Dark Web Crawler

Python Project

Team Members:

Sumukh Kesarla - 200031101611057
Umang Chaudhary - 200031101611013
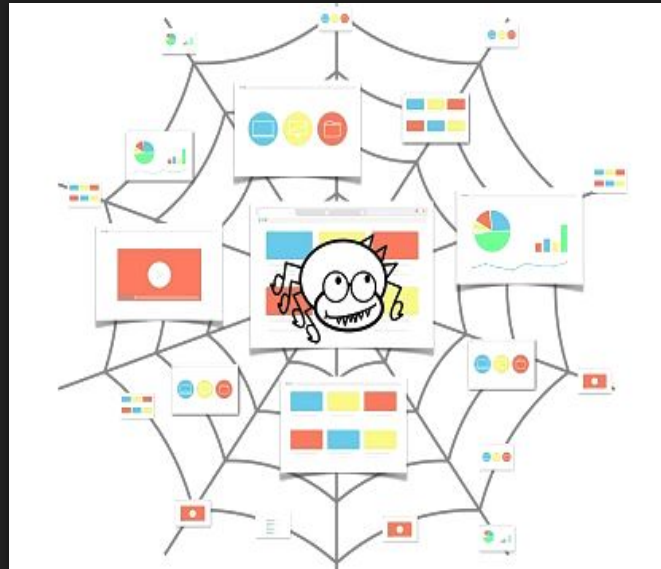Anshika Rajkumar - 200031101611007
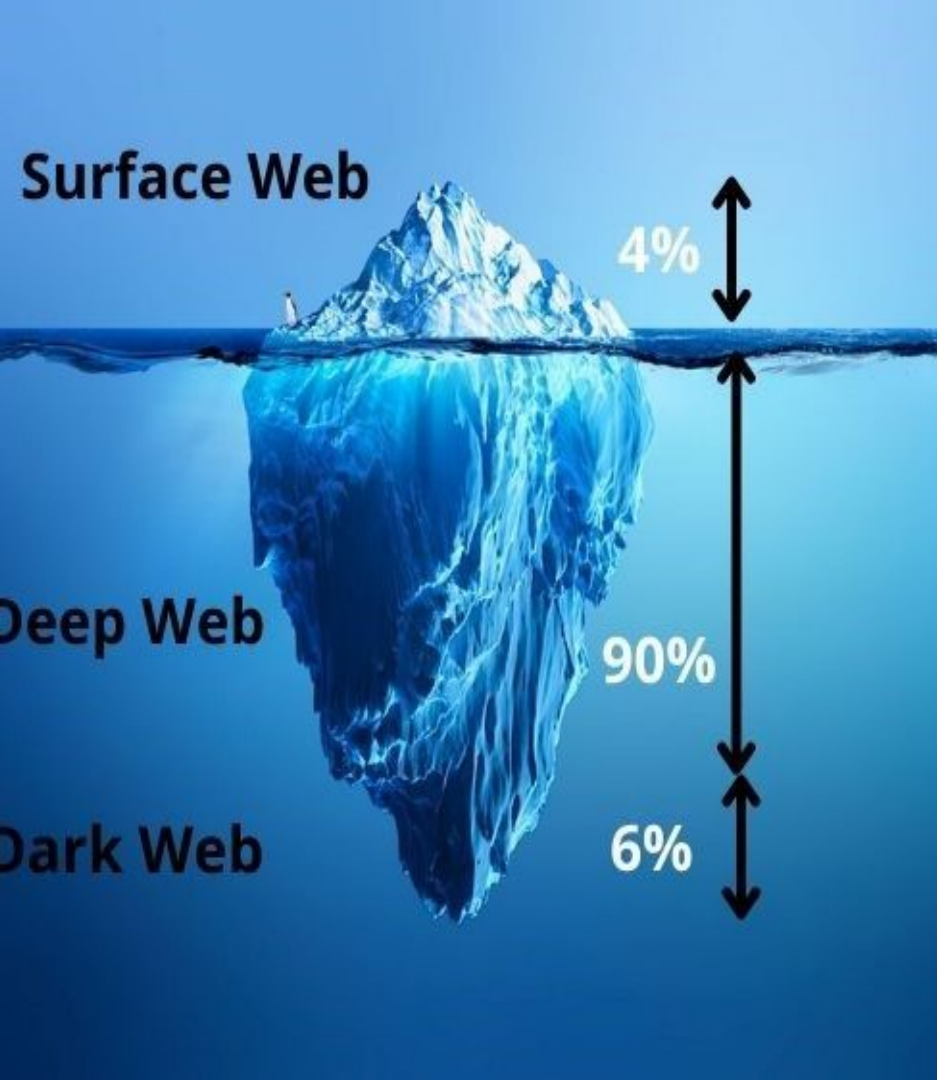Nishtha Shah - 200031101611048
Pooja Dinesh Fefar - 200031101611042

What is Web Crawler?

