manual

SET UP:

1.to install under the code button on https://github.com/huzai786/fix-script click on it and download the zip file.

2. Unzip it.

3.To create a virtual environment,

type command:

> virtualenv env (creates virtual environment)
> . env/Scripts/activate (activate virtual environment)   [its actually a dot]
> pip install -r requirements.txt (install all dependencies)

(recommended to create virtual env to seperate the dependencies from globle settings)

--SET UP DONE--

USAGE:

1.to run the script type the following:

> python main.py

and every thing is self explainatory.

THINGS TO KEEP IN MIND:

All the errors that are based on the server response when data is requested are handled by simply evading them.

1. some website doesnt allow simple requests from python because js is not enabled, which is inturn give error. those errors are handled.

2. you will see 2 type of error, max retries, read timeout, they are caused when the url were trying to request data from is lazy or redirecting the request again and again.

STRUCTURE:

1.there are 3 folders and 1 file main.py that is relavent to the user

2.file folder will contain files with urls of the corresponding keyword. which will be used to iterate overr and extract emails.

3. once a keyword is done, in the output folder search_word_email.txt will be created which will contain the emails that are scraped from the urls.