ACCESSING THE INVISIBLE WEB: ISSUES AND CONCERNS

Supported by tetfund TERTIARY EDUCATION TRUST FUND

Odey John A.* and Okoro Anthony T.
Department of Computer Science, University of Calabar, Cross-River State, Nigeria
*Corresponding author: johnodey@unical.edu.ng

Received: May 23, 2021 **Accepted:** July 14, 2021

Abstract: Timely and secured access to reliable information on the web is a key to the advancement of knowledge in all fields of endeavour. The World Wide Web has, since the emergence of the internet in the 1960s, served as the bedrock for real-time information access and retrieval. Interestingly, only about 4% of internet resources are available easily and accessed through conventional search engines. The remaining 96% are found in the inner layers of the web known as the deep web and dark web or collectively known as the invisible web. Efficient access to resources of the internet at the sub-surface levels requires the use of specialized anonymized tools such as The Onion Router (Tor) browser but this is not without its issues and other concerns. This paper examines issues and concerns inherent in the deep and dark web as well as the potential measures towards the safe and convenient usage of these invisible web domains.

Keywords: Invisible web, deep web, dark web, Tor network

Introduction

The World Wide Web (www) is an online information space that has grown over the years to provide resources and easy means of communication among users. It was created by Sir Tim Berners-Lee in 1989 and made public in 1991 as a platform for billions of internet users to interact and seamlessly share resources across the board (Crawley, 2019). The www is classified into the surface web, deep web, and dark web with the surface web being the most common area of the web that is publicly accessible by users through conventional search engines like Google, Yahoo, and Bing (Bedi *et al.*, 2020; HackerNoon, 2018). The ability to access a greater percentage of internet resources on the surface web depends on the techniques used by search engine crawlers to extract information on the World Wide Web (Bedi *et al.*, 2020).

The invisible web is not indexed by regular search engines. The contents are huge and sit beyond the surface web, which for various technical reasons are not indexed by search engines (Chertoff & Simon, 2015; Weimann, 2015). The dark web is a portion of the deep web that relies on darknets (a network that permits communication only with trusted peers) for security (Ciancaglini *et al.*, 2015). It is intentionally hidden and only accessible through specialized web browsers such as the Tor (The onion routers) browser, Tails OS, I2P (Invisible Internet Project), Freenet, and Subgraph Operating System (Subgraph OS) (Steve, 2016). Figs. 1 and 2 depict the structure and basic activities on the surface web, deep web, and dark web, respectively.

Research has shown that while the surface web provides users with information from web pages that are indexable by search engines, over 96% of the internet content is invisible on conventional search engines but is domiciled in the deep web. In other words, what most computer and internet users know and access is information indexed and displayed on conventional search engines, which constitutes a mere 4% of the internet's available resources. Incidentally, information on the deep web is stored in databases and cannot be indexed by these conventional search engines thus making it practically impossible for users to access this information (Digital.com, 2018). This is because search engines depend on the web's linkages to identify contents on the Web (Bergman, 2001).

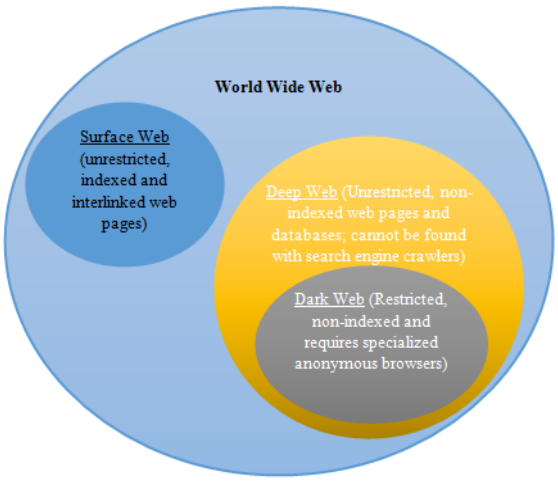

Fig. 1: Structure of the World Wide Web (adapted from (Ciancaglini *et al.*, 2015; Parker *et al.*, 2017)

Fig. 2: Parts of the web (Hacker Combat, 2017)

