# Internship Application Organizer — Work Plan & Documentation

Internship Application Organizer — Work Plan & Documentation

Project Summary

----------------

Build a personal tool to automatically ingest, organize, and track internship/job applications from incoming emails and application portals. The tool will parse confirmation emails, extract structured fields (company, title, job id), store the entries in a database, let the user track statuses, and export to CSV/Excel.

MVP Features

------------

- Connect to email (Gmail via OAuth) and fetch recent messages.

- Auto-detect “application” emails using keyword rules + lightweight classifier.

- Extract structured fields: company, role/title, job id, application date, platform, source link, attachments.

- Dashboard to view entries, edit status, export to CSV.

Architecture

------------

- Frontend: React (minimal scaffold included).

- Backend: FastAPI (app/backend).

- Database: SQLite (dev) / Postgres (prod).

- Optional workers: Celery/RQ for scraping and heavier tasks.

- Gmail API for ingestion; Playwright for portal scraping (optional).

Mermaid Flowcharts

------------------

Flowchart (Mermaid)

```

flowchart LR

A[User Email Provider (Gmail/IMAP)] -->|OAuth/API| B[Ingestion Worker]

B --> C{Heuristics / Classifier}

C -->|match| D[(Postgres/SQLite)]

C -->|no match| E[Manual Review Queue]

D --> F[Web UI / Dashboard]

F --> G[Export CSV / Excel]

F --> H[Status Tracking / Notes]

D --> I[Optional: Portal Scraper / API Poller]

```

Sequence Diagram (Mermaid)

```

sequenceDiagram

participant U as User

participant A as App (FastAPI)

participant G as Gmail API

participant DB as Database

U->>A: Click "Connect Gmail"

A->>G: OAuth flow (user consents)

G-->>A: Token

A->>G: Fetch messages

A->>A: Parse & extract

A->>DB: Store applications

U->>A: View dashboard / export

```

Resume-ready bullets

--------------------

- Built an email ingestion and job-application tracking tool that automatically detects application confirmation emails, extracts structured fields (company, role, job ID, platform) and stores them for quick review and export.

- Implemented parsing heuristics (regex) and created a path to improve recall with a TF-IDF + Logistic Regression classifier.

- Developed a REST API (FastAPI) and lightweight React dashboard to review entries, update statuses, and export CSV/Excel.

- Integrated Gmail OAuth (readonly) for secure ingestion and followed privacy best practices.

Tech stack

----------

- Python (FastAPI, SQLModel), JavaScript (React), Gmail API (google-api-python-client), SQLite/Postgres, scikit-learn (optional), pandas, Playwright (optional).