Chapter 02: Footprinting and Reconnaissance

Introduction

Footprinting serves as the initial phase in assessing the security posture of a target organization's IT infrastructure. By engaging in footprinting and reconnaissance, one can collect a wealth of information about a network or computer system, as well as any devices connected to that network. Essentially, footprinting creates a detailed security profile blueprint for an organization and should be performed systematically.

This chapter begins with an overview of footprinting concepts and offers insights into the methodology associated with footprinting. The chapter concludes with a summary of footprinting tools and countermeasures.

At the end of this chapter, you will be able to:

- Describe footprinting concepts
- Perform footprinting through search engines and using advanced Google hacking techniques
- Perform footprinting through web services and social networking sites
- Perform website footprinting and email footprinting
- Perform Whois, DNS, and network footprinting
- Perform footprinting through social engineering
- Perform footprinting tasks using advanced tools and AI
- Apply footprinting countermeasures

Footprinting Concepts

This phase serves as a foundational step for the attacker, who must collect extensive information to identify potential entry points into the target network.

Reconnaissance

Reconnaissance, often referred to as footprinting, is the initial stage in which an attacker aims to collect extensive information about a target before executing an attack. A critical component of footprinting is assessing the risk level linked to the organization's publicly accessible information. Footprinting, the first stage in ethical hacking, involves gathering details about a target network and its surrounding environment. Through the footprinting process, various opportunities to penetrate and evaluate the target organization's network can be identified.

Once you systematically complete the footprinting phase, you will have a comprehensive outline of the target organization's security profile. In this context, the "blueprint" indicates the distinct system profile of the organization obtained through footprinting.

There is no universal approach to footprinting, as information can be discovered through multiple methods. Nevertheless, this activity is vital since it enables the collection of all essential information regarding the target organization before progressing to the hacking phase. For this reason, it is important to conduct footprinting in a well-organized manner. The data gathered during this step assists in revealing vulnerabilities present in the target network and in recognizing various ways to exploit those vulnerabilities.

Types of Footprinting/Reconnaissance

Footprinting can be classified into passive footprinting and active footprinting.

Passive Footprinting

Passive footprinting refers to the process of collecting data about a target without directly engaging with it. This method is especially valuable when the aim is to keep the information-gathering activities undetected by the target. Conducting passive footprinting poses technical challenges, as no active traffic is directed towards the target organization from a host or anonymous services over the Internet. Instead, we rely on gathering archived and stored information about the target through search engines, social media platforms, and similar sources.

It involves:

- Open-source Intelligence (OSINT) gathering
- Proprietary databases and paid services
- Sharing intelligence with partner organizations or industry groups

Active Footprinting

Active footprinting entails collecting information about the target through direct engagement. In active footprinting, the target may become aware of the information-gathering activities since you are openly interacting with the target network. Active footprinting demands more planning compared to passive footprinting, as it can leave traces that might notify the target organization.

It involves:

- DNS interrogation
- Social engineering
- Network/port scanning
- User and service enumeration

Information Obtained in Footprinting

The primary goals of footprinting involve gathering information about the target's network, systems, and organization. By performing footprinting at various network levels, you can obtain details like network ranges, individual IP addresses, and information about employees, among others. This type of information can assist attackers in accessing confidential data or executing different types of attacks on the targeted network.

Organization Information

Details regarding an organization can be found on its website. Additionally, you can check the target's domain name in the Whois database to gather useful information.

The information collected includes:

- Employee details (employee names, contact addresses, designations, and work experience)
- Addresses and mobile/telephone numbers
- Branch and location details
- O Partners of the organization
- Web links to other company-related sites
- Background of the organization

- Web technologies
- News articles, press releases, and related documents
- Legal documents related to the organization
- Patents and trademarks related to the organization

Attackers can gain access to organizational data and utilize that information to pinpoint essential personnel, enabling them to carry out social engineering attacks aimed at obtaining sensitive information about the organization.

Network Information

You can collect network information by performing Whois database analysis, trace routing, and so on. The information collected includes:

- Domain and sub-domains
- Network blocks
- Network topology, trusted routers, and firewalls
- IP addresses of the reachable systems
- Whois records
- DNS records and related information

System Information

You can gather system information by performing network footprinting, DNS footprinting, website footprinting, email footprinting, and so on. The information collected includes:

- Web server OS
- Location of web servers
- Publicly available email addresses
- Usernames, passwords, and so on

Objectives of Footprinting

To formulate a hacking strategy, attackers need to collect information regarding the network of the target organization. They utilize this information to pinpoint the easiest methods for breaching the organization's security defenses. As previously mentioned, the footprinting approach simplifies the process of gathering information regarding the target organization, which is crucial in the hacking methodology.

Footprinting outlines the security framework, including the locations of firewalls, proxies, and various other security measures. Attackers can evaluate the footprinting data to uncover weaknesses in the target organization’s security setup and subsequently devise a hacking plan.

By employing a mix of tools and methods, attackers can take an unfamiliar entity (for instance, XYZ Organization) and narrow it down to a specific array of domain names, network blocks, and individual IP addresses of systems directly connected to the Internet, along with additional information related to its security framework.

A comprehensive footprint delivers extensive information about the target organization, enabling the attacker to spot vulnerabilities in the target systems and choose suitable exploits. Attackers can create their own information repository concerning the target organization's security flaws. This repository can aid in determining the most vulnerable aspect of the organization's security defenses.

Footprinting Threats

The following are various threats that can arise from footprinting:

Social Engineering: Hackers can, directly and indirectly, acquire information through persuasion and other tactics without engaging in any intrusion methods. They extract vital details from unsuspecting employees who are unaware of the hackers' true intentions.

System and Network Attacks: Footprinting allows an attacker to execute system and network assaults. Consequently, attackers can collect data about the target organization's system setup, the operating system in use, and more. With this data, they can identify vulnerabilities in the target's system and exploit them. This enables them to gain control over a target system or the entire network.

Information Leakage: The risk of information leakage is a significant concern for any organization. If attackers access sensitive information, they can either execute an attack based on that information or exploit it for financial gain.

Privacy Loss: Hackers can use footprinting to infiltrate an organization's systems and networks and even elevate their privileges to administrative levels, leading to a loss of privacy for both the organization and its individual employees.

Corporate Espionage: Corporate espionage poses a serious threat to organizations, as competitors frequently attempt to obtain sensitive information through footprinting. This method allows competitors to launch similar products, adjust pricing, and generally weaken the market standing of the targeted organization.

Business Loss: Footprinting can significantly impact organizations, particularly online businesses, e-commerce platforms, and those in the banking and finance sectors. Each year, billions of dollars are lost due to malicious attacks orchestrated by hackers.

Footprinting Methodology

The footprinting methodology refers to a systematic approach to gathering information about a target organization from various sources. This process includes collecting details about the target organization, such as URLs, physical locations, establishment information, employee count, specific domain name ranges, contact details, and other pertinent information. Attackers obtain this data from publicly available resources like search engines, social media platforms, Whois databases, and other similar sources. The diagram below depicts the common techniques employed to gather information about the target organization from diverse sources.

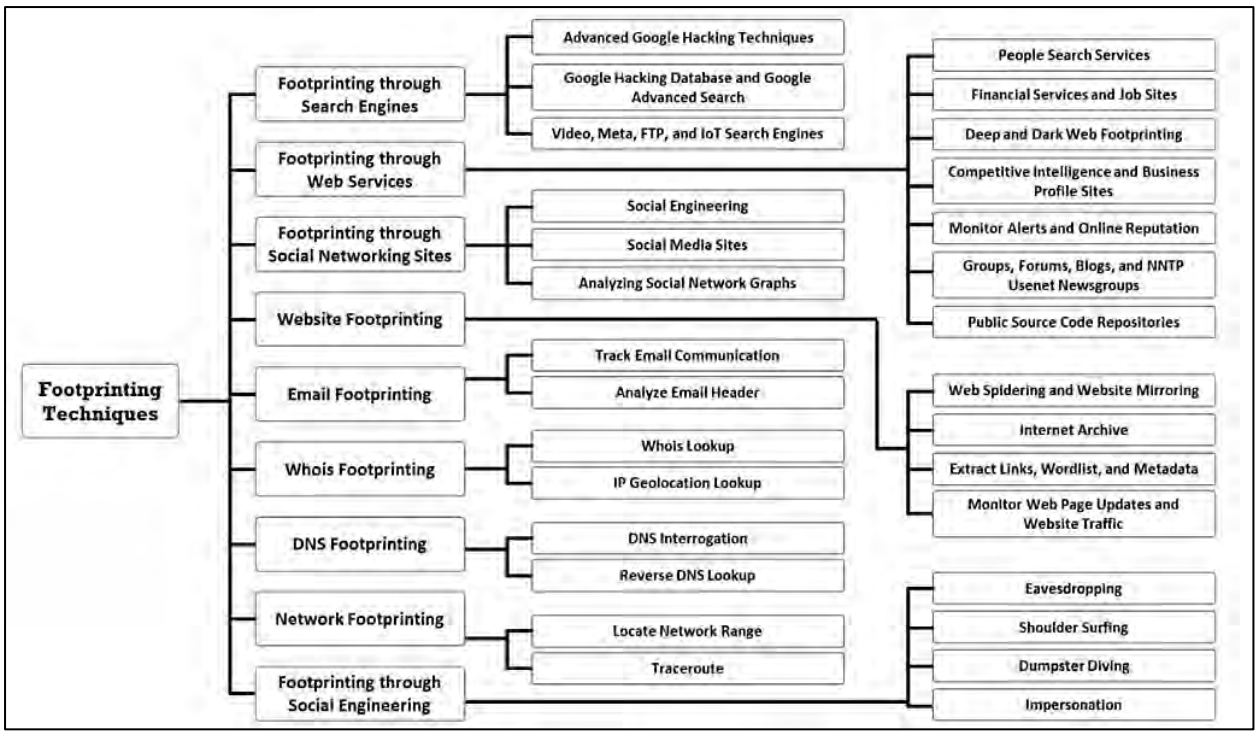

Figure 2-01: Footprinting Techniques

Footprinting through Search Engines

Search engines serve as the primary sources for crucial information regarding a target organization. They are essential for extracting important details about a target from the web. Search engines utilize automated programs, known as crawlers, to constantly scan active websites and accumulate the findings into their search engine index, which is then stored in a vast database. When a user performs a query, the search engine index provides a list of Search Engine Results Pages (SERPs). These results encompass web pages, videos, images, and various file types, all ranked and presented based on their relevance. Numerous search engines can gather information about a target organization, including technology platforms, employee information, login pages, intranet portals, contact details, and more. This information helps an attacker in executing social engineering and various sophisticated system attacks.

A Google search might uncover posts in forums made by security personnel, revealing the brands of the firewalls or antivirus software employed by the target. This knowledge assists the attacker in pinpointing weaknesses in such security measures.

For instance, take an organization like Microsoft. If you enter Microsoft in a search engine's Search box and hit Enter, you will see results that contain details about Microsoft. Reviewing these results can often reveal essential information such as the physical address, contact information, available services, employee count, and more, which could serve as a significant resource for hacking.

Figure 2-02: Major Search Engines

Examples of major search engines include Google, Bing, Yahoo, Ask, Aol, Baidu, WolframAlpha, and DuckDuckGo.

Attackers can utilize advanced search operators provided by search engines to craft intricate queries that help them locate, filter, and categorize specific data about their target. Additionally, search engines can be employed to discover other publicly available information sources. For instance, entering "top job portals" can lead you to significant job portals that offer essential details pertaining to the target organization.

As an ethical hacker, if you come across any deleted pages or information about your company in Search Engine Results Pages (SERPs) or the search engine's cache, you have the option to request the search engine to eliminate those pages or information from its indexed cache.

Footprinting using Advanced Google Hacking Techniques

Google hacking involves utilizing sophisticated search operators to create intricate queries aimed at uncovering sensitive or concealed information. Attackers subsequently use the information obtained to identify vulnerable targets. Footprinting with advanced Google hacking methods entails finding specific text strings within search results using specialized operators on the Google search engine.

Advanced Google hacking is the skill of crafting detailed search queries. These queries can pull out important data regarding a target company from Google's search results. Through Google hacking, attackers seek out websites that may be susceptible to exploitation. They can leverage the Google Hacking Database (GHDB), which contains a collection of queries, to discover sensitive information. Google operators assist in pinpointing the desired text while filtering out irrelevant results. By employing advanced Google operators, attackers can identify specific text strings, such as particular versions of vulnerable web applications. When a search query is input without advanced operators, Google examines the search terms throughout any part of the webpage, which includes the title, text, URL, and digital files. To refine a search, Google provides advanced search operators, which help in narrowing down the query to yield the most pertinent and precise results.

The syntax to use an advanced search operator is as follows:

operator: **search_term** **Note:** Do not enter any spaces between the operator and the query.

Some well-known Google advanced search operators are:

site: This operator limits search results to a specific site or domain.

For instance, the [games site: www.certifiedhacker.com] search yields information about games from the certifiedhacker site.

allinurl: This operator confines results to pages that include all the specified query terms in the URL.

For example, the [allinurl: google career] search returns only pages with both "google" and "career" in the URL.

inurl: This operator narrows results to pages that contain the specified word in the URL.

For instance, the [inurl: copy site:www.google.com] search will display Google pages whose URLs include the term "copy."

allintitle: This operator restricts results to pages that have all query terms specified in the title.

For example, the [allintitle: detect malware] search shows only pages with "detect" and "malware" in the title.

intitle: This operator limits results to pages that feature the specified term in the title.

For example, the [malware detection intitle:help] search provides pages that include the word "help" in the title, as well as "malware" and "detection" anywhere else on the page.

inanchor: This operator filters results to pages that contain the query terms found in the anchor text of links to the page.

For instance, the [Anti-virus inanchor: Norton] search retrieves pages with anchor text containing "Norton" that also mentions "Anti-virus."

allinanchor: This operator restricts results to pages that have all specified query terms in the anchor text on links to them.

For example, the [allinanchor: best cloud service provider] search yields pages with anchor text that includes "best," "cloud," "service," and "provider."

cache: This operator shows Google's cached version of a webpage instead of the live version.

For instance, [cache:www.eff.org] will display the cached version of the Electronic Frontier Foundation homepage.

link: This operator searches for sites or pages that include links to a specific website or page.

For example, [link:www.googleguide.com] identifies pages that link to Google Guide's home page.

Note: As per Google's documentation, "you cannot combine a link: search with a regular keyword search."

Additionally, it is important to note that when combining a link with another advanced operator, Google may not display all matching pages.

related: This operator showcases websites that are similar or related to a specified URL.

For example, [related:www.microsoft.com] generates search engine results for websites akin to microsoft.com.

info: This operator seeks information about a specified webpage.

For instance, [info:gothotel.com] reveals details about the GotHotel.com homepage, a national hotel directory.

location: This operator helps find information pertaining to a specific location.

For example, [location: 4 seasons restaurant] will provide results relevant to "4 seasons restaurant."

filetype: This operator facilitates searching for results based on the file extension.

For instance, [jasmine:jpg] will return jpg files related to jasmine.

What can a Hacker do with Google Hacking

An attacker can formulate intricate search-engine queries to sift through extensive search results to find information pertinent to computer security. By employing Google operators, the

attacker can identify specific strings of text within the search results. Consequently, the attacker can not only uncover websites and web servers that may be susceptible to exploitation but also discover private and sensitive information regarding the target. After pinpointing a vulnerable site, attackers may seek to execute various types of attacks, including buffer overflow and SQL injection, which undermine information security.

Instances of sensitive information that an attacker can retrieve from public servers using Google Hacking Database (GHDB) queries encompass:

- Error messages that reveal sensitive data
- Documents containing passwords
- Confidential directories
- Pages that contain host logon portals
- Pages with network or vulnerability information, such as IDS, firewall logs, and configurations
- Advisories and details about server vulnerabilities
- Software version details
- Source code for web applications
- Unprotected connected IoT devices, along with their control interfaces
- Hidden web pages, such as intranet and VPN resources

Example:

Employ Google Advanced Operator syntax:

```
[intitle : intranet inurl : intranet +intext : "human resources"]
```

Use the above syntax to locate sensitive information regarding a targeted organization and its personnel. Attackers utilize the gathered data to carry out social engineering attacks.

Figure 2-03 illustrates a Google search results page showcasing the outcomes of the above mentioned query.

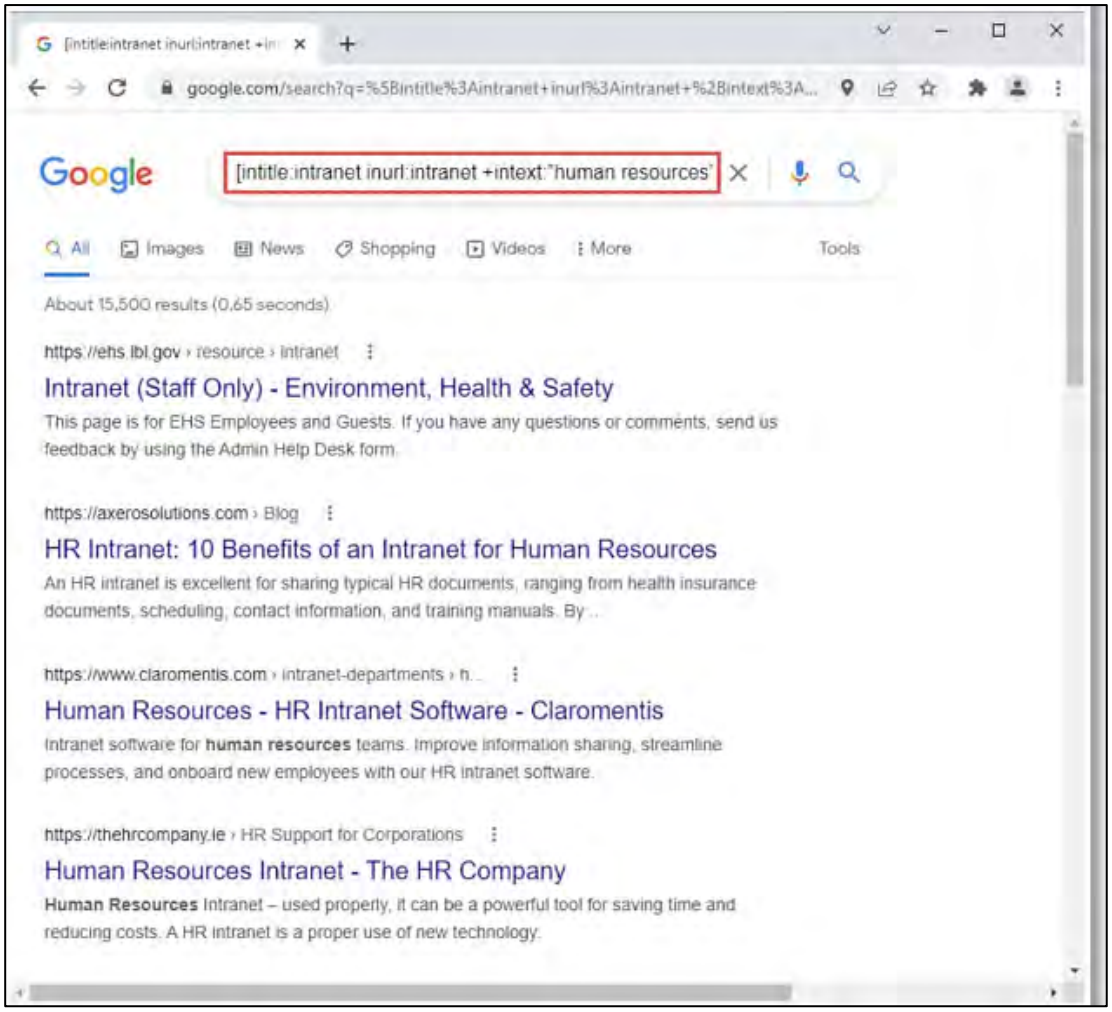

Figure 2-03: Search Engine Results for given Google Advance Operator Syntax

Footprinting using Advanced Google Hacking Techniques with AI

Attackers can utilize AI-driven technologies to improve and automate their reconnaissance efforts. With the support of AI, they can easily employ sophisticated Google hacking methods to gather important information about their intended targets.

For example, an attacker may leverage AI-enabled tools like ShellGPT for this purpose by inputting a suitable prompt such as:

"Use the filetype search operator to find PDF files on the target site eccouncil.org and save the output in the recon1.txt file."

The following shell command is crafted to perform advanced Google hacking. It utilizes the "filetype" operator to specifically search for PDF files within the eccouncil.org domain. The command saves the results it retrieves to a file called "recon1.txt":

```
lynx --dump "http://www.google.com/search?q=site:eccouncil.org+filetype: pdf" | grep "http" | cut -d "=" -f2 | grep -o "http[^&]*" > recon1.txt
```

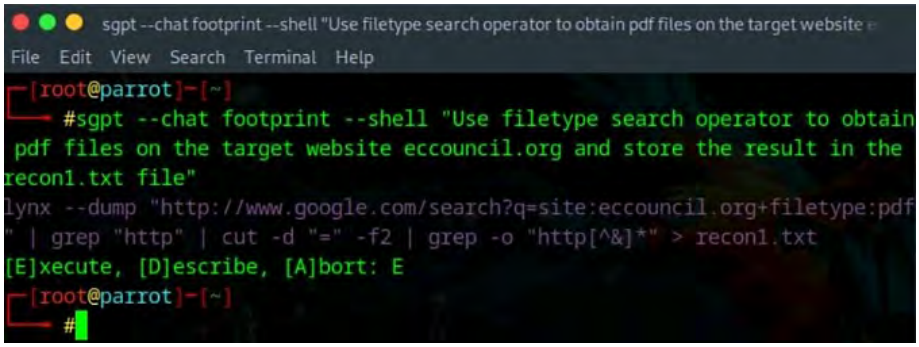

Figure 2-04: Prompt for Advanced Google Hacking with AI

Table 2-01 describes each option used in the command.

Command Breakdown	Description
lynx --dump "http://www.google.com/search?q=site: eccouncil.org+filetype: pdf"	Initiates the Lynx web browser in dump mode to access Google's search results for PDF files within the eccouncil.org domain
| grep "http"	Filters out lines containing the string "http" from the Lynx output
| cut -d "=" -f2	Splits each line using the "=" delimiter and selects the second field
| grep -o "http[^&]*"	Searches for patterns starting with "http" followed by any characters except "&"
> recon1.txt	Redirects the final output to a file named "recon1.txt" for storage

Table 2-01: Command Description to Perform Advanced Google Hacking

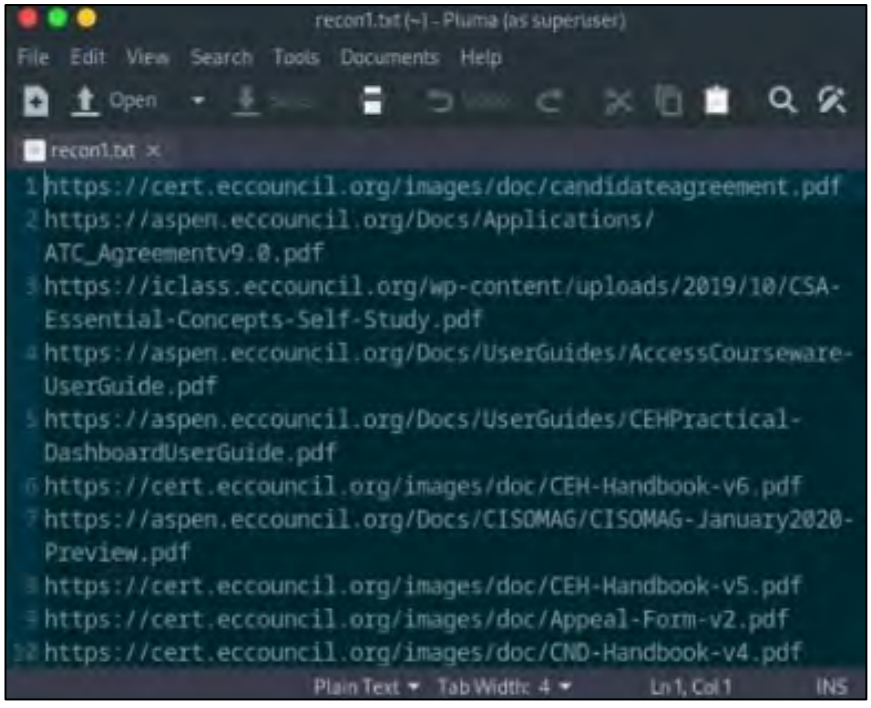

Figure 2-05: Output for Advanced Google Hacking with AI

Google Hacking Database

The Google Hacking Database (GHDB) serves as a key resource for exploring the continually expanding capabilities of the Google search engine. Within the GHDB, users can discover search keywords for files that include usernames, exposed servers, and even documents with passwords. The Exploit Database is an archive of public exploits and related vulnerable software that complies with Common Vulnerabilities and Exposures (CVE), created for penetration testers and researchers focused on vulnerabilities. By utilizing GHDB dorks, attackers can swiftly pinpoint all publicly accessible exploits and vulnerabilities within the target organization's IT framework. Attackers leverage Google dorks alongside advanced search operators to gather sensitive information about their target, including vulnerable servers, error messages, confidential files, login pages, and websites.

Google Hacking Database Categories

- Network or Vulnerability Data
- Vulnerable Files
- Pages Containing Login Portals
- Vulnerable Servers
- Various Online Devices
- Error Messages
- Advisories and Vulnerabilities
- Footholds
- Files Containing Juicy Info
- Files Containing Usernames
- Files Containing Passwords
- Sensitive Directories
- Sensitive Online Shopping Info
- Web Server Detection

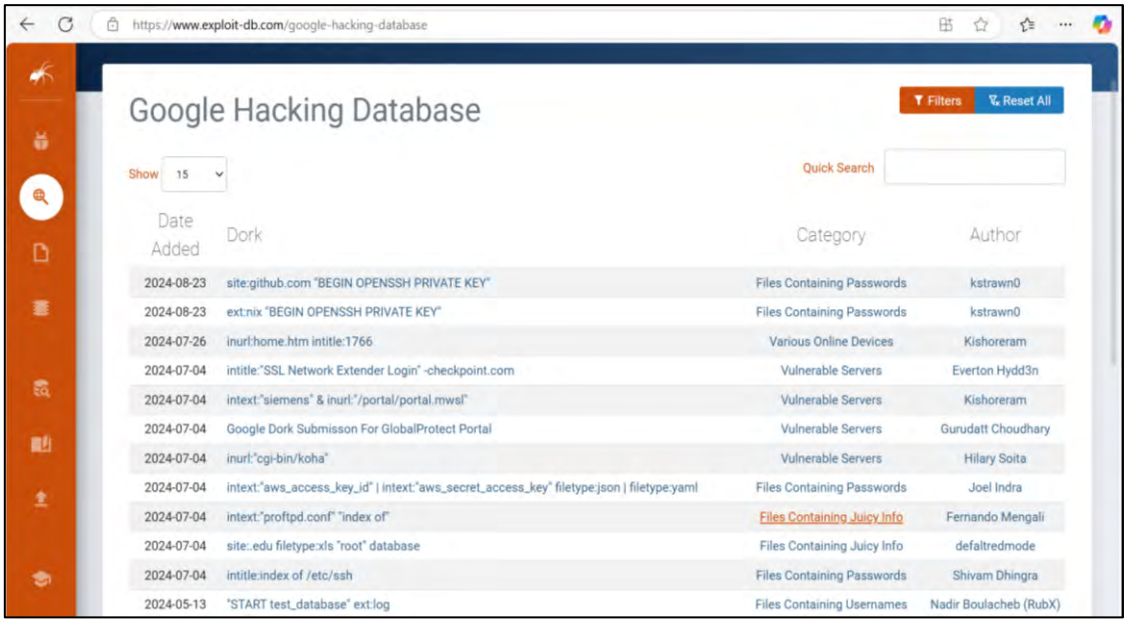

Figure 2-06: Google Hacking Database Screenshot

Attackers can utilize the GHDB in multiple ways to discover and exploit vulnerabilities:

- **Reconnaissance:** Attackers employ GHDB queries to collect information about potential targets, which may include exposed files, directories, and devices that can be taken advantage of.
- **Exploiting Misconfigurations:** By detecting sensitive information revealed through improperly configured web servers or services, attackers can take advantage of these misconfigurations to gain unauthorized access.
- **Finding Vulnerable Systems:** Through GHDB, attackers can identify systems that are operating outdated or vulnerable software versions, which provides a basis for further exploitation.
- **Credential Harvesting:** The sensitive data acquired from GHDB queries can encompass usernames and passwords, which attackers may use for credential stuffing or brute force attacks.
- **Identifying Open Ports and Services**: Certain GHDB queries can uncover open ports and services on a network, offering attackers a guide to potential entry points.