A web crawler, spider, or search engine bot downloads and indexes content from all over the Internet.

They're called **"web crawlers"** because crawling is the technical term for automatically accessing a website and obtaining data via a software program.

Dark Web ?

The dark web refers to sites that are not indexed and only accessible via specialized web browsers.

Significantly smaller than the tiny surface web, the dark web is considered a part of the deep web. Using our ocean and iceberg visual, the dark web would be the bottom tip of the submerged iceberg.

Direct Traffic through Tor Browser

1

```
Python310/python.exe "c:/Users/hp/OneDrive/Desktop/rru
SUCCESS: Connection Established via Tor Browser !!
```

Feeder Links for Beginning of Search

2

Request the Feeder Links to begin Searching via various Search Engines

3

```
Begining to Collect Links...
SUCCESS: 1084 Links are Collected
```

Trim the links to proper link format

4

```
SUCCESS: 1084 Links are Trimmed
```

Remove Duplicates

5

```
SUCCESS: Duplicates are Removed. 851 links are Collected
```

Print timely status messages on Console

6

```
SUCCESS: Duplicates are Removed. 851 links are Collected
```

Check if all the links are valid or not

7

```
http://tmszmggfizxri3kobr2qrnclstn4fmfqqgcr6cbevwu2rhu2pepajzid.onion/reviews is Valid. No of Links Scanned: 1
http://weedpysa7je6nhwy2iepqet4umtccigqvurp3syrfnjwp3wuwkdnread.onion/ is Invalid. No of Links Scanned: 2
http://idjhuib3xtycpj72sjojvvtgsubjalxyaqgvzcmxt4l7mufjsniz62yd.onion is Valid. No of Links Scanned: 3
http://bsfjn5eqoae2a5zjq4u7udtvmfhxv73jifz2q2hscsbvvbbmt7bl36yd.onion is Invalid. No of Links Scanned: 4
http://tortaxlw2hfwctl47r2vntznhnfk3c77avu4fei2shaj232aapuubtad.onion is Valid. No of Links Scanned: 5
http://2vrhkx57gvaz6h6fnevdmaihbnhdct3txeio74bdx4pf6pzxr6au2had.onion/reviews is Valid. No of Links Scanned: 6
http://deeeepv4bfndyatwkdzeciebqcwwlvgqa6mofdtsvwpon4elfut7lfqd.onion/Page/11/ is Valid. No of Links Scanned: 7
http://drugsxeuauw4xt6eoqwztfkpa2ul5453q4ao4kqrtedhfoy6jmui5ryd.onion is Invalid. No of Links Scanned: 8
http://aykz7vrbttbhvzcn3ine5lw6q3mgfrpvacjwfrzw7tgujuol6ilorqyd.onion/postview.php?id=42 is Valid. No of Links Scanned: 9
http://atvlvvrseabiphshnci2js7pme2xmebxqlolllvkt5dyrxqvawbuvlyd.onion/reviews.html is Valid. No of Links Scanned: 10
http://darkn5g756epq6t6fkqdo3mvkxlvwjtgabsbh3x47yhqg3uguiy5z5yd.onion is Valid. No of Links Scanned: 11
http://whonixv6cvcmddwnvrftgr5sarampxr7nxqgvl6mplgkc75c2leebnyd.onion/Free-client-support-vs-paid-client-support.html is In
valid. No of Links Scanned: 12
http://losf3msiygfmizgshddciot35lxxlf3huhme3ab3eme27e5kcicty4ad.onion/2021/03/05/lsd-drug/index.html is Invalid. No of Link
s Scanned: 13
http://aapd2gkc7azk5qlw3r4ugy3hm4lahc5k7p76ltkllzolwwt7ncc6oyad.onion/buy6.html is Invalid. No of Links Scanned: 14
http://itspecizgwhav2ijohcfajhghi2zi55azoldg5sdsi52xpcie52lriad.onion/Free-client-support-vs-paid-client-support.html is In
valid. No of Links Scanned: 15
http://abvl2y35kzgrhjowqup7til2z77x7vpingaerkusrv2dzkw6wtexmkqd.onion/reviews is Invalid. No of Links Scanned: 16
http://c5xoy22aadb2rqgw3jh2m2irmu563evukqqddu5zjandunaimzaye5id.onion is Valid. No of Links Scanned: 17
http://u2izrmre7yqzhv5gr3tbbsoyriwoxmc77rfieuz7xdjzjpvk37bhahqd.onion/reviews.htm is Invalid. No of Links Scanned: 18
http://digkc2nutat35igzdzqar57ieczp4zil7axwbsl5bn5c5gtbe2iaqeqd.onion/reviews is Invalid. No of Links Scanned: 19
http://oql5ppahfhgvasx4mfrw77ncgm65pipfbktrav7tulydvq4s3zmms5id.onion/ is Invalid. No of Links Scanned: 20
http://3okvzje5tk6elnzqqxgwabf7eh37a7eloafylvldtyyyoybjj7v5wjqd.onion is Valid. No of Links Scanned: 21
http://kfmrx2krsfijai5lrmt222euou3ckcfyzcnyuaw7rww33oqqgmwdhbqd.onion/weed.php is Invalid. No of Links Scanned: 22
http://darkbayx7a4sosoo4hqvoljqelgkusjlrqmt237ls6hndbplmel55oad.onion/offer/859 is Valid. No of Links Scanned: 23
http://btcw3q2t75d4tvpu7p5rmqoioqjabjdtpmwil2f5zvam6pk5swcd4syd.onion is Invalid. No of Links Scanned: 24
http://tlink2vs7rwo7pmnvzcfbjjxmt777ixnnah4k54rreeu6ntafmtvsrqd.onion is Valid. No of Links Scanned: 25
http://tw6gk6ze4bwoyzbtyeqhxnwx42ubxi2zbmgsyv7ducgldoaxdh3ejyad.onion/8 is Valid. No of Links Scanned: 26
```

