Beneath the Surface: Exploring the Dark Web and its Societal Impacts

Hasan Saleh
Abstract

The Dark Web is a hidden part of the Internet that has gained attention due to its illegal activities and potential impact on society. This thesis aims to explore the structure of the Dark Web and its actors. Moreover, this thesis covers the effects the Dark Web has had on individuals and society. A comprehensive literature review, interviews with experts, and explorations of the Dark Web was used to gather information. The findings reveal that the Dark Web consists of hidden services that are only accessible using specialised software and tools which helps individuals remain anonymous. Different actors operating on the Dark Web are identified and categorised into two different categories, lawful and unlawful based on the activities carried by them. The thesis aims to categorise and analyse the motives and behaviours of these actors. Anonymity provided by the Dark Web serves different kinds of purposes and can facilitate illegal activities such as drug trafficking and cybercrime while also providing a platform for individuals to be able to express their thoughts freely. The study concludes that the Dark Web influences various aspects of society such as privacy, security and criminal justice. The research seeks to unveil both the potential benefits and risks associated with the Dark Web and which challenges it poses for law enforcement agencies. Moreover, the study calls for methods which can be used to combat the negative impact the Dark Web has on society.
Acknowledgments

I would like to express my deepest appreciation to my reviewers, Karl Marklund and Aletta Nylén who made this work possible. Their advice, willingness to invest their time and effort into reviewing my work, engaging in thoughtful discussions, and offering guidance has been truly remarkable and helped me through all the stages of writing my thesis. The combined contributions and efforts of my reviewers have greatly enhanced my research experience and enabled me to successfully reach this significant milestone in my academic journey.
Table of Contents

1 Introduction 6
   1.1 Purpose ........................................... 7

2 Methodology 7
   2.1 Literature Study .................................. 7
   2.2 Expert Analysis .................................. 9
   2.3 Accessing the Dark Web .......................... 9

3 The Internet in a Nutshell 10
   3.1 Internet Overview ................................ 10
   3.2 OSI and Protocols ................................ 10
       3.2.1 Network Layer ................................ 11
       3.2.2 IP addresses ................................. 11
       3.2.3 Transport Layer ................................ 12
       3.2.4 Application Layer ............................ 13

4 The World Wide Web 13
   4.1 Accessing a Website .............................. 14
   4.2 Security ........................................... 16

5 Unveiling the Internet’s Infrastructure 18
   5.1 The Surface Web .................................. 18
       5.1.1 Crawling ....................................... 18
       5.1.2 Indexing ...................................... 19
   5.2 The Deep Web ..................................... 19
   5.3 The Dark Web ..................................... 19

6 RQ1: A Deeper Dive into the Dark Web 19
   6.1 Tor network ....................................... 20
       6.1.1 Onion routing in a Nutshell .................. 20
       6.1.2 Vulnerabilities ............................... 23
   6.2 Onion sites and Tor hidden service .............. 24
   6.3 Accessing the Dark Web .......................... 27

7 RQ2: Unveiling the Actors of the Dark Web 27

8 RQ3: Analysing the Impact 39
   8.1 Impact on Individuals and Businesses .......... 39
   8.2 Drug Dealers and Markets ....................... 40
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Purpose of usage for different entities on the Dark Web</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Services and products offered on different Dark Web websites</td>
<td>35</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Literature Review Methodology</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>The OSI model</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>A three-way handshake. Source: [1].</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Code showing an HTTP GET request.</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>Hijacking an HTTP connection.</td>
<td>21</td>
</tr>
<tr>
<td>6</td>
<td>The different layers of encryption applied on a message using onion routing. Source: [2].</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>Onion routing. Source: [3].</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>Data visible to eaves-droppers when combining Tor with HTTPS.</td>
<td>24</td>
</tr>
<tr>
<td>9</td>
<td>First step of setting up a hidden service (Dark Web website).</td>
<td>25</td>
</tr>
<tr>
<td>10</td>
<td>TorBot crawling’s results on the HiddenWiki</td>
<td>28</td>
</tr>
<tr>
<td>11</td>
<td>A Dark Web website which offers free media and news.</td>
<td>31</td>
</tr>
<tr>
<td>12</td>
<td>The New York Times official SecureDrop Dark Web website.</td>
<td>32</td>
</tr>
<tr>
<td>13</td>
<td>A Dark Web website which sells stolen paypal accounts, ebay accounts, and credit cards.</td>
<td>33</td>
</tr>
<tr>
<td>14</td>
<td>A Dark Web website which allows users to hire a hacker.</td>
<td>34</td>
</tr>
<tr>
<td>15</td>
<td>A Dark Web drug marketplace.</td>
<td>36</td>
</tr>
<tr>
<td>16</td>
<td>Top markets on the Dark Web sorted by revenue in 2022.</td>
<td>37</td>
</tr>
<tr>
<td>17</td>
<td>Monthly sales of drugs through different Dark Web markets.</td>
<td>41</td>
</tr>
<tr>
<td>18</td>
<td>Proportion of surveyed Internet users using drugs in the past year who purchased drugs over the Dark Web. Source: [6].</td>
<td>42</td>
</tr>
<tr>
<td>19</td>
<td>The impact that market closures have had on individuals.</td>
<td>42</td>
</tr>
<tr>
<td>20</td>
<td>Dark Web’s impact on different entities.</td>
<td>44</td>
</tr>
<tr>
<td>21</td>
<td>The results of Operation Pacifier. Source: [7].</td>
<td>48</td>
</tr>
</tbody>
</table>
1 Introduction

The Internet has made a huge impact on the world and society by providing communication abilities between people all over the world. The evolution of the Internet has made accessing information and communication easier than ever before. As of January 2023, there are about 5.16 billion Internet users worldwide [8].

The Internet continues to evolve and shape our world in ways that we could never have imagined. Connecting people across the globe, making it easier to research, learn, and stay informed about a wide range of topics, and providing the convenience of shopping from anywhere in the world are just a few possibilities that are provided to society by the Internet. The Internet is being actively monitored and protected from activities that could harm its users [9]. The usage of IP addresses makes it possible to identify and monitor activities occurring on the Internet. For example, websites, servers, and online services often log IP addresses of visitors and users. Internet Service Providers (ISPs) and network administrators can monitor and log the IP addresses associated with devices on their network. This is important for managing network traffic, identifying issues, and maintaining security. However, in reality, security is only ensured for a limited portion of the publicly accessible Internet.

The Internet in general, has different layers where privacy and accountability varies for each one of them. In this thesis, we consider the Internet as divided into the Surface Web and the Deep Web. The Surface Web, which is also known as the visible web is the part of the Internet that is easily accessible by anyone [10]. The Surface Web is publicly accessible by using standardised search engines [10]. This part of the Internet is actively monitored and protected, thus accountability can be enforced since its users are identifiable and their activities are controlled by law through a combination of national and international legal frameworks, regulations, and agreements that govern various aspects of online activities.

Another layer of the Internet is called the Deep Web. Accessing the Deep Web with search engines alone is not possible since websites located on the Deep Web are not indexed. The Deep Web contains a huge amount of data and information that is not easily accessible to the public. Take for example an academic research paper published on a website where paying a subscription fee is required to access contents. Only subscribers are able to access this research paper, thus the fact that the paper is not publicly available makes it part of the Deep Web. Websites that require any authentication credentials
such as email addresses or passwords are also a part of the Deep Web.

The Dark Web on the other hand is a specific part of the Deep Web that is also hidden yet is only accessible through specialized software and tools. Activities are conducted anonymously on the Dark Web, making it a shelter for criminal activities such as cybercrime and selling drugs. The anonymity provided on the Dark Web makes it difficult for law enforcement to detect illegal activities and creates significant challenges for them to be able to provide safety and security on the Dark Web \[^{10}\]. Since users are not identifiable on this platform, the Dark Web lacks accountability.

The aim of this thesis is to provide an in-depth understanding of the Dark Web, discussing its characteristics and functionalities. The thesis also concluded that the Dark Web is not a very researched subject since there was a limited availability of articles focusing on the Dark Web and its impacts.

1.1 Purpose

This thesis mainly focuses on informing the reader about the structure of the Dark Web and what it comprises. Moreover, the thesis is going to explore the feasibility and potential implications of mitigating the impact of the Dark Web. In summary, the three main questions that the research is going to cover are:

**RQ1.** What is the Dark Web?

**RQ2.** Which are the primary actors that operate on the Dark Web?

**RQ3.** What role does the Dark Web play in society?

2 Methodology

This thesis was mainly based on a literature study and expert analysis where experts from different fields were interviewed to supply information surrounding the Dark Web. To be able to explore the different actors that operate on the Dark Web and provide information about them, the thesis also relied on both accessing and browsing the Dark Web and utilizing a Dark Web crawler called TorBot to output data for analysis.

