Email Scrapper Algorithm

Christian Barlow

[cbhnn@mail.umkc.edu](mailto:cbhnn@mail.umkc.edu)

***Main Script:***

While running == true

user input file text

read text file

use url lib module to go to websites

use search function

slice out emails using mailto

user input output file

write to output file

ask user to run again

if user enters no

running = false

***User input:***

Input is invalid

While invalid input

enter file name

Return filename

***Read text file:***Assign user input to variable

Read text file

Return text file in readable format

***Url lib function:***

For all url in text file

If url invalid

display error

Else

go to website

return readable format

***Search and slice function***

Find email addresses using mailto in html

Slice from mail to anything that doesn’t include .@&#; 0-9 or a-z upper or lower

Return email adresses

***Write to output file:***

Get input from search and slice

Open output file

Write results to output file