University of Wisconsin – Milwaukee

Shedding Light on the Dark Side of the Internet

The Dangers and Mysteries of the Deep Web in Web 2.0

Omie R. Walls

INFOST 110

Introduction to Information Science

Professor Michael Zimmer

May 14, 2014

1 Introduction

For many of us, Google is likely the hub of our Internet lives and homepage of our browsers. A simple search can yield 5.27 billion results in about a third of a second, though most of us never discover what's beyond the second page of search results. It is also likely, most of us never discover what's beyond the search engine itself. That is what we call the *deep web*.

1.1 Description

Also known as the "dark web" or "hidden web", the deep web is any website that is not reported by a search engine. Search engines use algorithms to create "spiders" that crawl through web pages in search for new links. The search engine in return creates a cache of the page for their own inventory of the web. It is believed that search engines encompass about ten percent (some claim upwards of forty percent) of the Internet while the deep web contains, well, everything else (forbes).

1.2 Content

The size of the deep web is built upon a lot of speculation, because we *all* contribute to it. Those cute pictures of Aunt Mae's kittens in your private photo album labeled '*Summer to Remember'* are part of the deep web, along with your bank account information, database-driven websites and anything else that requires a password to access. However like many, what typically comes to my mind and many others when the deep web is mentioned, is the subsection of anonymous websites dedicated to the nefarious operations of drug lords, murderers, and human traffickers. That is also there as well.

2 Human Factors

What use is there for the hidden web? Is everything you'll ever need located in your search engine? The deep web contains a digital divide all its own between the "savviest" of Internet users on the deep web, and those who are too far less engaged in the Web to ever cultivate deep web interests. Could we be missing out on something important when it comes to the invisible web? A few factors come into play as to why we might think we aren't.

2.1 Access

The way the deep web has to be accessed, is the main reason there is a disproportion between regular web users and deep web users. Users must use a dedicated software to access the deep web. The most popular of these is Tor, while also notable are the alternatives I2P and Freenet. Tor is an acronym for "The onion router", a system implemented to guarantee online anonymity as part of a project backed by the US Naval Research Laboratory from 2004 to 2005. This was then successively supported by the Electronic Frontier Foundation (source 2). When asking the average person if they would trust downloading a specific software to access the deep web, they would likely feel less inclined to pursue it.

2.1.1 Features of Tor

Many of the websites in the deep web don't end in ".com", but instead, ".onion," and usually the site name is usually several characters long and something bizarre that the average person would not remember. Bookmarking would be the most useful in those situations.

An issue with the deep web is that when looking for information about the deep web in a regular web search engine, information provided about it is either very redundant or primitive or only tailored to professional and academic researchers. And occasionally, you will find a list of sites comprised by someone who doesn't know much about the deep web. To anyone given a list of sites in an open forum along with the warning, "This might contain CP [child pornography], I'm not sure," their curiosity about the deep web just might dwindle into nothingness. I know mine almost did. There's a guttural fear that engulfs you as you tread lightly and pray you never incidentally trip into the dreaded dark areas of 'CP.'

2.1.2 Assessing Scope of the Deep Web

For those with access to the deep web, benefits are plenty. As previously mentioned, search engines to not acquire everything the Internet has to offer, and is ineffective in cataloguing several categories of websites:

- *dynamic content* which can only be accessed through submission of information in a text field;
- *unlinked content* that is not attached to any page;
- *private websites* that are password restricted;
- *contextual websites* which are often recognized as *Web 3.0*, where information is determined or personalized by our complex personal psychology, location and social groups (source 3);
- *limited access content* which is limited in a technical way (i.e. CAPTCHAs) or content that prevents browsers from creating cached copies;
- *scripted content* that is only accessible through links provided by dynamic client-side scripting language;
- and content that is listed outside of HTTP or HTTPS.

The majority of this kind of content can however be accessed in the deep web and not just through the desktop PC. People prefer internet mobility more every day. "Much of their connectivity is at the expense of personal computers (PCs) as more people incorporate wireless appliances into their daily work (Sierkowski)." One great feature of the deep web is that Tor is actually accessible on virtually any device whether it is a tablet, smartphone, or iPad.