Polarization and unregulated transactions of resources on these internet domains tend to portray the deep web and specifically the dark web as a habitat for all web-related illegal activities such as drug deals, hacking, conspiracies, etc. (HackerNoon, 2018). It is also worthy of note that beyond the surface web, internet users see the dark web as a haven for free speech and users seeking extreme privacy especially among those working within strict government regulations like journalists, military, law enforcement agencies, human right activists, and political dissidents (Hadjimatheou, 2017;

Lightfoot, 2017). With the advantage of being unregulated and no compulsion to satisfy government or institutional standards, many use the dark web as a testbed for experimentation of innovations; more like a trial-and-error, which if successful such ideas are often transferred to the surface web (Thomaz *et al.*, 2020). This is to say, there exist a plethora of resources for the good and bad intentions of a user on the deep web. Thus, the use of the internet beyond the surface web raises security concerns for both users and the contents on the internet. Without proper security measures and consideration, internet surfing beyond the surface web can be dangerous and fatal. This paper is therefore aimed at addressing issues and concerns inherent in the deep and dark web.

Problem Definition

The invisible web is a web of anonymity where users' identities and locations are protected by encryption technologies routing user's data through many servers and peer to peer networks across the globe. Is it illegal to go on the invisible web? Simply put, No, it is not illegal to access the deep and dark web. In fact, some uses are perfectly legal and support the value of the "dark web" (Kaspersky 2021). On the invisible web, users can seek out clear benefits from its use:

- User privacy and anonymity considerations
- Services and web locations cannot be traced
- A platform to mitigate illegal activities by law enforcement agencies.

While privacy considerations and anonymity expectations are paramount benefits and justifications for the deep web, protecting whistle-blowers, privacy advocates, and political dissidents should not come at the expense of empowering illegal activities injurious to the norms of the societies like arms trafficking, dangerous literatures, child abusers, human trafficking, and sales of illicit drug. Therein lies the challenge: to devise approaches that walk the fine line of protecting liberal principles in an age of information control while identifying and eradicating the most insidious activities on the dark web. When viewed through this lens, the dark web's legality is based on how you as a user engage with it (Kaspersky 2021). Hence the wins in addressing the issues and concerns of the invisible web will accelerate positive innovations, freedom of information dissemination, and other inherent salient benefits.

According to Michael Bergman (Bergman, 2001), who is attributed as the originator of the terminology 'deep web', searching the internet is akin to spreading a fishnet across the surface of the ocean. There are chances that while many fishes may be caught, a greater swarm will be missed. Similarly, a greater percentage of internet resources are buried far down on the deep web. The superfluity of these resources has attracted several users to the domain. In this section, an in-depth review of related literature on the trends and topical issues including activities within each domain and security measures is discussed. To begin with, is a summary of the characteristics of the three web domains as presented by Sheils (2020) in Table 1.

Table 1: Comparison of the surface web, deep web, and the dark web (Sheils, 2020)

	The surface web	The deep web	The dark web
How to Access	Traditional search engine.	Requires password, encryption, or specialty software like Tor.	Requires Tor Project or similar to the view
Includes	All indexed web pages	All unindexed webpages	The subset of unindexed webpages inside the deep web.
Size	Approximately 4.47 billion pages	Massive, likely 4-5X was larger than the Surface web.	A subset of the Deep Web, but unmeasurable in size.
Uses	Email, social media, video, legitimate business websites, etc.	Usually used for legit purposes that requires anonymity.	Sometimes used for illegal activities.
Who uses it?	Anyone with an internet connection.	Whistleblowers, journalists, etc.	Hackers, sellers, and buyers of illegal merchandise.
Can be browsed anonymously?	No, nearly all activity can be seen by your Internet Service Provider (ISP).	Usually, especially if you use a Virtual Private Network (VPN) to access.	With precautions, yes.

There have been extensive researches in the past years on these domains. The major focus has been to understand the various operations that take place on the invisible web and development of strategies to curb illegal activities that are found.

