O-1A Assessment App: Design Overview

1. Purpose

We created this application to **quickly gauge** how likely someone is to qualify for an O-1A visa. It checks a CV for keywords related to the eight O-1A criteria and then provides:

1. **A list of criteria matched** (with evidence from the CV).
2. **A rough rating** (low, medium, or high).

2. Architecture at a Glance

1. **File Parsing**
    - Reads .txt or .docx files.
    - Converts each file to plain text.
2. **Keyword Extraction**
    - We use a **simple keyword-based approach** for identifying evidence under each O-1A criterion (Awards, Membership, Press, etc.).
3. **Scoring**
    - Counts how many criteria have matches:
        - **0–2 matches** → low
        - **3–5 matches** → medium
        - **6+ matches** → high
4. **API (FastAPI)**
    - Single endpoint /evaluate_cv takes the CV file and returns JSON with:
        - **Evidence** for each criterion.
        - **Overall rating**.

3. Why These Choices?

- **Ease of Use**: Simple keywords are fast to implement and understand.
- **Extendability**: If we need better accuracy later, we can swap in more advanced NLP models.
- **Maintainability**: Each step—parsing, extraction, scoring—is a separate module, so changing one doesn't break the others.

4. How We Evaluate Output

1. **Manual Checks**

- We test with known CVs containing specific keywords (e.g., “Award,” “Member,” “Invented”), then see if the app’s output is correct.

5. Future Improvements

- **More Intelligent NLP**: Use entity recognition or machine learning for fewer false positives.
- **Weighted Scoring**: Not all evidence is equal (e.g., a Nobel Prize vs. a local prize).
- **Feedback Loop**: Let real users correct outputs, then learn from that data.