

NoteSmoosh

Build Status

NoteSmoosh is an application that is aimed to help law school students (and others) by compiling study products from various formats and writing it into one cohesive document.

Functionalities include: - Text extraction from PDF, DOCX, PPTX, TXT, JPG and PNG - Text comparison and compilation based on similarity - Extraction for textbook images and handwritten text

Libraries

NoteSmoosh uses a number of open source projects to work properly:

- Textract - text extracting tool for multiple file types.
- appJar - GUI Interface for Python programming language
- Path - For reading and copying files
- sklearn - imported TFIDF Vectorizer which is used in text comparisons
- NLTK - Natural Language Language Processing (NLP) for lexical similarities
- Time - displays system time on the interface

Installation

NoteSmoosh requires either Windows or Linux (Ubuntu 16.04 or greater) to run.

NoteSmoosh also requires Python 3.5.0 or greater

For Linux

If you do not have the source directory already, you may get it from my GitHub which is <https://github.com/andrewcgarber/NoteSmoosh.git>

Once the files are pulled onto your machine navigate to the directory where NoteSmoosh_v2.py is located.

```
$ cd /your_directory_here/NoteSmoosh
```

Then, install all required libraries in order to run NoteSmoosh

```
$ pip install requirements.txt
```

Final step, run the program:

```
$ python3 NoteSmoosh_v2.py
```

Congrats! You are now running NoteSmoosh on Linux!

For Windows

If you do not have the source directory already, you may get it from my GitHub which is <https://github.com/andrewcgarber/NoteSmoosh.git>

Once the files are pulled onto your machine, use CMD or other command line tool navigate to the directory where NoteSmoosh_v2.py is located.

```
$ cd /your_directory_here/NoteSmoosh
```

Then, install all required libraries in order to run NoteSmoosh

```
$ py -3.5 -m pip install requirements.txt
```

Final step, run the program:

```
$ py -3.5 NoteSmoosh_v2.py
```

Congrats! You are now running NoteSmoosh on Windows!

Todos

- Merge on Headers full integration
- Web hosted application

License

MIT