

Kabelo's TAX Consulting Doc App

Kabelo's TAX Consulting Doc App

A desktop-friendly tool to convert legal/finance/tax documents into:

- **Plain-English rewrites** (faithful, easy to understand)
- **Bullet-point summaries** (by section, preserving legal context)

The app supports batch processing, parallel LLM calls, and a modern GUI.

Features

- **Input formats:** PDF, DOCX, TXT
- **Outputs:** Plain-English + Summary in DOCX and/or TXT (optional PDF)
- **Batch mode:** Process all files in a folder (recursive)
- **Parallel processing:** `--workers N` for speed
- **Modern GUI:** Dark/light themes, toasts, drag-and-drop (optional)
- **Robust prompts:** Section-preserving, safe chunking around token limits
- **Logging:** Per-run log in your output folder

Requirements

- **Python 3.10+**
- A valid **OpenAI API key**
- Python packages (installed via `requirements.txt` or automatically by the launcher):
 - ``openai``, ``python-docx``, ``pdfplumber``, ``PyPDF2``, ``reportlab``, ``tqdm``
 - GUI extras: ``ttkbootstrap``
- For macOS GUI: Tkinter is included with most Python.org builds

Folder Layout (recommended)

'''

your-project/

├─ legal_converter.py

├─ Myinterface.py

├─ run_gui.sh # macOS/Linux launcher

├─ launch_gui.py # cross-platform Python launcher

├─ requirements.txt

├─ .env # your API key (NOT committed)

├─ in/ # put your input docs here

└─ out/ # outputs will appear here

'''

Configure your API key (two options)

Option A — `.env` file (recommended)

Create a file named `.env` **in the project root** with:

'''

OPENAI_API_KEY= Past the API Key here

'''

Ensure your code loads it (already included in recent scripts):

```python

from dotenv import load\_dotenv

load\_dotenv()

'''

Install helper once:

```
```bash  
  
pip install python-dotenv  
  
```
```

#### Option B — Environment variable (macOS/Linux)

```
```bash  
  
echo 'export OPENAI_API_KEY="sk-REPLACE_WITH_YOUR_KEY"' >> ~/.zshrc  
  
source ~/.zshrc  
  
```  

```



#### 1) Create and activate a virtual environment

**\*\*macOS / Linux\*\***

```
```bash  
  
python3 -m venv .venv  
  
source .venv/bin/activate  
  
```
```

**\*\*Windows (PowerShell)\*\***

```
```powershell  
  
py -m venv .venv  
  
.\.venv\Scripts\Activate.ps1  
  
```
```

## 2) Install dependencies

```
```bash
```

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

GUI :

```
pip install ttkbootstrap tkinterdnd2 python-dotenv
```

```
```
```

```

```

## Run the App (GUI)

### Option 1 — Modern GUI (recommended)

```
```bash
```

activate venv first

```
python3 Myinterface.py
```

```
```
```

- Pick **Input** and **Output** folders.
- Select **Model** (default: `gpt-4o-mini`), **Formats** (`docx,txt`), and **Workers** (e.g., 3).
- Click **Run Conversion**.

### Option 2 — Shell launcher (macOS/Linux)

```
```bash
```

```
chmod +x run_gui.sh
```

```
./run_gui.sh
```

```
```
```

### Option 3 — Universal Python launcher

```
```bash
```

```
python3 Myinterface.py
```

...

Works on macOS/Windows. Creates `.venv` if missing, installs requirements, launches the GUI.

Quick Smoke Test

Create a tiny input file:

```
```bash
mkdir -p in out
printf "SECTION 1. TERM. This Agreement begins on the Effective Date." > in/demo.txt
python3 Myinterface.py
```

### choose in=./in, out=./out, then Run

...

Outputs: `out/demo/demo\_plainEnglish.docx/.txt` and  
`out/demo/demo\_summary.docx/.txt` (names may vary).

---

### Command-line (batch) mode

You can also run the converter without the GUI:

```
```bash
python3 legal_converter.py --input "/absolute/path/to/in" --output
"/absolute/path/to/out" --model gpt-4o-mini --formats docx,pdf --workers 3
```

...

****Notes:****

- Input is ****recursive****; all `.pdf`, `.docx`, `.txt` inside are processed.
- Use absolute paths to avoid “folder not found” issues.

Key Options

- `--input`` : Input folder (recursive discovery)
- `--output``: Output root folder
- `--model`` : LLM model (e.g., `gpt-4o-mini``, `gpt-4o``)
- `--formats``: `docx,txt`` (and `pdf`` if `reportlab`` installed)
- `--workers``: Parallel file workers (1–16; beware of API rate limits)

Troubleshooting

****“API key is not set / missing”****

- Ensure `.env`` exists and contains `OPENAI_API_KEY=...``
- Or export in the shell: `export OPENAI_API_KEY="sk-..."``
- Restart the app/Terminal after changes

****“No input files found”****

- Use absolute paths; confirm the folder really has `.pdf/.docx/.txt``
- macOS/Linux: `find ./in -type f \(-iname "*.pdf" -o -iname "*.docx" -o -iname "*.txt" \)``

****“quote>” prompt in Terminal****

- You’re missing a closing quote in a command; press ****Ctrl+C**** and retype with matching quotes.

****Permissions on macOS****

- `chmod +x run_gui.sh``
- If a downloaded file is blocked: `xattr -d com.apple.quarantine <file>``

****Tkinter / ttkbootstrap missing****

- Install: ``pip install ttkbootstrap``
- If ``_tkinter`` errors on macOS, install Python from python.org or update your Tk.

Support

If you hit an error, copy the ****last 20 lines**** from the log panel or ``out/processing.log`` and share them with me [email: Ratshefola.kt@outlook.com]. Include:

- Your OS (macOS/Windows)
- Python version (``python3 --version``)
- The exact command you ran