In 2012, The Federal Bureau of Investigation (FBI) in the United States of America carried out an investigation tagged "Operation Torpedo" using network investigative technique (NIT) to unveil the Internet Protocol (IP) addresses of at least twenty-five individuals who visited child pornography websites on the dark web. NIT was delivered to computers that accessed the illegal sites, allowing for delayed notification to the targets for thirty days and within two weeks, the FBI was able to collect the IP addresses of visitors to the malicious sites (Vogt, 2017). Interest in the invisible web grew more in October 2013 with the arrest of Ross William Ulbricht and subsequent shutting of his marketplace called Silk Road by the FBI (Nabki *et al.*, 2017). Silk Road, launched in February 2011, was an online marketplace for the trading of all sorts of contrabands, particularly illegal drugs using Tor Browser for anonymity and Bitcoins, a cryptographic currency for the method of payment (Bojarski, 2015; Sui *et al.*, 2015).

Similarly, in a raid called *Operation Onymous* in November 2014, law enforcement agents successfully brought down Silk Road 2.0, which sprung up immediately after the shutdown of Silk Road, in addition to 27 other Dark Net drug market sites (Gingerich, 2014). In 2014, the U.S. Defense Advanced Research Projects Agency (DARPA) launched a search engine known as *memex* to create a search index that would assist in fighting potential human trafficking and illegal operations carried out on the dark web. The application works by scraping and indexing almost the entire internet resources that are ignored by traditional search engines. It analyzes these resources, identifying their pattern and relationships but yet maintains anonymity through the concealment of the IP addresses of suspicious connections (Vogt, 2017). Cyber and national security experts have continued to develop frameworks including data mining tools that will aid in the identification and monitoring of suspicious activities on the dark web. One such investment is the development of an Automated Tool for Onion Labeling (ATOL) system that crawls, analyzes, and labels content thematically in the public Tor Hidden Service (HS) ecosystem, then indexes the content from these onion sites into a large-scale data repository, called LIGHTS that houses more than 100M pages. The system analyzes onion sites and labels them thematically to identify malicious sites (Ghosh *et al.*, 2017).

Apart from tracking and shutting of many illegal activities as profiled above, studies have been carried out to understand the technological innovations inherent in the invisible web. A study by the Trend Micro Forward-Looking Threat Research Team using a robust deep web analyzer (DeWA) was carried out to aid investigations into various cybercrime activities on the dark web. For over two years, DeWA tracked malicious online activities and explored new threats on the deep web. Analysis of the data extracted by software showed that most of the activities involved benign content, alongside illicit activities which include illicit drug deals, child exploitation, bitcoin

laundering services, assassination service advertisement, amongst others. The authors posit that taking down marketplaces on the deep web seems not to be a lasting solution in preventing illegal deals on the dark web, as there continue to be a proliferation of online shops and fora where these illicit deals still take place rather security defenders need to keep monitoring these activities and curtailing it, especially as the role of the deep web on the internet grows. The paper concludes that anonymity in the Deep Web will continue to be a major issue of discourse and of interest for both law enforcers and Internet users who want to circumvent government surveillance and intervention (Ciancaglini *et al.*, 2015).

Technologically, the invisible web has continued to experience advancements in terms of robustness, anonymity, and size. There is now increasingly availability of crimeware, advanced secure/anonymous web hosting services, and cryptocurrency/dark wallet, not forgetting the contributions of ubiquitous computing, distributed/cloud computing, mobile computing, and sensor networks (Sui *et al.*, 2015). It has created an enabling platform for applications such as the directory of the U.S. Library of Congress (www.loc.gov), FreeLunch.com, Census.gov, Copyright.gov, PubMed, Web of Science, WWW Virtual Library, Directory of Open Access Journals, FindLaw, and Wolfram Alpha, Westlaw and LexisNexis, Twitter, Facebook, and Instant Messaging (IM) applications. These databases are not indexed in conventional search engines. Access to them is usually through some sort of registration and authorization checks, subscriptions, Application Programming Interface (APIs), and PayPal services (Sui *et al*., 2015).

Issues and Constraints

Undoubtedly, the deep web though has numerous advantages, also has issues that pose great constraints and concerns for users. On the positive side, it houses a lot of entire internet resources that are far hidden from the surface web and can only be accessed by specialized anonymous cyber-tools; guaranteeing the privacy of users. These cyber-tools shield users from online surveillance hide their identity and do not permit the trailing of the user's online activity (Bellaby, 2018). In the ensuing discussions, issues, and constraints to safe and easy utilization of the deep web technology are highlighted. Issues around speed, cybersecurity, deep web coverage, ethics, and complexity of the access to internet resources beyond the surface web are also presented.