Additionally, attackers can utilize SearchSploit, a command-line search utility for Exploit-DB that enables them to take a copy of the Exploit database for offline use. It allows attackers to execute thorough offline searches through their locally checked-out version of the repository. This feature is especially beneficial for conducting security assessments in isolated or air-gapped networks without access to the Internet.

VPN Footprinting through Google Hacking Database

Google hacking operators, commonly known as Google dorks, can be utilized to map out Virtual Private Networks (VPNs). They can reveal details such as pages featuring login interfaces and directories containing keys for VPN servers.

Table 2-02 summarizes various Google hacking operators or Google dorks that are employed to gather specific information for VPN footprinting.

Google Dork	Description
inurl:"/sslvpn_logon.shtml" intitle:"User Authentication" "WatchGuard Technologies"	Finds pages containing login portals
inurl:/sslvpn/Login/Login site:vpn.*.*/ intitle:"login"	Finds VPN login portals
site:vpn.*.*/ intext:"login"intitle:"login"	Retrieves various VPN login pages
inurl:/weblogin intitle:"USG20-VPN" OR intitle:"USG20W-VPN"	Finds hosts with the Zyxel hardcoded password vulnerability
intext:"Please Login SSL VPN inurl:remote/login intext:FortiClient"	Finds Fortinet VPN login pages
intitle:"index of" /etc/openvpn/	Retrieves juicy information and sensitive directories
"----BEGIN OpenVPN Static key V1----" ext:key	Finds OpenVPN static keys
intitle:"index of" "vpn-config.*"	Retrieves juicy information about the vpn-config file
Index of / *.ovpn	Finds OpenVPN configuration files, some certificates, and keys
inurl:"/vpn/tmindex.html" vpn	Finds Netscaler and Citrix Gateway VPN login portals
intitle:"SSL VPN Service" + intext:"Your system administrator provided the following information to help understand and remedy the security conditions:"	Finds Cisco Adaptive Security Appliance (ASA) login web pages

Table 2-02: Google Search Queries for VPN Footprinting

VPN Footprinting through Google Hacking Database with AI

Attackers can utilize AI-driven technologies to improve and automate their footprinting efforts. By using AI, these individuals can easily conduct VPN footprinting and gather important information.

For example, an attacker could utilize ChatGPT to accomplish this by inputting a suitable prompt like:

Use inurl search operator to obtain the Fortinet VPN login pages

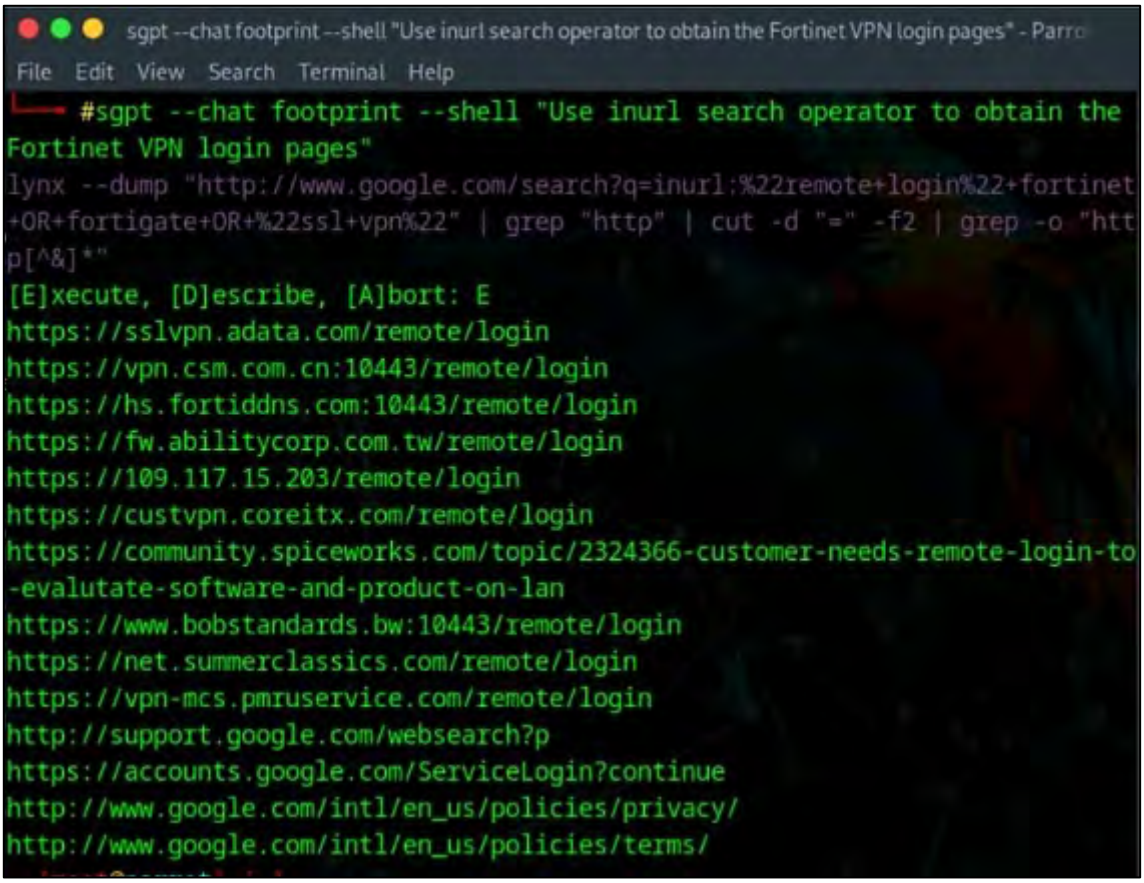

Figure 2-07: inurl Search Operator to obtain the Fortinet VPN Login Pages

The following command aims to locate Fortinet VPN login pages by utilizing Lynx, a web browser that operates in text mode, combined with Google search queries.

```
lynx -dump
“http://www.google.com/search?q=inurl:%22remote+login%22+fortinet+OR+fortigate
+OR+%22ssl+vpn%22 | grep "http | cut -d "" -f2| grep -o "http[^&]*"
```

It retrieves Google search results based on the provided query, extracts URLs from those results, and subsequently filters and formats them for display. The main objective is to locate Fortinet VPN login pages among the search results.

Table 2-03 describes each option used in the command.

Command Breakdown	Description
lynx –dump "http://www.google.com/search?q=inurl:%22remote+login%22+fortinet+OR+fortigate+OR+%22ssl+vpn%22"	Uses Lynx to search for web pages with "remote login" in the URL and "Fortinet," "Fortigate," or "ssl vpn" in the content.
| grep "http"	Pipes the output of the Lynx command to the `grep` command, which filters out lines containing the string "http".
| cut -d "=" -f2	Pipes the output of the previous command to the `cut` command, which splits each line using the "=" delimiter and selects the second field.
| grep -o "httP[^&]*"	Sends the output from the previous command to another 'grep' command, which looks for patterns that begin with "http" and are followed by any characters that are not "&".

Table 2-03: lynx Command Description with Google Search Queries

Footprinting through the SHODAN Search Engine

Shodan is a search engine that allows attackers to carry out footprinting at multiple levels. It is utilized to identify devices and networks that have vulnerabilities. Conducting a search in Shodan for VoIP and VPN footprinting can yield a variety of results, which can assist in collecting information related to VPNs and VoIP.

Figure 2-08 displays some of the search results for VPN and VoIP footprinting acquired via Shodan:

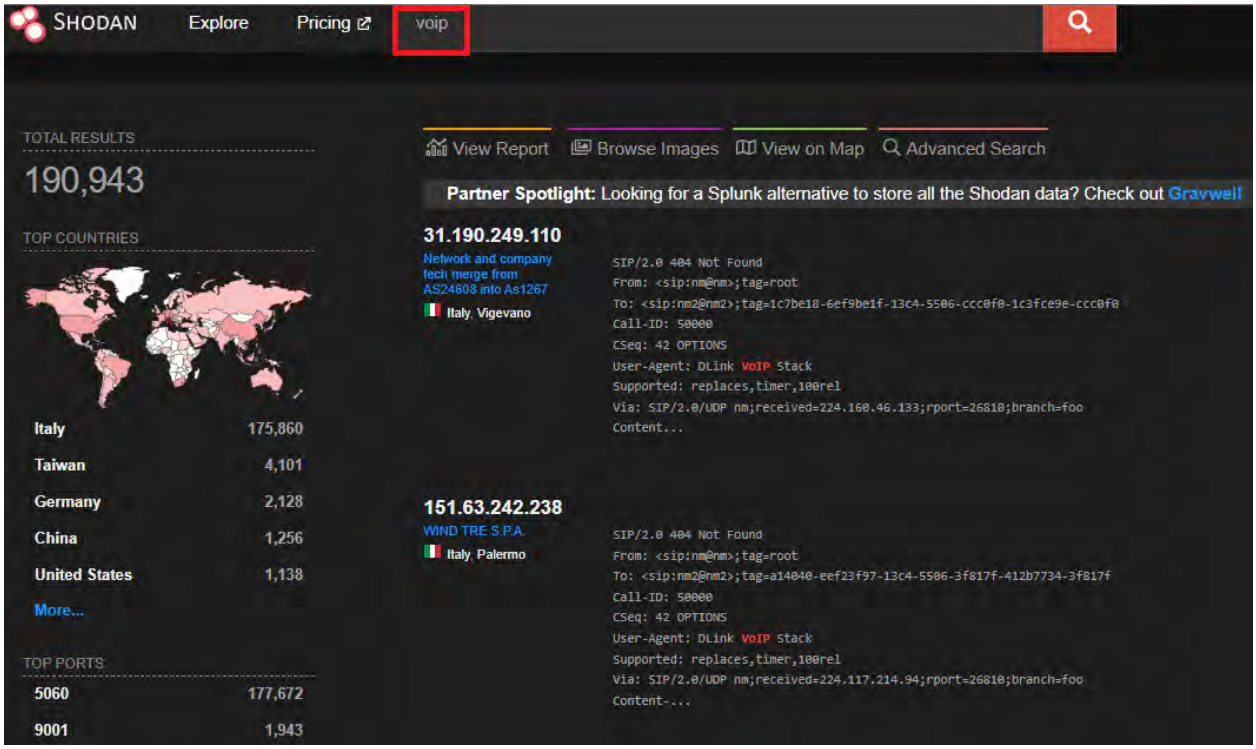

Figure 2-08: SHODAN Engine showing VoIP Results

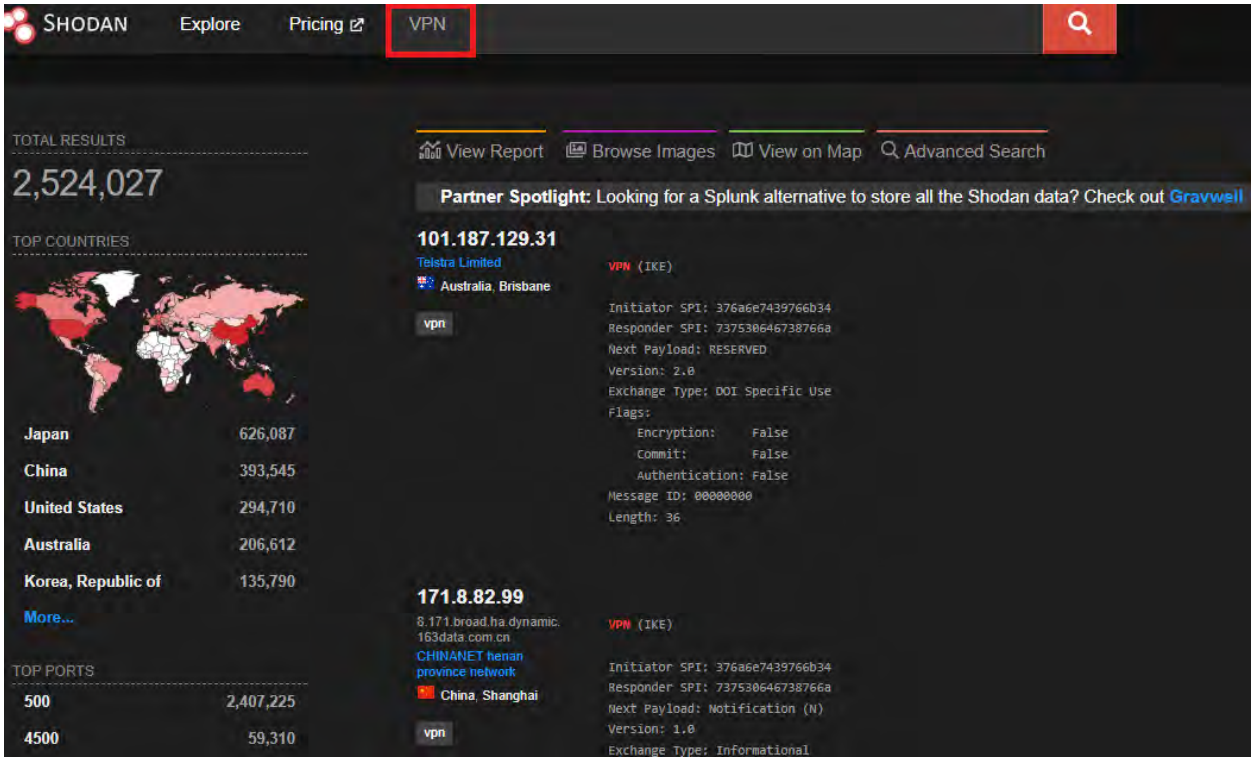

Figure 2-09: SHODAN Engine showing VPN Results

Other Techniques for Footprinting through Search Engines

Some other techniques of footprinting using Search engines include:

Gathering Information Using Google Advanced Search, Advanced Image Search, and Reverse Image Search

An attacker may not be able to easily gather information from a site rich in data using just a standard search box. A complex search involves multiple related criteria.

Google's Advanced Search feature enables an attacker to conduct detailed web searches. With Google Advanced Search and Advanced Image Search, one can navigate the web more accurately and precisely. These search features can achieve the same level of precision as the advanced operators without requiring you to remember or type the operators. By utilizing Google's Advanced Search option, you can identify sites that may link back to the targeted organization's website. This assists in obtaining information about the target's partners, vendors, clients, and other affiliations. Likewise, Google Advanced Image Search can be employed to find images relevant to the target, such as photos of its location, employees, and more.

To execute an advanced search in Google, click on Settings, located at the bottom-right corner of the Google homepage, then select Advanced Search from the menu or directly enter **https://www.google.com/advanced_search** into the address bar. The Advanced Search function allows you to set various criteria that your search must fulfill, as this function expands upon the search box by providing additional search options. To do this, select a field, then input the desired search string into the text box of that field and click on the Advanced Search button. By default, the different values are combined using "and" (indicating that all must apply), except for sets, blocks, and formats, which are linked with "or" (indicating that any one of them may apply).

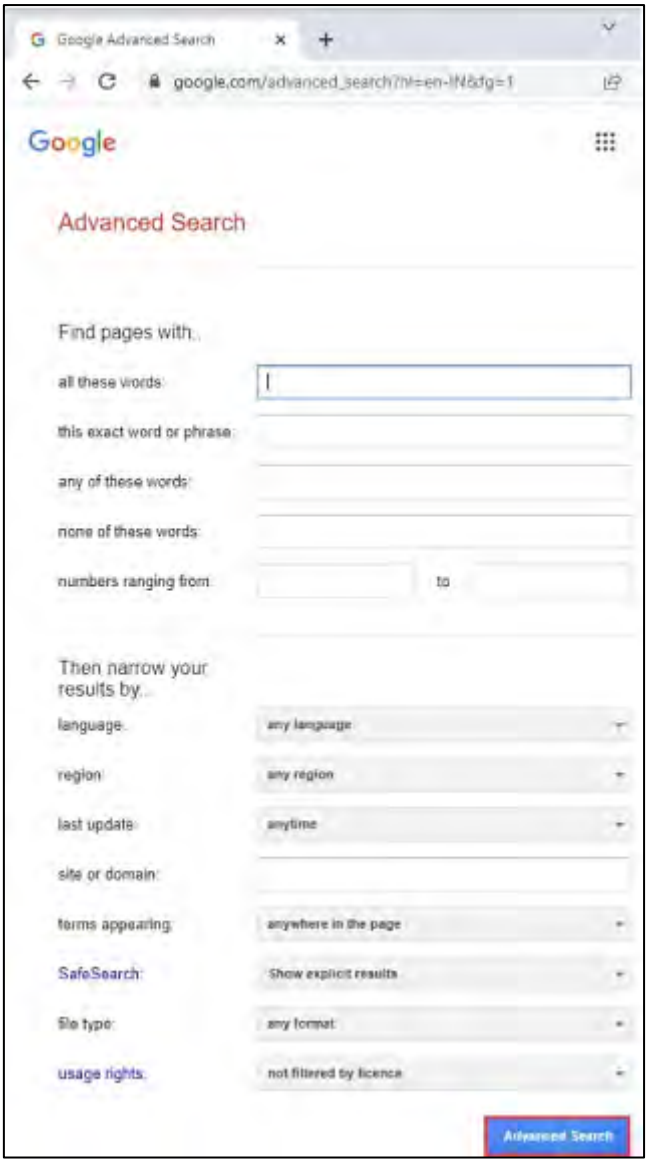

Figure 2-10: Google Advance Search

To conduct an advanced image search on Google, enter **https://www.google.com/advanced_image_search** into the address bar. The advanced image search feature enables you to refine your image queries in various ways. You can filter your search by image color, domain, file format, size, keyword, and more. To accomplish this, select a specific field. Then, input the term you wish to search for in the text box of that field and click the Advanced Search button.

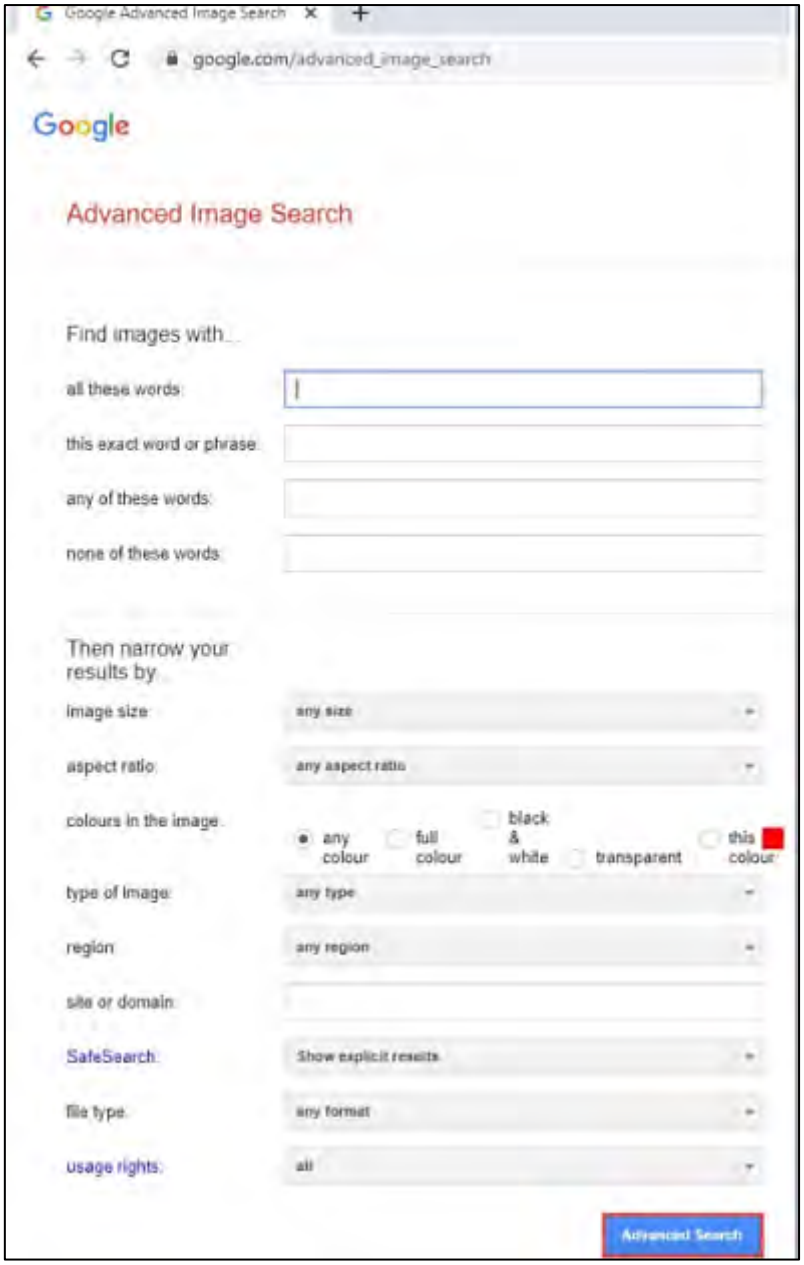

Figure 2-11: Google Advance Image Search

To conduct a reverse image search using Google, enter **https://www.google.com/imghp** in the address bar. This feature allows you to use an image as your search query. You have the option to either upload an image or paste its URL into the reverse image search engine. The engine checks its index and shows all the web locations where the image is found on the search results page. The results you receive can assist you in finding the source and information about the images, including photographs, profile pictures, and memes.

EXAM TIP: Attackers conduct reverse image searches using online tools such as Google Image Search, TinEye Reverse Image Search, Yahoo Image Search, and Bing Image Search.

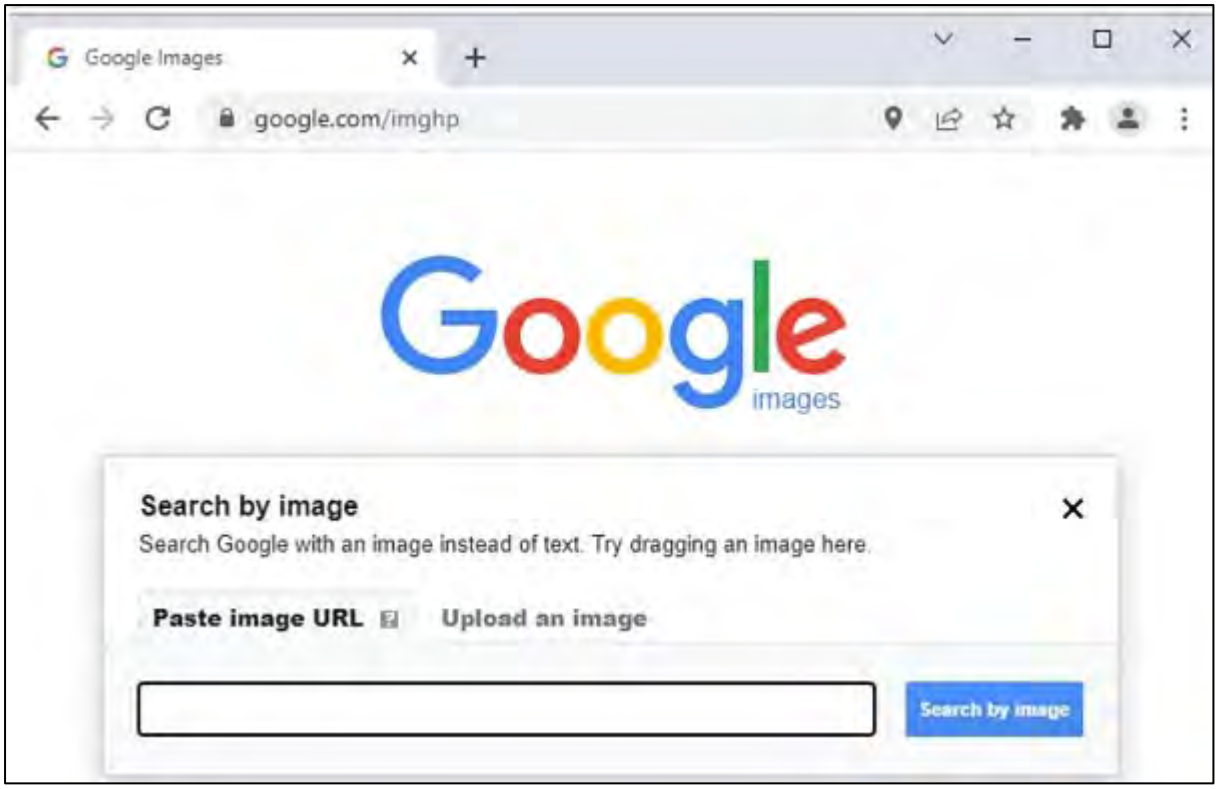

Figure 2-12: Reverse Image Search using Google

Gathering Information from Video Search Engines

Video search engines are online platforms that explore the internet for video material. These search engines either enable users to upload and store video content on their servers or analyze video content hosted on other platforms. The video material sourced from these engines is highly valuable, as it can be utilized to gather intelligence about a subject. Video search engines like YouTube, Google Videos, Yahoo Videos, and Bing Videos enable users to search for video content according to format and length.

After conducting searches for videos connected to the target via video search engines, an attacker can delve deeper into the video content to uncover concealed details like the video's date and time, as well as its thumbnail. By utilizing video analysis tools like YouTube Metadata, YouTube DataViewer, EZGif, and VideoReverser.com, an attacker can reverse a video or transform it into text and other formats to retrieve essential information regarding the target.