2.1 Literature Study

Identifying good and reliable literature surrounding the Dark Web included reviewing academic articles, reports and scholarly resources. One primary
A keyword-based search approach was conducted to provide comprehensive information surrounding the Dark Web. The first step used to initiate the literature search was the identification of keywords. A broad of relevant keywords were chosen such as "dark web", "deep web", "Tor", and "hidden services" which aimed to provide different aspects of the Dark Web. Furthermore, to be able to understand the different activities that take place on the Dark Web, keywords such as "cybercrime", "markets", "paedophilia", "messaging", "anonymity" were used too.

After identifying relevant keywords, a selection of databases was conducted. These databases consisted of academic databases such as Scopus and Google Scholar. Both databases provided a fair amount of different academic literature and studies which helped capture studies from various disciplines.

Search execution was then done on the selected databases using the identified keywords. The search provided different scholarly articles published that reviewed the Dark Web and its functionalities. Not all articles found were included in the research since a vast majority of them didn’t have the Dark Web as the main focus. Thus, those which mostly focused on the Surface Web and general Internet usage were excluded.

Data extracted from relevant articles were analyzed to be able to identify common patterns and trends.
2.2 Expert Analysis

To be able to further understand how the Dark Web works in general and what effects it has had on society, an expert evaluation methodology was also used. Identifying relevant experts was done by utilizing academic databases, research institutions and recommendations from other experts. Experts which had a good understanding of anonymity networks and cybercrime were included and were namely experts which work on providing security at police departments in Sweden, experts which has researched different topics of the Dark Web, and experts which have helped build the Tor project which is a nonprofit organization that develops and maintains the Tor network [11]. After identifying experts, each of them was individually contacted and invited to participate in the evaluation. Individual evaluation sessions were conducted with each expert either in person or via online meetings. Consent was obtained from the experts that contributed with information and the obtained data was used strictly for research purposes.

2.3 Accessing the Dark Web

Another approach that was taken to gain knowledge and insights about the Dark Web actors was accessing the Dark Web. Accessing the Dark Web involved utilizing specialized software called Tor to navigate through hidden networks and anonymously browse various websites and forums. With the help of the Tor browser, different observations on the Dark Web were documented such as descriptions of websites, forums and other online spaces. A Dark Web crawler, called TorBot was also utilized to crawl specific categories of Dark Web websites and forums relevant to the study’s research questions. TorBot is an open source intelligence tool developed in python which aims to collect data from the Dark Web with the help of data mining algorithms [12]. This tool helped collect Dark Web links which where then utilized to investigate the different actors and activities that take place on the Dark Web.

It is important to note that this process was only conducted for research purposes and complied to legal and ethical guidelines. The activities conducted during this process complied with local and international laws. The process did not involve distributing or buying illegal contents on the Dark Web and respected copyright and trademark laws. The author refrained from engaging in or facilitating cybercriminal activities while conducting this research and did not contribute to human rights violations, or harm to individuals.
3 The Internet in a Nutshell

To be able to understand the structure of the Dark Web, one should firstly understand how the Internet in general works and how it is divided into different layers.

3.1 Internet Overview

The most important characteristic of the Internet is that it has no control center, in other words, there is no single computer or organization that serves as a sole point of vulnerability \[^{13}\]. Instead, the Internet consists of many interconnected machines forming a network. A network is a group of connected devices which can share data. The Internet is formed by interconnected networks working together as a cohesive system.

Computers in networks are typically physically connected using cables, fiber optics, or wireless signals. Data being transferred consists of bits which travel at different speeds and gets interpreted when received by the computer.

Providing the capability of communication between different computers and devices in a network is a challenge. The same thing applies to us humans, two different human beings should be able to understand the same language so that they can communicate with each other. Thus to solve this, protocols for networking were created, which are a set of rules that should be followed across all devices that are trying to communicate over the Internet. Networking follows a set model called the OSI model which is used to guide technologies and computers to interoperate. Figure 1 provides an overview of the different layers of the OSI model which provides a standard for different computer systems and devices to be able to communicate with each other \[^{13}\].

3.2 OSI and Protocols

The OSI model has seven different layers where each layer handles a specific job and follows different sets of rules to ensure a secure and stable connection between devices. The main focuses here will be the application, transport, and network layers, as understanding the basics of those layers would help provide an easier understanding of the differences between the Surface Web and the Dark Web.
3.2 OSI and Protocols

3.2.1 Network Layer

The network layer is responsible for providing a logical communication between devices, in other words, it set up the routes that data should take to be able to arrive to its destination. This is done by assigning each data segment a source IP address and a destination IP address [13]. IP addresses are unique addresses which identifies a device on the Internet and allows information to be sent to any other device on the Internet as long the IP address of the other device is provided. Finding the best path that data can take is also a task of the network layer and is called routing which is path decisions made by routers to provide the fastest delivery of data. Other things are also taken into consideration in the network layer, for example network congestion and link reliability.

3.2.2 IP addresses

IP addresses are utilized to assign unique identifiers to devices within a network, enabling routing and communication across the Internet. There are currently two different versions of IP addresses used today, IPv4 and IPv6. IPv4 stands for Internet protocol version 4 while IPv6 stands for Internet protocol version 6 [13]. More specifically, IPv4 addresses are a string which consists of a 32-bit number split into four different groups of 8-bit numbers. Each 8-bit value is then transformed into decimal form to arrange an IP address. An example of an IP address can be 192.158.1.54. On the
other hand, IPv6 addresses consists of 128-bit numbers which are expressed in hexadecimal form [13]. One might wonder why there are two different versions of IP addresses, this is due to the fact that IPv4 addresses are 32-bit binary numbers which can totally express 4.3 billion unique addresses. This was enough addresses when the protocol was firstly introduced in the 1980s, however the exponential growth of the Internet was unexpected and thus resulted in a shortage of IPv4 addresses, that's why IPv6 was introduced as a solution.

3.2.3 Transport Layer

The network layer only provides best effort delivery service, in other words, it does not guarantees data delivery nor data integrity. Reliable data transfers are introduced in the transport layer, where its main function is to control the flow of data, error correction of corrupted data, and congestion control. The main concern of this layer is to provide communication between different application processes being run on different hosts while the function of the network layer is to provide communication between the hosts themselves [13]. The transport layer ensures that the data sent has been received that the data is processed in the same order as it was sent.

Another primary function of the transport layer is to control the flow of data, where the data flowing from one end to another is being constantly monitored and examined so that no overloading takes place which might lead to network congestion and failure. There are mainly two primary transport protocols that are frequently used over the Internet: Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) [13]. To understand the difference between both, one should firstly understand the difference between connection-oriented and connectionless services. TCP offers connection-oriented services which is establishing a reliable and secure connection between the host and the receiver before transmitting data and keeping the connection alive all the time until no further data needs to be transmitted. TCP uses what is called a three way handshake technique which allows the host and the receiver to establish a connection between each other before transmitting data. Data that is lost is oftenly re-transmitted by TCP.

UDP on the other hand offers connectionless services without offering reliability. UDP does not offer a connection between the host and the receiver, instead, data is sent without achieving any connection. No error detection or correction techniques are used in UDP, and packets lost are not re-transmitted. UDP is mainly faster than TCP as it requires no overhead and just sends the data as is, while TCP is mainly safer than UDP and is used in for example
web-browsing and file transfers.

### 3.2.4 Application Layer

The application layer provides network services for the applications running on devices communicating with each other. This layer enables different applications to send and receive data over the network. Keep in mind that the application layer is not the application itself, instead, it is a component within the application that allows network communications to other devices [13]. The application layer is also responsible for providing error handling and security services between different applications to achieve a secure and reliable communication.

Security is a very important factor when it comes to communicating over the Internet, that’s why the application layer offers different security services. Authentication and authorization are two services offered by this layer that help verify the identity of every communicating entity. Moreover, this helps control how and by who network resources should be accessed to avoid congestion. The application layer can encrypt data that is transmitted over the network so that unauthorised parties would not be able to collect any information about the data transmitted. Encryption is typically done using different algorithms such as RSA and AES.

HTTP and SMTP are two examples of protocols that operate on the application layer. SMTP stands for Simple Mail Transfer Protocol which is used to handle the action of sending and receiving mails over the Internet. This protocol is mainly used for email services. HTTP on the other hand, stands for Hypertext Transfer Protocol which is a protocol used to transfer data over the Internet [13]. HTTP is used by web browsers and servers to be able to transfer data over the web. The next section describes in detail how a website is accessed on the Internet and dive into more details about the HTTP protocol.