Save results in a .CSV file

8

A1 | 0

	A
1	
2	http://tssa3yo5xfkcn4razcnmdhw5uxshx6zwzngwizpyf7phvea3gccrqbad.onion/public/stories/war_es_nur_ein_traum.html
3	http://7dfi4qizg5k4os44qwj3s3wnvq7cmogrcm2zpnk7rnefqyt2fodtkaid.onion/reviews
4	http://locjhy4ac643ol5yvapcubsoirbal7jbow63twovafszafclw5nk3eqd.onion/
5	http://b4qy5tfoimjaqalvkxiri6g3prb5tzx4u65ttfxd4gsvu3okctnek5qd.onion/product/heroin-lab-tested/?ivrating=4
6	http://mfjj65nurutalrivvb2s5zmow4c62qusp4sia4utxht3qffym2vypsqd.onion/product/HCaS8zh7v3SsMbfHIYqyZXJGOIoOdMOX
7	http://tssa3yo5xfkcn4razcnmdhw5uxshx6zwzngwizpyf7phvea3gccrqbad.onion/public/stories/jane_and_wisdomm_the_tiger.html
8	http://do6umu6gkwiwcyaucnxql6mht36hsk2bxdrunbfl5xrnkyrldpphj3id.onion/8
9	http://oxpgl32nws3znenlqcg5hrigmm7x7vognjv2wse4ynwaslaqf44jehyd.onion/product/buy-codeine-15mg-online/
10	http://deeptyspkdq3nfvqvyzbkgwhtok4qoyhypsyiuo24wux4jnb6e3nyiqd.onion/index.php?qa=274&qa_1=gold-200mg-rick-250-mg-mdma-mdcrystal-lsd-chapeleiro-bike-100&show=278
11	http://jvn7bywnbnu3do3hp7rvqrubsoakudng2a4x3jdny4dhfm5t5yfxuvad.onion
12	http://2qhfojw2badyzdihcrpmh7xmj6enxo3paatxtad2fqtg342vi2g34oqd.onion/bank-logins/index.html
13	http://orc52yt3qnatcmwwbiht4n6qvbxdeapzzvp6xaiwnuta2c2lqn4afdad.onion
14	http://5z57imdllhxv4lssagb2b3vrmbju3rt6kvjl3raakgtdzrxfh4obxdad.onion/index.php?threads/hey_15/post-26=
15	http://vlywdn4jpie6ytv4kc5qlwhpzpygidf5fcdynsn6cu7cgiooz74d7yqd.onion/2022-cardable-sites-list-updated-using-live-cc-non-vbv-cc/index.html
16	http://tmknbkslnscj76qryv2rdfo76txx5ykgcunrec7ukghrwox6q73lurad.onion/
17	http://zlutlerepfdjgtfl6qnblvu4mrwzmlrfhlfdipzk4pltfyblizefdaad.onion/product/120.html
18	http://tmszmggfizxri3kobr2qrnclstn4fmfqqgcr6cbevwu2rhu2pepajzid.onion/reviews
19	http://7au74367cn6peqgxh3t2ldiioaa3boc6kc3af5f7rnbwzpc72ajtrwad.onion/index.html
20	http://juhanurmihxlp77nkq76byazcldy2hlmovfu2epvl5ankdibsot4csyd.onion/search/?q=weed&d=1
21	http://mx6z7r45benxczwl5odr3w64so5oflwzx6kkmg5d3yivpgvhkad4hvyd.onion/category/9.html
22	http://jr2wuylrnuurvcshc5b6242sbbiik6pcn3xlsktkk5jiwnukqmux7qid.onion