Anonymity on the invisible web: Anonymous surfing of the internet can be of good or bad intentions as there exist diverse resources on the web, so are numerous internet users with varying search/transaction interests. A typical example is human rights activists and journalists who use these anonymous networks to track illegal dealings. Similarly, whistleblowers and groups in hostile political and war zones also leverage on the solid layer of encryption and anonymity offered by the deep web technology to traverse the internet and report the true state of events (uncensored) without being tracked by pieces of machinery opposed to such exposures.

On the other hand, fraudsters can also capitalize on the anonymity of the technology to facilitate cybercrime (Crawley, 2019; Holland, 2020). *Silk Road*, for instance, was an illicit version of eBay and was used by internet fraudsters to perpetrate all forms of illegal transactions on the dark web using Bitcoins and virtual currencies that are hidden within the anonymous networks. These crimes include sales of illegal drugs, weapons, hiring of assassins, trafficking, and child pornography, amongst others (Laverty, 2020).

These internet activities require privacy and protection of users and web resources; thus are essential in guaranteeing confidence and continuous reliance on any web resource. This is regardless of the part of the web that is accessed. On the surface web, protocols such as the HyperText Transfer Protocol Secure (HTTPS), Secure File Transfer Protocol (SFTP), Secures Shell (SSH), Internet Protocol Security (IPsec), Secure Sockets Layer (SSL) are known secure internet protocols that provide these security features against fraud and data theft (Weedmark, 2020). However, anonymous surfing of the internet beyond the surface web, specifically on the dark web, requires the deployment of highly encrypted anonymizing proxy networks such as Tor or I2P to prevent the users from being tracked.

In addition to the use of these anonymized tools, further safety measures are encouraged considering that all security and privacy of the information of internet users are paramount and should not be handled with levity. The reason is that some of the activities on the deep web like checking emails and e-banking transactions among others are legitimate interactions that most often result in the exposure of user access codes like passwords, no matter how strong, unique, and/or hard-to-guess the access codes might be. To safeguard these confidential data, it is precautionary not to divulge key information or grant requests based on trust. It is also recommended to cover the webcam of the user's computer before accessing the deep web to prevent spying eyes from having visuals of your face and environs. More so, turning off the plugins and scripts of the chosen browser is a step towards hiding the IP address of the user. These smart practices can help a user stay safe while surfing the internet.

Complexity issues of the invisible web: Web contents on the deep web are not indexed on conventional search engines. Though some of these contents are rendered through HTTP(s) protocols, they are invisible on conventional browsers. Most of the web contents on the deep web are innocuous and not illegal content as often perceived. It also comprises contents that are too deprecated and obscure which at times requires the use of web archives such as Wayback Machine (http://web.archive.org/) to access these contents on the regular web browsers (Crawley, 2019). Generally, the deep web deals with big data that are uncatalogued; lacking structure.

The nature of website design also contributes to how complex the website may be seen. Website owners might intentionally design the site with a limitation for security and privacy reasons or it could also be as a result of the constraints built into the site. For instance, contents on dynamic web pages, blocked sites, websites requiring user long-in authentication, or CAPTCHA constraints might be difficult to be indexed since the crawlers are denied access until the authentication requirements are supplied (HackerNoon, 2018). It could also be that the webpage is designed to be accessed only for a specified number of times, of which it becomes unavailable to WebCrawler if access to it is made at the expiration of the allowed attempts. A stringent measure could also be set on the website's robots.txt file, exclusively exempting it from being crawled by search engines (Hawkins, 2016). Robots.txt is usually a file embedded in the website that defines what file and web page should be accessible by web crawlers.

Configuration and accessibility settings as highlighted above contribute to the complexity of accessing resources on the deep web. Interestingly, there might be cogent reasons why a user needs access to 'perceived' private information on the internet. This could be the search for a lost relative, old friend, business contact, etc. In some cases, the user's needs may be satisfied but not in most other cases.