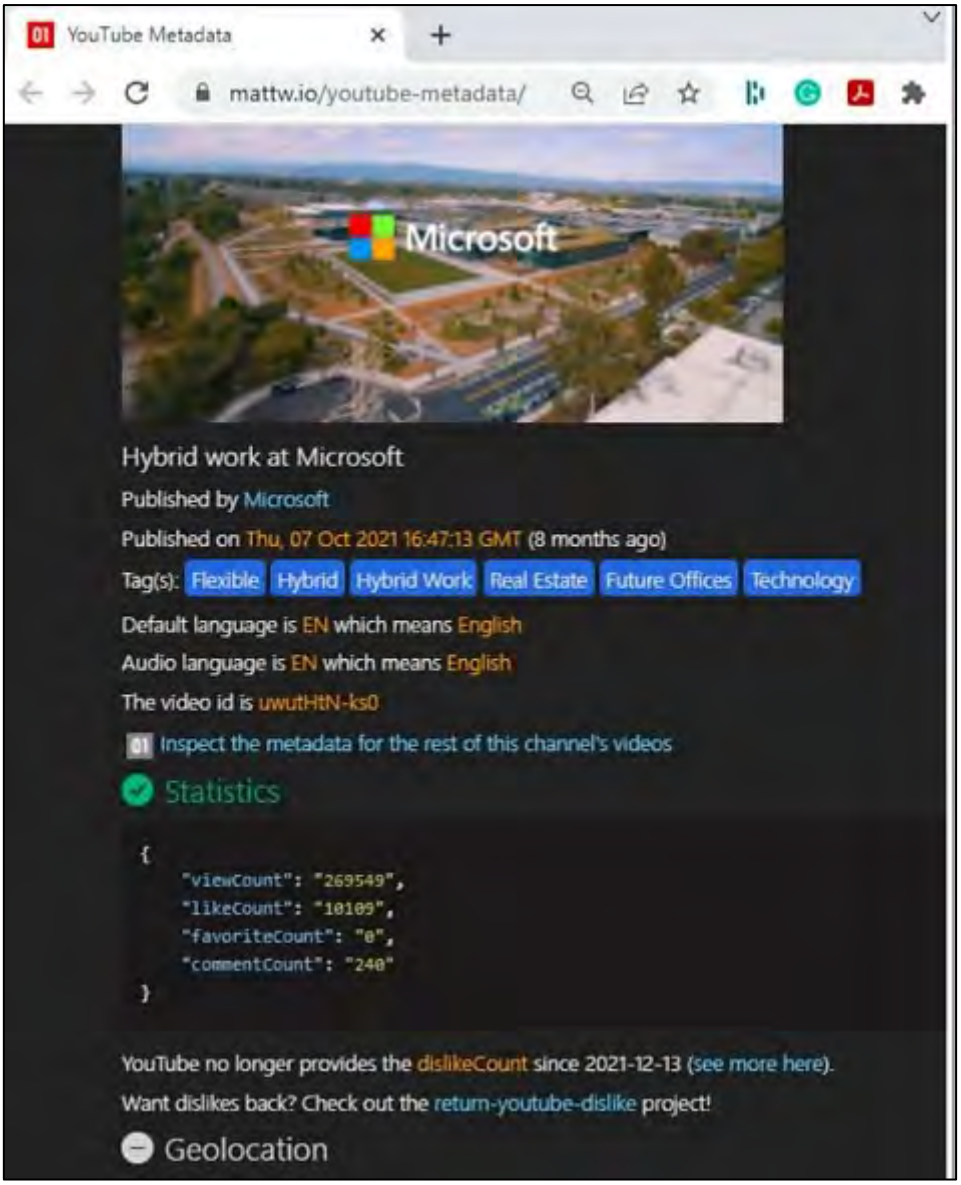

Figure 2-13: YouTube Metadata showing Video Analysis Results

Gathering Information from Meta Search Engines

Meta search engines represent a distinct category of search engines that leverage the results from other search engines (like Google, Bing, Ask.com, etc.) to generate their own results from the Internet in a very brief period. These search engines do not maintain their own search indexes; rather, they receive user inputs and concurrently dispatch queries to various third-party search engines to gather results. After sufficient results are collected, they are ranked based on their relevance and displayed to the user through a web interface. Additionally, meta search engines offer a feature that filters out duplicate search results, ensuring that if a user searches for the same query again, identical results will not be shown twice. Compared to basic search engines, a meta search engine is beneficial as it can obtain more results with the same level of effort.

By using meta search engines like Startpage, MetaGer, and eTools.ch, users can send numerous search queries to multiple search engines at once and compile highly detailed information, including data from shopping websites (such as Amazon and eBay), images, videos, blogs, news

articles, and various other sources. Moreover, meta search engines also enhance the privacy of the users by concealing their IP addresses.

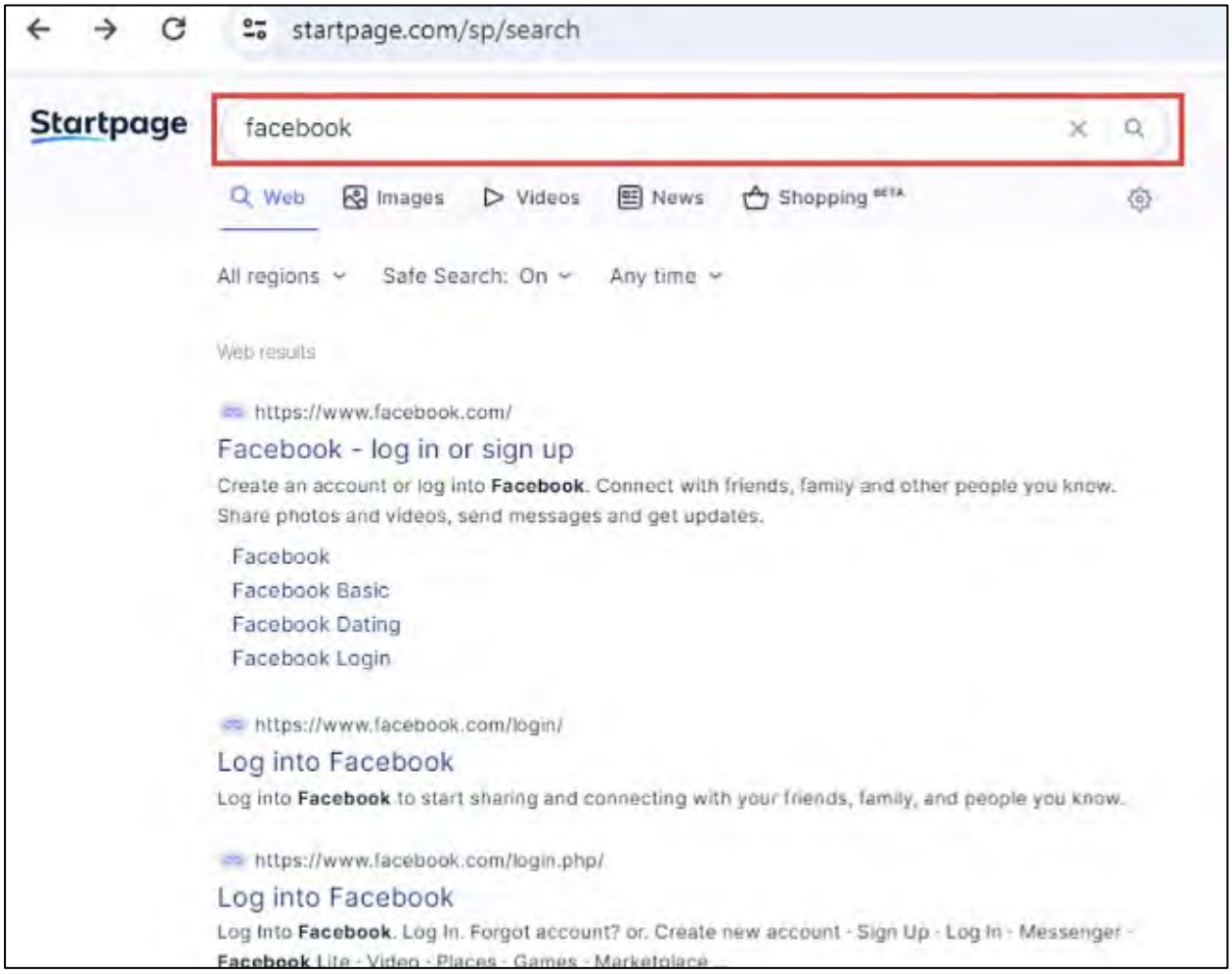

Figure 2-14: Meta Search Engine StartPage.com showing search results for Facebook

Gathering Information from File Transfer Protocol (FTP) Search Engines

FTP search engines allow users to locate files stored on FTP servers that hold significant information about the target organization. Various industries, institutions, companies, and universities utilize FTP servers to maintain extensive file archives and other software that is shared among their staff. A specialized client like FileZilla (https://filezilla-project.org) can be utilized to access FTP accounts; it also offers features such as uploading, downloading, and renaming files. While passwords typically protect FTP servers, numerous servers remain unprotected and can be accessed directly through web browsers.

By employing FTP search engines like NAPALM FTP Indexer, FreewareWeb FTP File Search, Mamont, and Globalfilesearch.com, attackers can look for crucial files and directories that contain sensitive information, such as business plans, tax documents, personal employee data, financial records, licensed software, and other confidential materials.

Some of the important advanced Google search queries for finding FTP servers are listed in Table 2-04.

Google Dork	Description
site:.in	.com
intitle:"index of" "*/ftp.txt"	Finds files containing juicy information
intext:"index of" "ftp"	Finds files containing juicy information
inurl:WS_FTP.log	Finds files containing juicy information
intitle:index.of /cut /robots.txt	Finds files containing passwords
intitle:"Index of ftp passwords"	Finds files containing passwords
inurl:/ftp intitle:"office"	Detects the web server
inurl:/web-ftp.cgi	Detects the web server
site:ftp../ intext:"login" intitle:"server login"	Finds pages containing login portals
intitle:"Index Of" ws_ftp.ini	Finds the "ws_ftp.ini" file, which contains usernames and passwords of FTP users
inurl:ftp -inurl:(http	https) intext:"@gmail.com" intext:subject fwd
allintitle:"CrushFTP WebInterface"	Detects various pages of CrushFTP WebInterface, including login portals and password reset/recovery pages
"ws_ftp.log" ext:log	Finds sensitive directories
intitle:"Monsta ftp" intext:"Lock session to IP"	Shows websites that use the FTP service of Monsta FTP
"index of" /ftp/logs	Finds potential log files
intitle:"index of" inurl:ftp intext:admin	Lists admin folders on FTP servers

Table 2-04: Google Search Queries to find FTP Servers

As illustrated in Figure 2-15, attackers can utilize NAPALM FTP Indexer, a web-based tool, to locate essential files and documents associated with the target domain.

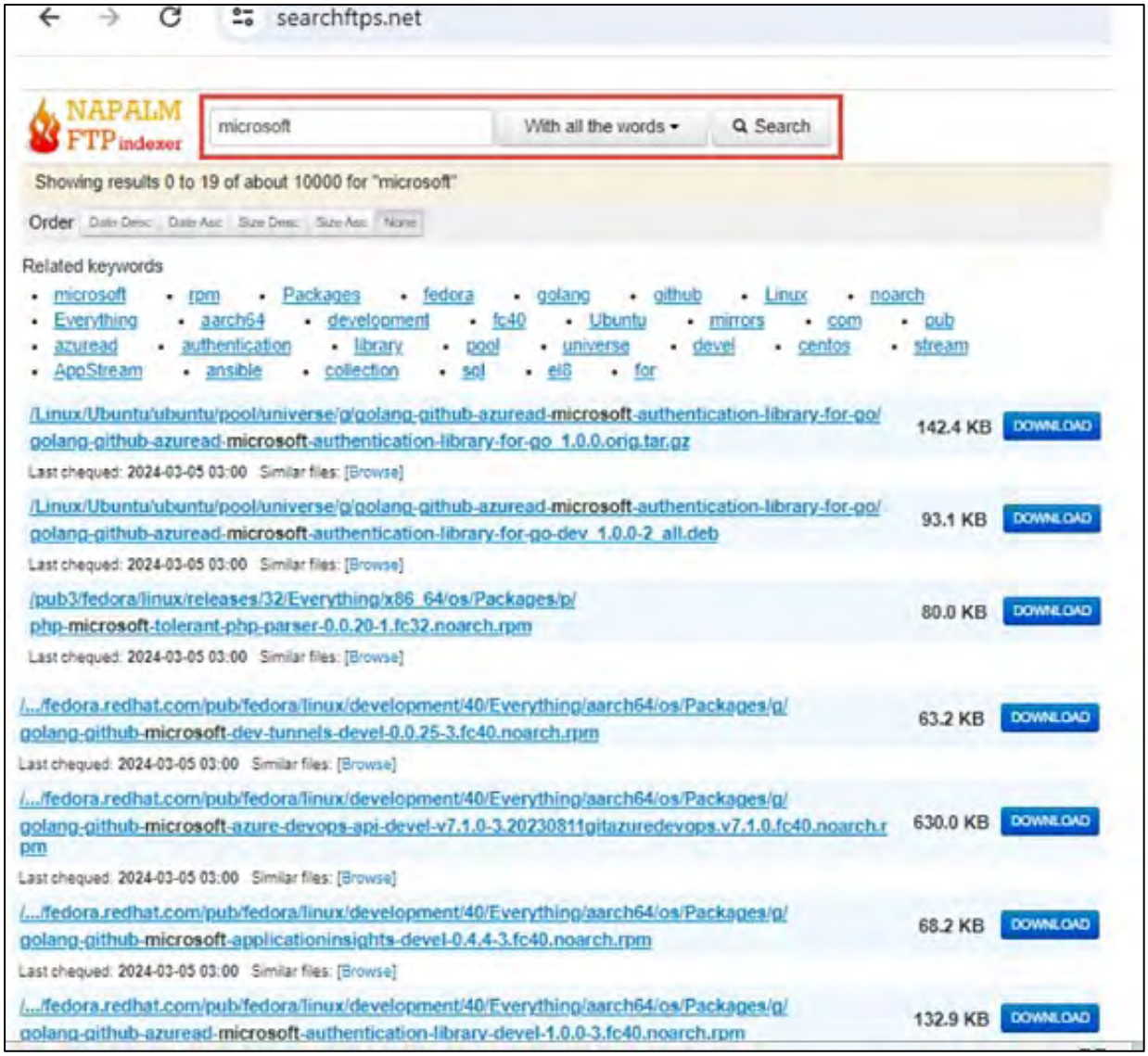

Figure 2-15: FTP Search Engine NAPALM FTP Indexer showing Search Results for Microsoft

Gathering Information from IoT Search Engines

Search engines designed for the Internet of Things (IoT) scan the web for publicly accessible IoT devices. A simple query on these search engines allows an attacker to take control of Supervisory Control and Data Acquisition (SCADA) systems, traffic management systems, internet-connected household appliances, industrial equipment, CCTV cameras, and more. Many of these IoT devices lack proper security measures, meaning they either do not have passwords or are still using default credentials that attackers can easily exploit.

With the help of IoT search engines like Shodan, Censys, and ZoomEye, attackers can gather information such as manufacturer details, geographic location, IP address, hostname, and open ports related to the targeted IoT device. By utilizing this information, an attacker can create a backdoor to the IoT devices, gaining access to initiate further attacks.

As illustrated in Figure 2-16, attackers can use Shodan to identify all IoT devices within the target organization that have open ports and services.

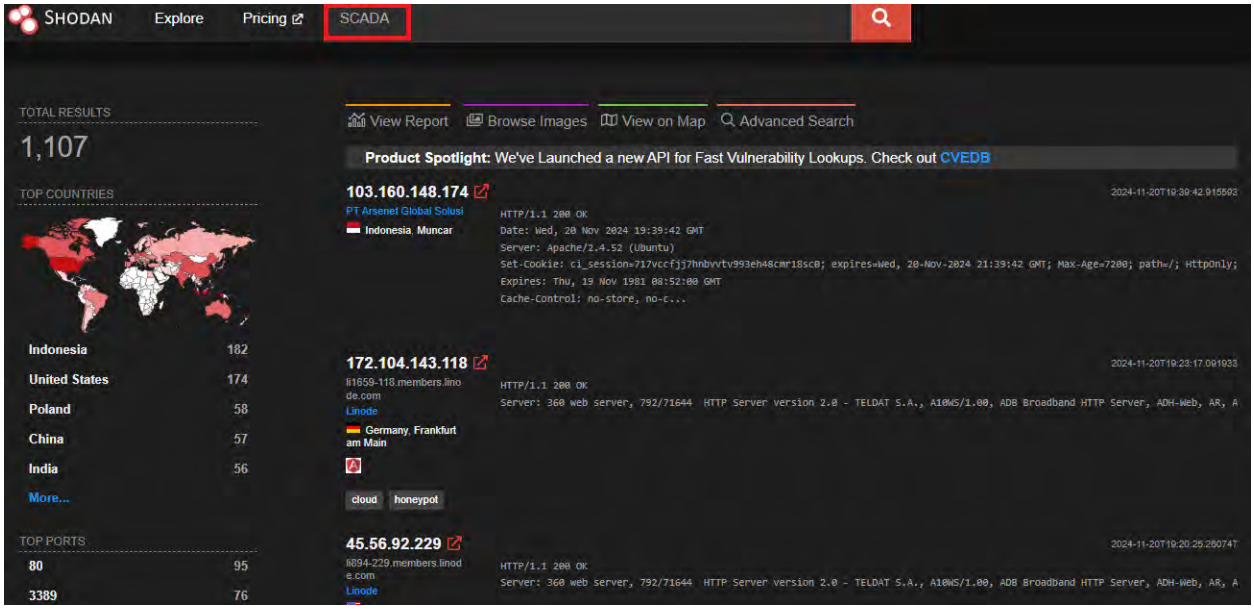

Figure 2-16: Shodan showing Search Results for SCADA Devices

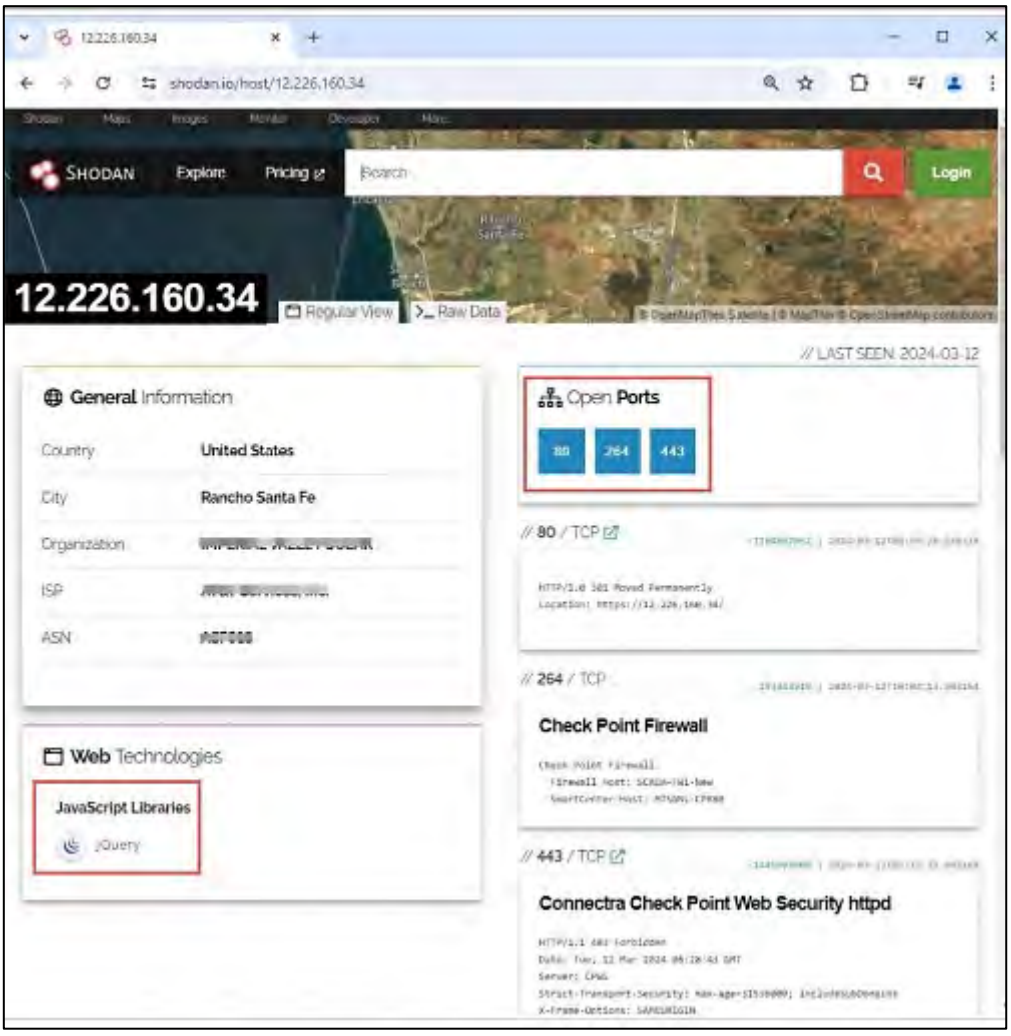

Figure 2-17: Shodan showing Open Ports and Services of a SCADA System

Footprinting through Internet Research Services

Internet research services, including people search services, can offer confidential information about the individual in question. Services such as people search, alert notifications, financial platforms, and job boards provide details about a subject, including infrastructure specifics, physical address, and employee information. With this data, an attacker might devise a hacking plan to infiltrate the target organization’s network and execute various forms of sophisticated system attacks.

Finding a Company's Top-Level Domains (TLDs) and Sub-domains

A company's Top-Level Domains (TLDs) and sub-domains can reveal a significant amount of valuable information to an attacker. A public website is intended to display the organization's presence on the Internet and is accessible to the public for free. Its purpose is to draw in customers and partners. It might include details such as the organization's background, the products and services offered, and contact information. The external URL of the target organization can be found using search engines like Google and Bing.

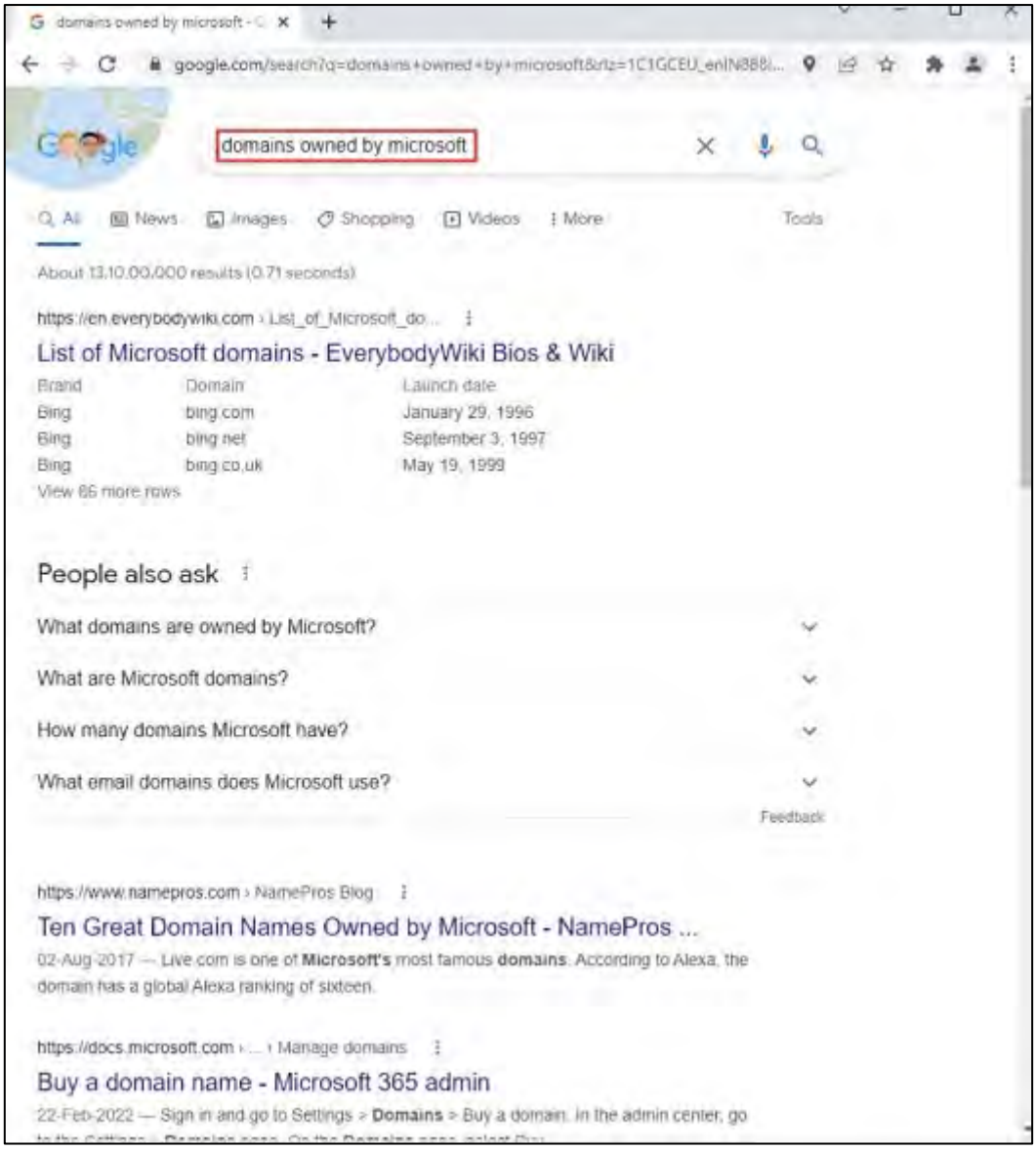

Figure 2-18: Google Search Engine showing Search Results for Domains owned by Microsoft

The sub-domain is restricted to a select group of individuals. This group may include employees of a company or members of a specific department. In numerous organizations, web administrators set up sub-domains to evaluate new technologies prior to implementing them on the main site. Typically, these sub-domains are under testing and are unsecured, making them more susceptible to various forms of exploitation. Sub-domains offer insights into the different departments and business divisions within a company. Identifying such sub-domains can uncover vital information about the target, including the website's source code and files on the web server. Access limitations can be enforced based on IP addresses, domains or subnets, usernames, and passwords. The sub-domain facilitates access to the private functionalities of an organization. Most organizations tend to follow standard formats for their sub-domains. As a result, a hacker familiar with a company's external URL can often find the sub-domain through trial and error or by utilizing a service like Netcraft. You can also leverage the advanced Google search operator provided below to locate all sub-domains of the target:

```
site:microsoft.com -inurl:www
```

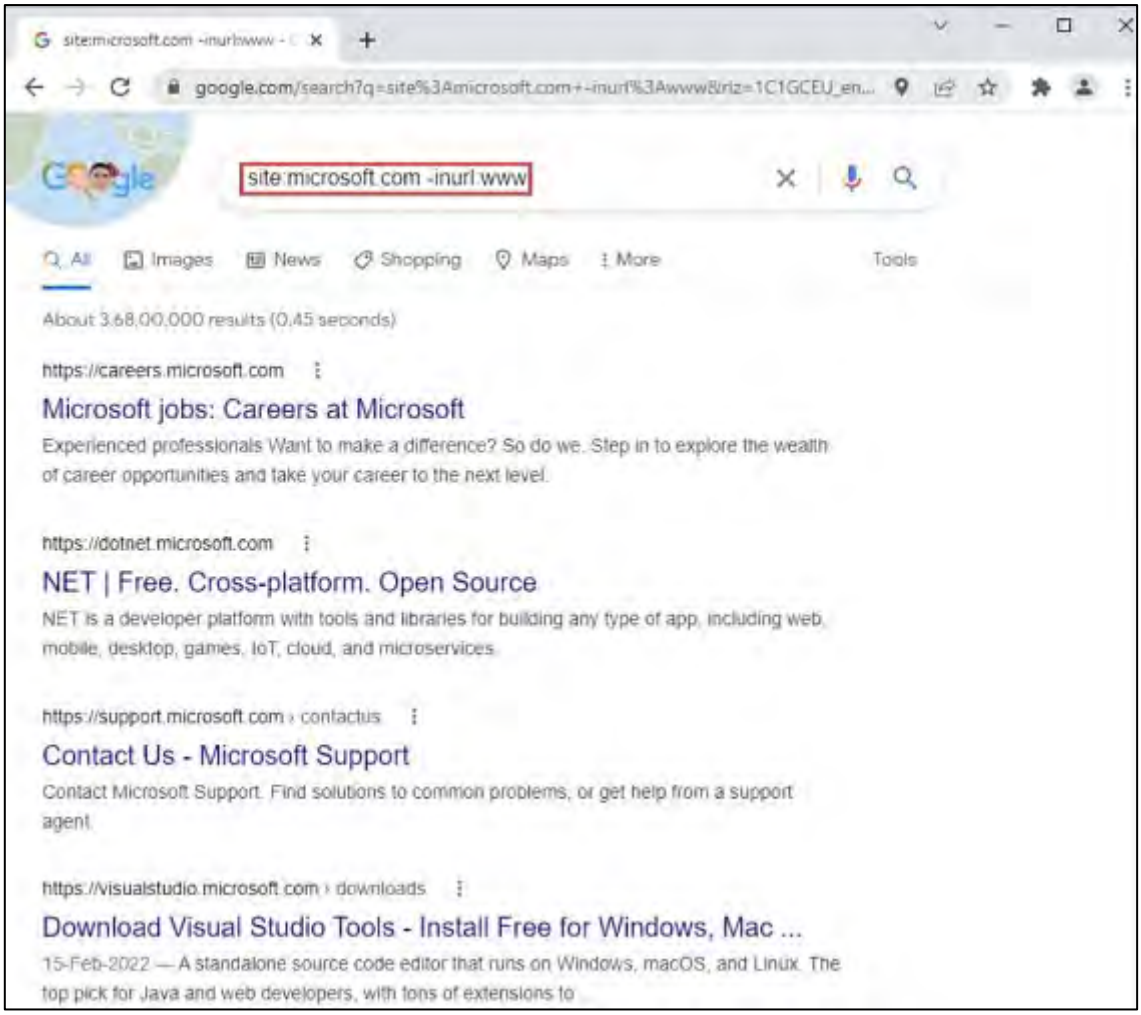

Figure 2-19: Finding sub-domains using Google Advanced Search Operator

Tools to Search Company's Sub-domains

Netcraft

Netcraft offers a range of Internet security services, such as anti-fraud and anti-phishing solutions, application assessments, and PCI compliance scanning. Additionally, they evaluate the market share of web servers, operating systems, hosting companies, SSL certificate authorities, and other Internet-related metrics.