### 4 The World Wide Web

The World Wide Web (WWW) is an Internet-based application. The World Wide Web, commonly referred to as the web, serves as an organizational framework for accessing information through the Internet. The Internet can be used without using the web, this can be done for example via sending an Email which uses the SMTP protocol and works without a web browser.
4.1 Accessing a Website

Taking a look at how websites are accessed provides a better understanding about what is happening in the hidden background and gives a better comprehension of the OSI model and how different layers are tied together.

Every website on the Internet is stored on a web-server, which is a computer in the network that is used to store the contents of a website, such as text, images, links, etc. Not only does it store the contents of a website, it is used to deliver and process websites to clients that request it using the HTTP protocol.

Assuming that Wikipedia is the website a client is going to access, the first action that should be taken is opening a web browser. Web browsers can be seen as interpreters which reads the code of a web-page and display it for a user. Google Chrome is an example of a widely used web browser.

After opening up the web-browser, a user should supply the browser with a URL for the website. In this case, the URL is `https://en.wikipedia.org/`.

As mentioned in Section 3.2.4, IP addresses are used on the Internet to be able to identify certain devices and get access for certain content, in this case, one of the main reasons why URLs (which stands for Uniform Resource Locators) are used instead of IP addresses is because they are more human-readable and are easier for people to remember than a series of numbers separated by dots.

However, the Internet does not understand how URLs work. They are purposely made to provide easier accessibility for users, thus, to be able to establish a connection on the Internet, the standard protocol should be followed and thus a URL should firstly be converted to an IP address. Converting the URL into an IP address is done by sending a UDP request containing the URL to a DNS server [13].

A DNS stands for Domain Name System and is a server which is responsible for finding the correct IP address for websites. The IP address supplied by the DNS server is then used by the client to communicate with the actual web server [13]. DNS can be considered as a phone book for the Internet where the server checks firstly if the supplied URL has any associate IP address, if it does then the IP address is sent back to the client, if it doesn’t, communication with other DNS servers is needed.

Once the IP address of the website has been retrieved, a connection between the client and the webserver should be established. Browsers typically use TCP to establish a connection and transfer data reliably. As described in
Section 3.2.3 TCP ensures that data is being transferred and resends data that has been lost. To be able to initiate a reliable TCP connection between the server and the client, a three-way handshake is used firstly. Figure 3 demonstrates how a three-way handshake gets established.

![Figure 3: A three-way handshake. Source: [1]](image)

Theoretically speaking, data can now be exchanged between the browser and the server after establishing a TCP connection. This is done by following the HTTP protocol which consists of two types of messages, requests and responses. HTTP requests are generated by the browser itself and are used to retrieve information and contents from the server. When accessing Wikipedia for example, the browser will send out an HTTP request message to the server to be able to retrieve the HTML code and load the website for the user. HTTP responses are answers to HTTP requests.

An HTTP GET request is generated and sent away to the server when trying to load Wikipedia in the web-browser. Figure 4 above displays an example of an HTTP GET request generated when trying to access Wikipedia. The GET method is used to retrieve the resources of the web-page specified, which in this case is Wikipedia. Other fields specify different things, for example the Host field states the domain name, while the User-Agent field specifies the web-browser being used by the client, which in this case is Mozilla Firefox.

```
GET /wiki/ HTTP/1.1
Host: en.wikipedia.org
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:88.0) Gecko/20100101 Firefox/88.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate, br
Connection: keep-alive
Upgrade-Insecure-Requests: 1
```

![Figure 4: Code showing an HTTP GET request.](image)
4.2 Security

The HTTP protocol itself does not provide any type of security or protection for the data being transmitted. Sensitive data such as usernames and passwords are not encrypted and can thus be retrieved by hackers while being sent to the web-server. The HTTP requests and responses are simply sent in plain text. An extension of HTTP called HTTPS was created to be able to solve this problem. Notice how the URL of Wikipedia (https://en.wikipedia.org/) starts with https instead of http.

HTTPS comes supplied with encryption and verification methods which uses TLS (Transport Layer Security) to be able to provide these features [13]. TLS is a cryptographic protocol designed to provide communications security over a computer network including emails, web browsing and file transfers [13].

TLS mainly provides three major features:

- The client can be sure that the data exchanged between it and the server is not being read by anyone else.
- The client can be sure that the data exchanged between it and the server is not being changed by anyone else before arriving to the server or back to it.
- The client can be sure that it is communicating with the intended server.

The main goals here are security, authentication and integrity which are solved by encrypting data and signing it. The whole process of TLS is built on using public key cryptography and digital signatures.

Public Key Cryptography

Public key cryptography uses two cryptographic keys, a public key and a private key. The public key, which is available to anyone, is used for encryption of data while the private key, which is kept secret and never shared, is used for decryption of data. Data can only be decrypted using the private key. The keys can be generated using different mathematical techniques which will not be covered in this thesis.

Digital Signatures

A digital signature is used to validate the authenticity and integrity of a digital document. These are typically created by using hashing the data itself. The person who creates the digital signature uses a private key to
encrypt the hash, and then the public key is used by the other end to verify the signature.

After establishing a TCP connection between the client and the server, the client issues a TLS handshake to be able to authenticate and secure the connection. A web server must have a TLS certificate to be able to establish secure connections with its clients. Certificates are used to allow the server to prove its identity when communicating with its clients. These certificates are mainly sold by certificate authority organizations which validates the domain and the owner details before giving out the certificate for the website. An example of a certificate authority is Amazon Trust Services. TLS certificates has a maximum validity period of 13 months. A certificate typically contains the name of the CA company, the name of the domain, the servers public key, and the CA’s digital signature. The CA signs the certificate using its own private secret key that no one else knows. Anyone with access to the CA’s public key is able to verify that the digital signature was initiated by the CA itself [14]. Every modern browser comes preinstalled with the public keys of certified certificate authorities nowadays, which makes it possible for the client to use one of the public keys it already has depending on which CA the certificate came from. The browser computes the hash of the certificate and decrypts the digital signature by using the public key it already has, if both hashes match, then this means that the certificate was really issued by the CA and the client can be sure that the public key the server sent is really the server’s and not somebody else’s [14]. The following steps provide a very brief explanation of the basics of TLS handshakes, do note that these steps differ depending on which version of TLS is being used:

1. The handshake starts with the client sending a hello message to the server asking to initiate a TLS connection.
2. The web-server receives the message and sends back its certificate and its public key back to the client.
3. Before using the server’s public key, the client needs to verify first that this message was indeed sent from the server. This is done by decrypting the signature in the certificate.
4. After the verification process, the client generates a random key which will be used by both the server and the client as a symmetric key to encrypt/decrypt messages they will be exchanging with each other.

The symmetric key is then encrypted with the server’s public key and sent back to the server. Since only the server has the private key, it is the only one that can decrypt the message and thus the key generated by the client
will only be known by the client itself and by the server.

5  Unveiling the Internet’s Infrastructure

Usual Internet activities like emailing, googling and accessing social media and different platforms are only a small proportion of what the Internet can deliver to its users. There are different layers of the Internet which have different characteristics and functionalities. This section will explore three different parts of the Internet examining their features and how they differ from each another.

5.1  The Surface Web

The surface web, usually called the visible web, is the visible part of the Internet that most people use on a daily basis. One could describe it as the tip of the iceberg which can be seen by everyone. Everything that users can find on the WWW (World-Wide-Web) using search engines belongs to the surface web. Search engines such as Google have two primary functions: Crawling and Indexing which makes it possible to access different websites on the Internet.

5.1.1  Crawling

Every search engine uses different search algorithms and favour web-pages based on content quality and users experience. Crawling is the a process which sends out a team of robots known as spiders to find out new and updated content [15]. The goal of a spider is to learn what every web-page on the web consists of and retrieves its information. When a user provides a search query for the search engine, the search engine firstly use web crawlers to scour the Internet and build a database consisting of relevant websites that are related to the search query the user supplied. Web crawlers mainly start searching from a specific specified seed or a know URL list. Crawling these websites leads them to other websites via hyperlinks for example. Sometimes different websites are excluded because the hosting web-server itself denies access to these crawlers and thus these websites are not crawled and not shown to public.

In summary, crawling is about discovering and finding relevant URLs on the web. Web scraping on the other hand, is the process of extracting data from websites and can be done by using different tools to provide the user with the contents of specific websites.
5.2 The Deep Web

5.1.2 Indexing

The information that has been found by crawlers are arranged and categorised in this step. Indexing is about processing and analyzing the contents of the web-page. Indexing stores relevant websites in its huge database called the index. These databases are kept fresh since crawling and indexing is done continuously to provide the users with the best results. If a website is not indexed, it is simply not stored in the search engines database and can thus not be viewed by others.

According to wordlwidewebsize.com, the Indexed Web contains at least 7.26 billion pages [16].

