**### DIRTY WORD SEARCHER**

**## Overview:**

1. This tool is intended to be used against any set of data. The user has the power to generate their own wordlists, target their own data.

**## Current Functionality:**

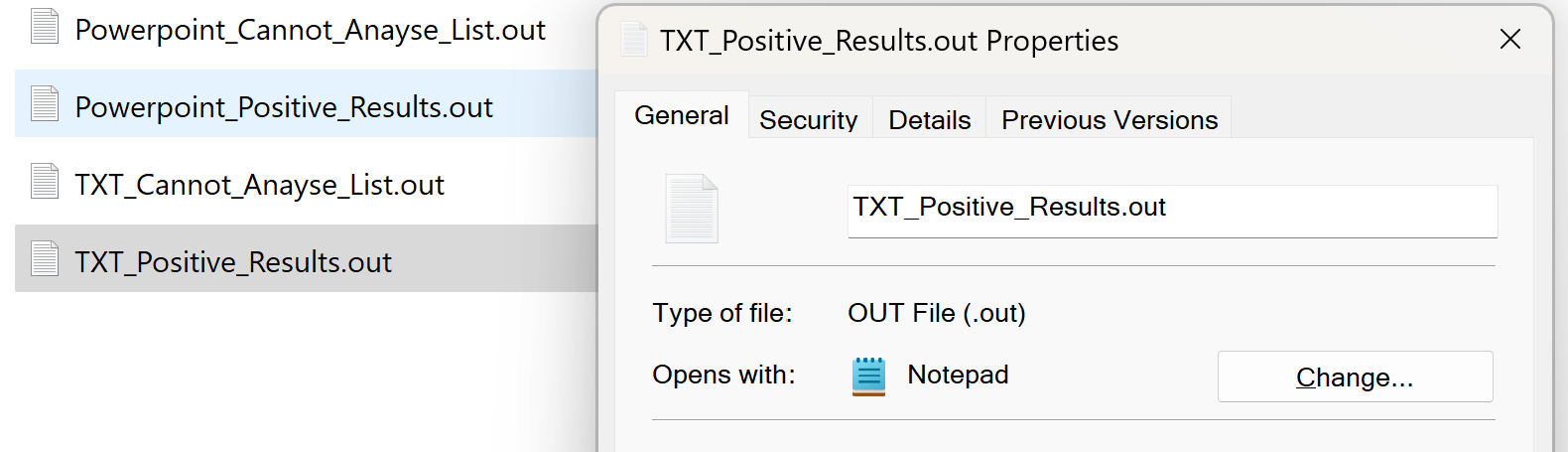
1. This tool takes a number of user inputs, including a dirty word list, a starting directory where the scan will start its search for files, and a destination directory where the user will specify they want the output to be targeted.
2. The following extensions are ones analysed by the dirty word searcher at time of writing this:

* Powerpoint documents (.ppt, .pptx)
* Word documents (.doc, .docx)
* Excel documents (.xls, .xlsx)
* PDF documents (.pdf)
* Text documents (.txt)

1. Warning: when alaysing `.txt` files, if you begin this at `C:\`, it may flag a lot of false positives depending on the dirty word list supplied.
2. The following are also searched for however no functionality to scan for dirty words at this time:

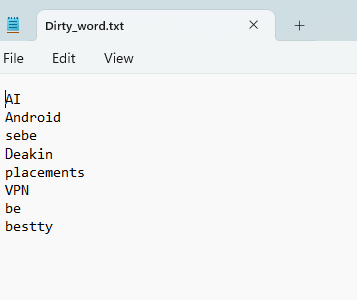
* .jpeg / .png / .bmp / .tiff / .mp3 / .mp4 / .avi / .mov / .msg

1. When searching for these file types, the script will make a text will that displays the full file path to all files of that file type, making it convenient for the user to see all files within the data set.
2. The output will be directed to the directory specified by the end user, and a new `RESULTS` directory will be created.
3. The files will be outputted as a `.out` file to avoid collision with the .txt scanner that searches for the `.txt` extension. When opening the files, you may need to specify which application to open with, `Right Click > Properties > Opens With > Notepad`