View final result

9

```
http://tmszmggfizxri3kobr2qrnclstn4fmfqqgcr6cbevwu2rhu2pepajzid.onion/reviews is Valid. No of Links Scanned: 1
http://weedpysa7je6nhwy2iepqet4umtccigqvurp3syrfnjwp3wuwkdnread.onion/ is Invalid. No of Links Scanned: 2
http://idjhuib3xtycpj72sjojvvtgsubjalxyaqgvzcmxt4l7mufjsniz62yd.onion is Valid. No of Links Scanned: 3
http://bsfjn5eqoae2a5zjq4u7udtvmfhxv73jifz2q2hscsbvvbbmt7bl36yd.onion is Invalid. No of Links Scanned: 4
http://tortaxlw2hfwctl47r2vntznhnfk3c77avu4fei2shaj232aapuubtad.onion is Valid. No of Links Scanned: 5
http://2vrhkx57gvaz6h6fnevdmaihbnhdct3txeio74bdx4pf6pzxr6au2had.onion/reviews is Valid. No of Links Scanned: 6
http://deeeepv4bfndyatwkdzeciebqcwwlvgqa6mofdtsvwpon4elfut7lfqd.onion/Page/11/ is Valid. No of Links Scanned: 7
http://drugsxeuauw4xt6eoqwztfkpa2ul5453q4ao4kqrtedhfoy6jmui5ryd.onion is Invalid. No of Links Scanned: 8
http://aykz7vrbttbhvzcn3ine5lw6q3mgfrpvacjwfrzw7tgujuol6ilorqyd.onion/postview.php?id=42 is Valid. No of Links Scanned: 9
http://atvlvvrseabiphshnci2js7pme2xmebxqlolllvkt5dyrxqvawbuvlyd.onion/reviews.html is Valid. No of Links Scanned: 10
http://darkn5g756epq6t6fkqdo3mvkxlvwjtgabsbh3x47yhqg3uguiy5z5yd.onion is Valid. No of Links Scanned: 11
http://whonixv6cvcmddwnvrftgr5sarampxr7nxqgvl6mplgkc75c2leebnyd.onion/Free-client-support-vs-paid-client-support.html is Invalid. No of Links Scanned: 12
http://losf3msiygfmizgshddciot35lxxlf3huhme3ab3eme27e5kcicty4ad.onion/2021/03/05/lsd-drug/index.html is Invalid. No of Links Scanned: 13
http://aapd2gkc7azk5qlw3r4ugy3hm4lahc5k7p76ltkllzolwwt7ncc6oyad.onion/buy6.html is Invalid. No of Links Scanned: 14
http://itspecizgwhav2ijohcfajhghi2zi55azoldg5sdsi52xpcie52lriad.onion/Free-client-support-vs-paid-client-support.html is Invalid. No of Links Scanned: 15
http://abvl2y35kzgrhjowqup7til2z77x7vpingaerkusrv2dzkw6wtexmkqd.onion/reviews is Invalid. No of Links Scanned: 16
http://c5xoy22aadb2rqgw3jh2m2irmu563evukqqddu5zjandunaimzaye5id.onion is Valid. No of Links Scanned: 17
http://u2izrmre7yqzhv5gr3tbbsoyriwoxmc77rfieuz7xdjzjpvk37bhahqd.onion/reviews.htm is Invalid. No of Links Scanned: 18
http://digkc2nutat35igzdzqar57ieczp4zil7axwbsl5bn5c5gtbe2iaqeqd.onion/reviews is Invalid. No of Links Scanned: 19
http://oql5ppahfhgvasx4mfrw77ncgm65pipfbktrav7tulydvq4s3zmms5id.onion/ is Invalid. No of Links Scanned: 20
http://3okvzje5tk6elnzqqxgwabf7eh37a7eloafylvldtyyyoybjj7v5wjqd.onion is Valid. No of Links Scanned: 21
http://kfmrx2krsfijai5lrmt222euou3ckcfyzcnyuaw7rww33oqqgmwdhbqd.onion/weed.php is Invalid. No of Links Scanned: 22
http://darkbayx7a4sosoo4hqvoljqelgkusjlrqmt237ls6hndbplmel55oad.onion/offer/859 is Valid. No of Links Scanned: 23
http://btcw3q2t75d4tvpu7p5rmqoioqjabjdtpmwil2f5zvam6pk5swcd4syd.onion is Invalid. No of Links Scanned: 24
http://tlink2vs7rwo7pmnvzcfbjjxmt777ixnnah4k54rreeu6ntafmtvsrqd.onion is Valid. No of Links Scanned: 25
http://tw6gk6ze4bwoyzbtyeqhxnwx42ubxi2zbmgsyv7ducgldoaxdh3ejyad.onion/8 is Valid. No of Links Scanned: 26
```

THANK YOU