Internet speed on the invisible web: When compared to the surface web, accessing the information on the deep web is much slower. For a search query to return the requested result, the deep web search engine will need to search and analyze all searched page contents to ensure that the query result is highly relevant in line with the desired search string. On the contrary, conventional search engines access information on the surface easily and faster because the crawlers can find and access the web pages which are indexed and interlinked. For the deep web, the databases are not indexed. Hence, the speed of access is dependent on the speed at which the databases are found and accessed (Laverty, 2020). On the surface web, the browser will need to send all web requests through the user's public IP address to the relevant server, and requested information is retrieved, clearly a straightforward process. In contrast, the random bouncing buffering of data through several nodes on the Tor or I2P network means that web requests will have to pass through these nodes and the resultant effect is a decline in the speed of access and retrieval of web contents.

Depth of the invisible web: Deep web technology is ubiquitous and often built into the security features of some organizations like the paywalls of organizations in such a way that it tracks unauthorized access to information which when caught, the intruder risks litigation or arrest for copyright infringement and violation of the site's terms of use (Laverty, 2020). The deep web is also remarkable in terms of the depth and accuracy of its results. Comparatively, it is believed that the surface web occupies about 19 Terabytes of the web's storage capacity while over 7500 Terabyte web contents are domiciled in the deep web (Deep web, 2019). The content of the deep web is about 500 times larger than the surface web. Surprisingly, these huge resources are rarely accessed by conventional search engines since most of the data are private and personal information that is not indexed. In other cases, the resources are either deprecated, outdated, or are tucked away in databases in such a way that they are hidden beyond the reach of the crawling capacity of search engines (Bischoff, 2018). Files in cloud storage servers, private social media profiles, and academic journals, among others also constitute web contents that are not necessarily indexed by conventional search engines. Normally, queries on deep web search engines are well-constructed strings that cover far more ground, thereby streamlining the results to a more efficient, higher quality, and relevant content (Lynch, 2020) and with

over 96% coverage of the entire internet resources, the deep web has an inner layer called the dark web.

The dark web is termed 'dark' because of its ability to provide greater privacy on the internet (Wolford, 2018). Oftentimes, this privacy is achieved through the set-up of fully encrypted anonymizing proxy networks that run on specialized search engines such as Infomine, Complete Planet, Deep Dive, and TechXtra. These tools allow the user to locate these hidden and unindexed databases on the internet. Nevertheless, they require the use of additional specialized browsers like the *.onion* browser to view the web contents. It is also worthy of mention that the depth of anonymity found in the deep web is dependent on the strength of the encryption, the size of the network, the number of concurrent users, and the internal architecture of the system (Mann, 2019).

Ethical issues with invisible web: Being an unregulated zone with untraceable internet activities, questions on the confidentiality, ethics, and legality of the deep web becomes imminent. Lack of ethical guidelines leads to unauthorized access to sensitive personal information from normally restricted databases. Consequently, this creates dilemmas that make individuals susceptible to fraud and identity theft (Lynch, 2020). Sadly, authorities with mandates to tackle crime on the invisible web are most times in a fix of taking decisions that will curb the crime and at the same time ensure that the fundamental human rights of the victims and even suspects are preserved. Ethical issues such as this have affected the impact on the fight against illicit activities on the web. In an attempt to solve this problem, government anticrime agencies use undercover agents to infiltrate the networks of anonymous internet users, some of which might be criminals, to gain knowledge about their activities and subsequently arrest those with criminal motives (Hadjimatheou, 2017). While this seems ideal, at least from the point of eradicating digitized crime, the action raises ethical issues bothering the rights of innocent users. One clear implication of this monitoring activity is that the privacy of a legitimate user who chose to be anonymous on the internet is bridged by the prying eyes of the anti-cybercrime agencies. It also undermines the trust of users who see the invisible web as a safe place devoid of incessant cyber-attacks and political persecution (Hadjimatheou, 2017).

Ethical considerations as they relate to interstate and boundary jurisdiction is another issue worth considering in the fight against cybercrime on the deep web. For a well-established and investigated cybercrime case, the security agents at times are constrained with boundary jurisdiction imbroglio – a situation where they do not have the authorized and legal right to pursue fraudsters across boundaries. This is now mitigated through the adequate collaboration of security agencies via Interpol services across the globe.

