RAG on My Data - Instructions

Instructions Starting from "Ingesting your files"

Ingesting your files

```
python ingest.py --folder /path/to/your/files --rebuild
```

Supported file types:

- Text: .txt, .csv, .md

- Documents: .doc, .docx, .pdf, .xls, .xlsx

- Images: .jpg, .jpeg, .png (using Moondream if available)

Searching your indexed data

```
python search.py --query "What is project management?"
```

You will see matched files and short text snippets.

Access via ChatBox (Mac -> Debian)

Ensure that your Debian server (Chatreey) has port 11434 open, then use ChatBox pointing to:

```
http://<ip_of_chatreey>:11434
```

Mounting FritzBox network drives

Install necessary tools:

```
sudo apt install cifs-utils
```

Create mount point:

RAG on My Data - Instructions

sudo mkdir -p /mnt/fritzbox

Create credentials file /etc/samba/credentials-fritzbox:

username=your_fritzbox_user

password=your_fritzbox_password

Secure it:

sudo chmod 600 /etc/samba/credentials-fritzbox

Add to /etc/fstab:

//fritz.box/FRITZ.NAS /mnt/fritzbox cifs credentials=/etc/samba/credentials-fritzbox,vers=2.1,iocharset=utf8,uid=1000,gid=1000,file_mode=0777,dir_mode=0777 0 0

Mount:

sudo mount -a

Future plans

- Add OCR for scanned PDFs
- Add scheduler for continuous background ingestion
- Add export of results in CSV