MEDIA SONAR

Untangling the Web: Where to Get Started with Online Investigation

By Danielle Gautier

WHITEPAPER

PURPOSE

Distinguish between the Surface Web, the Deep Web and the Dark Web.

Learn the ways the Dark Web can be used to create public and private risk.

Learn how online investigations can help security teams assess risk and investigate threats and attacks.

INTRODUCING

The Internet has been called the great democratizer and never before has it been so challenging to safeguard people and assets. The dual use of the Internet, both as a conduit from criminal activity, and as a tool for law enforcement and corporate security teams, is forcing the industry to dissolve the distinction between online and physical worlds.

Today, more often than not, the first news about a crisis will come from social media. Active shooter situations play out on Twitter. In a crisis, people share their status with friends and family on social media sites. On the other hand, when engaging in criminal and malicious activity, the Dark Web is an increasingly essential tool in communicating and transacting covertly. Where corporate crime is concerned, the Dark Web is especially useful. Businesses are prime targets, not only from external attacks, but also from insider threats. Obtaining the right technology skills and tools is needed in order to ensure ongoing protection for people, places and things from those who would use the Web to inflict damage and harm.

SEARCH ENGINES AND THE SURFACE WEB

To distinguish between the Surface, Deep and Dark Web, it is important to start with an understanding of how search engines and the search engine index work in relation to the Internet as a whole. Search engines build their database of websites, know as the index, through a process called crawling. Google, Yahoo and other search engines find new pages to index using “bots” or “spiders” to crawl the Web. The bots start with a list of URLs determined by previous crawls, and when bots detect new links on these pages, they are added to the list of sites to index moving forward. Then, search engine algorithms produce a ranked list of Web pages from their index based on search terms users provide. Search engines will obtain some of their listings when authors submit their own Web pages, but bots do the bulk of the work.

The goal of all search engines is to find and organize distributed data found on the Internet. Before search engines were developed, the Internet was a collection of File Transfer Protocol (FTP) sites. Users would navigate the FTP sites to find specific shared files. As the number of World Wide Web users grew, so did the need to easily find and organize the files. Search engines were developed to meet that need, and many eager players participated in the search engine wars that ensued.

Google, the current victor in those wars, is by far the most popular search engine, powering 77% of all Surface Web searches in the world. Google processes over 40,000 search queries every second on average, which translates to over 3.5 billion searches per day and 1.2 trillion searches per year worldwide.

These numbers are impressive at first glance, but there are limits to relying on search engines such as Google. The most important being that what is presented within search engine results does not constitute the entirety of the Internet, and that search engine results show a biased view of what is important.

the Google interface the user gets the impression that the search results imply a kind of totality. In fact, one only sees a small part of what one could see if one also integrates other research tools."

Eric Schmidt, while serving as Google’s Chief Executive Officer, told the Wall Street Journal: "I actually think most people don't want Google to answer their questions, they want Google to tell them what they should be doing next."

The importance of easily finding and organizing files on the Web is without question, but search engines, however useful for the average Web user, are inadequate for security investigators and journalists alike. More robust research and intelligence tools are required to access the vast amounts of files and data that lie beneath the surface.

UNDERSTANDING THE DARK WEB

WARNING – Always exercise caution when accessing the Dark Web. While the Dark Web was developed for positive applications and it is not inherently a bad place, there are many malicious Dark Web users looking to take advantage of you. Possibilities include phishing for credit card numbers, assuming your identity, stalking, blackmailing, obtaining illegal information and ransoms. Additionally, some content on the Dark Web can be disturbing.

Thred software allows Web investigators to find and organize content on the Deep Web and the Dark Web. It is a safe and anonymous alternative to accessing the Dark Web directly.

The Dark Web forms a small subset of the Deep Web. Originally created in the mid-1990s by the United States government as a means for intelligence operatives to communicate securely, the Dark Web is a chaotic place where everyone is anonymous.

In addition to being used by the intelligence community, the Dark Web is used by journalists to contact their sources, whistleblowers as well as by individuals living under authoritarian regimes where free speech, or Internet access, is outlawed or limited.

There are many useful and positive applications for darknets but, at the same time, many scams and illegal activities play out on the Dark Web. Researchers at King's College in London classified the contents of 2,723 live Dark Web sites over a five-week period in 2016 and found that 57 percent host illicit material.

HEADING 1

HEADING 2

HEADING 3

Heading 4

Heading 5

MEDIA SONAR

mediasonar.com

H1701