As demonstrated in Figure 2-20, attackers can utilize Netcraft to gather all subdomains associated with the target domain.

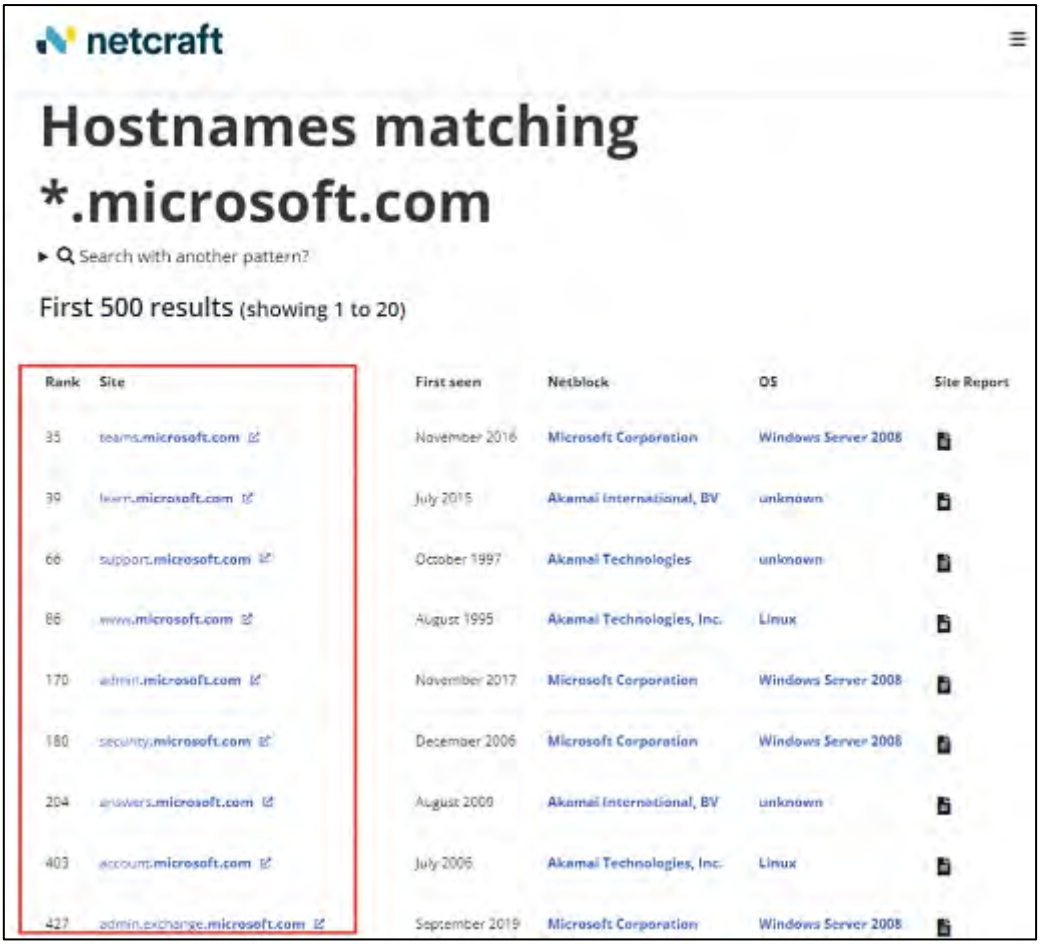

Figure 2-20: Netcraft Displaying Subdomains of microsoft.com

DNSDumpster

DNSdumpster.com functions as a domain research tool that malicious actors can utilize to identify hosts associated with a specific domain. In the screenshot presented, attackers are looking up subdomains connected to microsoft.com to gather essential information about the target company's domain, including subdomains, IP addresses, and utilized DNS servers, among other details.

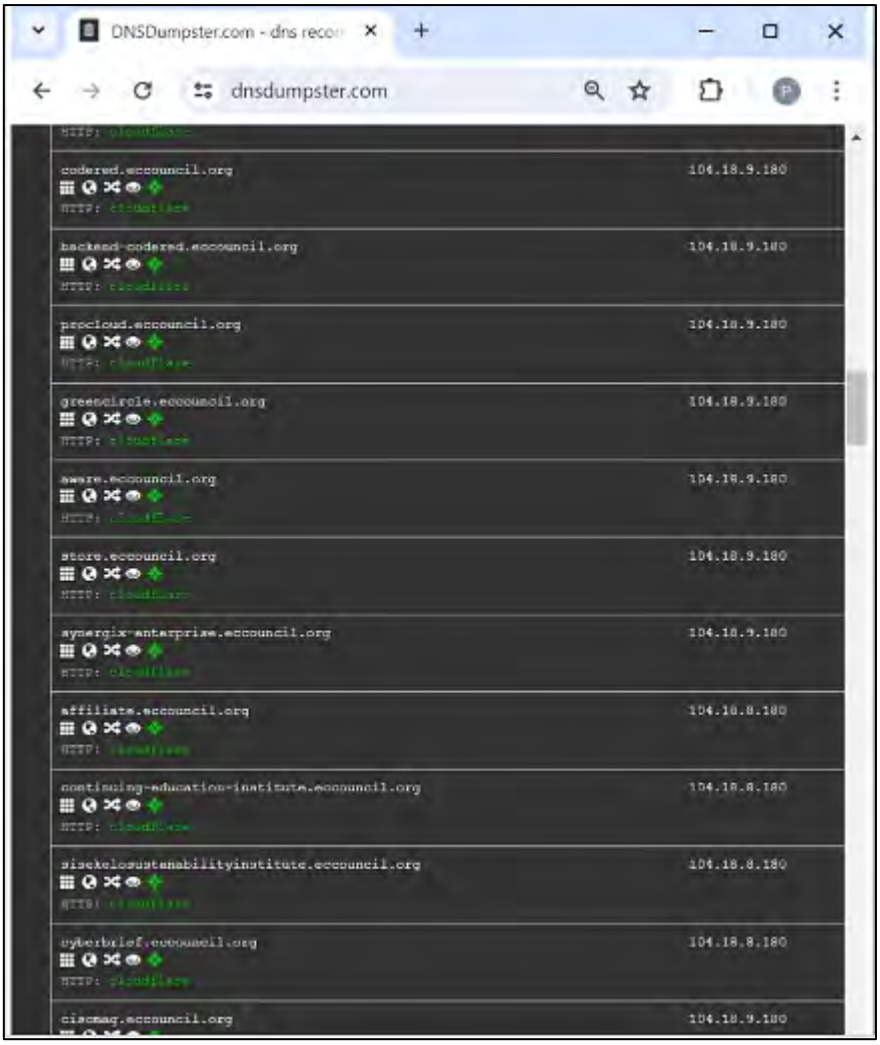

Figure 2-21: DNSdumpster Tool displaying Subdomains of eccouncil.org

Pentest-Tools Find Subdomains

Pentest-Tools Find Subdomains is an online resource designed for identifying subdomains and their corresponding IP addresses, along with network details and HTTP server information. In Figure 2-22 attackers look for subdomains associated with microsoft.com to gather essential information about the target company's domain, including subdomains, IP addresses, operating systems, utilized servers, employed technologies, web platforms, and page titles.

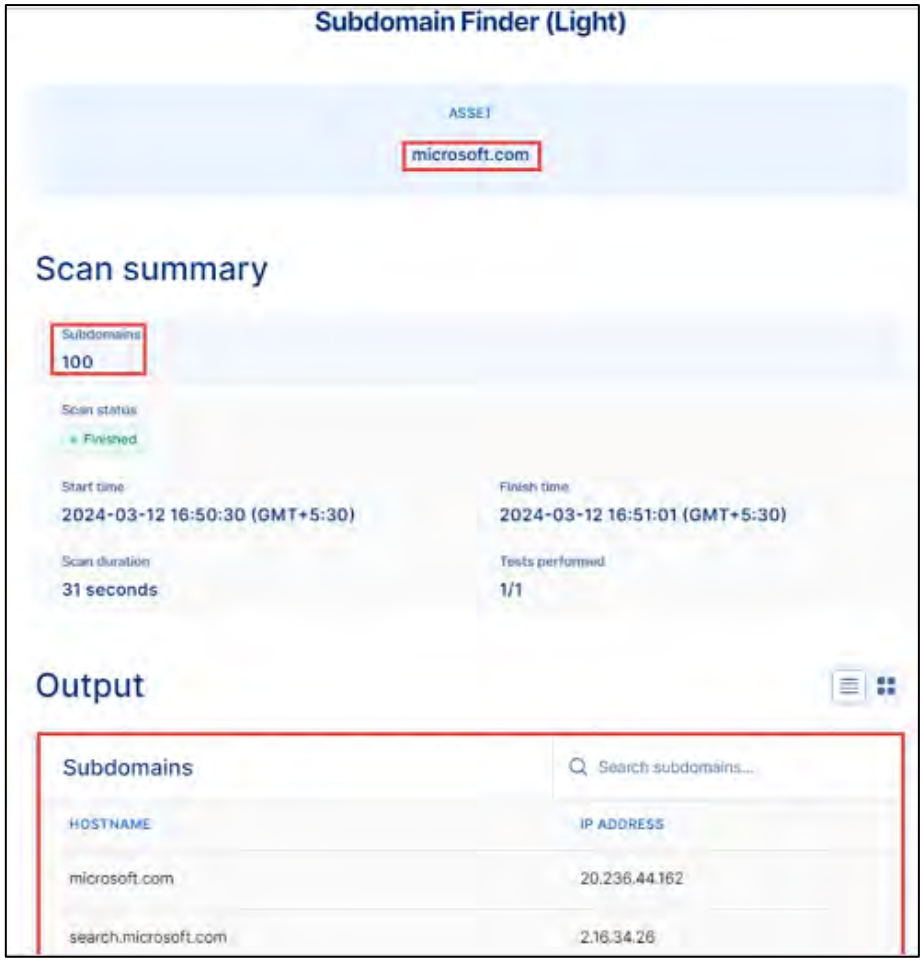

Figure 2-22: Pentest-Tools displaying Subdomains of microsoft.com

Finding a Company's Top-Level Domains (TLDs) and Sub-domains with AI

Attackers can utilize AI-driven technologies to improve and streamline their footprinting activities. By employing AI, they can easily identify the main domains and subdomains associated with their target.

Example 1:

A hacker can utilize ChatGPT to accomplish this task by crafting a suitable prompt like:

"Discover all the subdomains of 'google.com' using dig command"
The following command is utilized:
dig +short google.com NS | xargs -I{} dig +nocmd +noall +answer @{} google.com ANY | grep -E '(CNAME|A|AAAA)'

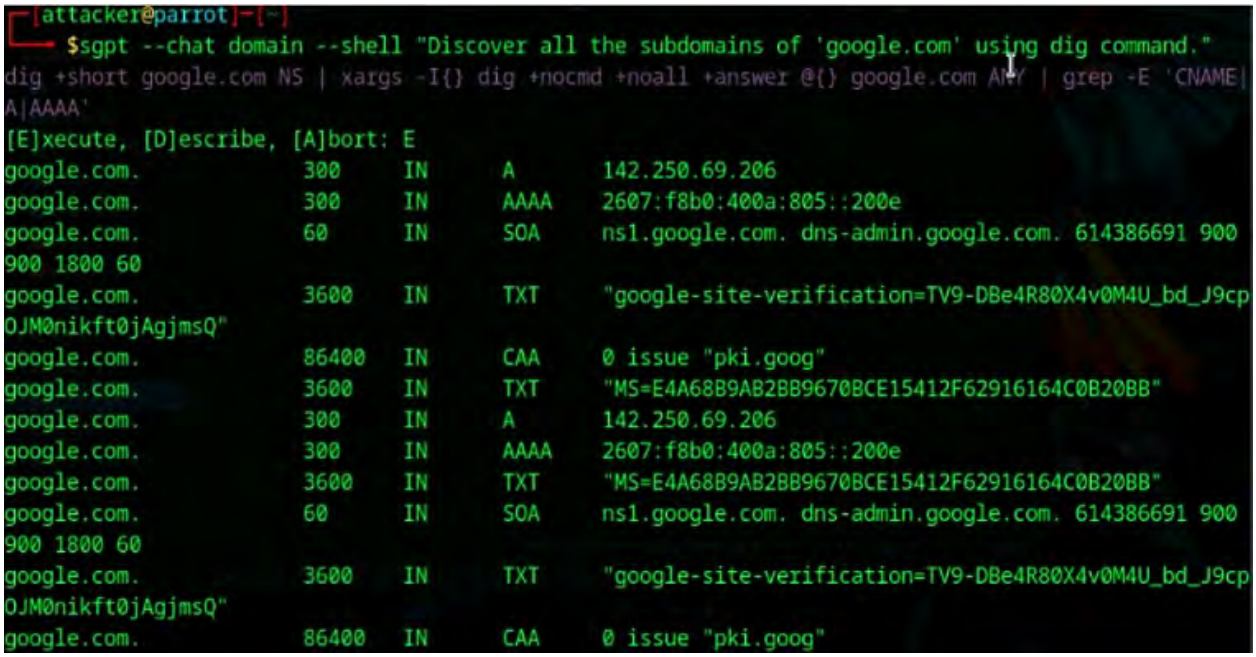

Figure 2-23: Discover all the Subdomains of 'google.com' using dig Command

Table 2-05 describes each option used in the command.

Command Breakdown	Description
dig +short google.com NS	Retrieves the Name Server (NS) records for "google.com" in a shortened format.
| xargs -I{}	Takes each line of output (representing a name server) and supplies it as an argument for the next `dig` command.
dig +nocmd +noall +answer @{} google.com A +noall +answer	Performs a DNS lookup for the A records of "google.com" using each name server obtained from the previous command.

Table 2-05: dig Command Description to Discover Subdomains

Example 2:

A hacker can utilize ChatGPT to accomplish this task by crafting a suitable prompt like:

> Use Sublist3r to gather a list of subdomains of the target organization eccouncil"
> The following command is utilized to run the Sublist3r tool with designated parameters:
> sublist3r -d eccouncil.org -o eccouncil_subdomains.txt

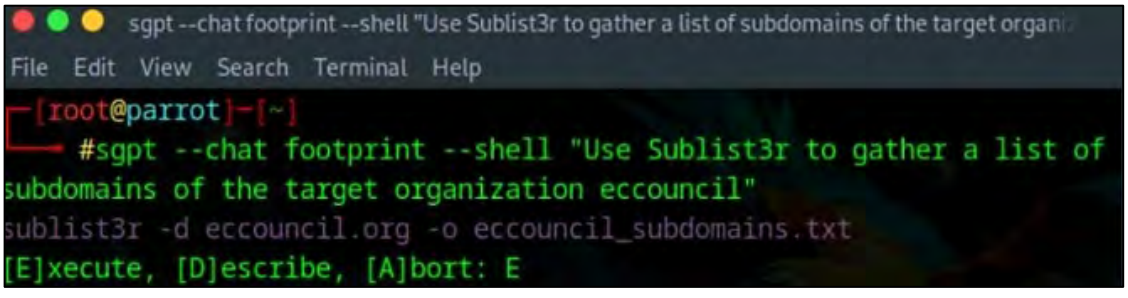

Figure 2-24: Use Sublist3r to Gather a List of Subdomains of the Target Organization eccouncil

Table 2-06 describes each option used in the command.

Command Breakdown	Description
sublist3r	Start the Sublist3r tool
-d eccouncil.org	Specifies the target domain (eccouncil.org), for which we want to enumerate subdomains.
-o eccouncil_subdomains.txt	Indicates the output file where the found subdomains will be stored (eccouncil_subdomains.txt).

Table 2-06: Sublist3r Tool Command Description

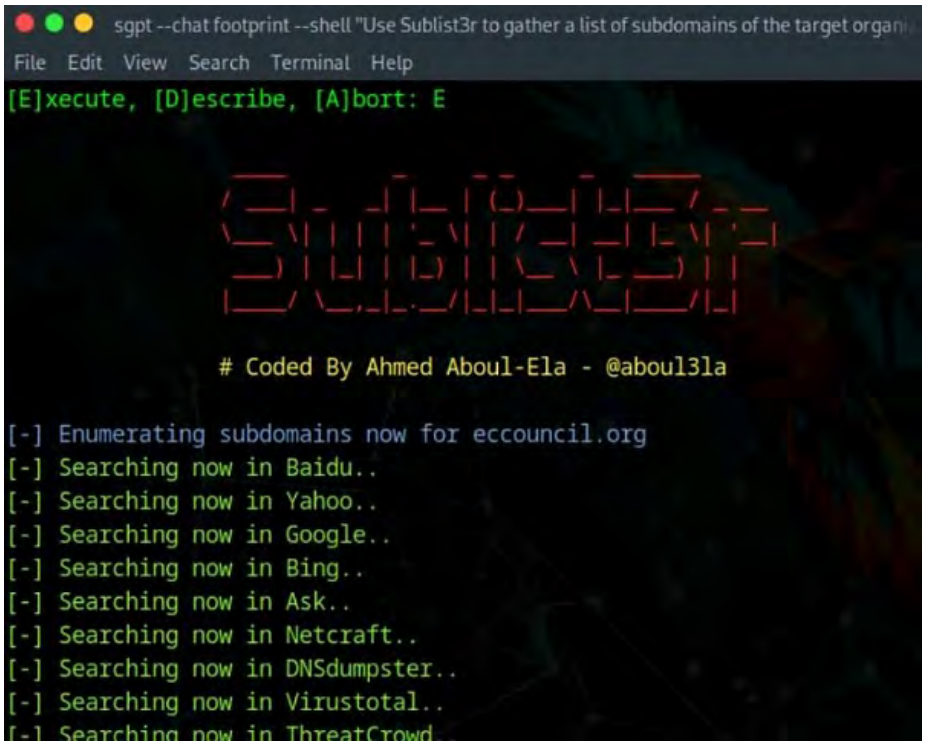

Figure 2-25: Subdomains Associated with Target-1

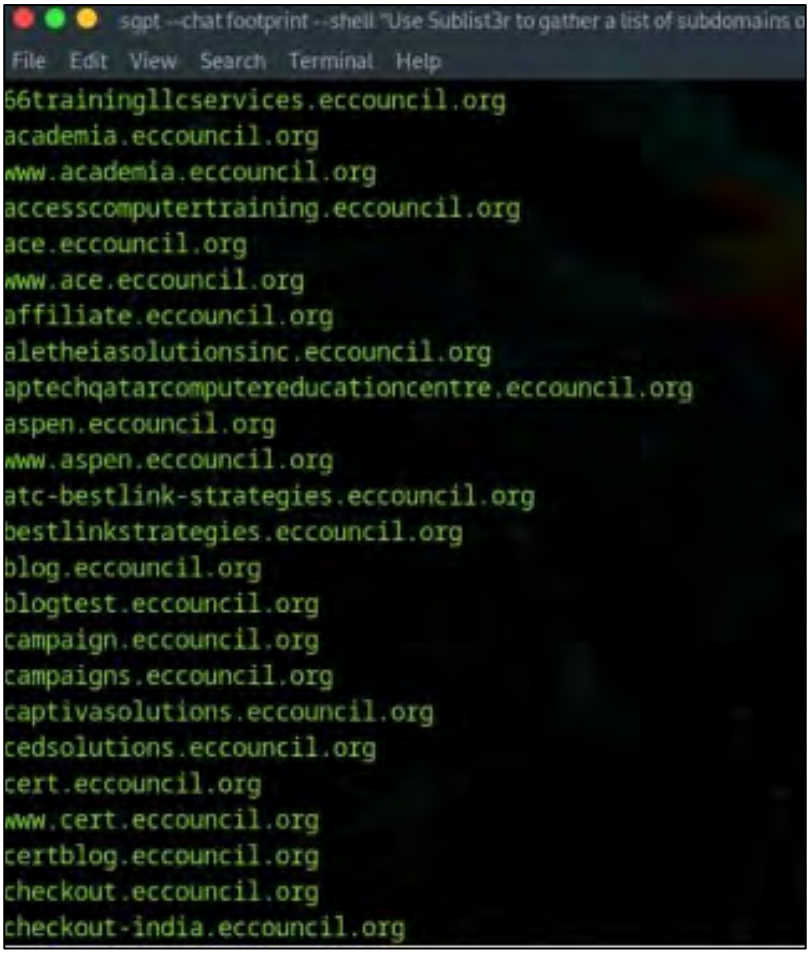

Figure 2-26: Subdomains Associated with Target-2

In general, this command directs Sublist3r to look for subdomains linked to the "eccouncil.org" domain and store the findings in a text file titled "eccouncil_subdomains.txt".

Extracting Website Information from https://archive.org

The Archive is an Internet Archive Wayback Machine that allows users to investigate past versions of websites. This investigation enables an attacker to collect data on an organization's web pages from the moment they were created. Since the website https://archive.org archives web pages from their inception, an attacker can access information that has been deleted from the target website, including web pages, audio files, video files, images, text, and software programs. Attackers utilize this data to execute phishing and other forms of web application attacks against the targeted organization.

