

# Assignment 3: Entity Extraction for Hospital Knowledge Graph

---

## How to Run

### Prerequisites

- Python 3.8+
- OpenAI API key

### Setup Instructions

#### 1. Install dependencies:

```
pip install openai python-dotenv pandas scikit-learn
```

#### 2. Configure API key: Create a `.env` file in this directory with:

```
OPENAI_API_KEY=your-openai-key-here
```

#### 3. Run the notebook:

- Open `Assignment_3_Entity_Extraction.ipynb` in Jupyter or VS Code
- Run all cells in order

## What the Notebook Does

**Task 1 (Cell 3-5):** Tests the entity extractor on sample text

- Extracts Equipment and Department entities
- Shows extraction results

**Task 2 (Cell 6-8):** Evaluates on 12 real hospital texts

- Loads `evaluation_dataset.json`
- Calculates Precision, Recall, F1 scores
- Saves results to `evaluation_results.csv`

**Task 3 (Cell 9-11):** Tests LLM-as-a-judge

- Evaluates extraction correctness
- Shows sample judgments

## Expected Output

After running all cells, you should see:

- Overall F1 Score: ~0.75
- Equipment F1: ~0.90
- Department F1: ~0.60
- Detailed results in [evaluation\\_results.csv](#)