5.2 The Deep Web

Websites which are not crawled or indexed are not displayed to users through search engines. These websites mainly belong to the Deep Web which contains web pages that are inaccessible to public users for a reason and requires certain actions or a direct URL link to be accessed. Contents on the Deep Web includes private intranet such as those at universities and government agencies. It also includes everything that requires login credentials to access data for example bank accounts and personal accounts. It is estimated that the deep web consists of 95% of the whole web [17].

5.3 The Dark Web

The last layer which is the main focus of this thesis is the Dark Web which belongs to the Deep Web and is a part of it. The Dark Web and the Deep Web are not the same and have many differences. The main difference is that the Dark Web requires special software to be accessed. Dark Web websites are also not indexed either, meaning they can’t be accessed by using standardized search engines. The Dark Web mainly contains illegal content, such as drug trafficking and other illegal activities which will be covered later on. The next section provides a deeper explanation about the Dark Web and covers its usage and properties.

6 RQ1: A Deeper Dive into the Dark Web

As noted before, the Dark Web consists of websites which are intentionally hidden from the public and not indexed by regular search engines. Every activity that is held on the Dark Web is done anonymously and users cannot
be identified mainly because different encryption services are being used. The Dark Web itself exists on what is called the Darknets network which is an overlayed network on the Internet [18]. A Darknet network refers to a portion of the Internet that is intentionally hidden and inaccessible through standard web browsers and search engines. It operates on encrypted and anonymized communication protocols, allowing users to access online content, services, and resources with a high degree of privacy and confidentiality. Before diving into how Dark Web websites can be accessed, one should firstly understand how the networks that enable access to Dark Web websites work.

There are a couple of different services that offer access to Darknets such as I2P, Freenet, Zeronet, and GNUnet, however this thesis entirely focuses on the most popular service called Tor. This section aims to answer the research question RQ1: What is the Dark Web?

6.1 Tor network

Tor network allows users to anonymously access content on the Internet. Tor refers to "the onion router" which uses different techniques to be able to anonymize web traffic. It uses onion routing which was developed in the mid-1990s by United States Naval Research Laboratory employees to protect American intelligence communications online [4]. Tor provides security and anonymity by implementing further encryption in the application layer of the OSI model. As described in Section 4.1, when using HTTPS to access a web-server, the server knows who is trying to access it and anyone who is sniffing the connection between the user and the server can track the user’s activity. Tor is made to make surfing on the Internet private and thus anyone watching the connection between the user and the server is unable to track your Internet activity. While Tor provides a high degree of anonymity, it doesn’t completely hide your identity from the websites you are trying to access. This is because it remains susceptible to various attacks, which will be explored further in this section. Users can access sites via the Tor Network using a special browser developed by the Tor Project.

6.1.1 Onion routing in a Nutshell

Onion routing is the technique used to provide the anonymity features in the Tor Network. As described in Section 4.1, using HTTP alone to access websites is not safe as data travelling through the connection is not being encrypted. It can be read by conducting a man in the middle attack on the connection the user has established. Conversely, with HTTPS, information is encrypted, making it impossible for anyone intercepting the data to comprehend
its contents. This is due to the requirement for the interceptor to possess the symmetric key, established between the client and server, in order to decrypt the data.

When sending a packet with data on the Internet, the packet also contains the source IP address and the destination IP address. Someone sniffing an HTTPS connection cannot see the contents of the data in the packet, however, the client’s IP address and the server’s IP address are still visible. Hiding these IP addresses would violate the Internet protocol since the source and the destination IP addresses should be known on the Internet in order for the packets to be forwarded or else the packet is discarded.

Onion routing is a clever technique to solve this issue and anonymise connections to servers making it nearly impossible to know which site a client is trying to communicate with. The website itself does not know who is communicating with it either. Tor network itself consists of a group of volunteer-operated servers which are oftenly called relays. Tor relays are routers or nodes that receives your traffic on the Internet and pass it along. A client initiates a connection on this network by connecting through a series of relays rather than making a direct connection. The default number of relays used by the connection is three and are called Entry Node, Middle Node and Exit Node.

When a client establishes a TCP connection using Tor, Tor finds three different nodes out of its operated servers. The main focus of the algorithm is that not any single node knows the entire path. The Entry Node knows who you are but not who you are communicating with, the Middle Node knows neither, and the Exit Node knows who you are communicating with but not who you are. Having three nodes in the circuit makes it difficult to correlate incoming and outgoing traffic. This even helps prevent a single point of failure.

Figure 5: Hijacking an HTTP connection.
in case one node is compromised or malicious. Once three distinct nodes are identified, three unique symmetric shared keys are established between the client and these three relays. As a result, the client possesses all three keys, while each router holds only one of the keys, excluding the others. Now assume that the client has a packet of data that should be sent to the server. The client firstly encrypts the packet starting with the Exit Node’s key, then the Middle Node’s key and lastly the Entry Node’s key. The packet is then sent by the client to the Entry Node which is the only node able to see your IP address. The Entry Node uses its key to decrypt the first layer of the packet and then knows that the packet should be forwarded to the Middle Node. The Middle Node cannot see your IP address, it only sees where the packet came from and where the packet is headed when it decrypts the second layer with its key. The Middle Node then forwards the packet to the Exit Node which finally decrypts the whole packet with its key and sends the data to the server. The destination server will thus never know your IP address. Figure 6 illustrates how different layers of encryptions are applied to a message before getting forwarded by the source.

The same process is applied but in reversed order when receiving data from the server, the packet is encrypted on the way back and then decrypted by the client since the client has access to all the keys. Further documentation and explanation of the Tor design can be found on Tor’s official documentation.
6.1 Tor network

Two main vulnerabilities can be noted from the implementation of the Tor network. The first being that the actual data is being decrypted by the Exit Node and sent directly to the server with no protection. This of coarse is not secure since an attacker can easily listen to the last channel and be able to read the data. The Tor browser, utilizing the Tor network, addressed this problem by incorporating a solution that involves employing HTTPS within its implementation. Consequently, alongside onion routing, HTTPS is employed to guarantee that the Exit Node remains unaware of the data transferring between the client and the server. Additionally, the data transferred from the Exit Node to the server is made secure by utilizing TLS encryption, ensuring that the message remains encrypted.

Another vulnerability that might arise is if sniffing was done on both the channel between the client and the Entry Node and the channel between the Exit Node and the server. Sniffing the channel between the client and the Entry Node can provide eavesdroppers with the IP address of the client. This is not a concern since eavesdroppers will only know that the client is using Tor and not what the client is trying to access. However, combining this

Figure 7: Onion routing. Source: [3].
information with the information that can be sniffed on the channel from the Exit Node to the server (mainly the servers IP address), it might be possible to figure what the client is trying to access. This is difficult to execute since these relay nodes are not only acting as relay nodes for one client but also serve other clients as well. Packets are always traveling through and out of them the whole time, thus trying to identify which packets belongs to the client is quite difficult and requires special algorithms.

6.2 Onion sites and Tor hidden service

Tor hidden services are website servers which can only be accessed using the Tor network. These servers are not subject to eavesdropping and provide total anonymity for the server itself so that the IP address of the server is not publicly known and the server is hard to track. Websites hosted on Tor hidden service end with a .onion in there URL and are Dark Web websites. The Dark Web namely consists of a network of hidden services. Dark Web websites solve the vulnerability mentioned in Section 6.1.2 which is that sniffing can be done on exit nodes to learn where the packet is travelling and thus provides server information for the eavesdropper.
When clients try to access normal websites on the Tor network, the IP address of the server is publicly known. However Tor hidden services aims to hide the identity of the server which makes it impossible for people to trace or know the actual location of the server, in other words, neither the client nor the server know anything about each other.

The main difference between onion routing and Tor hidden services is the fact that packets do not leave the Tor network when getting forwarded to the hidden service making exit nodes attacks impossible. Tor hidden services provides three main benefits to its users: Location hiding which hides the location of the server and allows it to offer TCP connections to clients without disclosing its IP address, end-to-end authentication which means that a person visiting an onion website knows for sure that the content they are seeing can only come from that website itself, and end-to-end encryption which offers encrypted traffic from the client to the server without using HTTPS [19].

For this particular section assume that Bob is the hidden service and Alice is the client trying to access this server. The first step in the whole process starts when Bob calculates its key pairs which consists of a public key and a private key. Bob then picks 3 random introduction points in the Tor network and creates a Tor circuit informing them to act as introduction points for it and supplies them with its public key. Introduction points are normal onion routers operated on the Tor network and act as contact points and introduce clients to the hidden service without making the server get involved.

![Figure 9: First step of setting up a hidden service (Dark Web website).]
in the process. This provides location-hiding for Bob since Bob’s location is not revealed to any of these introduction points (this is achieved by the use of the Tor circuits). Access to the hidden service will only be allowed through these introduction points. After this step, Bob creates a hidden service descriptor which contains Bob’s public key and which introduction points Bob has chosen. The descriptor is then signed with Bob’s private key and uploads the descriptor to an onion directory server which is part of the Tor network and is a distributed hash table [19].