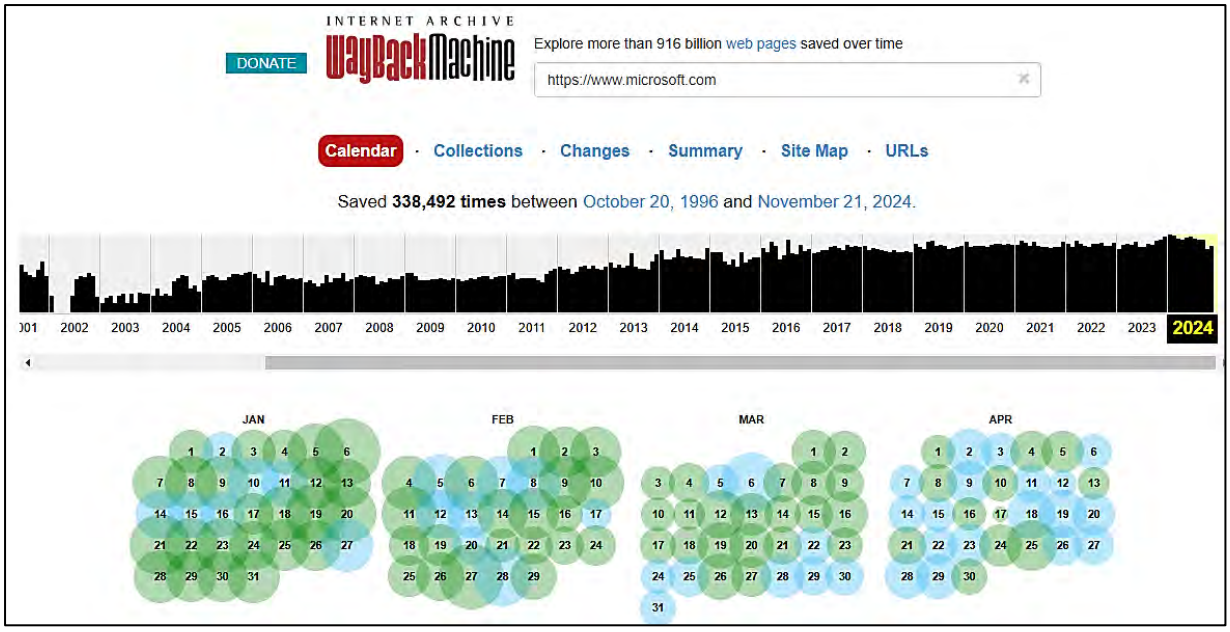

Figure 2-27: Screenshot of Archive showing Archived Versions of microsoft.com

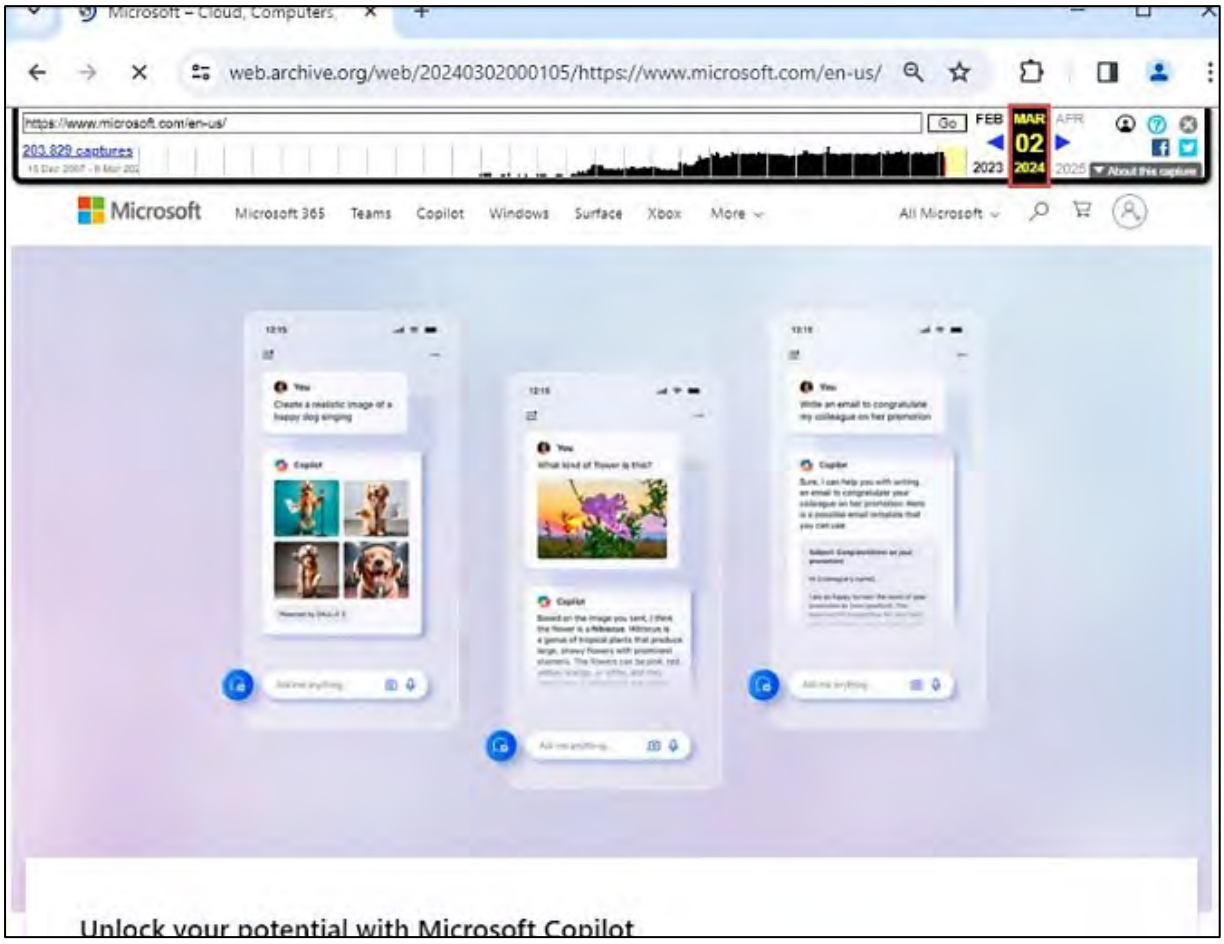

Figure 2-28: Screenshot of Archive showing Archived Web Pages of microsoft.com

Attackers may utilize tools like Photon to obtain saved URLs of the targeted website from archive.org. Execute the command below to fetch the archive.org links for the specified website:

```
photon.py -u <URL of the Target Website> -l 3 -t 200 –wayback
```

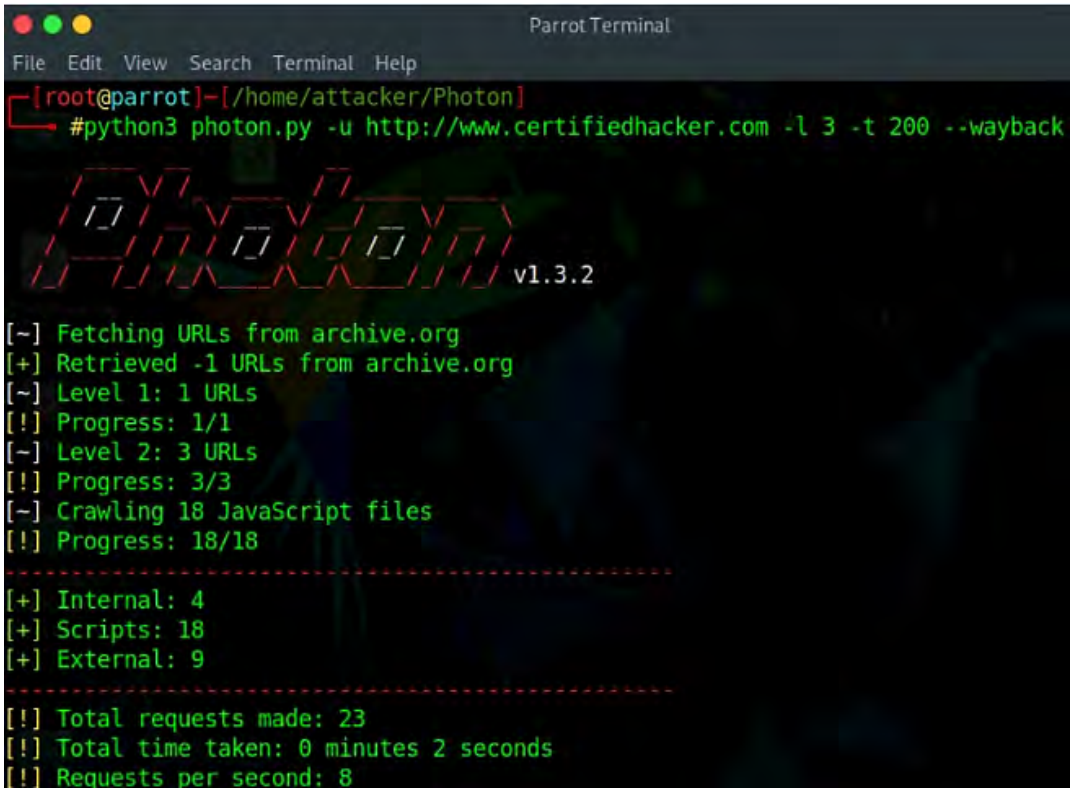

Figure 2-29: Photon Output for the Command to Retrieve archive.org Links

Execute this command to obtain archived URLs from the specified website:

```
python photon.py -u <URL of the Target Website> -l 3 -t 200 –only-urls
```

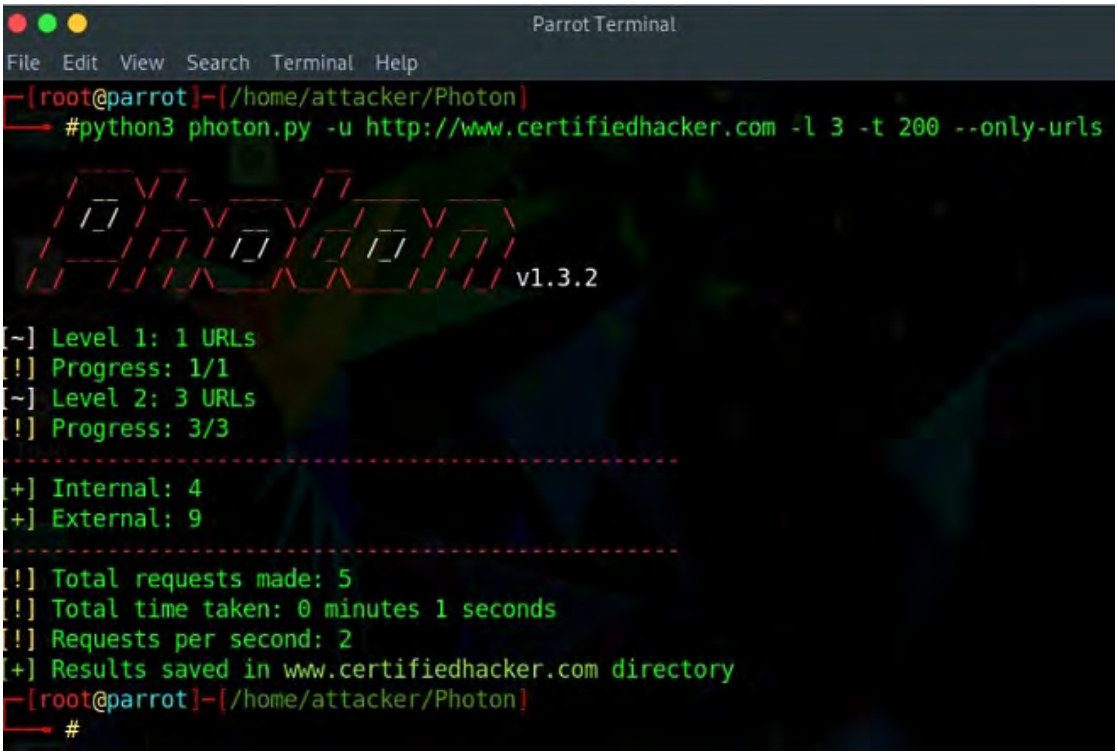

Figure 2-30: Photon Output for the Command to Retrieve Archived URLs

Footprinting through People Search Services

You can utilize public record sites to gather details about email addresses, phone numbers, residential addresses, and more. Numerous individuals turn to online people search platforms to discover information about others. Typically, services like Spokeo, Intelius, Pipl, BeenVerified, Whitepages, Instant Checkmate, and PeekYou offer details on people's names, addresses, contact information, birth dates, images, videos, careers, information about their relatives and friends, social media profiles, property details, and optional insights into criminal background checks. Additionally, these online people search services may also disclose a person's occupation, businesses they own, projects in the pipeline, working conditions, personal websites and blogs, phone numbers, key dates, company emails, personal cell numbers, fax numbers, and individual email addresses. With this information, an attacker might attempt to gain access to bank information, credit card data, and historical records, among other things. Such details can be very advantageous for attackers to execute their schemes. There are numerous online people search services available that assist in gathering information about individuals.

People Search Service – Spokeo

Attackers can utilize the Spokeo people search online platform to look up individuals associated with the targeted organization. By employing this service, attackers can gather details, including phone numbers, email addresses, previous addresses, age, date of birth, relatives, social media profiles, and legal records.

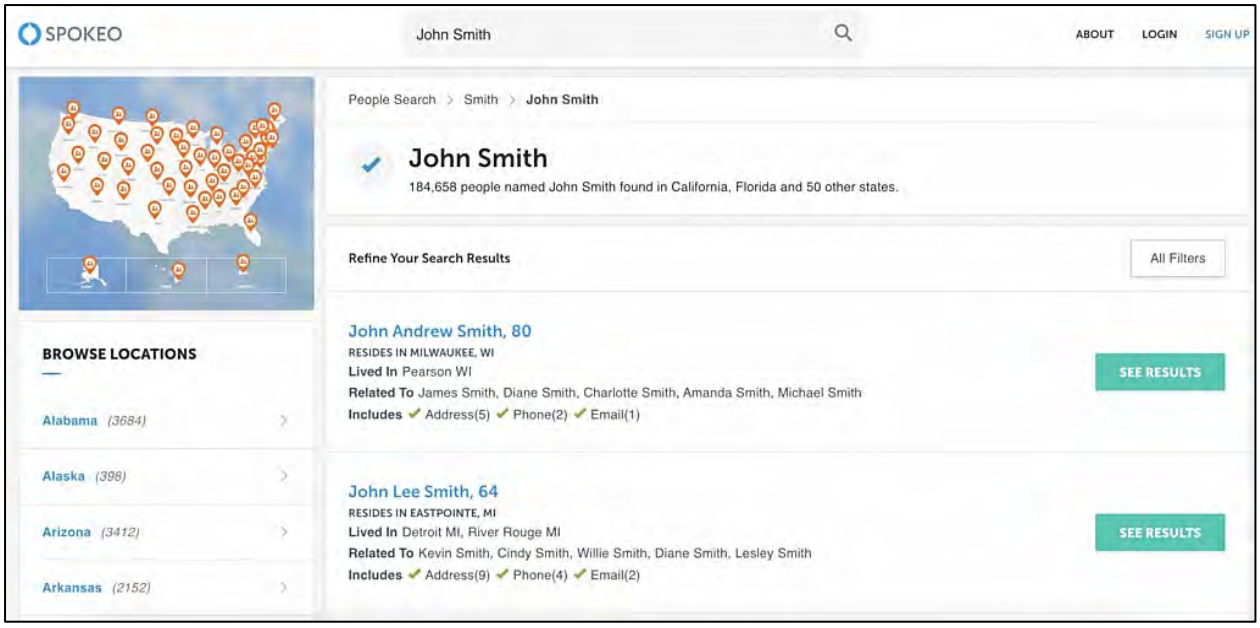

Figure 2-31: Spokeo People Search Service

Footprinting through Job Sites

Attackers can collect crucial data regarding the operating system, software versions, the organization's infrastructure, and database schema by examining job sites using various methods. Many corporate websites share recruitment details on their job posting sections, which subsequently disclose information about hardware and software, network-related elements, and the technologies utilized by the company (e.g., firewall, type of internal server, OS in use, network devices, hypervisors, VMs, etc.). Additionally, the site may display a list of key employees along with their email addresses. This type of information can be advantageous for an attacker. For instance, if a company posts a job for a Network Administrator, it includes the qualifications required for the role.

Furthermore, attackers can review employee resumes available on job sites and derive information such as individuals' expertise, educational background, and employment history. An employee's job history can disclose technical insights about the targeted organization. Attackers can leverage the technical details gathered from job platforms like Dice, LinkedIn, Glassdoor, and Simply Hired to identify potential vulnerabilities in the target's IT infrastructure.

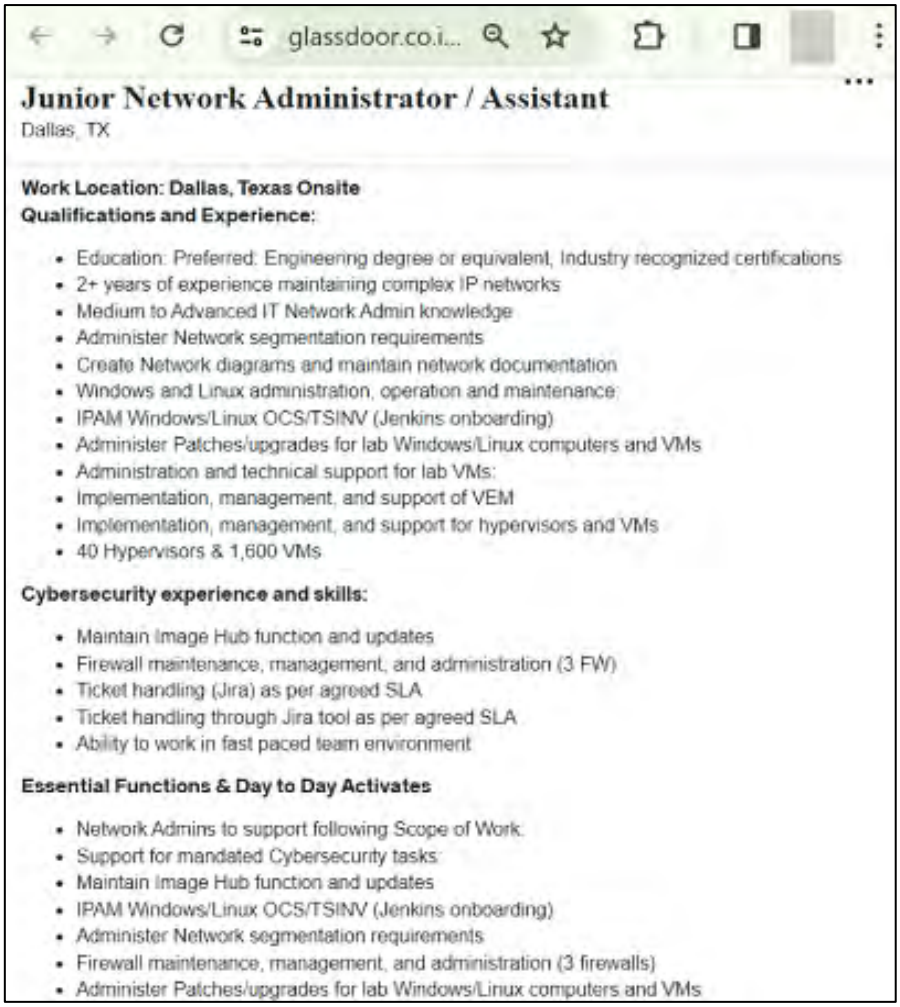

Figure 2-32: Screenshot of Job Posting showing Valuable Information

Dark Web Footprinting

The surface web represents the outer layer of the internet, enabling users to discover web pages and content through standard web browsers. Search engines utilize crawlers, which are automated bots designed to access and retrieve web pages. Users can navigate the surface web using browsers like Google Chrome, Mozilla Firefox, and Opera.

The deep web refers to the layer of cyberspace that includes web pages and content that remain hidden and not indexed. This content is not discoverable via conventional web browsers or search engines. The scale of the deep web is immeasurable and encompasses nearly the entire World Wide Web. Basic search engines cannot access the deep web through the crawling process. This layer consists of official government or federal databases and various organizational information. One can navigate the deep web using search engines such as Tor Browser and the WWW Virtual Library. It can serve both legal and illegal purposes.

The dark web, also known as the Darknet, constitutes a deeper aspect of cyberspace, allowing individuals to browse anonymously without leaving traces. Accessing the dark web necessitates specialized tools or darknet browsers. Attackers mainly exploit the dark web to conduct reconnaissance on target organizations and execute attacks. The dark web can be explored with search engines such as Tor Browser and ExoneraTor.

Attackers may employ dark web search tools like Tor Browser, ExoneraTor, and OnionLand Search engine to collect sensitive information about targets, including credit card details, passport information, identification card details, medical records, social media accounts, and Social Security Numbers (SSNs). With access to this information, they can carry out further attacks on the targets.

Tor Browser

The Tor Browser is utilized for accessing the dark web, functioning as a default VPN for users by routing the network IP address through multiple servers before engaging with the web. Malicious actors employ this browser to reach concealed content, unlisted websites, and encrypted databases located in the dark web.

As illustrated in Figure 2-33, attackers can gather more comprehensive and concealed information regarding the target organization by using the Tor Browser.

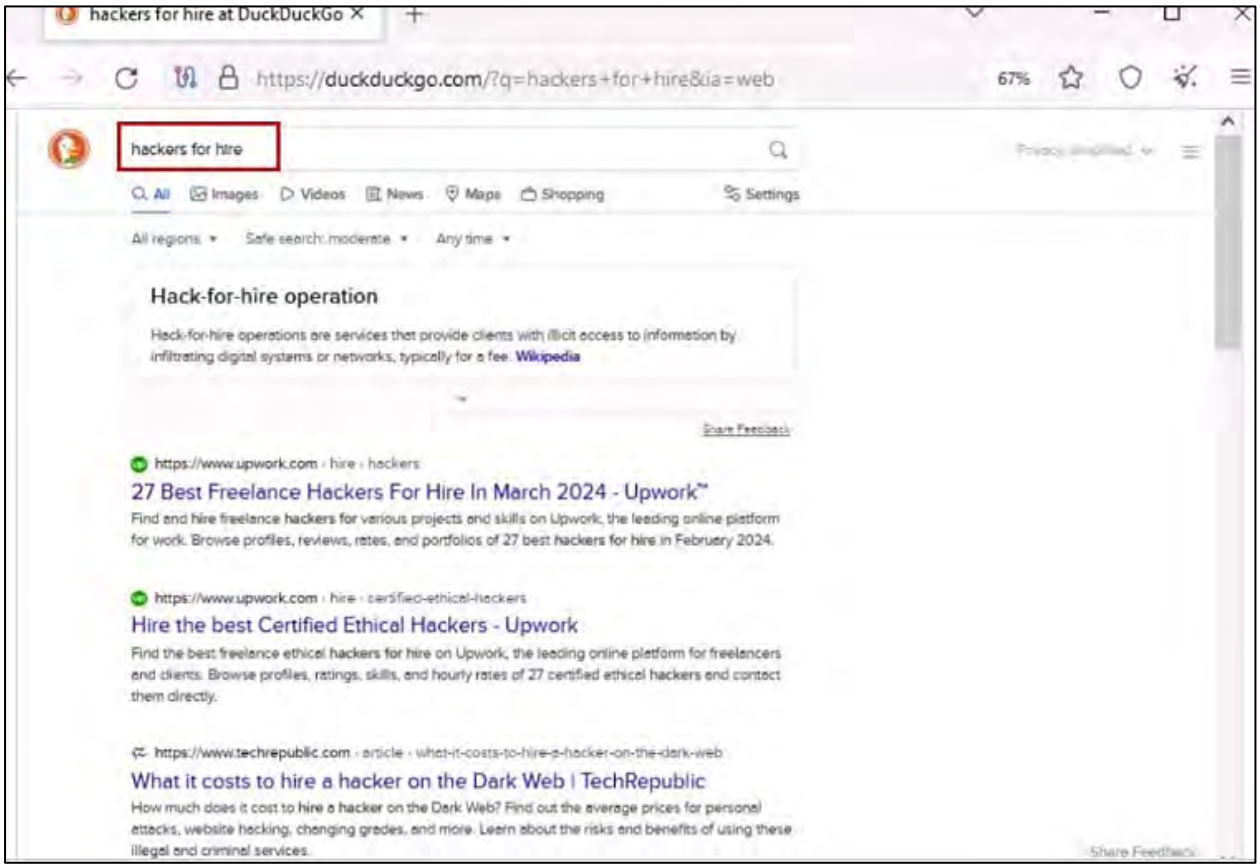

Figure 2-33: Screenshot of Tor Browser

Searching the Dark Web with Advanced Search Parameters

Attackers can employ sophisticated search parameters to narrow down their searches in the Deep Web to locate specific information utilizing Dark Web search tools like the Tor Browser. They can uncover vital data using advanced search methods such as filtering by file type or applying domain restrictions.

For example, attackers might leverage advanced search parameters to discover sensitive documents, including financial records or login information. These refined search techniques allow attackers to swiftly identify and access valuable information, which they can then exploit for malicious purposes.

Attackers can enhance their searches and concentrate on the particular data they are looking for by employing the following parameters:

- **Personal profiles:** Look for details related to the victim's personal profiles, such as social media accounts or personal websites.
  For instance, "John Doe" site:facebook.com OR site:linkedin.com
- **Scientific publications:** Search for works on specific topics, such as academic or scientific research articles and papers.
  For example, "John Doe" site:scholar.google.com
- **Court records:** Look for legal documents associated with court cases or records.
  For instance, "John Doe" court records
- **Member directories:** Search for directories of individuals employed by the organizations.
  For example, "John Doe" site:example.com "employee directory"
- **Medical records:** Seek out medical information or health history related to the victim.
  For example, "John Doe" medical records
- **Location records:** Look for location details such as the victim's location history or GPS data.
  For instance, "John Doe" location history

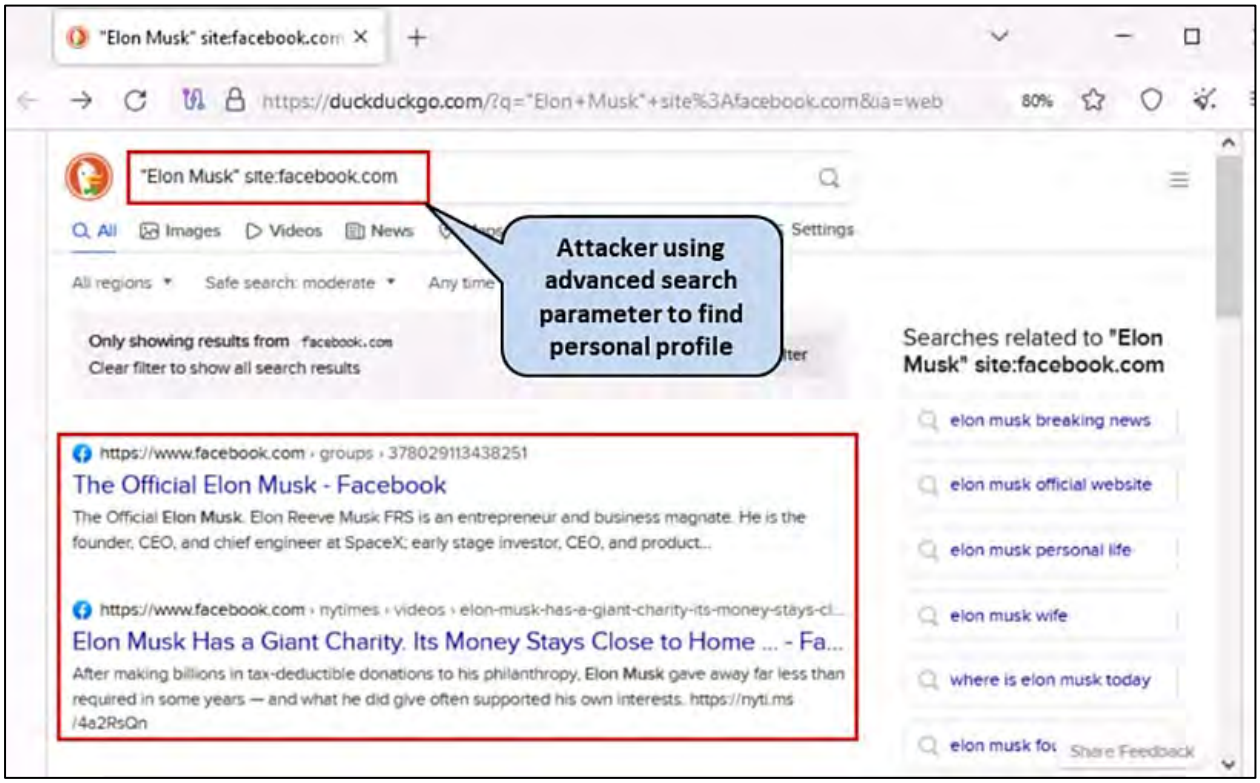

Figure 2-34: Tor Browser showing Advanced Search Parameter Results

Table 2-07 illustrates advanced search queries designed for footprinting on the dark web to find sensitive information while using the Tor Browser. These queries utilize specific search parameters and operators that are suitable for dark web search engines.

Type of Information	Search Query	Explanation
Sensitive PDFs	filetype:pdf site:onion confidential	Finds PDF documents marked as confidential on .onion sites.
Passwords in Config Files	inurl:config filetype:txt password	Searches for text files in configuration URLs containing passwords.
Financial Documents	filetype:xlsx site:onion financial	Locates Excel files related to financial data on .onion sites.
Database Dumps	filetype:sql site:onion dump	Finds SQL database dump files on .onion sites.
Email Lists	filetype:csv site:onion email	Searches for CSV files containing email lists on .onion sites.
Login Credentials	intitle:"login credentials" filetype:docx	Locates Word documents with login credentials in the title.
Server Configurations	filetype:xml inurl:config server	Finds XML files related to server configurations.
Private Keys	filetype:key site:onion private	Searches for private key files on .onion sites.
Medical Records	filetype:pdf site:onion "medical records"	Locates PDF documents containing medical records on .onion sites.
Business Plans	filetype:ppt site:onion "business plan"	Finds PowerPoint files with business plans on .onion sites.
Source Code	filetype:py site:onion "def "	Searches for Python source code files on .onion sites.
Legal Documents	filetype:docx site:onion "legal document"	Locates Word documents related to legal matters on .onion sites.
Bank Statements	filetype:pdf site:onion "bank statement"	Finds PDF documents containing bank statements on .onion sites.
Intellectual Property	filetype:pdf inurl:patent confidential	Searches for patent documents marked as confidential in PDFs.
Security Vulnerabilities	filetype:txt inurl:exploit "security vulnerability"	Finds text files detailing security vulnerabilities and exploits.