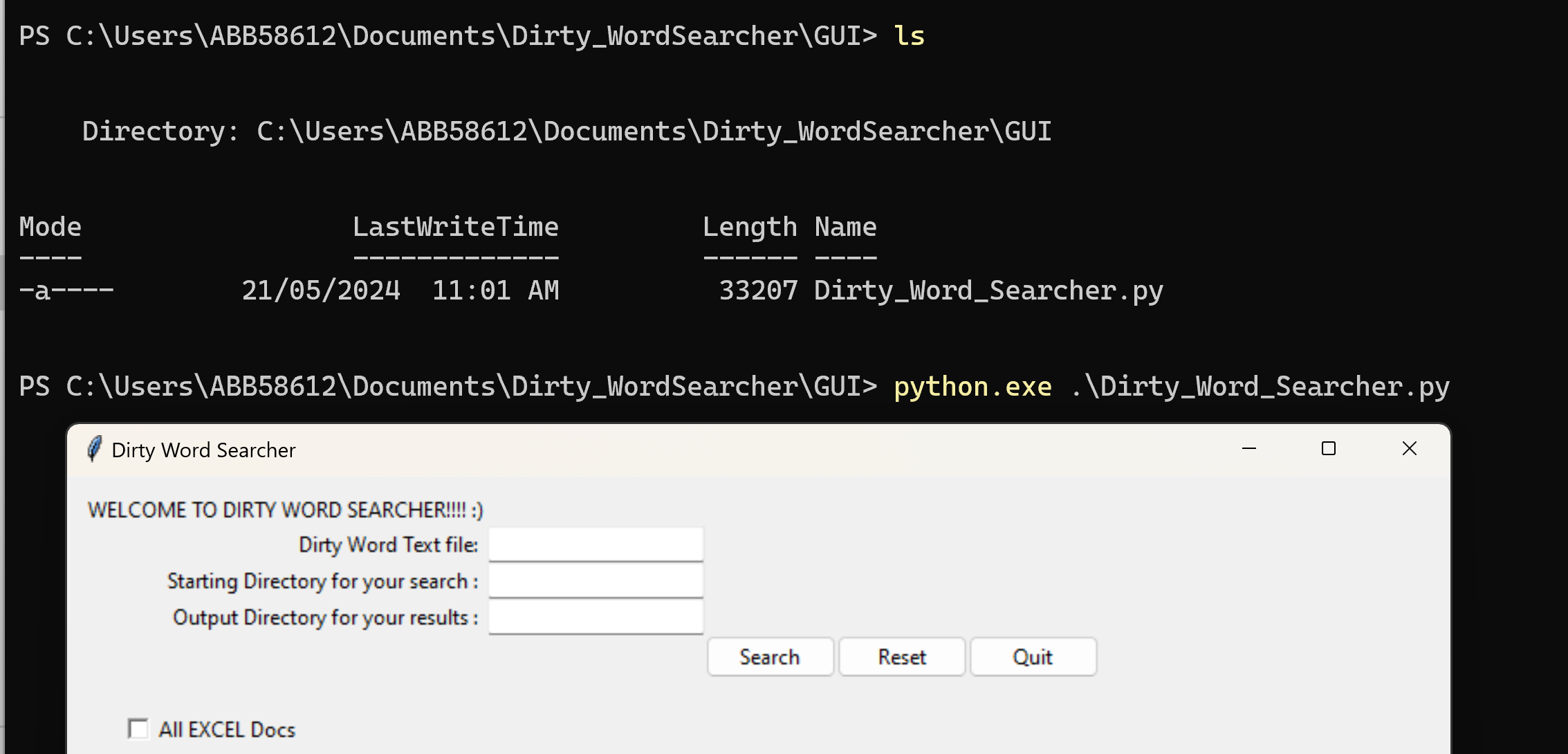


**## Usage:**

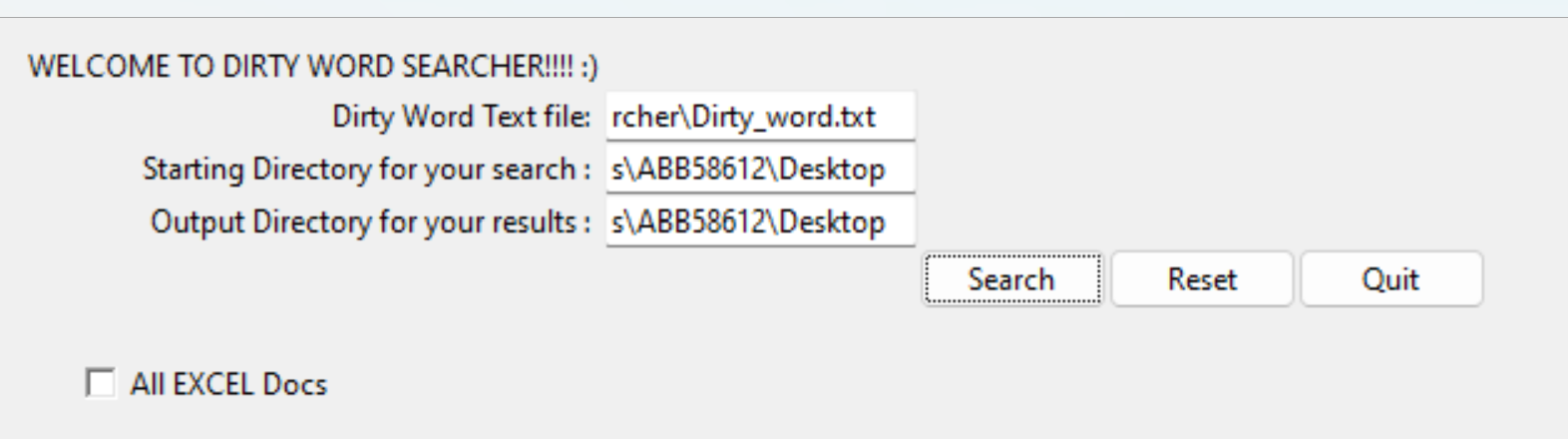
1. Create a Dirty Word List.



