SURFACING HIDDEN VALUES THROUGH DEEP WEB

*Spurthy S, Asst. Professor
SJR College for Women, Rajajinagar, Bangalore-10.

ABSTRACT

Deep web search engine is to explore the hidden internet. There is a vast section of the internet which is hidden and not accessible through regular search engines and web browsers. This part of Internet is known as the Deep Web, and it is about 550 times the size of the Web we know. Deep Web is referred to the data which are not indexed by any standard search engine such as Google or Yahoo.

The 'Deep Web' refers to all web pages that search engines cannot find, such as user databases, registration-required web forums, webmail pages, and pages behind paywalls.

1.   INTRODUCTION

The deep Web is qualitatively different from the surface Web. Deep Web sources store their content in searchable databases that only produce results dynamically in response to a direct request [1]. But a direct query is a "one at a time" laborious way to search. BrightPlanet's search technology automates the process of making dozens of direct queries simultaneously using multiple-thread technology and thus is the only search technology, so far, that is capable of identifying, retrieving, qualifying, classifying, and organizing both "deep" and "surface" content[2].

Public information on the deep Web is currently 400 to 550 times larger than the commonly defined World Wide Web. The deep Web contains 7,500 terabytes of information compared to nineteen terabytes of information in the surface Web.The deep Web contains nearly 550 billion individual documents compared to the one billion of the surface Web.More than 200,000 deep Web sites presently exist.Sixty of the largest deep-Web sites collectively contain about 750 terabytes of information — sufficient by themselves to exceed the size of the surface Web forty times.

On average, deep Web sites receive fifty per cent greater monthly traffic than surface sites and are more highly linked to than surface sites; however, the typical (median) deep Web site is not well known to the Internet-searching public.The deep Web is the largest growing category of new information on the Internet.Deep Web sites tend to be narrower, with deeper content, than conventional surface sites.Total quality content of the deep Web is 1,000 to 2,000 times greater than that of the surface Web.

Deep Web content is highly relevant to every information need, market, and domain.More
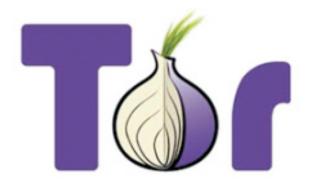
than half of the deep Web content resides in topic-specific databases. A full ninety-five per cent of the deep Web is publicly accessible information — not subject to fees or subscriptions. To put these findings in perspective, a study at the NEC Research Institute [3].

3. TO ACCESS DEEP WEB

Hidden below the virtual waterline lies a tangled and secretive network known as the Deep Web. Unindexed by search engines, and accessible only with special browsers such as The Onion Router (Tor), the Deep Web is made up of peer-to-peer connections, which allow users to share files directly (and secretly).[4]

 Technically, this is not a difficult process. You simply need to install and use Tor. Go to www.torproject.org and download the Tor Browser Bundle, which contains all the required tools. Run the downloaded file, choose an extraction location, then open the folder and click Start Tor Browser.

The core principle of Tor, "onion routing", was developed in the mid-1990s by United States Naval Research Laboratory employees, mathematician Paul Syverson and computer scientists Michael G. Reed and David Goldschlag, with the purpose of protecting U.S. intelligence communications online. [5]

4. DEEP WEB GROWING FASTER THAN SURFACE WEB:

Comparative Deep and Surface Web Site Growth Rates

Use of site domain registration as a proxy for growth has a number of limitations. First, sites are frequently registered well in advance of going "live." Second, the domain registration is at the root or domain level (e.g., www.mainsite.com). The search function and page — whether for surface or deep sites — often is introduced after the site is initially unveiled and may itself reside on a subsidiary form not discoverable by the who is analysis.

The best way to test for actual growth is a time series analysis. BrightPlanet plans to institute such tracking mechanisms to obtain better growth estimates in the future.[6]

However, this limited test does suggest faster growth for the deep Web. Both median and average deep Web sites are four or five months "younger" than surface Web sites (Mar. 95 v. Aug. 95). This is not surprising. The Internet has become the preferred medium for public dissemination of records and information, and more and more information disseminators (such as government agencies and major research projects) that have enough content to qualify as deep Web are moving their information online. Moreover, the technology for delivering deep Web sites has been around for a shorter period of time. [7]

6. ADVANTAGES

A Deep Web search engine's chief advantage is the depth and thoroughness of its results. Standard Internet searches through sites like Bing, Google and Yahoo barely scrape the

surface of the Web. Deep Web search engines such as Infornine, TechXtra and Yippy issue direct queries to access otherwise hidden content, covering far more ground and retrieving results from a much wider data pool. The more content searched, the more likely you are to find what you need. [8]