Table 2-07: Table of Search Queries for Footprinting on the Dark Web

Determining the Operating System

Attackers utilize a range of online resources like Netcraft, Shodan, and Censys to identify the operating system utilized by the target organization. These tools scan the Internet to find connected devices, such as routers, servers, and IoT devices associated with the target organization. By leveraging these tools, attackers gather details such as the city, country, latitude/longitude, hostname, operating system, and IP address of the target organization. This information helps attackers pinpoint possible vulnerabilities and discover effective exploits to execute various attacks on the target.

Netcraft

The method of gathering details about the operating system of a target network is known as OS fingerprinting. Open **https://www.netcraft.com/tools/** in your browser and enter the URL of the target website in the field labeled "What's that site running?"

Attackers utilize the Netcraft tool to discover all the websites linked to the target domain and identify the operating system used at each site.

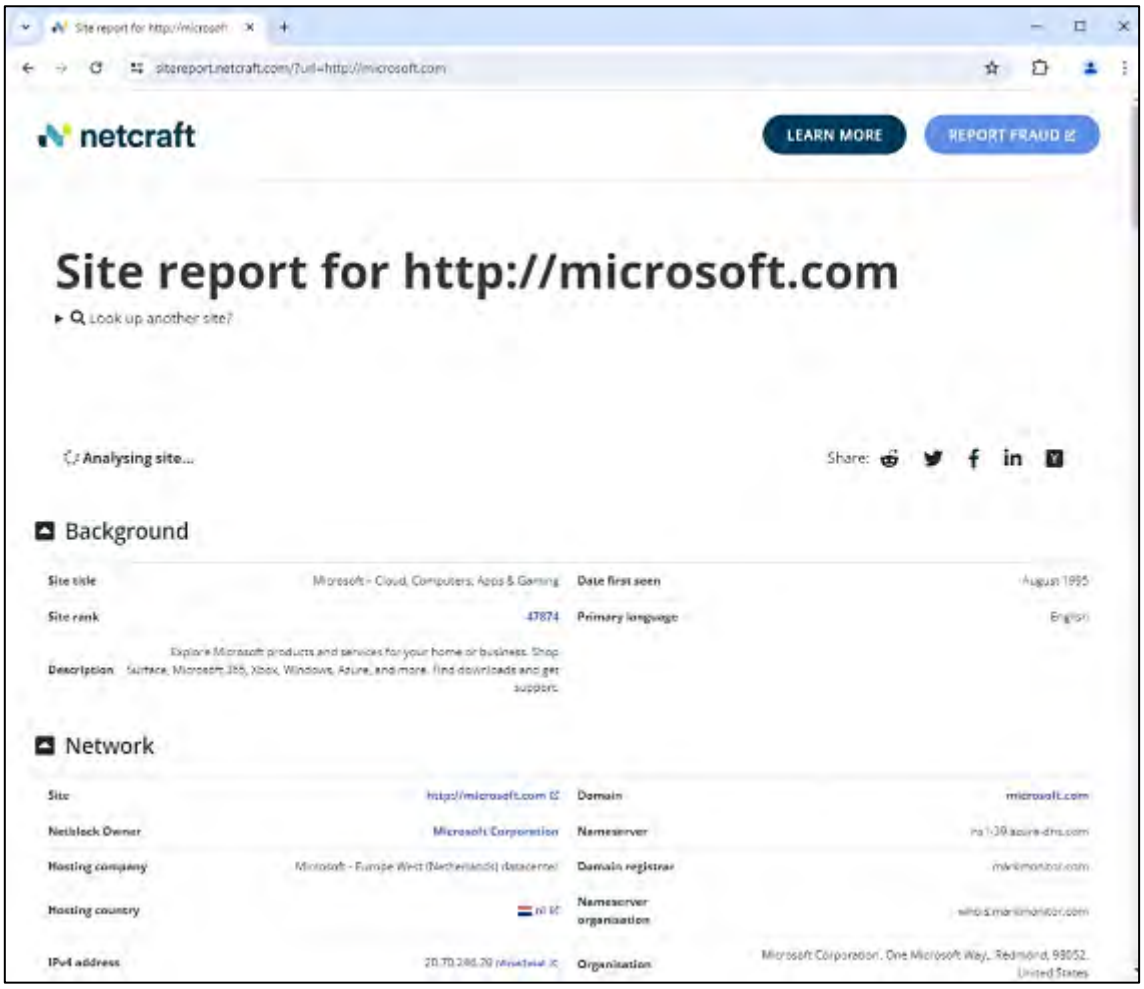

Figure 2-35: Netcraft showing Results for Microsoft

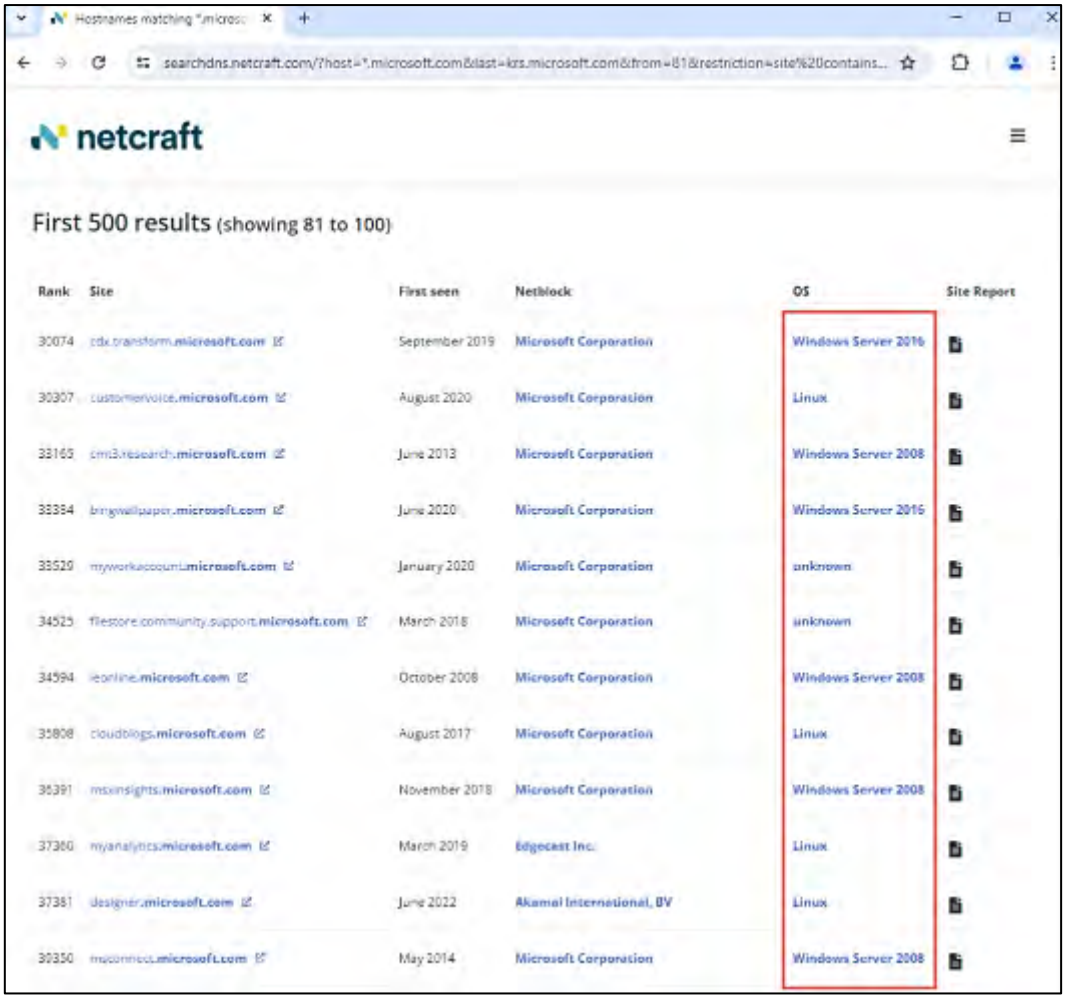

Figure 2-36: Netcraft showing the Target Operating System

SHODAN Search Engine

Shodan is a search engine designed to find devices connected to the Internet, such as routers, servers, and IoT devices. It enables users to identify which devices are online, their locations, and the individuals using them. This tool allows attackers to monitor all devices on a target network that are reachable via the Internet. Additionally, it helps them to locate devices by city, country, latitude/longitude, hostname, operating system, and IP address. Moreover, it assists attackers in searching for known vulnerabilities and exploits across various platforms like Exploit DB, Metasploit, CVE, OSVDB, and Packetstorm through a unified interface. Figure 2-37 illustrates how attackers utilize this tool to identify different target devices that are connected to the Internet and the operating systems they are running.

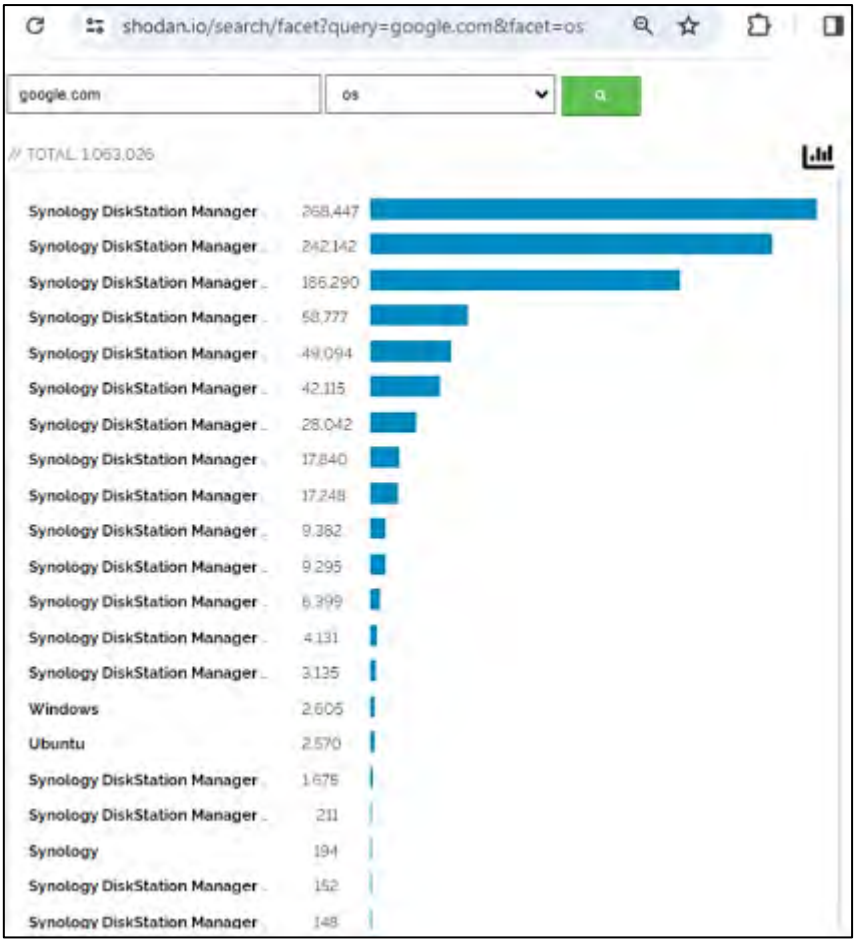

Figure 2-37: SHODAN Search Engine showing Target Operating System

Censys

Censys keeps track of online infrastructure and identifies unrecognized assets throughout the Internet. It offers a comprehensive overview of all servers and devices that are accessible online.

Attackers utilize this tool to observe their target's IT infrastructure, allowing them to find different devices linked to the Internet, along with information such as the operating system in use, IP address, protocols employed, and geographic location.

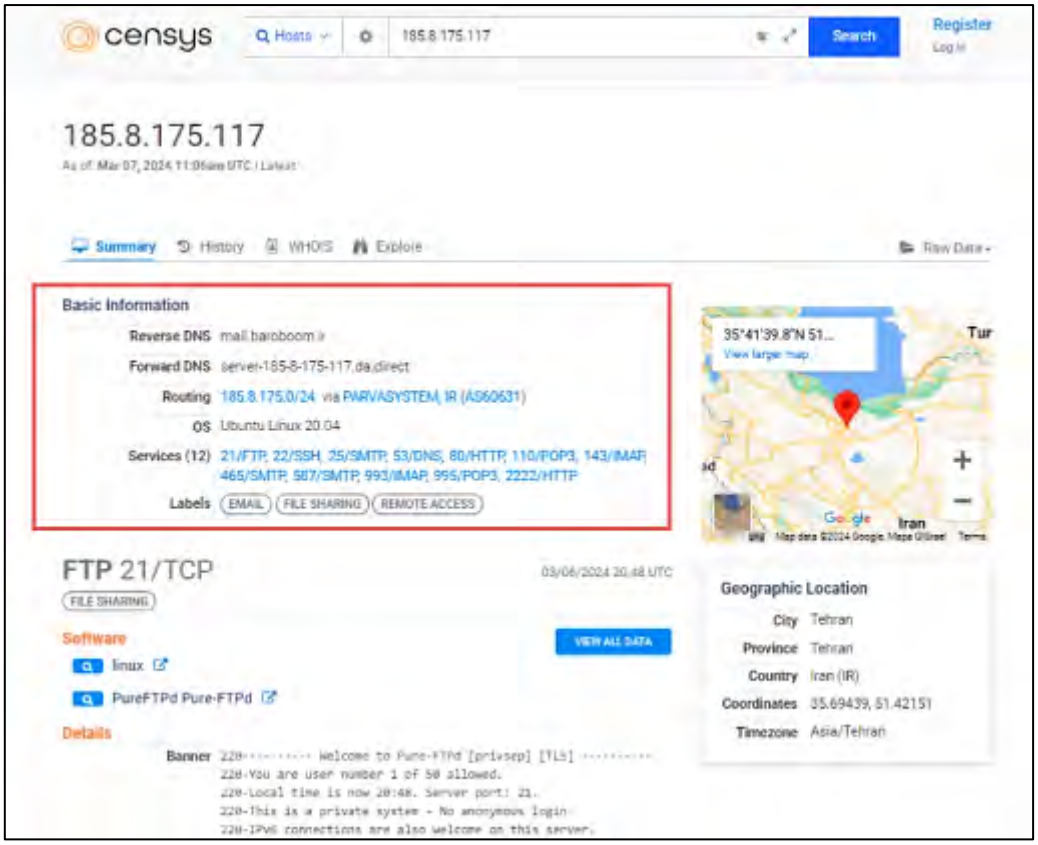

Figure 2-38: Censys Search Engine showing Target Operating System

Competitive Intelligence Gathering

Gathering competitive intelligence refers to the process of identifying, collecting, analyzing, verifying, and utilizing information about competitors from various sources, including the Internet. Competitive intelligence involves gaining insights into other businesses to enhance competitiveness. This approach is more discreet and indirect as opposed to overt methods such as hacking or industrial espionage. It emphasizes the outside business landscape. In this process, professionals collect information ethically and legally rather than in a covert manner.

Competitive intelligence assists in understanding:

- What actions are competitors taking?
- How are competitors marketing their products and services?
- What feedback do customers provide regarding the strengths and weaknesses of competitors?

Organizations conduct competitive intelligence by either hiring individuals to seek out information or by employing commercial database services, which are typically more cost-effective. The information collected can aid company managers and executives in making informed strategic decisions.

Sources of Competitive Intelligence

Competitive Intelligence collection can be executed through either a direct or indirect method.

Direct Approach: The direct method acts as the main source for collecting competitive intelligence. Techniques for the direct method encompass obtaining information from trade exhibitions, employing social engineering with employees and customers, and similar methods.

Indirect Approach: Utilizing an indirect method, data about competitors is acquired through various online resources. Techniques for the indirect method consist of:

- Company websites and employment ads
- Support threads and reviews
- Search engines, Internet, and online database
- Social media postings
- Press releases and annual reports
- Trade journals, conferences, and newspapers
- Patent and trademarks
- Product catalogs and retail outlets
- Analyst and regulatory reports
- Customer and vendor interviews
- Agents, distributors, and suppliers
- Industry-specific blogs and publications
- Legal databases, e.g., LexisNexis
- Business information databases, e.g., D&B Hoovers
- Online job postings
- Financial filings
- Technology solutions, e.g., Crunchbase
- Intellectual property analysis

Competitive Intelligence - When Did this Company Begin? How Did it Develop?

Collecting documents and records from competitors enhances productivity and profitability, which subsequently fosters the company's growth. It aids in uncovering answers to the following:

- **When did it begin?**
  By utilizing competitive intelligence, businesses can obtain the history of a specific company, including its founding date. At times, they acquire vital information that is not easily accessible to others.
- **How did it develop?**
  What strategies does the company employ? Development intelligence may encompass advertising tactics, customer relationship management, and similar aspects.
- **Who leads it?**
  This knowledge enables a company to understand the decision-makers of its rivals.
- **Where is it located?**
  Competitive intelligence also involves the geographic location of the business and insights regarding its various branches and their operations.

Attackers can leverage the information acquired through competitive intelligence to craft a hacking strategy.

Information Resource Sites

EDGAR Database

The Electronic Data Gathering, Analysis, and Retrieval system (EDGAR) automates the collection, validation, indexing, acceptance, and transmission of submissions from companies and other parties mandated by law to file with the U.S. Securities and Exchange Commission

(SEC). Its main goal is to enhance the efficiency and fairness of the securities market to benefit investors, corporations, and the economy by speeding up the reception, acceptance, distribution, and analysis of time-sensitive corporate information submitted to the agency.

D&B Hoovers

D&B Hoovers utilizes a commercial database containing 120 million business records and analytics to provide a sales intelligence solution that allows sales and marketing professionals to target the most suitable prospects, enabling them to achieve immediate business growth.

LexisNexis

LexisNexis offers workflow solutions that are content-enabled and tailored specifically for professionals in legal, risk management, corporate, government, law enforcement, accounting, and academic sectors. It operates an electronic database that contains information related to legal and public records. The platform allows users to access documents and records from legal, news, and business sources. This is advantageous for corporations and government entities looking for data analytics that support compliance, customer acquisition, fraud detection, health outcomes, identity solutions, investigations, receivables management, risk assessment, and optimization of workflows.

Business Wire

Business Wire specializes in the distribution of press releases and regulatory disclosures. This company disseminates complete news releases, images, and other multimedia materials from different organizations worldwide to journalists, news outlets, financial markets, investors, information websites, databases, and the general public. It operates its own patented electronic network for the distribution of news.

Factiva

Factiva is an international news database and a provider of licensed content. It serves as a tool for business information and research, aggregating information from both licensed and free sources and offering features such as searching, alerting, sharing, and management of business information. Factiva's products grant access to over 33,000 sources, including licensed publications, notable websites, blogs, images, and videos. Its resources are accessible from almost every country around the globe in 28 languages, encompassing more than 600 newswires that are continuously updated.

Competitive Intelligence - What Are the Company's Plans?

Information resource sites that aid attackers in obtaining a company's business strategies include:

MarketWatch

MarketWatch monitors market trends for active investors. The platform is a pioneer in delivering business news, personal finance insights, real-time analysis, and investment tools and data, with journalists producing headlines, articles, videos, and market summaries.

The Wall Street Transcript

The Wall Street Transcript functions as both a website and a subscription-based publication that offers industry analyses. It reflects the opinions of fund managers and equity analysts across various sectors. The site also features interviews with company CEOs.

Euromonitor

Euromonitor delivers strategic research capabilities for consumer markets. It offers reports on industries, consumers, and population demographics, providing market analysis and surveys tailored to organizational needs.

Experian

Experian supplies insights into competitors' strategies in search, affiliate, display, and social marketing, enhancing the effectiveness of marketing campaigns. It allows users to:

- Evaluate the success of existing customer acquisition strategies
- Identify the factors contributing to competitors' achievements
- Utilize historical consumer data to predict future trends and swiftly adapt to behavioral changes
- Assess website performance in comparison to industry standards or specific competitors

The Search Monitor

The Search Monitor offers competitive intelligence to track brand and trademark usage, affiliate compliance, and rival advertisers across paid search, organic search, local search, social media, mobile, and shopping engines globally. It assists interactive agencies, search marketers, and affiliate marketers in monitoring ad rankings, ad content, keyword reach, click rates and CPCs, monthly advertising expenditures, market share, trademark use, and affiliate activities.

USPTO

The United States Patent and Trademark Office (USPTO) delivers information related to patent and trademark registrations. It provides general guidance regarding patents and options for accessing patent and trademark databases.

Competitive Intelligence - What Expert Opinions Say About the Company?

Information resource sites that assist attackers in gathering expert insights about a target company include:

SEMRush

SEMRush serves as a competitive keyword research platform. It offers a compilation of Google keywords and AdWords related to any website, in addition to a list of competitors within both organic and paid Google search results. This tool enables a comprehensive understanding of what rivals are advertising and how they allocate their budgets to various online marketing strategies.

ABI/INFORM Global

ABI/INFORM Global is a comprehensive business database. It provides up-to-date business and financial information for researchers. Users of ABI/INFORM Global can analyze business conditions, management techniques, industry trends, management practices and theories, corporate strategies and tactics, as well as the competitive landscape.

SimilarWeb

SimilarWeb collects data from diverse sources to estimate a company's website and mobile app traffic, geographic distribution, and referral information. Additionally, it offers a browser extension that creates a panel for further refining other data sources by anonymously tracking browsing behavior across millions of users globally.

SERanking

SERanking is an online tool for competitor analysis that offers a detailed overview of the website traffic trends of the target organization. This tool aids businesses or users in examining their primary competitors, identifying new entrants, and comparing their keywords with those of rivals. It supports approaches for conducting Pay-Per-Click (PPC) competitive research to analyze competitors' strategies, thereby enhancing one's own advertising tactics.

Other Techniques for Footprinting through Internet Research Services

Finding the Geographical Location of the Target

Details like the geographic location of a company are crucial in the hacking process. Hackers who are aware of an organization's location may engage in techniques such as dumpster diving, monitoring, social engineering, and various other non-technical tactics to collect additional information.

Tools for Finding the Geographical Location

The tools available for determining geographical positions enable users to discover and examine various places on the planet. They offer details like photographs of structures and their environments, including nearby Wi-Fi networks. These resources feature interactive maps, outline maps, satellite images, and guidance on how to engage with and create custom maps. Tools like Google Maps, Apple Maps, and Waze deliver driving directions, current traffic conditions, landmarks, and comprehensive address and contact details. Malicious actors can exploit this information to gain unauthorized access to buildings, both wired and wireless networks and systems.

Attackers may utilize tools such as Google Earth, Google Maps, and Wikimapia to identify or locate building entrances, security cameras, gates, hiding spots, vulnerabilities in perimeter fences, and utility resources like electricity connections, as well as to assess distances between various objects and more.

Google Earth

Attackers utilize the Google Earth tool to pinpoint the precise location of a target. This tool enables them to access 3D images that represent a significant portion of the populated surface of the Earth with great clarity. The level of detail allows them to gather street views, altitude data, and even geographic coordinates.

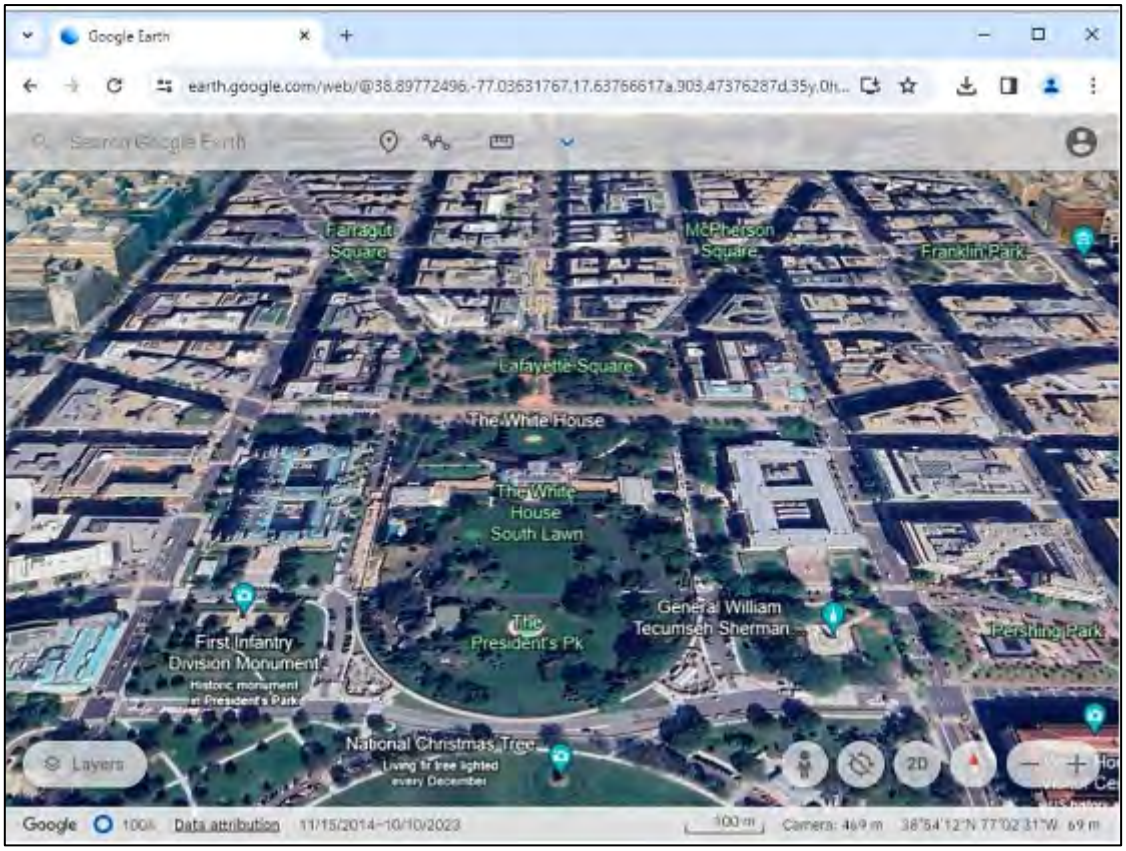

Figure 2-39: Screenshot of Google Earth

Gathering Information from Financial Services

Individuals attempting to access personal or financial information frequently focus on financial data, including stock quotes, charts, financial news, and investment portfolios. Platforms such as Google Finance, MSN Money, Yahoo Finance, and Investing.com offer a wealth of valuable information, including the current market value of a company's shares, company profiles, competitor information, stock exchange rates, corporate press releases, financial reports, along with news and blog articles about companies. The type of information available differs from one platform to another. Financial institutions depend on web services to conduct transactions and allow users to access their accounts. Attackers can gather sensitive and private information about these institutions by employing malware, taking advantage of software vulnerabilities, undermining authentication systems, conducting service flooding attacks, and executing brute force or phishing attacks.