2.1.3 Search Engines for Universal Appeal

"The Web presented unparalleled opportunities for the creation and publication of materials, but as yet there was no corresponding system to filter and sort these materials (caulfield)." That was the case for the surface web until Google arrived, and now this is the case for the deep web. The issue seems to be that the deep web doesn't sweep up more users because of the lack of commonalities like search engines. Over the last 5 years, this issue has been rapidly changing. Just recently, developers released the first "Googlized" search engine for better

access of the deep web. Its layout virtually mocks that of Google, even down to the "I'm feeling lucky" button that I never press (I don't gamble with search engines).

2.2 Usefulness

Lastly, many people do not know what is useful to them on the deep web. In a time where privacy is a major concern and total anonymity is virtually unheard of nowadays, the deep web flourishes. Pierluigi Paganini writes in "The good and the bad of the Deep Web" that deep web is a tool to bypass hierarchical regimes:

"But consider also that the Deep Web is the privileged channel used by governments to exchange documents secretly, for journalists to bypass censorship of several states and also dissidents to avoid the control of authoritarian regimes (source 2)…"

Censorship through the Tor network is attained by the program encrypting data as it passes through nodes within the network.

"Tor client software routes Internet traffic through a worldwide volunteer network of servers hiding user's information eluding any activities of monitoring (source 2)."

The deep web has massive amounts of unregulated information that is uncontrolled. If there is ever such a place where information can exist in freedom and move about without restriction, it is on the deep web.

People can express their true feelings without punishment, spend untraceable[i] money, and buy anything their heart desires. Nevertheless, not everyone operates as nobly within the free realm of the deep web. The freedom of the deep web does come with a real price.

3 Legal & Ethical Issues

Though the Internet was developed assuming all intents and purposes were honorable, the dark side of the web has emerged with serious drawbacks and showed the world real problems like; spam, malware, hacking, phishing, denial of service attacks, click fraud, invasion of privacy, defamation, frauds, violation of digital property rights, etc. It seems that the dark web is the epitome of Faustian bargains seeing that it brings human trafficking and child pornography right to our doorsteps and puts a mask on the world's most dangerous criminals. The responses to the dark side of the Internet have been overall effective in keeping the crime that exists contained to an expectable level (source 4).

3.1 Piracy, Illegal Solicitation & Insider Trading

Piracy and disregard for digital rights is a major issue on the deep web. In what appears to be the digital underground, you can find just about anything under the shadow.

"The Tor browser gives employees a marketplace to exchange insider information on their firms for Bitcoin or insider information on other public firms. Individuals can trade with such information, profiting illegally without a paper trail (source 5)."

With that said, many see insider trading as fraud, and that is a matter of perspective. Fraud is the act of deception for the purpose of a gain, and with insider trading, no one is being

deceived however there is a gain by acquirement of knowledge. Whether right or wrong in the eye of the trader, one thing is certain- it is illegal and perspective holds no bearings in defense of such acts.

Because of the anonymity, widow of opportunity in the hidden web is pretty wide open when it comes to gang activity, hiring assassins/murderers, and selling every type of drug and/or people in every size of quantity.

3.2 Psychological Effects

In Postman article "Informing Ourselves to Death," Postman raises the concern of the plausibility of having too much information and our inability to process large amounts of it, resulting in feelings of incapability to create solutions for current issues. In having too much information, it decreases our self-view of the world. We become smaller and see ourselves as unable to change the way things are. Many of the drawbacks that are pressed in the article are even truer about the dark web. The dark web, surface web, and the web itself is often viewed metaphorically. The surface web is only the tip of the iceberg, and the deep web would represent everything the iceberg is underneath the water (which is a massive chunk of ice that is vastly larger than the appearance the iceberg gives off above sea level). When uncovering the depths of the web, there may come even more a feeling of information overload than a search engine could ever provide (Postman).

4 Conclusion

The deep web can be a fascinating place, and a superbly sized domain of discovery. However, if not careful, exposure to the wrong things can ruin an experience. Does the deep web's good outweigh the bad? Or is there more to be said about piracy and evildoers. The perception of the deep web differs from person to person. It seems to move in its own direction and not in the direction the rest of the web is going. Web 3.0 appears to be a less-privatized and more personalized Internet experience, while the deep web runs off something different – pure, unrestricted capitalism, free speech, and power in anonymity.

[i] I personally do not feel Bitcoin is that untraceable. Technology is one of those things that have a way with catching up to trends and evolving with the demand. I also think it's foolish to trade money with nefarious strangers over the web, but I digress…