Once the hidden service has been setup, the service publishes its onion service address which is a 56 character hash name and ends with .onion and are based on the public key of the hidden service. Onion addresses are not publicised over the whole Tor network, they can be found through accessing private communities on the Internet or simply provided through communication with someone that has an onion address of a website. There are public websites on the Internet that provide users with different onion addresses (dark web websites). An onion address can look like the following: hashvalueofpublickeyofhiddenservice.onion

Once Alice has been supplied with an onion address, a connection is made to the distributed hash table (the directory) and the hidden service descriptor is provided to Alice. Once Alice has the hidden service descriptor, the signature is verified using the encoded public key in the onion address. This provides end-to-end authentication for the protocol. The client (Alice) now has the server’s public key and information about the introduction points the hidden service is using. Before Alice makes a connection to one of the introduction points, she firstly establishes a connection to a random onion router on the Tor network to act as a rendezvous point (RP). Alice supplies the rendezvous point with a one-time secret code which is called a rendezvous cookie. The cookie is used to make Alice recognize Bob when a connection is established.

After achieving a connection with a rendezvous point, Alice sends its one time secret again with the RP address to one of Bob’s introduction points over a Tor circuit. The message is encrypted with the hidden service’s public key and then passed over to the introduction point which forwards it to Bob. Bob decrypts the message and decides to allow the connection. Bob now establishes a Tor circuit with the rendezvous point and supplies it with the one time cookie again. The rendezvous point compares the two cookies, if they match then the client is informed that a connection has been successfully established.

Do note that a Diffie-Hellman handshake also takes place between Alice and Bob in the process above so that end-to-end encryption is supplied.
6.3 Accessing the Dark Web

Accessing the Dark Web (and the Tor network in general) can be easily done via downloading the Tor browser provided by the Tor project team. The Tor browser is similar to any other web-browser available for free on the Internet. The difference is that the Tor browser connects its clients to the Tor network and does not store browser history. Once the Tor browser have been downloaded, one could access the Dark Web’s websites by simply acquiring the link of an onion website first and then pasting the link in the Tor browser.

7 RQ2: Unveiling the Actors of the Dark Web

The anonymity that is provided by the Dark Web can serve both good and bad purposes. Anonymity online makes it possible for users to express opinions and impressions with no limits since their online identities are not connected to the real world. To be able to further investigate the different actors that operate on the Dark Web and what activities it offers, accessing and crawling the Dark Web was done via the help of the Tor browser and an Open Source Intelligence (OSINT) tool called TorBot which gathers data from the Dark Web’s websites. Open source intelligence involves the gathering of information from sources that are openly published or otherwise accessible to the public [20]. Since the Dark Web is openly available for the public and can be accessed by anyone using the special software that is needed, it is possible to develop OSINT tools that can scrape different websites on the Dark Web and gather information about them.

TorBot is an OSINT tool which is used crawl the Dark Web’s websites and provides the title and a short description of each website crawled. The tool was mainly used to scrape websites which serve as main pages offering a directory of links to Dark Web websites. Two popular websites were crawled, HiddenWiki and Torlinks. Both of these websites serve as a URL list for Tor hidden services. Furthermore, to be able to access these websites and explore their content, the Tor browser was used to access the Dark Web’s websites. Figure 10 illustrates some of the results obtained by the TorBot crawling tool. Do note that the description of TorBot in the figure states that it is an OSINT tool for the Deep Web, however, as previously discussed in Section 5.3, the Dark Web is part of the Deep Web.
This section aims to investigate the different actors on the Dark Web and analyze their activities by exploring the Dark Web and crawling different websites with the TorBot tool, providing an answer to the research question **RQ2**: Which are the primary actors that operate on the Dark Web?. Furthermore, the actors are divided into different groups: lawful and unlawful. To differentiate between lawful and unlawful entities on the Dark Web, we establish specific criteria that guide our categorization process. The primary factor considered is the adherence to existing legal frameworks. Entities operating within the boundaries of the law are classified as lawful, while those actively violating established laws fall under the classification of unlawful actors. The categories of the actors are summarized in Table 1 and then described in detail further on in this section.
Table 1: Purpose of usage for different entities on the Dark Web

<table>
<thead>
<tr>
<th>Actors</th>
<th>Purpose of Usage</th>
<th>Lawful</th>
<th>Unlawful</th>
</tr>
</thead>
</table>
| Ordinary Users                | • Protection of privacy.  
|                               | • Overcome identity thefts.  
|                               | • Private communication. | X       |          |
| Journalists & Researchers     | • Studying illicit activities.  
|                               | • Tracking emerging trends.  
|                               | • Protection of privacy when publishing non-state controlled articles.  
|                               | • Writing and reading controversial topics.  
|                               | • Exchanging information without revealing their identities. | X       | X        |
| Activists & Whistle-blowers   | • Bypassing restrictions and communicating freely.  
|                               | • Leaking documents.  
|                               | • Exposing wrongdoing and corruption. | X       | X        |
| Organizations & Businesses    | • Monitor the Dark Web to identify potential threats to their cooperations.  
|                               | • Improve security to prevent data leakage or thefts. | X       |          |
| Hackers & Malware Developers  | • Selling hacking tools and services  
|                               | • Strategizing large scale attacks  
|                               | • Selling stolen data. | X       |          |
| Market Operators              | • Selling drugs and illegal services for profit  
|                               | • Offering stolen credit cards and ID documents  
|                               | • Selling weapons and firearms | X       |          |
| Terrorists                    | • Disseminating propaganda and spreading their ideologies  
|                               | • Raise funds to buy equipment and conduct operations  
|                               | • Selling stolen items and goods | X       |          |
| Other offenders               | • Hosting websites which offer child pornography  
|                               | • Experimenting on humans  
|                               | • Offering assassination services | X       |          |

**Civilians**

The Dark Web offers protection of personal privacy which helps conceal users’ identities and makes it hard to identify them. Being able to express different ideas and opinions without getting identified can provide many benefits to online users [21]. Users around the world may harbor apprehensions regarding potential consequences such as political or economic retaliation, harassment or life threatening threats and may thus refer to use the Dark Web to overcome these alarms. The Dark Web is primarily sought after by regular members of the public who wish to browse the Internet anonymously,
avoiding any tracking.

The HiddenWiki which was accessed using the Tor browser offers links to different popular sites that have Dark Web versions of their websites. Popular websites sometimes provide Dark Web versions of their platforms to offer an alternative access point for users concerned about privacy and anonymity. By hosting a Dark Web version, these websites aim to cater to individuals who wish to browse and engage with their content without revealing their identities. TorBot facilitated the availability of Dark Web versions of popular websites like Facebook, Twitter, BBC News, Reddit, The New York Times, and Deutsche Welle.

The crawling’s results on the HiddenWiki that were supplied by TorBot provided a list of websites which offer anonymous and secure e-mail services that can be used by individuals to protect their privacy while communicating with others; these websites were:

- KeyBase: Provides secure messaging and file sharing
- TorMail: Provides secured anonymity mail service
- ChatTor: A Tor-Based chat that provides privacy
- CTemplar: An encrypted tor email service

According to the results supplied by TorBot and the investigation of different websites on the Dark Web, Dark Web websites help individuals to overcome different theft activities that could take place on the Surface Web such as identity theft. Identify theft refers to crimes where personal information’s of users have been compromised. Having online privacy contributes to psychological well-being for individuals since knowing that private and sensitive information is protected reduces users concerns about the invasion of their private lives. Moreover, online privacy helps avoid discrimination and judgements of users based on their characteristics such as race and religious beliefs for example.

The Tor project identifies the primary motivations behind individuals using the Dark Web as protecting their privacy from marketers and identity thieves, seeking secure and private communication, and safeguarding their geolocations [22].

Researchers and Journalists

The Dark Web provides potential benefits for journalists and researchers. The online anonymity that the Dark Web provides offers journalists and researchers the ability to explore sensitive subjects such as drug trafficking,
cybercrime and terrorism. The investigation that was carried on the different Dark Web websites concluded that there are a vast amount of websites which offer resources for journalists and researchers that could be used to explore different subjects and court cases which might be relevant to journalists. A website named Judicial review that was found on the HiddenWiki offered a huge database of court cases. The Dark Web provides platforms, including news and media forums, that can be utilized by nations lacking unrestricted media access. This allows journalists and individuals to share political information within the confines of the Dark Web. Media organizations such as The Guardian and The Washington Post tend to use services on the Dark Web which allows them to securely exchange information without the fear of their identities being revealed or their communication being intercepted. SecureDrop was one of the websites that was found through TorBot that provided a secure service for media organizations to accept documents from anonymous services. Figure 11 demonstrates a Dark Web news site called ProPublica which can be used to read news about different political topics.

