Adobe Hackathon – Connecting the Dots

This document contains Round 1A and 1B task breakdowns with terminal-style output screenshots (blue text).

Round 1A – PDF Outline Extraction

Task: Extract H1, H2, H3 headings from a PDF with page numbers.

Execution must happen offline in a Docker container under 60 seconds using only CPU.

```
$ python outline_extractor.py DesignSystems.pdf

{
  "document": "DesignSystems.pdf",
  "outline": [
    { "level": "H1", "title": "Introduction to Design Systems", "page": 1 },
    { "level": "H2", "title": "Why Design Systems Matter", "page": 2 },
    { "level": "H1", "title": "Adobe Spectrum", "page": 3 }
  ]
}
```

Round 1B – Persona-based Retrieval

Task: Given multiple PDFs and a persona-task pair, extract and rank the most relevant sections.

Constraints: Offline, Docker, CPU-only, 60 seconds.

```
$ python semantic_ranker.py --persona "Product Designer starting out with Adobe tools"

{
  "task": "Understand basic UI design principles using Adobe XD",
  "persona": "Product Designer starting out with Adobe tools",
  "relevant_sections": [
    {
      "document": "AdobeXDGuide.pdf",
      "section_title": "Designing with Components",
      "score": 0.95
    },
    {
      "document": "DesignThinking.pdf",
      "section_title": "User-Centered Design",
      "score": 0.88
    }
  ]
}
```