

FileFlip



Show Image

FileFlip is a web application that converts PDF documents to CSV, XLSX, and other worksheet formats, specifically designed for accounting data that can be uploaded to Sage accounting software.

Features

- **PDF Table Extraction:** Automatically detect and extract tables from PDF documents
- **Multiple Format Support:** Convert to CSV, XLSX, or Sage-compatible formats
- **Data Preview:** View extracted data before conversion
- **Smart Column Mapping:** Intelligent detection of common accounting fields
- **Batch Conversion:** Process multiple tables at once
- **Modern UI:** Clean, responsive interface built with IBM Carbon Design System

Tech Stack

Backend

- Python 3.10+
- FastAPI for API endpoints
- pdfplumber and tabula-py for PDF data extraction
- pandas for data manipulation
- Docker for containerization

Frontend

- TypeScript
- React 18
- Tailwind CSS
- IBM Carbon Design System

- Vite build system

Getting Started

Prerequisites

- Docker and Docker Compose
- Node.js (v16+)
- Python (v3.10+)

Local Development

1. Clone the repository:

```
bash

git clone https://github.com/jamtax/FileFlip.git
cd FileFlip
```

2. Start the development environment using Docker Compose:

```
bash

docker-compose up
```

3. Access the application:

- Frontend: <http://localhost:3000>
- Backend API: <http://localhost:8000/docs>

Manual Setup (without Docker)

Backend

1. Navigate to the backend directory:

```
bash

cd backend
```

2. Create a virtual environment:

```
bash

python -m venv venv
source venv/bin/activate # On Windows: venv\Scripts\activate
```

3. Install dependencies:

```
bash  
  
pip install -r requirements.txt
```

4. Run the backend server:

```
bash  
  
uvicorn app.main:app --reload
```

Frontend

1. Navigate to the frontend directory:

```
bash  
  
cd frontend
```

2. Install dependencies:

```
bash  
  
npm install
```

3. Start the development server:

```
bash  
  
npm run dev
```

Usage

1. **Upload PDF:** Drag and drop or select a PDF file containing tabular data
2. **Review Tables:** Preview the detected tables and their content
3. **Configure Conversion:** Select output format and adjust settings as needed
4. **Download:** Convert and download the data in your chosen format

Development Guidelines

Code Structure

- **Backend:**
 - `app/api/` - API endpoints

- `app/services/` - Business logic
- `app/models/` - Data models
- **Frontend:**
 - `src/components/` - Reusable UI components
 - `src/pages/` - Application pages
 - `src/services/` - API integration
 - `src/hooks/` - Custom React hooks

Adding a New Format

To add a new output format:

1. Add the new format option to the `ConversionOptions` interface in `frontend/src/types/index.ts`
2. Implement the conversion logic in `backend/app/services/pdf_extractor.py` (in the `DataConverter` class)
3. Update the frontend UI in `frontend/src/pages/ConversionPage.tsx` to include the new format option
4. Add appropriate format-specific configuration options

Deployment

Using GitHub Actions (CI/CD)

The repository includes a GitHub Actions workflow that:

1. Runs tests for both backend and frontend
2. Builds Docker images
3. Pushes images to Amazon ECR
4. Updates the ECS service

To configure:

1. Set up the following GitHub secrets:
 - `AWS_ACCESS_KEY_ID`
 - `AWS_SECRET_ACCESS_KEY`

- `AWS_REGION`

2. Create the necessary ECR repositories and ECS services

Manual Deployment

For manual deployment:

1. Build Docker images:

```
bash

docker build -t fileflip-backend ./backend
docker build -t fileflip-frontend ./frontend
```

2. Tag and push to your container registry

3. Deploy to your infrastructure (AWS, GCP, Azure, etc.)

License

This project is licensed under the MIT License - see the LICENSE file for details.

Acknowledgements

- [pdfplumber](#) for PDF text extraction
- [tabula-py](#) for PDF table extraction
- [pandas](#) for data processing
- [FastAPI](#) for the API framework
- [IBM Carbon Design System](#) for UI components
- [Tailwind CSS](#) for styling
- [React](#) for the frontend framework

Contact

For questions or support, please contact:

- Email: jolean@jamtax.co.za
- Phone: 079 765 6234
- Website: jamtax.co.za