Testing Guide for Intelligent Document Analyst System

Quick Testing Instructions

Prerequisites

1. Install Docker on your system

2. Extract the zip file to a directory

3. Open terminal/command prompt in the extracted directory

Build and Test Steps

1. **Build the Docker image:**

2. **Test with provided sample data:**

3. **Check the output:**

   - Look for `challenge1b_output.json` in the `test_output_new` directory

   - Compare with the expected output in `test_output/challenge1b_output.json`

Testing with Your Own Data

1. **Prepare your input directory:**

2. **Run with your data:**

Expected Performance

- Processing time: < 60 seconds for 3-5 documents

- Memory usage: < 1GB

- CPU-only execution

Troubleshooting

- If Docker build fails, ensure you have internet connection for downloading dependencies

- If processing fails, check that PDF files are readable and not corrupted

- Verify input JSON format matches the examples provided

Sample Test Cases Included

1. **Travel Planning** (test_input/): South of France travel documents

2. **Collection Examples**: Three reference collections from the GitHub repository

The system has been tested and validated with these scenarios. You can use them as benchmarks for your own testing.