5.   DISADVANTAGES

Since there's more content to analyze, Deep Web search engines tend to be slower than standard search engines. Searching the Deep Web also requires a more precise search string. Deep Web searches should be reserved for serious, painstaking research, not for simple questions and basic Web surfing. Deep Web searches may also return sensitive personal information from normally restricted databases, creating ethical dilemmas and leaving individuals susceptible to fraud and identity theft. [9]

Exploring the Deep Web can be dangerous. Always think twice before clicking any returned search links. Because no ruling authority regulates the Deep Web, users enjoy complete anonymity and may partake in illegal activities such as trading copyrighted materials, selling narcotics, circulating child pornography and negotiating for other criminal acts. Exploring a search engine's results may lead you into the Internet's darkest corners, leaving you vulnerable to nefarious characters and sites designed to spread malware or hack your computer. [10]

1.   **Credit Cards:** There are people willing to steal credit card information for the right amount.

2.   **Hacking:** There is a whole bunch of sites where you can find people to hack anything for you, be it is your ex's account or a top secret government website. These guys are up for anything.

3.   **Thieves Burglars:** All sorts of illegal betting and match fixing goes down here. Gone are the days when a shady man with piece of paper was the only way you could bet money, here with click of a button you can do as you please.

4.   **Weapons:** Okay, before you start telling me that you could get a gun at the corner store, this one is not for 'Murica. For the people in Europe, if you want to get a gun this is probably where you'll get one.

5.   **Hitmen:** Getting someone killed is as you expect not cheap but from the looks of it, they'll get the job done.

6.   **Human Experimentation:** There are people doing experiments of all kinds on real live human subjects, which are mostly homeless people picked off streets. I came across this site (not for the weak hearted) where this guy used to pick up orphans and other people off streets and sold them as sex dolls. Here's where it starts to get disturbing. He cuts off their arms and legs, closes all openings that is the eyes, mouth, ears, you get the idea. It's not over yet he also offers the options of having hooks installed so that you can hang your sex doll anywhere.

7. Government Secrets: The deep web offers every conspiracy theory that any one has imagined. Also there are many government secrets out there. Recently some were taken down by the FBI.

8. Drugs: Perhaps the most famous thing about the deep web is the fact that you can get every drug ever made. Ranging from premium quality marijuana to pills and acid. You might have even heard of 'The Silkroad' which was the most credible source for some fine bud or anything else before it got taken down very recently.

6.   CONCLUSION

Serious information seekers can no longer avoid the importance or quality of deep Web information. But deep Web information is only a component of total information available. Searching must evolve to encompass the complete Web.

Directed query technology is the only means to integrate deep and surface Web information. The information retrieval answer has to involve both "mega" searching of appropriate deep Web sites and "meta" searching of surface Web search engines to overcome their coverage problem. Client-side tools are not universally acceptable because of the need to download the tool and issue effective queries to it. [11]

REFERENCES

[1]   The Hacker News

[2]   Bright Planet: Deep Web Harvest Engines vs. Search Engines – Finding Intel in a Growing Internet. www.brightplanet.com

[3]   http://www.completeplanet.com/ Tutorials/DeepWeb/index.asp

[4]   Tor : Hidden Services and Deanonymisation

[5]   Tor Flow, a dynamic visualization of data flowing over the Tor network

[6]   P. Lyman and H.R. Varian, "How Much Information," published by the UC Berkeley School of Information Management and Systems, October 18. 2000. See http://www.sims.berkeley.edu/research/projects/how-much-info/index.html. The comparisons here are limited to archivable and retrievable public information, exclusive of entertainment and communications content such as chat or e-mail.

[7].   As this analysis has shown, in numerical terms the deep Web already dominates. However, from a general user perspective, it is unknown.

[8].   Security Affairs: The good and the bad of the Deep Web . http://securityaffairs.co

[9]   World Crunch: Welcome to the Deep Web: The Internet's Dark and Scary Underbelly. www.worldcrunch.com

[10]   The Next Web: Mail-order drugs, hitmen & child porn: A journey into the dark corners of the deep web. hhtp://thenextweb.com

[11]   BrightPlanet's search tutorial at http://www.completeplanet.com/searchresources/tutorial.htm.

[12]   Smart IT emerging analysis on movement patterns of moving objects, International Conference on Development of Smart Cities: Interface, Governance and Technology at Dr.AIT, Bangalore, 2016/9.

[13]   'Marketing Strategies for perishable fruits and vegetables retail segment in supermarkets' published in AJMR-Refereed International Journal of management, March-2010