Google Finance

The Google Finance service provides headlines related to businesses and enterprises for various companies, highlighting their financial choices and key news occurrences. Additionally, stock information is accessible, along with price charts that feature indicators for significant news events and corporate activities. The platform also compiles articles from Google News and Google Blog Search concerning each company.

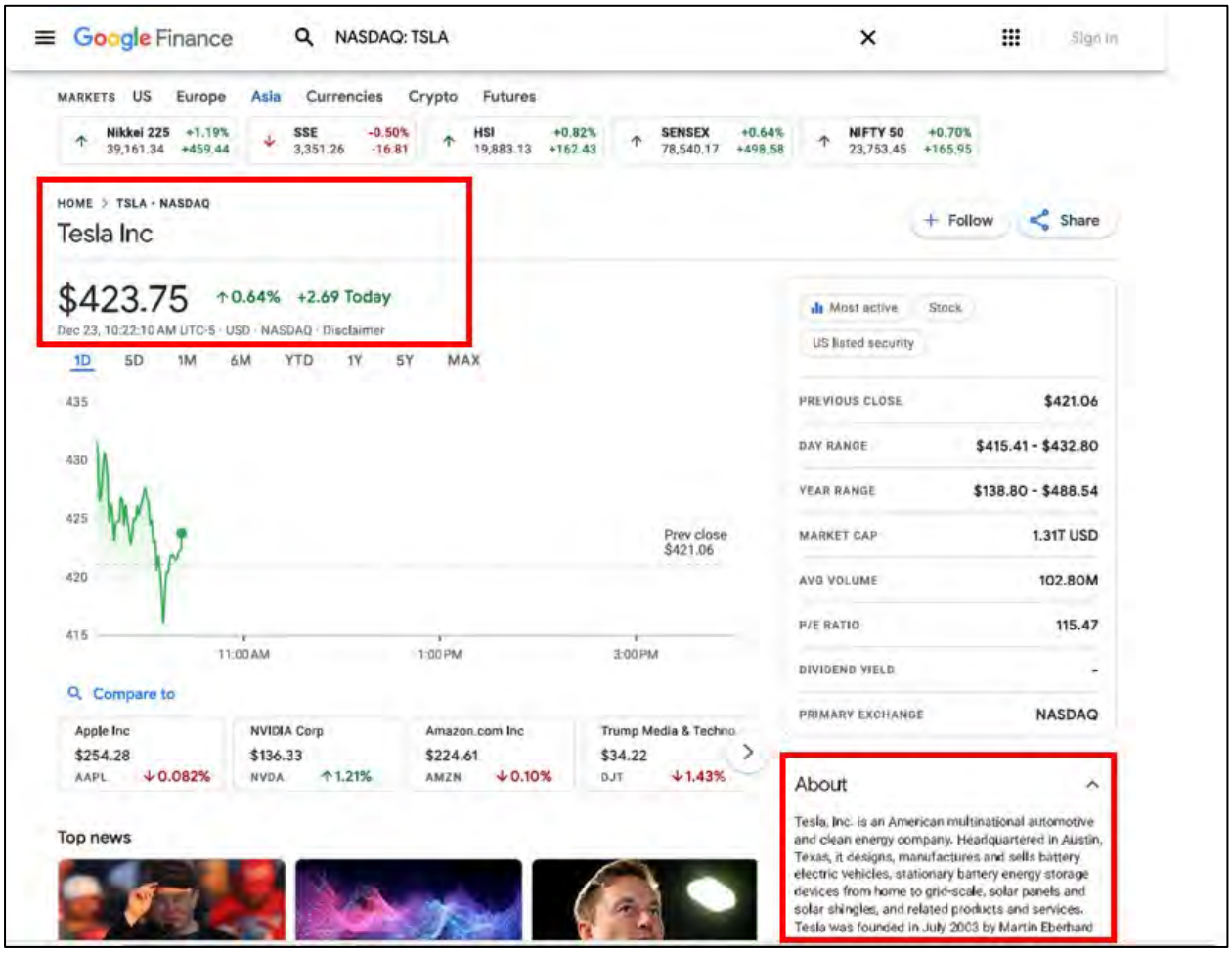

Figure 2-40: Google Finance Service

Gathering Information from Business Profile Sites

Gathering valuable information from corporate websites is an essential task during the information collection phase. These business profile platforms offer details about companies situated in a specific area along with their contact information, accessible to anyone.

Attackers utilize business profile websites like opencorporates, Crunchbase, and corporationwiki to collect crucial details about target organizations, including their locations, addresses, contact details (such as phone numbers and email addresses), employee directories, department titles, services offered, and industry type.

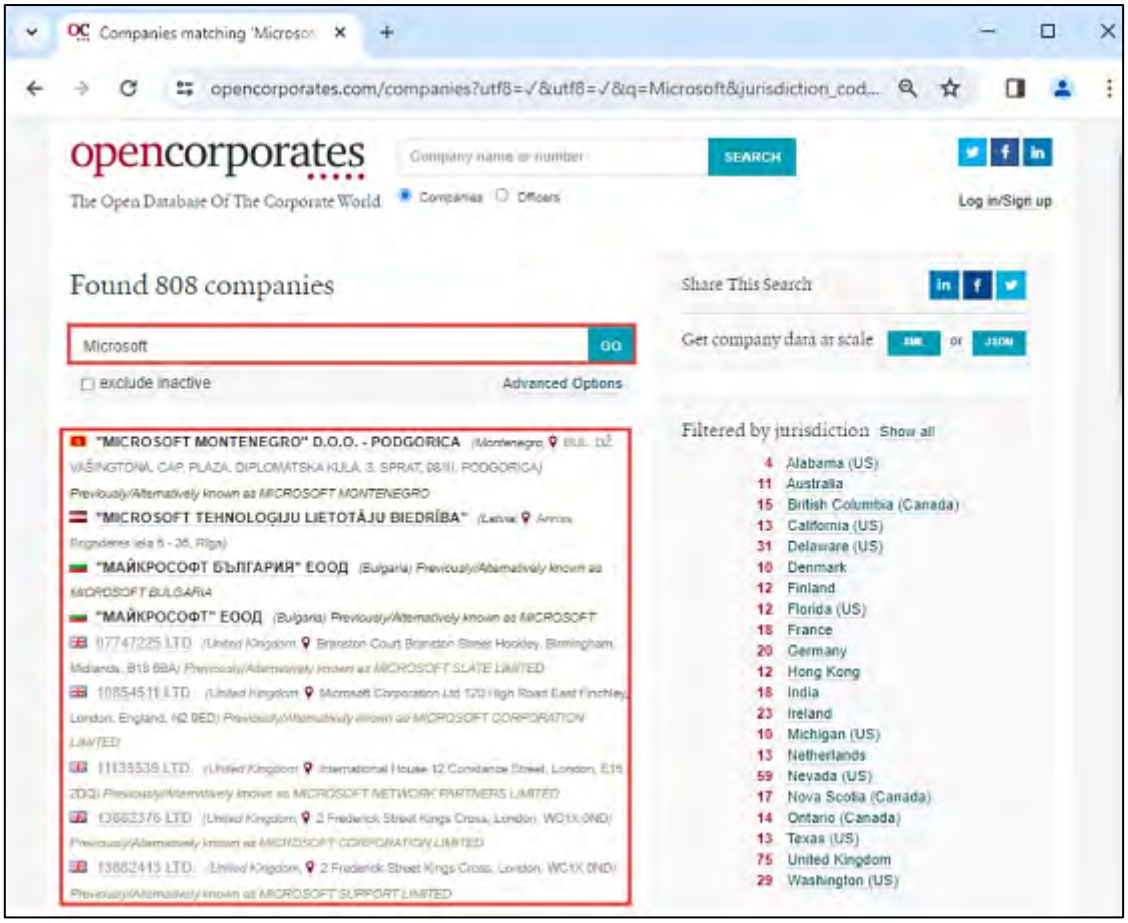

Figure 2-41: Screenshot of opencorporates showing Search Results of Microsoft

Monitoring Targets Using Alerts

Alerts are services that monitor content and deliver automated, current information according to user preferences, typically through email or SMS. To start receiving alerts, a user needs to sign up on the website and supply either an email address or a phone number. Online alert services automatically inform users when new content from news sources, blogs, and discussion forums contains specific search terms chosen by the user. These services offer current updates regarding competitors and the industry.

Tools like Google Alerts, X Alerts, and Giga Alerts assist individuals in tracking mentions of the organization's name, member names, website, or any important people or projects. Attackers can periodically collect updated information about the target through these alert services and utilize it for subsequent attacks.

Google Alerts

Google Alerts automatically informs users when fresh content from news sources, websites, blogs, videos, and discussion groups aligns with the search terms chosen by the user and saved by the Google Alerts service.

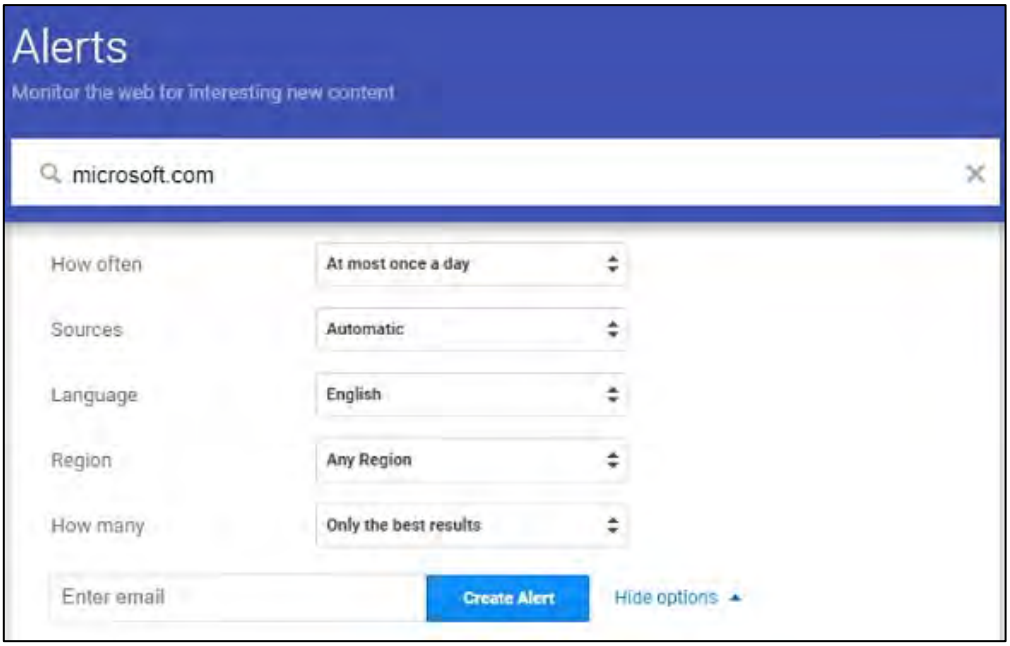

Figure 2-42: Screenshot of Google Alert

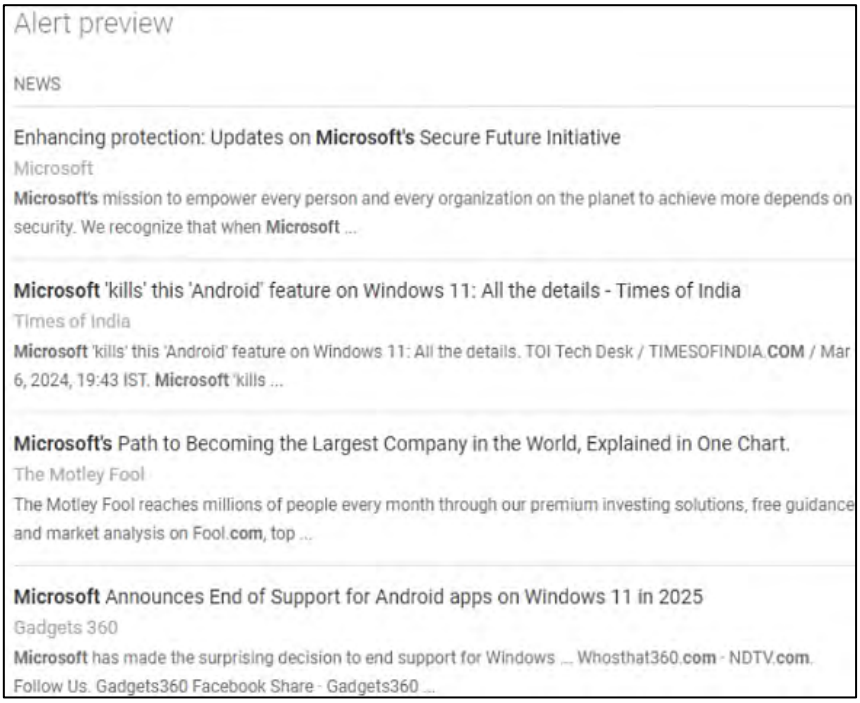

Figure 2-43: Screenshot of Google Alert Preview

Tracking the Online Reputation of the Target

Online Reputation Management (ORM) involves monitoring what appears when individuals search for your company's online reputation. ORM then implements strategies to reduce unfavorable search results or reviews. This approach aids in enhancing brand reputation. Businesses frequently utilize ORM tracking tools to monitor public feedback and subsequently take actions to bolster their credibility and maintain their customers' trust. To effectively manage their online reputation positively, organizations often aim for greater transparency online. This transparency may enable those with malicious intent to gather accurate information about the targeted organization.

Online Reputation Tracking Tools

Online reputation management tools enable us to find out what individuals are saying about the company's brand in real time across various platforms, including the web, social media, and news outlets. They assist in observing, evaluating, and controlling one's online reputation. Attackers can utilize tools like Mention, ReviewPush, and Reputology to monitor the online reputation of businesses.

An attacker may employ ORM tracking tools to:

- Monitor a company's online reputation
- Gather information regarding a company's search engine rankings
- Receive email alerts when a company is mentioned online
- Follow conversations
- Obtain social updates about the targeted organization

Mention

Mention is an online reputation monitoring tool that assists attackers in observing the web, social media, forums, and blogs to gain insights about the target brand and sector. As illustrated in the screenshot, this tool enables attackers to follow online discussions as they occur, no matter where they take place. With Mention, attackers can receive live, current reports sent to any email address instantly.

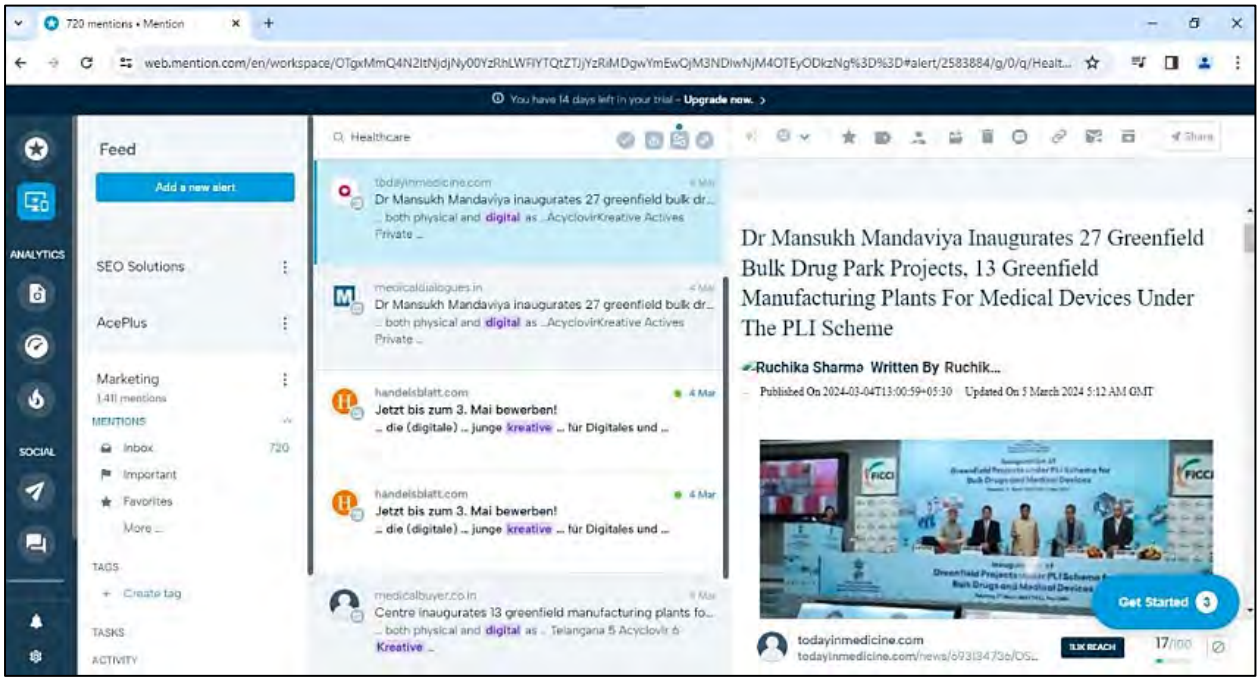

Figure 2-44: Screenshot of Mention

Gathering Information from Groups, Forums, and Blogs

Numerous Internet users participate in blogs, forums, and groups for sharing knowledge. Consequently, attackers frequently target these platforms to gather information about a specific organization and its personnel. Organizations typically neglect to supervise the information that employees disclose to others in group discussions, blogs, and forums. Attackers exploit this vulnerability to compile sensitive data regarding the target, including public network details, system information, and personal employee data. They can create fraudulent profiles on

platforms like Google Groups and LinkedIn Groups. Their goal is to join groups related to the target organization, where they can access personal and business information. Additionally, attackers can look for information in blogs, forums, and groups by using Fully Qualified Domain Names (FQDNs), IP addresses, and usernames.

The types of employee information that an attacker can collect from these online platforms may comprise:

- The employee's full name
- Location of employment and residence
- Home phone number, cellular number, or office number
- Personal and work-related email addresses
- Images of the employee's home or workplace that contain identifiable details
- Photos of employee accolades or upcoming objectives

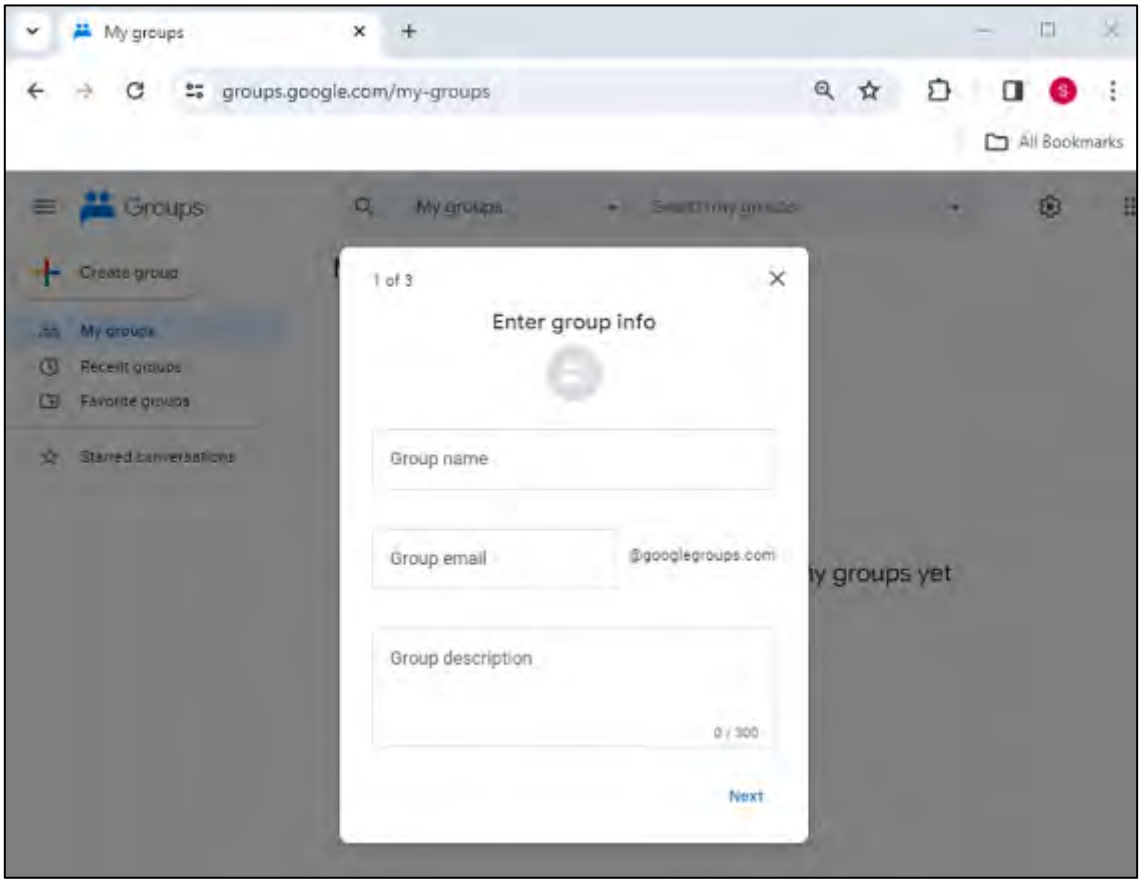

Figure 2-45: Screenshot of Google Groups

Gathering Information from Public Source-Code Repositories

Source code repositories are online platforms or tools that can be hosted on internal servers or third-party websites like GitHub, GitLab, SourceForge, and BitBucket. These platforms hold sensitive information such as configuration files, private Secure Shell (SSH) and Secure Sockets Layer (SSL) keys, source code, dynamic libraries, and software tools developed by contributors, which attackers can exploit to execute attacks on the target organization.

Software developers often store a vast amount of source code related to websites and applications online, either publicly or privately, for future reference. To adhere to deadlines and enhance productivity during product development, developers might tap into the data within

the repositories to quickly create or update their applications, thereby reducing development time and costs.

Websites hosting source code, like GitHub, may have security vulnerabilities that could enable attackers to target various applications. The applications stored on these public platforms might also include confidential files and information that help attackers locate and identify the developers and the technologies utilized.

The data gathered from these public repositories might not be enough for executing direct attacks; however, when combined with certain active footprinting techniques, it can enable attackers to carry out targeted spear phishing attacks on specific users or employees of the targeted organization. Additionally, vulnerabilities in these repositories could provide attackers with valuable information to conduct social engineering and infrastructure attacks against the target organization.

Attackers can utilize tools such as Recon-ng to discover public source-code repositories.

Recon-ng

Recon-ng is a comprehensive reconnaissance framework intended to create an effective space for performing web-based reconnaissance in a swift and detailed manner. It aids attackers in collecting data from publicly available source-code repositories.

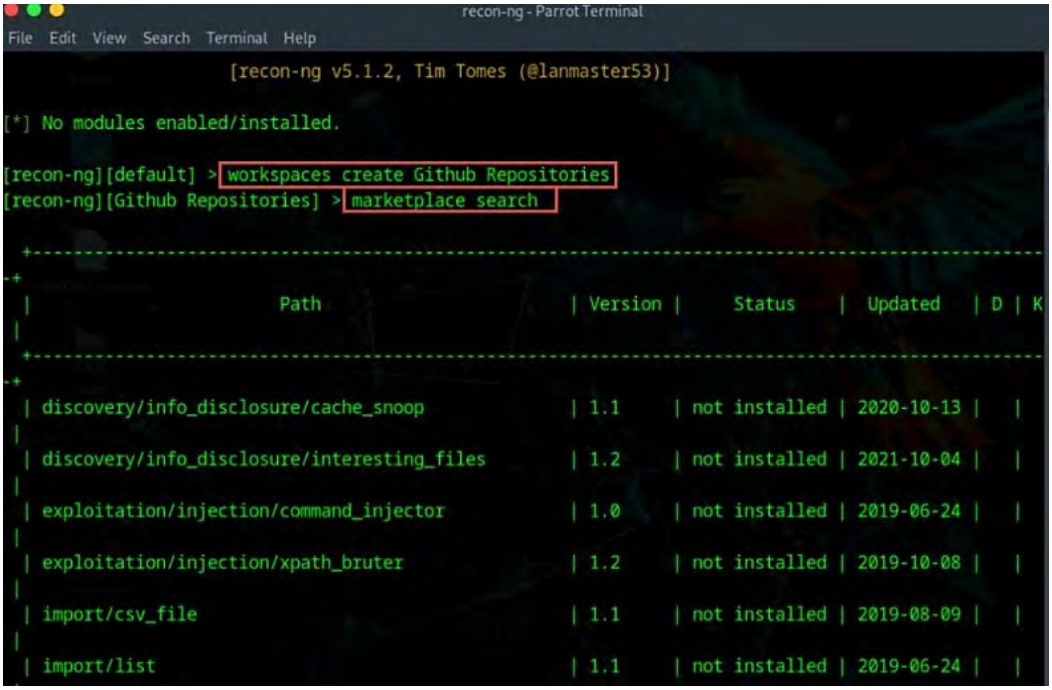

Figure 2-46: Screenshot of recon-ng showing the Creation of a Workspace and a Marketplace Search

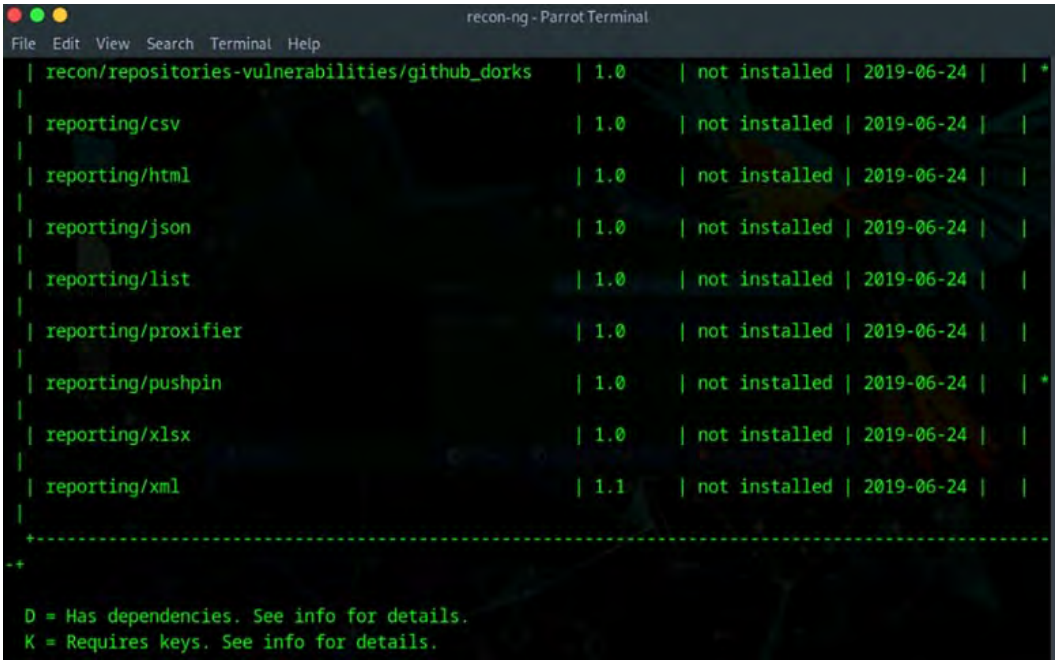

Figure 2-47: Screenshot of recon-ng displaying Source-Code Repositories

Footprinting through Social Networking Sites

When it comes to gathering information, footprinting via social networking sites differs from footprinting through social engineering in several ways. In social engineering footprinting, an attacker deceives individuals into disclosing data, while footprinting via social networking sites involves collecting information that is publicly accessible on those platforms. Additionally, attackers can leverage social networking sites to carry out social engineering attacks.