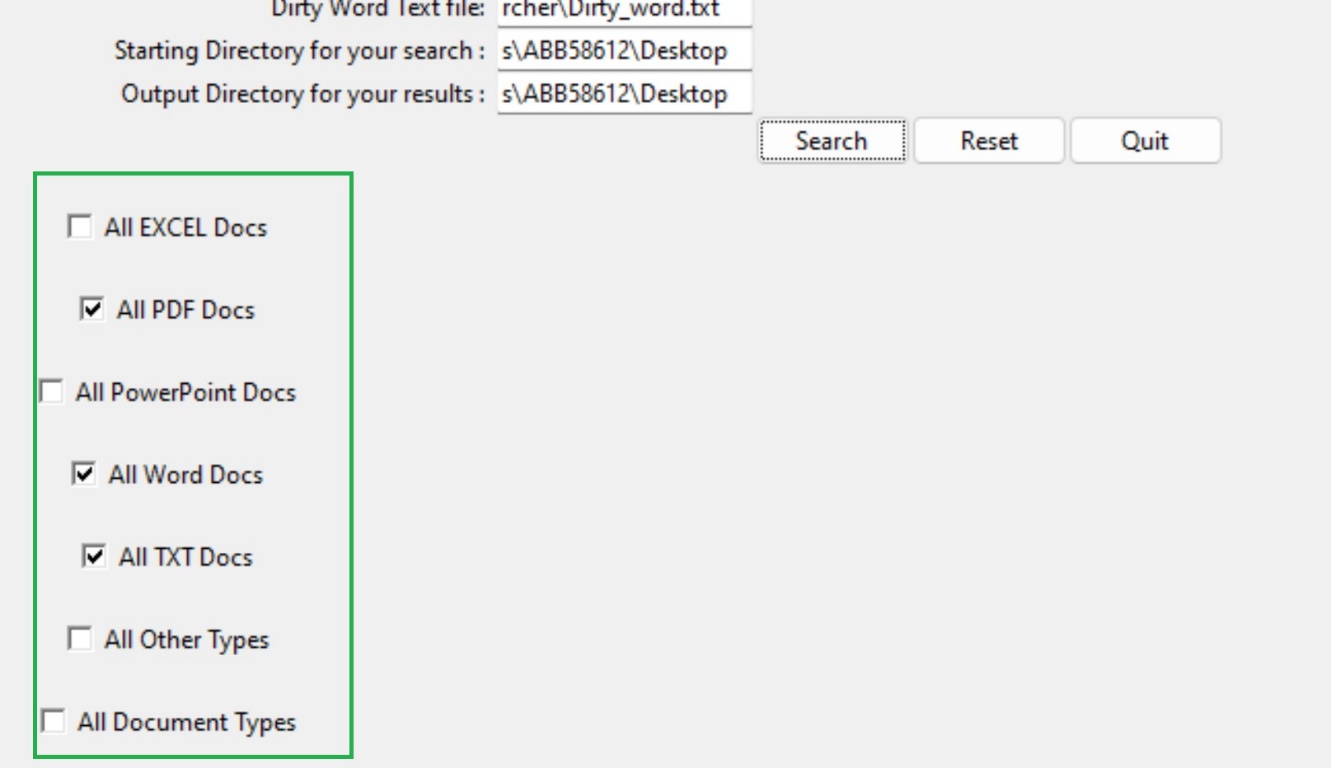
1. Identify starting directory path.
2. Identify a directory to output results.
3. Open PowerShell and execute the following command "python.exe "<path to epic\_gui.py>".



1. Copy file paths from step 1-3 and input them into respective fields.



1. Select files types you want to Scan.



1. Analyse output.

