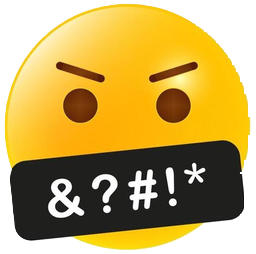
**README.DOCX**

**Dirty word Searcher**



Note: Prior to using the tool, it's essential to thoroughly review and comprehend its requirements, functionalities, and limitations. Any misuse of the tool leading to unexpected data outcomes will be the responsibility of the operator.

Contents

[1. Overview 3](#_Toc167796398)

[2. System Requirements 3](#_Toc167796399)

[3. User Options 3](#_Toc167796400)

[a. Dirty Word Text File 4](#_Toc167796401)

[b. Starting and Output directory 4](#_Toc167796402)

[c. File types 4](#_Toc167796403)

[4. Running 4](#_Toc167796404)

# Overview

DirtywordSearcher.exe is a Python-based executable designed to search for files containing specific words identified by the user. The program offers various fields and options that users must select to maintain control over the targeted artefacts and keywords of interest.

# System Requirements

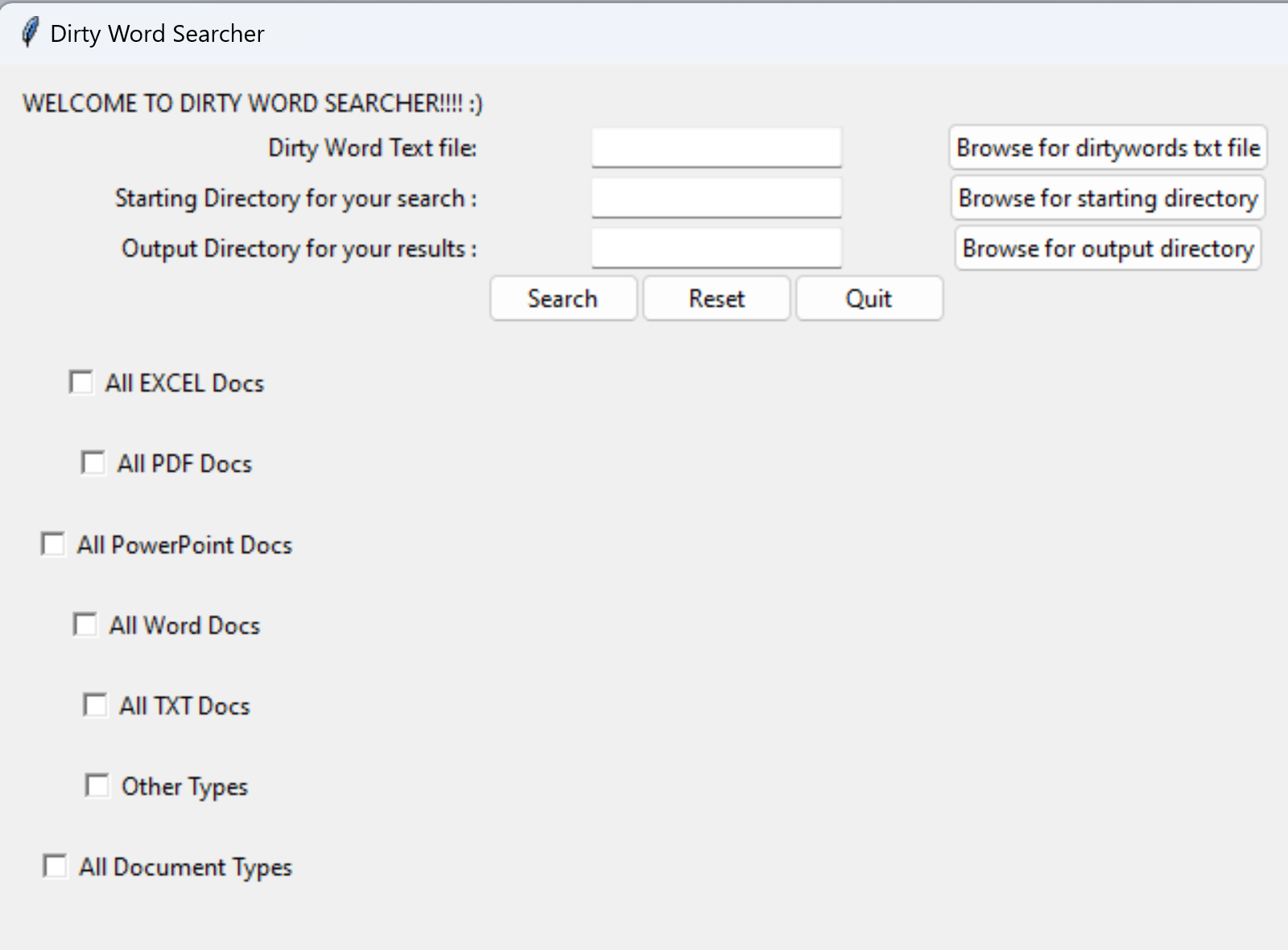
This executable is targeting windows based file systems and there are no requirements outside of the standalone executable and windows default program like notpade.exe.

Initial setup will require the Administrators password to ensure permissions are correctly set for the executable.

The program has no not been tested on any linux/unix like systems, and the team provides no level of assurance or confidence in the results it may/may not return.

# User Options

There a number of user inputs required by the user as seen bellow. User must have a valid txt file for the dirty words file, have a valid starting and output path.



## Dirty Word Text File

This is required to be a text file with a txt extension. For the results to return as expected ensure you use **EXACT** words and phrases, one per line. The program **will not** match partial words/phrases, so if there are a numerous variations of a word/phrase, please ensure you have all combinations present.

## Starting and Output directory

These directories can be manually inputted into their respective fields, or searched for with the browse buttons.

**Note:** The input requires at the minimum a drive letter (i.e C:\) and cannot do multiple drives simultaneously, this is due to a number of reasons, one being the performance of the program and volume of data it can ingest at any one time. If your device has multiple drives, you will be required to run the program against each drive letter.

## File types

The checkboxes down the left hand side give the users power to determine which artefacts they wish to interrogate.

The “Other Types” option will search for all other file types that have been requested to pass through the dirty word searcher. Due to the level of accuracy when interrogating movies and images for dirty words, the team as decided to highlight them for the user to investigate and **will not** be passed through dirty words text file. The files include the following extensions:

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

**Note:** when searching for “.txt” file types from C:\ depending on the dirty word list provided, the volume of files passed will be greater than others due to all the system documentation that resides at that level, increasing the chances of false positives and potential crashing of program.

# Running