Another ethical issue that raises concern on the use of the deep web is the informed consent seeking process that comes in form of cookies or *terms and conditions* agreement which must be accepted before access on a web page can be granted. This pervasive demand for consent without a clear understanding of its hidden details falls short of the standard informed consent seeking process; thus posing a threat to the sincerity of purpose of the website and the privacy of the user (Bellaby, 2018). Furthermore, the power of anonymity inherent in the deep web has put all users at liberty to trade on any area of interest, often not minding whose right to privacy and safety is affected. From the foregoing, there is a need to curtail this power overflow but it must be done with caution to ensure the liberal right to freedom of speech is maintained (Gehl, 2016). Be that as it may, it is not unlawful to use the deep web neither is it bad to remain anonymous on the web provided transparency, freedom, and rights of all are preserved. The goal standard is for government policymakers to develop more comprehensive law enforcement, regulatory and national security responses that will take into cognizance of the good and bad uses of the invisible web (Sui *et al.*, 2015).

Cybercrime on the invisible web: Surfing the internet anonymously is not without its attendant effect of exposing users to activities considered illegal in some country settings. This is because the deep web has the potentials to host increasingly high numbers of malicious services and activities. These cybercrime issues could range from gambling, illicit trading on weapons, drugs, credit card fraud, identity theft, and leaks of sensitive information to things like the hiring of assassins, terrorism, child trafficking, and pornography, illegal financial transactions using cryptocurrency among others (Chertoff & Simon, 2015). Deceptively and aesthetically, most of the web pages harboring these illicit intentions and contents still present login information and security checks as one would find in a normally trusted website. Genuine internet users are cautioned to be watchful for these cybercrime flags while on the invisible web to avoid being coerced into subscribing to any of the platforms.

The evolution of malware programs is another contending issue on the invisible web. Malware (also known as malicious software) are programs such as viruses, worms, ransom ware, spyware, Trojan, scareware, etc. that are intentionally designed to cause damage to a computer, server, client, or computer network. Precisely, they are built to disrupt or deny operations, gather the information that leads to loss of privacy or exploitation, gain unauthorized access to system resources, and other abusive behavior (Nash, 2005). There is no gainsaying the fact that the dark web is increasingly becoming the harbor for most malware programs. The dark web provides an enabling environment for seamless anonymous communication between malware servers. For example, Trojan malware such as Skynet, NionSpy, Vawtrack, and Dyre are now taking advantage of the vast network of computers and protection afforded by Tor and I2P networks to mine, steal and spread programs on the dark web. These unabated malware attacks are projected to increase considering the steady migration of malware programs from the surface web to the dark web (Cox, 2015). Consequently, users of the dark web, whether for good or bad intentions, despite having the benefit of being anonymous, are now susceptible to malware attacks, which if not carefully controlled with powerful antimalware programs, can lead to the loss of data or damage of the system used. On the long run, users will always be at crossroads, weighing the options of anonymity and safety from malware attacks.

Poor access control to the invisible web: Technologically, users of the deep web are constrained with access to a unified query interface for the deep web databases. This is because the deep web contains massive, dynamic, and heterogeneous resources that are usually tucked in an unindexed network of databases. Each of the databases provides a specific query interface for users to access (Liu *et al.*, 2014). For instance, a deep web query system for an airline booking system will present to users a single interface for querying multiple related databases in the system and then extracts and returns the relevant information from different web database sources that were queried for the users (Swami *et al.*, 2013). Now, for a large source of data, it becomes very difficult to traverse these databases. This will certainly remain a challenge until an integrated query interface in connection with specific domains for unified access to the data in the domain is effectively built (Liu, Xie & Chang 2014). It is also noteworthy that deep web interfaces are unpolished, less standardized, and would always require higher levels of user involvement albeit continued effort by researchers through the application of different pattern matching algorithms, filtering algorithms, and modeling with an overall goal of mitigating these challenges promises brighter future for ease of use of this domain (Thomaz *et al.*, 2020).

Conclusion

The deep web is made up of peer-to-peer connections, which allow users to share files directly (and secretly). It contains about 96% of the www content. Due to its strong appeal to privacy and anonymity features, internet fraudsters of all kinds now take advantage of the lack of tracking their identity to shield their anonymity from advertisers and officials alike despite its known potential benefits. Similarly, journalists, police, military, whistleblowers around the world now rely on it as a more secure alternative to the public web when searching for sensitive or dangerous information. The deep web is majorly composed of contents that are not indexed, contents with restricted access, private content, and information archived in searchable databases. This implies that not all information on the deep is deliberately hidden. Some might just be behind certain cyber-walls that require payments or authentication to access them.