Figure 11: A Dark Web website which offers free media and news.
Figure 12: The New York Times official SecureDrop Dark Web website.

Activists and Whistle-blowers

During the investigation, an observation was made that highlights the Dark Web’s potential as a refuge for activists and whistle-blowers seeking to expose wrongdoing and share sensitive information while maintaining anonymity. Many of the Dark Web websites that were found provided platforms for activists and whistle-blowers to expose information without revealing their identities. Figure 12 shows the SecureDrop website that can be used by whistle blowers to securely share information with The New York Times.

AfriLeaks was another website that was identified which helps whistle blowers share information with Africa’s journalists. In summary, activists and whistle blowers play a role in leaking documents and revealing corruptions taking place in their countries.

Organisations

Organisations and businesses tend to explore the Dark Web to protect their businesses from different threats that their companies could face or have faced. Corporations usually face many cyberthreats nowadays which can include DDoS and hacking attacks that could leak their customers private information and data. Different websites that sold leaked data were found on the Dark Web markets. TorBot for example provided a website which sold stolen credit cards, Paypal accounts, and eBay accounts. Companies
might sometimes not notice that they have been exposed to such attacks which creates a further risk. To be able to provide protection, businesses typically monitor the Dark Web and react to threats and stolen data found on marketplaces [21]. Figure 13 shows a Dark Web website which offers stolen Paypal accounts and eBay accounts which can be purchased using crypto-currency.

Figure 13: A Dark Web website which sells stolen paypal accounts, ebay accounts, and credit cards.

Hackers and Malware Developers

Hackers operating on the Dark Web engage in a variety of illegal activities, including offering hacking tools and exploits for sale. These tools encompass a wide range of malicious software, such as malware keyloggers, Trojans, exploit kits, and system-compromising software. They also develop and sell malware capable of evading cybersecurity defenses and infiltrating systems on the Dark Web. Additionally, hackers provide hacking services that can be acquired by individuals seeking to gain unauthorized access to email accounts, social media platforms, or websites [23].
Since the Dark Web offers anonymous communications between individuals, hackers typically communicate and strategize large scale attacks with little risk of being monitored or caught. Data stolen by hackers is securely stored on the Dark Web and sometimes sold by them to other individuals. These transactions typically take place on digital marketplaces hosted by the hackers on the Dark Web which look like any traditional e-commerce website. Stolen data can range from credit card data, login credentials, or intellectual property. The Dark Web website shown in Figure 14 offers a hire-a-hacker service where hackers can be hired to perform different kinds of operations such as stealing accounts.

Phishing is also a technique used by hackers to perform different types of financial fraud. Users wishing to access a specific Dark Web website are often forwarded to a similar looking website which steals users credentials when they try to perform a transaction.

**Market Operators**

Market operators create and manage underground marketplaces on the Dark Web. Just like any other marketplace found on the Surface Web such as Amazon, these market operators hosts similar websites on the Dark Web which sell illegal goods and resources. These goods can range from drugs including narcotics, opioids and steroids to weapons and firearms. Illegal goods and services have big demand which makes market operators seek for new clients. Almost any type of illegal product or service can be found on
the Dark Web. Silk Road was a widely used Dark Web market back in 2013. Silk Road had the same web design as any common shopping website, each product has a detailed description, a photograph and a price. Silk Road offered drugs, fake IDs, passports, and stolen credit cards. Credit cards are typically not used to purchase from such markets as that would be too easy to trace, instead, payments are done via crypto-currency. Silk Road was shutdown by U.S. Federal Bureau of Investigation (FBI) in 2013 [24] however there are still a vast amount of marketplaces available on the Dark Web [5]. Different kind of services were found while investigating the Dark Web, Table 2 provides an insight of the different products and services that could be purchased on the Dark Web’s marketplaces.

Table 2: Services and products offered on different Dark Web websites.

<table>
<thead>
<tr>
<th>Products and Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDos attacks</td>
</tr>
<tr>
<td>Stealing personal information services</td>
</tr>
<tr>
<td>Trojans and Malware</td>
</tr>
<tr>
<td>Drugs</td>
</tr>
<tr>
<td>Weapons (firearms, ammunition)</td>
</tr>
<tr>
<td>Stolen credit cards</td>
</tr>
<tr>
<td>Fake IDs</td>
</tr>
<tr>
<td>Froged documents (university degrees, diplomas)</td>
</tr>
<tr>
<td>Human trafficking services</td>
</tr>
<tr>
<td>Organ trade (illegal organ trafficking)</td>
</tr>
<tr>
<td>Counterfeit currency</td>
</tr>
</tbody>
</table>
Tom and Jerry Store

We have been active during the Agora, Evolution, Silkroad 2 era, then continued through Alphabay and Nucleus, and even the late Dream Market and also Wallstreet, with the same successful concept:
High quality drugs combined with an extremely discreet and fast shipping.
Extremely stealth shipping from the Netherlands!

High Quality Cocaine [90%]

We offer High Quality Cocaine 90%+ with FREE SHIPPING!
All orders that come in before 14:00 Dutch local time are shipped the very same day!
Shipping internationally!

<table>
<thead>
<tr>
<th>Product</th>
<th>Price</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2g High Quality Cocaine</td>
<td>90 EUR = 0.00358 B</td>
<td>1 x Buy now</td>
</tr>
<tr>
<td>5g High Quality Cocaine</td>
<td>200 EUR = 0.00705 B</td>
<td>1 x Buy now</td>
</tr>
<tr>
<td>10g High Quality Cocaine</td>
<td>350 EUR = 0.01392 B</td>
<td>1 x Buy now</td>
</tr>
<tr>
<td>25g High Quality Cocaine</td>
<td>750 EUR = 0.02983 B</td>
<td>1 x Buy now</td>
</tr>
<tr>
<td>50g High Quality Cocaine</td>
<td>1400 EUR = 0.05568 B</td>
<td>1 x Buy now</td>
</tr>
<tr>
<td>5g High Quality Crack Cocaine</td>
<td>220 EUR = 0.00875 B</td>
<td>1 x Buy now</td>
</tr>
</tbody>
</table>

Figure 15: A Dark Web drug marketplace.
Figure 16: Top markets on the Dark Web sorted by revenue in 2022. Adopted from: [5].

Figure 16 demonstrates the top Dark Web marketplaces provided by Chainalysis which conducted a research about the top Dark Web markets in 2022. Do note that some of these markets could have been shutdown and are no longer operating. The Hydra Marketplace which was the largest marketplace
for online narcotics and had the most revenue back in 2022 was shutdown by a U.S.-German operation in April 2022 [5]. On the other hand, the OMG!OMG! marketplace seems to be taking up the mantel for Hydra MarketPlace.

**Terrorists**

Terrorists typically use the Dark Web to discuss and plan unethical activities. Terrorist organizations utilize the Dark Web to disseminate propaganda and spread their ideologies and instruction materials. They have websites which are used to recruit new members. Terrorists mainly aim to achieve two goals to be able to fulfill their objectives, namely establishing online presence to be able to spread propaganda without being detected by law enforcement. Attack plans and other terrorist activities are discussed by them on the Dark Web which maintains their anonymity. Moreover, terrorists use the Dark Web to be able to survive. Money is needed to be able to conduct attack operations and buy equipment, this is fulfilled through the donations terrorists get through their supporters on their websites on the Dark Web and the different services terrorists sell on the Dark Web such as human organs and stolen items and goods [21]. The terrorist attack in Paris that took place in November 2015 was done via weapons and explosives that were purchased by terrorist groups through the Dark Web [21].

Table 1 provides a brief summary of this section. Law and unlawful categories illustrate if the entity involved uses the Dark Web for lawful or unlawful purposes. Ordinary users for example, which only use the Dark Web for the benefits it provides are not breaking any laws. One could certainly argue that journalists, researchers, activists and whistle-blowers are not breaking any laws by using the Dark Web, however, using the Dark Web to disclose classified or confidential information might be illegal in some countries. Unauthorized publication of sensitive materials can violate laws. Journalists tend to publish false and damaging information that could harm organisations reputation without proper verification which could also lead to defamation lawsuits. The following activities carried out by journalists on the Dark Web can be counted as unlawful depending on the laws that might vary in different countries:

- Promoting illegal activities.
- Violating data protection and privacy laws.
- Publishing classified government information.
8 RQ3: Analysing the Impact

The Dark Web has exerted a significant influence on society, yielding both positive and negative ramifications. The following section aims to perform an analysis on the societal impact the Dark Web has done, focusing on three main questions:

- How does the cybersecurity risks and illegal activities associated with the Dark Web impact individuals and businesses? What are the social implications that arise due to the spread of extremist ideologies on the Dark Web?
- How does the presence of the Dark Web contribute to the proliferation of illegal activities and criminal networks?
- How does the presence of the Dark Web affect law enforcement agencies?

