azure Functions |

**🔐 Bonus: Confidence-Driven Smart Routing**

You can also route documents based on OCR confidence:

* If Azure's confidence < 85%, send to OpenAI Vision for fallback.
* Or, show to a human validator via UI with highlighted low-confidence fields.

A **visual flowchart** of this hybrid architecture