The Tor network is among the different privacy tools that help users anonymously access internet resources on the deep web. Unfortunately, amidst the numerous advantages of a Tor network, research has shown that activities on the network can be tracked by internet service providers due to the Tor node IPs that are usually public. Again, its effectiveness can only be valued when configured properly and used with compatible external applications (Bojarski 2015). Bischoff (2018) reports that this challenge can be mitigated by setting up Tor privately using a Tor bridge or VPN to prevent the blockage of Tor exit nodes which are usually accessible by providers and can easily be blocked by them at ease.

In light of the above findings, a powerful (preferably licensed) VPN is strongly recommended for all users of the deep web because it is built to encrypt users' data and helps to ensure the privacy of the user online (Norton 2019). A combination of the Tor network and VPN is

a hybrid architecture that can maximize internet security and privacy. Additional security can be achieved by enabling the *"NoScript"* and *"Forbid Scripts Globally*" extensions on the Tor browser. The user can as well, change the value of the "javascript_enabled" config file from "true" to "false" (Deep web 2019).

This review paper highlights the importance of users willing to explore the benefits of deep web technology to be aware of the security issues inherent in it and therefore deploy appropriate precautionary measures in surfing the internet. These review findings show that doing this most likely will prevent internet users from nefarious characters and sites that are designed to spread malware or hack the computer.

Conflict of Interest

Authors have declared that there is no conflict of interest reported in this work.

References

Al Nabki W, Fidalgo E, Alegre E & de Paz I 2017. Classifying illegal activities on tor network based on web textual contents', *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, Valencia, Spain, pp. 35-43.

Bedi P, Gupta N & Jindal V 2020. Dark Web: A Boon or a Bane. In: Khosrow-Pour, M. D. Encyclopedia of Criminal Activities and the Deep Web, *Hershey, PA*, IGI Global, pp. 152-164.

Bellaby W 2018. Going dark: Anonymizing technology in cyberspace', *Ethics and Information Technology*, pp. 189–204.

Bergman M 2001. White Paper: The Deep Web: Surfacing Hidden Value. *Journal of Electronic Publishing*, 7(1).

Bischoff P 2018. *How to Access the Dark Net and Deep Web Safely - Step by Step Guide*, viewed 17 March 2020, <https://www.comparitech.com/blog/vpn-privacy/how-to-access-the-deep-web-and-darknet/>.

Bojarski K 2015. Dealer, hacker, lawyer, spy: Modern techniques and legal boundaries of counter-cybercrime operations. *The European Review of Organised Crime*, 2(2): 25-50.

Chertoff M & Simon T 2015. The impact of the dark web on internet governance and cyber security. *Global Commission on Internet Governance*.

Ciancaglini V, Balduzzi M, McArdle R & Rösler M 2015. Below the surface: Exploring the deep web. *Trend Micro*.

Cox J 2015. *The Dark Web Is Becoming a Safe Haven for Malware*, viewed 17 March 2020, <https://www.vice.com/en_us/article/ae35xe/malware-is-using-the-dark-web-to-stay-hidden>.

Crawley K 2019. *How to Prevent Crime on the Deep Web and Dark Web*, AT&T Cybersecurity, viewed 22 August 2019, <https://cybersecurity.att.com/blogs/security-essentials/deep-web-and-dark-web>.

Deep web 2019. *Deep Web Sites*, viewed 17 March 2020, <https://www.deepweb-sites.com>.

Digital.com 2018. *The Dark Web and Deep Web: How to Access the Hidden Internet Today*, viewed 19 November 2019, <https://digital.com/blog/deep-dark-web>.

Gehl W 2016, 'Power/freedom on the dark web: A digital ethnography of the Dark Web Social Network', *New Media & Society*, vol. 18, no. 7, pp. 1219-1235

Gingerich D 2014, *The Effectiveness of the Tor Anonymity Network*, viewed 7 March 2020, <http://www.cs.lewisu.edu/mathcs/msisprojects/papers/Tor_DavidGingerich.pdf>.