8.1 Impact on Individuals and Businesses

Since the Dark Web is associated with cybercrime such as data breaches and sale of stolen information, civilians in general may become victims of such activities which leads to financial losses and emotional distress. Data breaches that occur on the Dark Web expose people’s personal information which can lead to identity theft and financial fraud. In 2014, hackers stole data of 500 million users from Yahoo [25], including names, email addresses and passwords. UTorrent was also a victim of an attack were 400 thousand accounts were compromised [25]. In both cases, stolen data appeared on the Dark Web’s illicit marketplaces which resulted in both financial losses for the businesses involved and damaged their reputations. This aims to show how both companies and civilians can both be negatively impacted by the Dark Web threats.

The availability of hacking tools and different services on the Dark Web can put normal civilians life in danger. The fact that assassination and hacking services can be purchased on the Dark Web would impact individuals which may not be aware of the Dark Web’s role and might make them victims of such cybercrime.

While normal individuals may seek the Dark Web for its benefits, this does not guarantee that these individuals do not get exposed to the different marketplaces that offer drugs and counterfeit items and different websites which includes sensitive content. The exposure of such websites to these individuals might introduce them to criminal activities and creates potential
risks to their well-being.

The Dark Web is a good platform for extremist groups which aim to freely share and promote their ideologies. Propaganda, hate speech, and instructions for carrying out violent acts are just an example of what extremists tend to share and provide. This content contributes to the radicalization of individuals who might not have been exposed to such ideologies before. This posses a challenge to society as these ideologies which can fuel hatred and acts of violence can lead to a fractured community. Terrorists and extremist groups tend to use the Dark Web to recruit new members to their communities and introduce them to their ideologies. This can contribute to the growth of extremist movements which can lead to an increase in terrorism and violence.

8.2 Drug Dealers and Markets

The main advantage of Dark Web markets is the anonymity they provide for both their customers and their sellers. Being able to purchase drugs on these marketplaces requires no physical contact between the dealer and the buyer which reduces the inhibitions of some customers who might be reticent to interact personally with drug dealers. Moreover, individuals trying to purchase drugs do not need to risk going to dangerous places to buy them, since this can be easily done online. An analysis of goods and services listed on the Dark Web revealed that in 2017, 62 percent of offers on the Dark Web were drugs \[6\]. Data provided by The Global Drug Survey which is a sample of roughly 100,000 to 500,000 self selected people from more than 50 countries shows that individuals who use drugs and purchased them via a Dark Web market tripled from 4.7 per in January 2014 to 15 per cent to January 2020 \[6\].

Dark Web marketplaces are typically successful due to the fact that they overcome traditional market frictions of street drug dealing such as rip-offs and scams. Drugs being sold on the streets has lower quality and is oftenly mixed with other substances such as soda, sugar and starch which makes consumers not fully aware of the quality of the product. The Dark Web’s marketplaces overcome these issues by introducing a feedback system for every seller operating on the marketplace which mirrors feedback systems of illicit online markets such as eBay \[26\]. Moreover, the Dark Web platforms offer en escrow system which allows the system to hold the money paid by the user until the goods gets delivered. This solves scam issues which can happen between a seller and a buyer. Sellers operating on the Dark Web tend to be sellers who were previously selling drugs offline and decided to move to selling
8.2 Drug Dealers and Markets

online which lowers the risk of them being caught by law enforcement, lowers the chance of them being exposed to violence and increases their financial gains [26].

Figure 17: Monthly sales of drugs through different Dark Web markets. Source: [6].

The Dark Web’s marketplaces encounters huge risks for shutdowns. These shutdowns can happen for several different reasons:

- Voluntary shutdowns: The marketplace is voluntary shutdown by their administrators mainly because the market is unprofitable or because the administrators are fearing a seizure by law enforcement.

- Exit scams: The marketplace decides to shutdown to scam its users and keeps all the money that was in the escrow system.

- Hacked or raided: Sometimes marketplaces gets hacked by other users where hackers try to steal money and shutdown the market for certain reasons. On the other hand, the markets can get raided by law enforcement agencies and get seized.

Shutdowns of such markets can arguably have an impact on users who uses them. According to the Global Drug Survey 2018 [6], the shutdown of the market AlphaBay suggested that 15 percent of Dark Web users used such markets less frequently after the shutdowns, and 9 percent had stopped using the Dark Web for drug purchases [6].
Figure 18: Proportion of surveyed Internet users using drugs in the past year who purchased drugs over the Dark Web. Source: [6].

Figure 19: The impact that market closures have had on individuals. Source: [6].

The Dark Web is certainly known to have caused an impact on street crime
too. Individuals seeking to purchase drugs certainly use the market places offered by the Dark Web which leads to a decrease in street-level drug dealing as drug users always seek convenience and anonymity while purchasing drugs. Being able to purchase weapons and firearms via the Dark Web’s marketplaces can contribute to decrease street crimes involving weapons.

8.3 Law Enforcement

Law enforcement agencies tend to protect users and fight crime on the Dark Web. Since the Dark Web provides a platform for criminals to operate without disclosing their identities, this makes it difficult for law enforcement agencies to identify and track these individuals. Law enforcement agencies need specialized skills and tools to be able to investigate Dark Web operations since the Dark Web lacks centralized authority and people making transactions tend to use cryptocurrencies which makes it difficult to trace financial flows. A vast amount of resources and time is required so that law enforcement can continuously adapt and develop expertise in areas such as cybersecurity and digital forensics.
Figure 20: Dark Web’s impact on different entities.
8.4 Shutting down the Dark Web: Is it a Possibility?

Since the Dark Web is associated with many illegal activities and can potentially posses harm to society, shutting it down might be an option that would mitigate its impacts. However, due to various reasons, this objective remains unattainable. Firstly, the Dark Web itself is perfectly legal. As described in Section 6.2, the Dark Web just adds another layer of security and anonymity to browse the web which is not illegal. The Dark Web is certainly just like the Surface Web but uses mechanisms to hide users’ identities. Even-though these mechanisms can be utilized to perform illegal activities, this itself is not a compelling argument to shutdown the Dark Web considering that illegal activities are also prevalent on the Surface Web. What is legal and illegal varies from place to place which makes it a challenge to shutting down the Dark Web on a global scale. Almost anything in society can be used for illegal purposes, thus, criminals utilizing the Dark Web is just a negative consequence. Freedom of speech is facilitated by the Dark Web, thus banning the Dark Web solely based on illegal activities could violate individuals’ rights to express themselves securely and anonymously.

Secondly, shutting down the Dark Web is technologically impossible mainly because of its decentralized infrastructure. There is no central authority nor server that can be targeted to shutdown the entire Dark Web, hidden services operate and are distributed across multiple servers and locations which makes it nearly impossible to shut them down effectively.

On the other hand, law enforcement agencies tend to use different techniques to mitigate criminal activities on the Dark Web. These combating methods are discussed in the next section.

9 Mitigating the negative impact of the Dark Web

Combating crime on the Dark Web is a difficult task that requires a variety of methods and strategies that are being used by law enforcement agencies and cybersecurity firms. Prior to delving into the strategies employed to address criminal activities on the Dark Web, it is important to recognize the inherent advantages and disadvantages of taking action against crime on the Dark Web in any capacity.
9.1 Pros and Cons of Combating Crime on the Dark Web

Combating crime on the Dark Web aims to reduce criminal activities such as drug trafficking, weapon sales, and human trafficking. Websites that facilitate illegal activities like hiring a murderer or displaying pedophilic content pose a significant risk to communities, thus, shutting down these websites would in fact, increase public safety. Disrupting criminal operations and offenders would make it harder for criminals to operate anonymously.

Combating crime on the Dark Web would for instance enhance cybersecurity technologies. Since tools and methods should be developed to be able to fight cybercrime, this will contribute to strengthening the overall cybersecurity measures, making the Internet a safer place. Proving to the public that combating crime on the Dark Web is possible would make communities and organizations build trust and confidence in the justice system. This gives a perception for the public that engaging in illegal activities will result in repercussions, which may act as a hindrance for criminals.

The main disadvantage that arises when it comes to combating crime on the Dark Web is the fact that users who use the Dark Web for privacy reasons might get affected since some efforts to combat crime on the Dark Web might involve monitoring and data collection. Moreover, a lot of resources and time is needed to be able to develop and employ crime combating methods.

As law enforcement agencies become more adept at combating crime on the Dark Web, criminals are also going to adapt their strategies and techniques to evade detection. This might result in a cycle where law enforcement must continually evolve their methods to keep up with emerging criminal tactics.

