Open, Deep, and Dark: Differentiating the Parts of the Internet Used For Cybercrime

Thomas J. Holt, Professor, School of Criminal Justice, Michigan State University
holtt@msu.edu

ABSTRACT: The structure of the Internet allows individuals to engage in crime via the Open, Dark, and Deep Web. Open Web content can be accessed by anyone with a standard browser. The Deep Web requires either certain permissions, like passwords, or knowledge that the information exists, like that the website cannot be found via search engines. One needs special tools to access the Dark Web, where the users' activities are anonymized, making it more difficult to track their activities.

While the Internet is a resource enabling the exchange of data for communication between individuals in a rapid and decentralized fashion across the globe, there are distinctions in the ways that individuals can access information hosted online, based on the use of either common or specialized web browsing and encryption protocols. There are three ways to describe how information can be accessed over the Internet: the **Open Web**, the **Deep Web**, and the **Dark Web**. In some respects, it is similar to an iceberg, as there are parts that can be seen above the surface of the water, and a larger part hidden below.

The term **Open Web**, or **Surface Web**, is often used to refer to the body of content hosted on web servers that can be accessed through any sort of web browser (i.e. Microsoft Edge, Google Chrome). Additionally, this content can be indexed by search engines, like Google or Bing, as the site operators have not specified directions in the "robots.txt" for search engines not to display the webpage.

The term **Deep Web** is used to refer to content hosted and accessible through the Open or Surface Web, but it may not be accessible through search engines because of one of several reasons:

1) the content is proprietary, involves personally identifiable information (PII), or is regulated by law to restrict access (such as email accounts, tax records, payment systems, etc)
2) the information is password protected (as with a forum that requires users to register to observe content)
3) the content is behind a paywall (as with scientific journals and media content)
4) the site operators have disabled features that allow the url to be cached in search engine results

Lastly, there is the **Dark Web**, which is a portion of the Internet that can only be accessed via the use of specialized encryption software and browser protocols. Individuals can only access the Dark Web through the use of a service called **TOR**, which stands for the **The Onion Router.** TOR is a free proxy and encryption protocol that hides the IP address and location details of the user. TOR is a widely popular and relatively secure service that individuals download and install on their system. Once downloaded and activated, TOR encrypts an individual's web traffic and routes it through a network of other Tor users' systems that is randomized, making it difficult to locate the actual source of any user's computer. Additionally, websites and forums that are hosted on the Dark Web cannot be indexed by any search engines, nor can they be accessed through traditional web browsers. In fact, all sites hosted on the Dark Web end in the extension .onion, reflecting the way that Tor protocols operate.

Because of the security that Tor affords, a wide range of cybercriminals use this service to conceal their activities, including child pornography trading, drug markets, and sensitive information exchanges. In fact, this technology limits the ability of law enforcement agencies to eliminate illicit content because the hosting source cannot be identified through traditional means. At the same time, US laws have been amended to allow law enforcement to obtain warrants to use remote access if anonymized software has been used.