Ghosh S, Das A, Porras P, Yegneswaran V & Gehani A 2017. Automated categorization of onion sites for analyzing the dark web ecosystem', *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, New York, NY, United States, pp. 1793-1802.

Hacker Combat 2017. *Top 10 Deep Web Search Engines of 2017*, viewed 26 February 2019, <https://hackercombat.com/the-best-10-deep-web-search-engines-of-2017/>.

HackerNoon 2018. *Understanding the Deep and Dark Web*, HackerNoon, viewed 4 November 2018, <https://hackernoon.com/understanding-the-deep-dark-web-8e4cad356587>.

Hadjimatheou K 2017. Policing the Dark Web: Ethical and Legal Issues', *The University of Warwick and TNO*.

Hawkins B 2016. Under The Ocean of the Internet - The Deep Web', *SANS Institute: Information Security Reading Room*.

Holland J 2020. Transnational cybercrime: The dark web. In: Khosrow-Pour, M. Encyclopedia of Criminal Activities and the Deep Web, *Hershey, PA*, IGI Global, pp. 108-128.

Kaspersky 2021. *What is the Deep and Dark Web?*, viewed 22 May 2021, <https://www.kaspersky.com/resource-center/threats/deep-web>.

Laverty S 2020. *Advantages, Disadvantages, and Risks of Deep Web Search Engines*, viewed 17 March 2020, <https://smallbusiness.chron.com/advantages-disadvantages-risks-deep-search-engines-74087.html>.

Lightfoot S 2017. Surveillance and privacy on the deep web. *Researchgate*, DOI:10.13140/RG.2.2.21692.74889.

Liu Y, Xie C & Chang J 2014. Research on the integration of deep web query interfaces. *International Symposium on Computer, Consumer and Control*, 332-335.

Lynch W 2020. *Advantages, Disadvantages, and Risks of Deep Web Search Engines*, viewed 18 March 2020, <https://itstillworks.com/12758539/what-is-the-difference-between-metacrawler-a-search-engine>.

Mann B 2019. *What Is I2P & How Does It Compare vs. Tor Browser?*, viewed 12 February 2020, <https://blokt.com/guides/what-is-i2p-vs-tor-browser>.

Nash T 2005 An undirected attack against critical infrastructure: A case study for improving your control system security. *Technical Report, US-CERT Control Systems Security Center*.

Norton 2019. *How to safely access the deep and dark webs*, viewed 22 June 2020, <https://us.norton.com/internetsecurity-how-to-how-can-i-access-the-deep-web.html>.

Parker A, Sharma S & Yadav S 2017. Introduction to Deep Web. *Int. Res. J. Engr. and Techn.*, 4(6): 5650-5653.

Sheils C 2020. *The Deep Web And The Dark Web*, viewed 7 March 2020, <https://digital.com/blog/deep-dark-web>.

Swami D, Sonune G & Meshram B 2013. Understanding the Technique of Data Extraction from Deep Web', *Int. J. Comp. Sci. and Information Techn.*, 4(3): 533-537.

Steve 2016. *Surface Web, Deep Web, Dark Web -- What's the Difference?*, viewed 22 April 2016, <https://www.cambiaresearch.com/articles/85/surface-web-deep-web-dark-web----whats-the-difference>.

Sui D, Caverlee J & Rudesill D 2015. The Deep web and the darknet: A look inside the internet's massive black box. *Science and Technology Innovation Program*, Washington DC.

Thomaz F, Salge C, Karahanna E & Hulland J 2020. Learning from the Dark Web: Leveraging conversational agents in the era of hyper-privacy to enhance marketing. *J. Acad. Marketing Sci.*, 48: 43–63.

Vogt S 2017. The digital underworld: Combating crime on the dark web in the modern era. *Santa Clara J. Int. Law*, 15(1): 104-124.

Weedmark D 2020. *Secure Internet Protocols*, viewed 17 March 2020, <https://smallbusiness.chron.com/secure-internet-protocols-46719.html>.

Weimann G 2015. Going Dark: Terrorism on the Dark Web', *Studies in Conflict & Terrorism*, 39(3): 195-206.

Wolford B 2018. *What the dark is and how you can access it*, viewed 18 March 2020, <https://protonmail.com/blog/what-is-dark-web/>.