9.2 Crime Combating Methods

Being able to provide advanced techniques that could disrupt illegal operations on the Dark Web is a difficult task that requires a vast amount of resources and time. This section aims to delve into the various methods employed by law enforcement agencies to combat criminal activities on the Dark Web and provides an overview of various tools that are being used and developed by law enforcement authorities.

9.2.1 Approaches Employed by Law Enforcement

Law enforcement actively monitors the Dark Web to be able to gather information on criminal activities. Monitoring is mainly used to track down illegal
marketplaces and identify individuals involved in illicit activities. The process of monitoring customer data involves analyzing individuals’ web data to identify any associations with unconventional domains \[23\]. This is done without invading user’s privacy as only the destinations of the Web requests need to be monitored. Sites that are available for the public such as Pastebin which is used to exchange information and address for new hidden services are oftenly monitored by Law Enforcement and constantly observed to keep track of new Dark Web domains \[23\].

A well known tactic used by Law Enforcement Agencies to apprehend criminals on the Dark Web is performing sting operations \[7\]. Sting operations are operations conducted to catch individuals engaging in illegal activities on the Dark Web. These operations are typically performed by firstly identifying a specific target of individuals engaging in illegal activities on the Dark Web such as drug trafficking or illegal market operations. Once a specific target has been identified, Law Enforcement agents create fake personas to establish credible online identities on the Dark Web. By creating fake profiles, police agents access Dark Web sites and pretends to be offenders or victims to lure offenders into committing a crime. They may for example pose as buyers, sellers, or intermediaries and interact with suspects to gather evidence and build a solid case against the individuals involved. Law Enforcement agencies should of course ensure that the operation is conducted within the boundaries of law since these operations might sometimes provoke the commission of a crime by someone who would not otherwise have done so. Evidence gathered should also be admissible in court and the rights of the individuals involved should always be respected.

Another well known tactic that is oftenly used by Law Enforcement Agencies is called the honeypot trap. Honeypot traps are cybersecurity techniques that involve setting up a system or service with the purpose of attracting and capturing criminals and malicious actors \[27\]. Law Enforcement Agencies build up trap sites on the Dark Web which are related to illegal activity but setup to attract individuals seeking to buy or sell illegal stuff. A good example that illustrates a honey trap tactic that was used by Law Enforcement is the Dark Web market named Playpen that was discovered by the FBI. The operation performed was called Operation Pacifier. Playpen was a Dark Web market which distributed child pornography, however due to a flaw in the website \[7\], the FBI was able to breach the website and hijacked it. Instead of directly taking actions and arresting the website host, the FBI chose to create an exploit which would uncover visitors IP addresses. Figure 21 displays the results that were produced by this operation.
9.2 Crime Combating Methods

9.2.2 Tools

Law enforcement authorities use different tools and techniques to be able to track and monitor activities on the Dark Web. These tools can be used to identify threats and stop them. Even-though Tor provides a secure protocol to protect the identities of its users, its not invulnerable to attacks. In Section 6.1.2, a vulnerability of the onion routing technique was presented where a hacker sniffing the channel between the Entry Node and the channel between the Exit Node and the server would deanonymize Tor users. This specific attack is called an end-to-end confirmation attack which attempts to correlate the traffic entering and exiting the Tor network to be able to deanonymize users [21]. More specifically, end-to-end confirmation attacks performed by Law Enforcement agencies tend to take control of both the entry and the exit relays of the Tor network. By doing so, they are able to monitor the traffic entering and leaving the network which allows them to correlate the timing of packets. This would provide Law Enforcement agencies with the IP address of the client and the IP address of the server they are trying to access which would deanonymize the user.

Law enforcement agencies commonly employ a type of attack targeting the hidden service directory within the Tor network [21]. Hidden service directories in the Tor network are used to retrieve a list of the introduction points used by the server. Law enforcement agencies tend to compromise the directory which helps them monitor the activities of the hidden service.

Open Source Intelligence (OSINT) tools which are tools that allow the collection
of information and help in collecting and analyzing data to obtain insights, investigate individuals or organizations are usually used by Law Enforcement agencies to monitor and collect information from various sources on the Dark Web such as forums and marketplaces. These tools are able to crawl and index these sites, extracting relevant data related to criminal activities. They even assist in analyzing collected data from the Dark Web. OSINT tools uses techniques such as natural language processing and sentiment analysis to extract meaningful insights and patterns. They can even be used to track cryptocurrency transactions.

10 Discussion

The thesis provides valuable insights into the Dark Web, its actors and societal impact, however, by nature, the Dark Web is anonymous which made it hard to collect data about its actors and cover all of the activities that takes place on it, thus, this work might not cover all the aspects of the Dark Web comprehensively. One significant limitation of this thesis is the absence of available data that could directly illustrate the precise impact of the Dark Web. As an example, no direct research was conducted focusing on the relationship between the Dark Web and street crime. The decentralized nature of the Dark Web made it challenging to obtain data regarding its societal impacts. This limitation restricted the ability to provide information about the Dark Web’s influence on various aspects of society such as crime rates and economic factors.

Secondly, a challenge encountered during the research was the difficulty in finding and engaging with experts in the field. Additionally, some experts who were approached for interviews were hesitant or unwilling to disclose information. Since the Dark Web is notorious for facilitating illegal activities, experts, particularly those working in law enforcement may be bound by legal restrictions that prevent them from openly discussing specific details or sharing classified information. Fear for personal safety and security may also be one of their concerns. Moreover, since the Dark Web operates within a highly specialized and technically complex environments, some of the experts which denied to disclose information may not have an extensive knowledge of the Dark Web making them hesitant to discuss a topic they feel ill-equipped to address accurately.

Despite the limitations the study reveals that the encryption and anonymity provided by the Dark Web has yielded both positive and negative impacts to society. Utilizing the TorBot crawling tool made it possible to gather a wide
range of data from the Dark Web, which helped offer insights into various activities and communities that exist in this hidden environment. Accessing the Dark Web itself presented an opportunity to observe and analyze the operations and behaviors of its actors firsthand and made it possible to determine different actors on the Dark Web. These actors were divided into two different groups, lawful and unlawful based on whether the activities carried by these actors violate the law of the countries they are living in.

The distinction between lawful and unlawful actors on the Dark Web aimed to challenge the common perception that the entire Dark Web is a hub of criminal activity. By recognizing the presence of lawful actors, we acknowledge that not all activities conducted on the Dark Web are illegal or malicious. This recognition prompted us to consider the potential benefits and positive impacts that can arise from these lawful activities. Some actors may belong to both of these groups since what classifies as lawful and unlawful differs in countries.

When it comes to unlawful individuals, the anonymity offered by the Dark Web creates an advantageous environment for the illicit trade of prohibited items and services. These include drugs, weapons, stolen data, and various illegal services. Criminal networks which develop on the Dark Web pose a challenge to Law Enforcement agencies where different tactics and methods should be used to combat these activities. The proliferation of illegal activities on the Dark Web leads to harmful consequences for society such as drug addiction, violence and financial losses. On the other hand, the levels of privacy and anonymity provided by the Dark Web helps protect users who wish to hide their identities and enable freedom of speech.

When conducting the investigation on the Dark Web, the author carefully considered the ethics of all actions taken. In particular, the investigation did not contribute or perpetuate criminal behaviour.

11 Conclusion

The Dark Web provides a platform with strengthened anonymity offering both opportunities and challenges to society. Lawful actors utilize this space for legitimate purposes such as anonymous communication, research and privacy protection. On the other hand it is undeniable that the Dark Web also serves as a breeding ground for unlawful activities. Illicit actors take advantage of the anonymity provided, engaging in the sale of illegal goods
and services, including drugs, weapons, stolen data, and other illicit offerings. The sale of drugs, weapons, and stolen data not only fuels criminal activities but also contributes to social harm and addiction. The societal impact of the Dark Web’s unlawful actors extends beyond the virtual realm, influencing communities and individuals both online and offline.

By acknowledging the presence of both positive and negative actors, we move beyond the stereotype that the Dark Web is solely a breeding ground for illegal activities. The categorization of these actors was made based of the adherence to existing legal frameworks. Those actors who operate within the boundaries of the law were categorized as lawful while those violating established laws fall under the classification of unlawful actors. This categorization allows for a better understanding of the motivations, behaviors, and implications associated with different actors operating within this field.

Addressing the societal impact of the Dark Web requires a multi-faceted approach. Collaboration between law enforcement, technology companies, policymakers, and civil society organizations is important in developing effective strategies to tackle the illegal activities conducted on the Dark Web. Balancing the need for privacy and security with the prevention of criminal activities is a complex challenge that demands continuous adaptation.

References


REFERENCES