People Search on Social Networking Sites

Finding a specific individual on a social networking platform is quite simple. Social networking services are online platforms or websites designed to facilitate the creation of social connections or relationships among individuals. These platforms feature information supplied by users in their profiles. They enable direct or indirect connections between people through various areas, such as shared interests, workplace locations, and educational backgrounds.

Social networking sites permit individuals to quickly share information, as they can modify their details instantly. These platforms enable users to update information about upcoming or ongoing events, recent announcements, invitations, and more. Social networking websites serve as a valuable resource for locating individuals and their associated information. Many social networking sites permit visitors to search for individuals without having to register, making the task of searching for people simple and anonymous. A user can search for someone using their name, email, or address. Certain sites allow users to verify if an account is active, which provides insights into the status of the person being searched for.

Platforms like Facebook, Twitter, LinkedIn, and Instagram offer the ability to find people through their name, keywords, company, school, friends, colleagues, and individuals in close proximity. Looking for people on these platforms yields personal details such as name, job title, company affiliation, current location, and educational background. Furthermore, you can discover professional information such as company or business name, current location, phone number, email address, photos, videos, and more. Social networking platforms like Twitter are

utilized for disseminating advice, news, concerns, opinions, rumors, and facts. By searching for individuals on social networking services, an attacker can collect essential information that can assist in carrying out social engineering or other types of attacks.

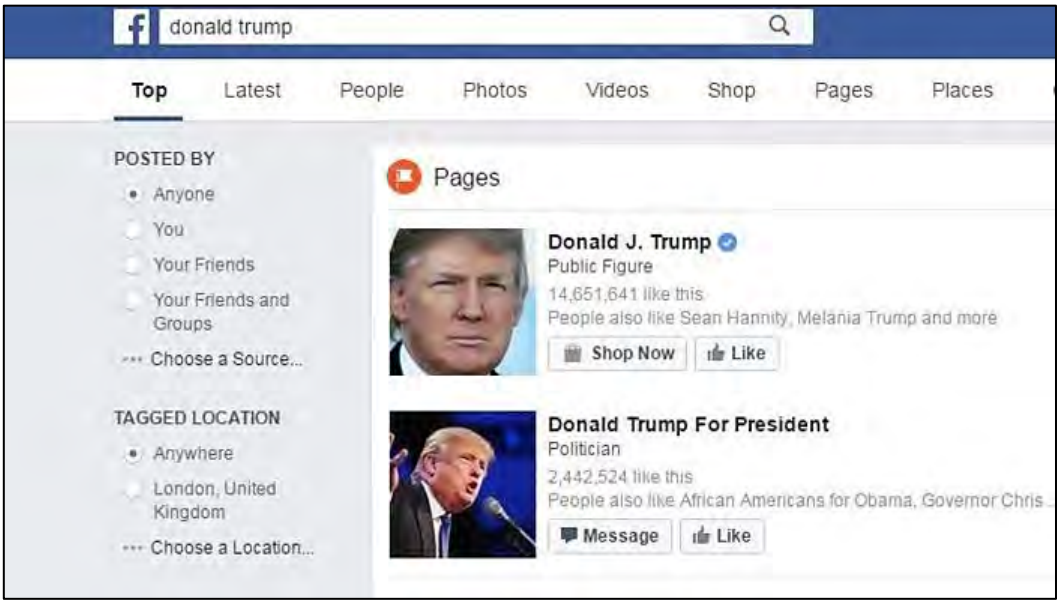

Figure 2-48: Screenshot of Facebook showing Search Results

Gathering Information from LinkedIn

LinkedIn is a networking platform designed for professionals. It links individuals worldwide to enhance productivity and achieve success. The site features personal details like name, job title, company name, current location, educational background, and more. Data collected from LinkedIn can assist an attacker in executing social engineering or various other forms of attacks.

Attackers can utilize theHarvester tool to collect information from LinkedIn related to the specified organization name.

theHarvester

theHarvester is a tool meant for use during the initial phases of a penetration test. It assists in collecting open-source intelligence and helps assess a company's external threat landscape online. Attackers utilize this tool to gather information from the LinkedIn platform to discover employees of the targeted organization along with their job roles. As illustrated in the provided screenshot, the attacker executes the following command to list users on LinkedIn:

```
theHarvester -d microsoft -l 200 -b linkedin
```

Table 2-08 describes each option used in the command.

Command Breakdown	Description
-d	Indicates the domain or company name to be searched.
-l	Denotes the number of results to retrieve.
-b	Specifies LinkedIn as the data source.

Table 2-08: theHarvester Command Description

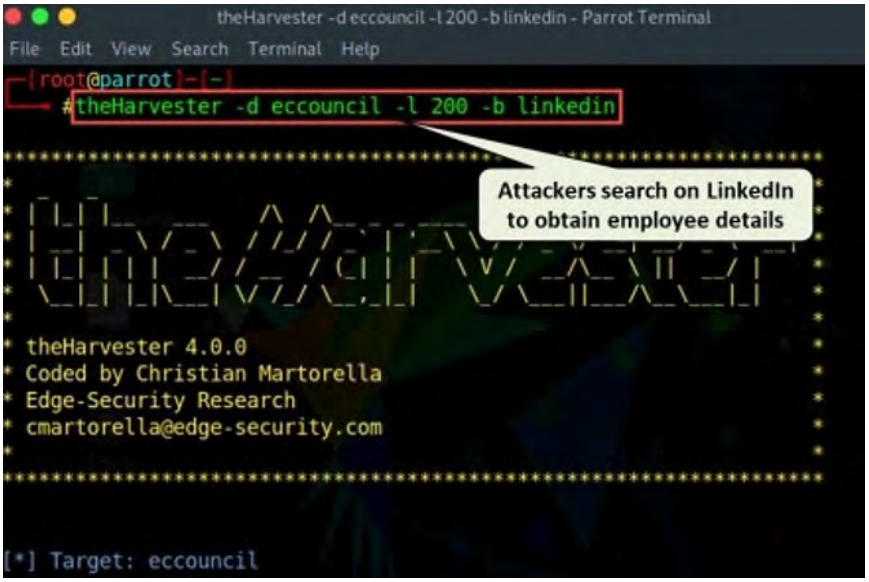

Figure 2-49: Screenshot showing theHarvester Command to Enumerate Users on LinkedIn

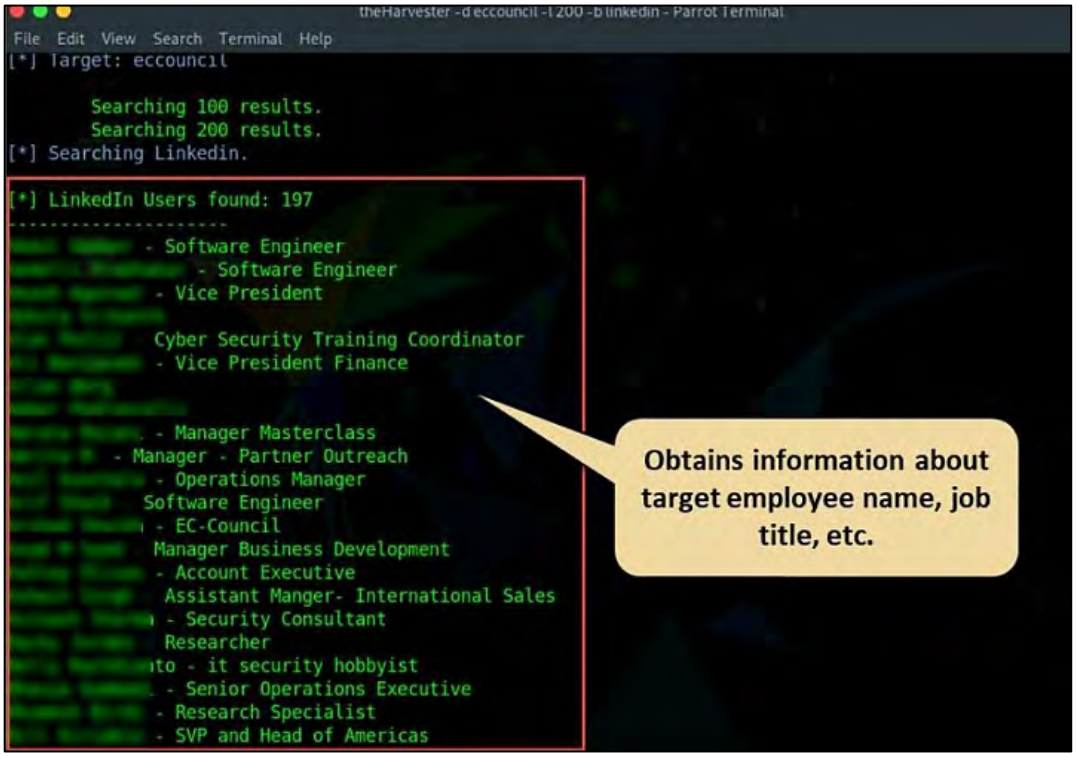

Figure 2-50: Screenshot showing theHarvester Search Results from LinkedIn

Harvesting Email Lists

Collecting email addresses associated with a target organization plays a crucial role in the later stages of hacking. Attackers may utilize automated software like theHarvester and Email Spider to gather publicly accessible email addresses of the organization's employees. These tools compile email lists related to a specific domain by searching through search engines such as Google, Bing, and Yahoo. Attackers leverage these email lists and usernames to execute social engineering and brute force attacks against the target organization.

theHarvester

Attackers utilize theHarvester tool to gather email addresses associated with the target domain. For instance, attackers execute the following command to collect email addresses from microsoft.com using the Baidu search engine:

```
theharvester -d microsoft.com -l 200 -b baidu
```

Table 2-09 describes each option used in the command.

Command Breakdown	Description
-d	Indicates the domain for email harvesting.
-l	Restricts the results to 200.
-b	Directs theHarvester to obtain results from the Baidu search engine.

Tale 2-09: theHarvester Command Description

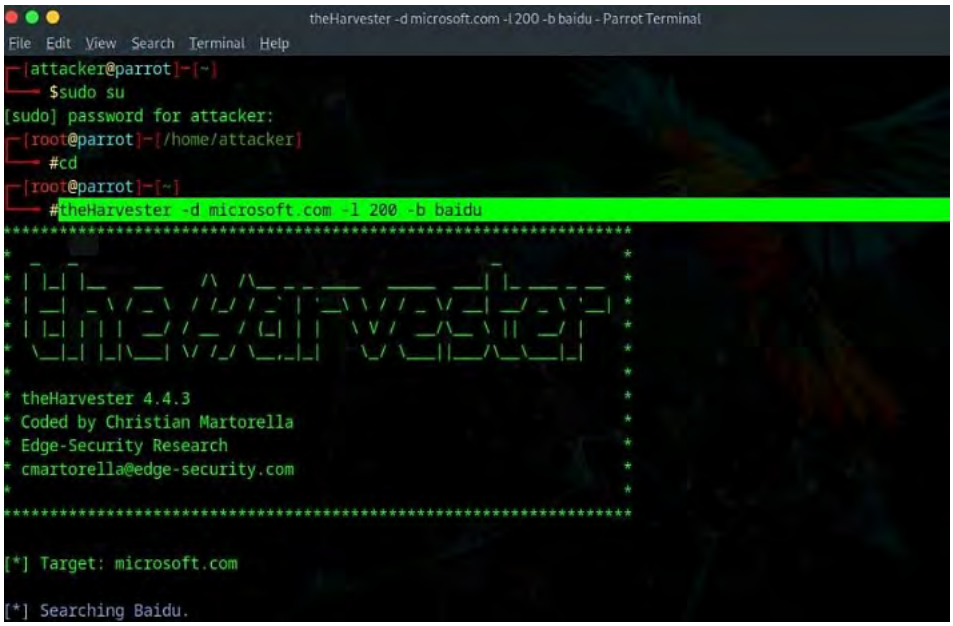

Figure 2-51: Screenshot showing theHarvester Command to Extract Email Addresses

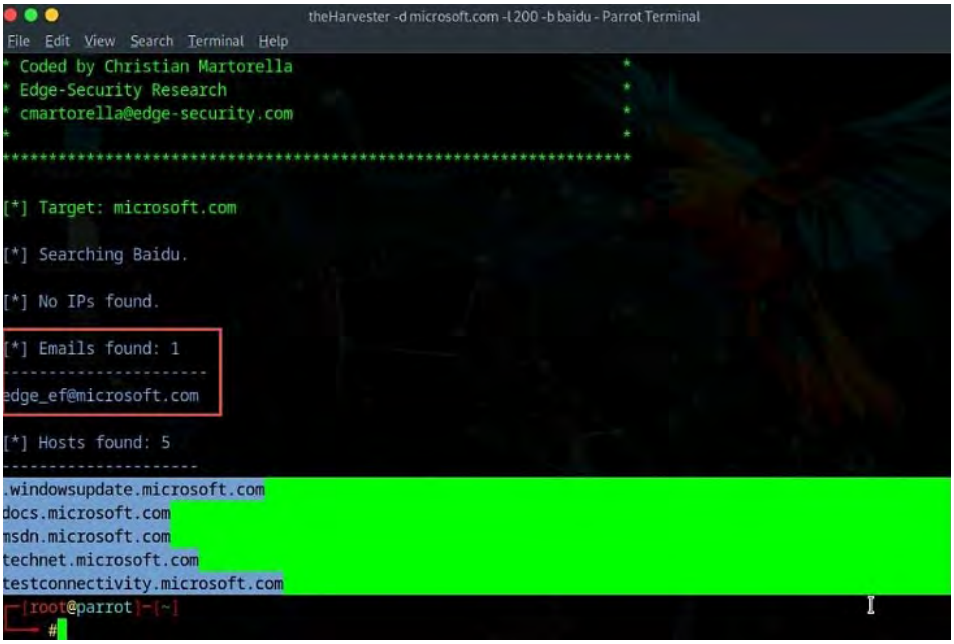

Figure 2-52: Screenshot showing the Email List extracted by theHarvester

Harvesting Email Lists with AI

Attackers can utilize AI-driven tools to improve and automate their reconnaissance efforts. With the support of AI, these attackers can easily uncover important email information of their intended target.

For instance, an attacker might employ ChatGPT to carry out this activity by utilizing a suitable prompt such as:

> "Use theHarvester to gather email accounts associated with 'microsoft.com', limiting results to 200, and leveraging 'baidu' as a data source"
>
> To collect email addresses linked to "microsoft.com" using theHarvester while limiting the results to 200 and using "Baidu" as the source of information, you can execute the following command:
>
> theHarvester -d microsoft.com -l 200 -b Baidu -f Microsoft_emails.xml

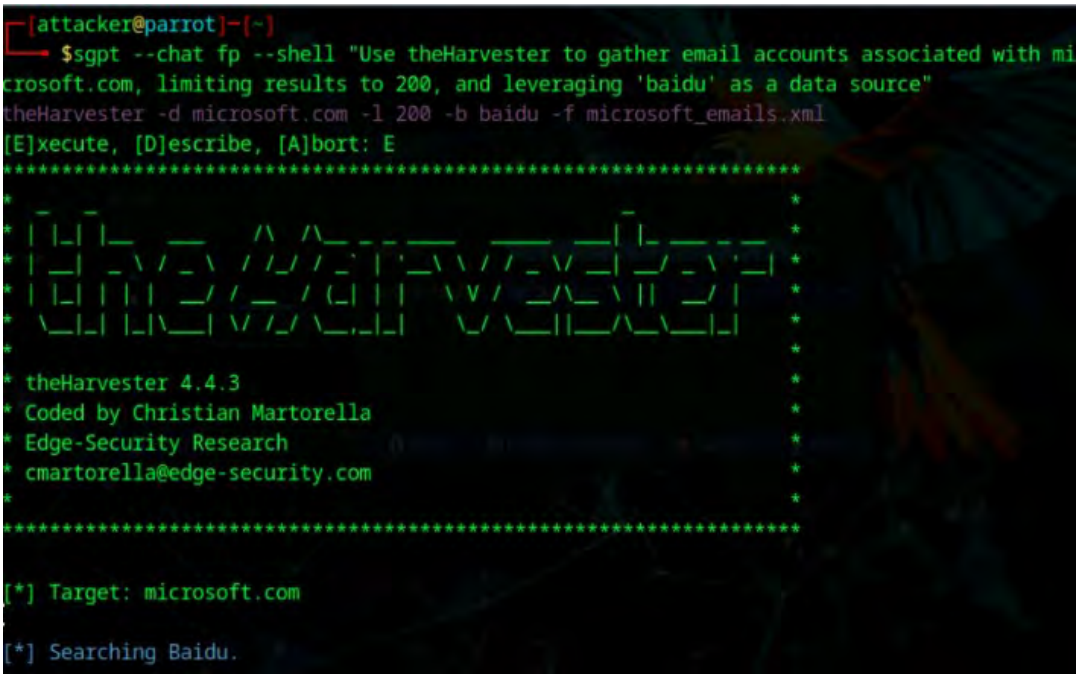

Figure 2-53: Use theHarvester to Gather Email Accounts Associated with 'microsoft.com'

Table 2-10 describes each option used in the command.

Command Breakdown	Description
theHarvester	Invoke theHarvester tool.
-d microsoft.com	Specifies the target domain, in this case, "microsoft.com", for which we want to gather email accounts.
-l 200	Limits the number of results to 200, ensuring a manageable output.
-b baidu	Specifies "baidu" as the data source to be utilized for gathering email accounts associated with "microsoft.com".
-f Microsoft_emails.xm	Specifies the output file name where the email accounts will be saved. In this case, the file is named "Microsoft_emails.xml".

Table 2-10: theHarvester Command Description

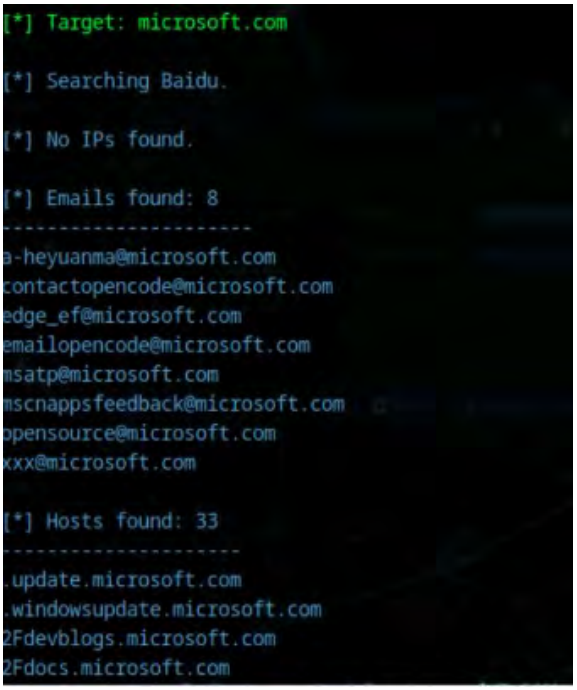

Figure 2-54: Gathered Email Accounts Associated with 'microsoft.com' using theHarvester

Analyzing Target Social Media Presence

Numerous online tools and resources exist to collect important information about a target from one or multiple social media platforms. These tools enable attackers to uncover the most shared content across social media platforms by utilizing hashtags or keywords, monitoring accounts and URLs on different social media services, and retrieving a target's email address, among other things. This gathered information assists attackers in executing phishing, social engineering, and other forms of attacks.

Attackers utilize tools like BuzzSumo, Google Trends, Hashatit, and Ubersuggest to find data on social media channels.

BuzzSumo

BuzzSumo's sophisticated social search tool identifies the content that has been most frequently shared about a specific topic, author, or website. It displays sharing activity from all major social platforms, including Twitter, Facebook, LinkedIn, Google Plus, and Pinterest.

As illustrated in Figure 2-55, attackers utilize BuzzSumo to observe the most popular content related to their target domain, allowing them to gather information such as social media account details, URLs, and email addresses.

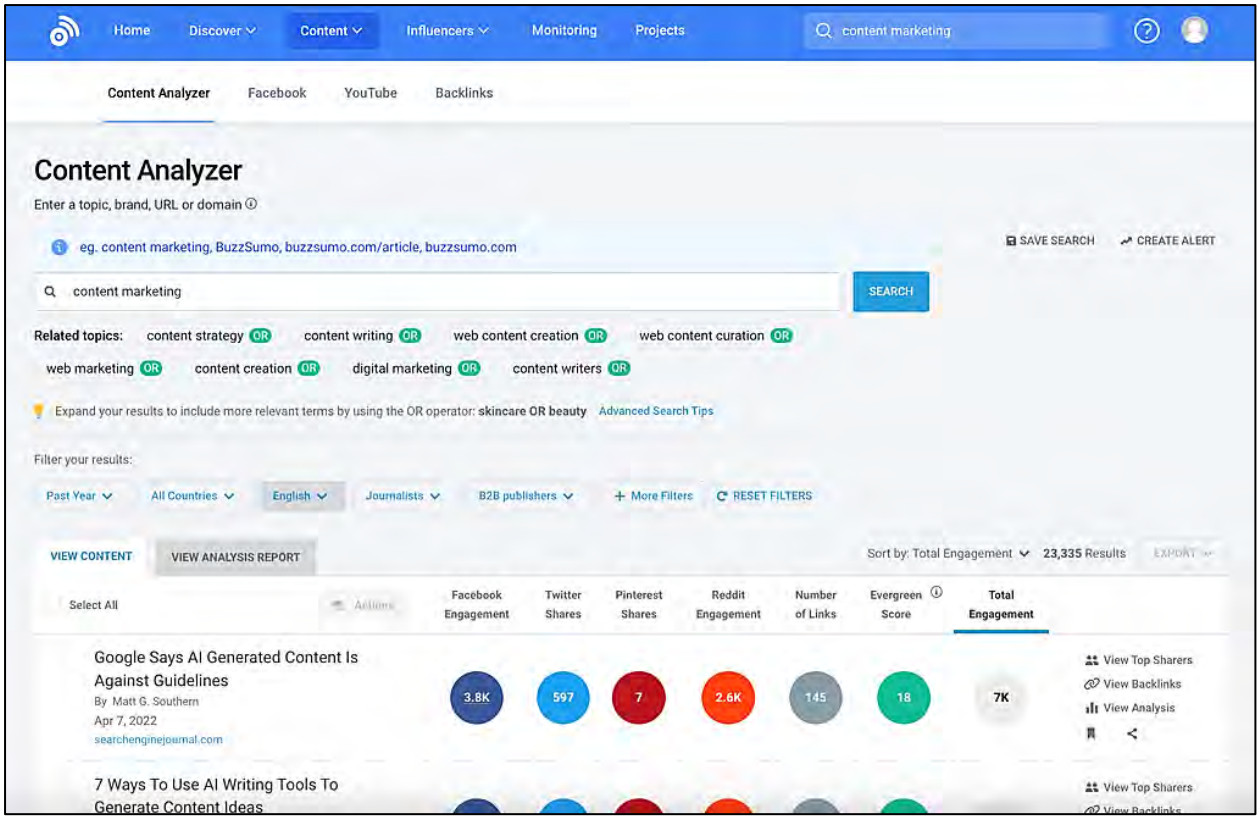

Figure 2-55: Screenshot of BuzzSumo showing the Shared Content

Tools for Footprinting through Social Networking Sites

Hackers employ different tools like Sherlock and Social Searcher to survey social networking platforms, including Twitter, Instagram, Facebook, and Pinterest, in order to collect sensitive information about their targets, such as date of birth, educational background, job status, names of family members, and details about the organization they are affiliated with, which may include business strategies, potential clients, and future project plans.

Sherlock

As shown in Figure 2-56, attackers utilize Sherlock to scan numerous social networking platforms for a specific username. This tool assists the attacker in identifying the target user across different social networking sites, along with the full URL.

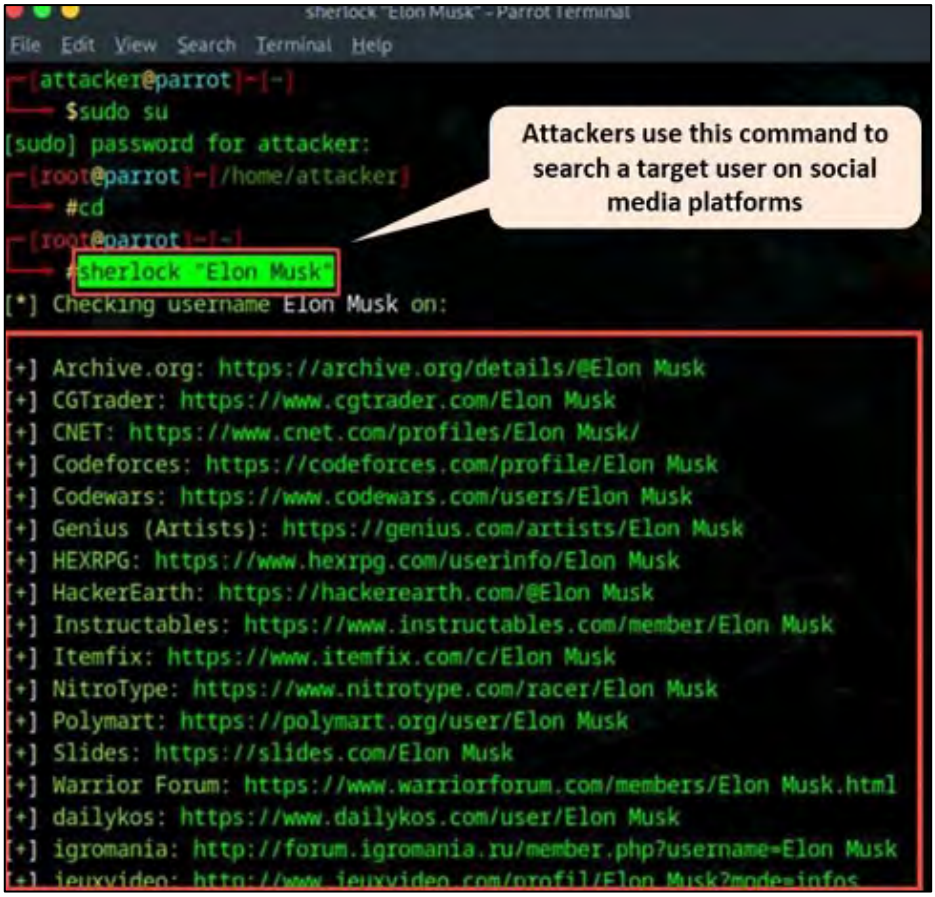

Figure 2-56: Result of Sherlock Tool

Social Searcher

Social Searcher enables individuals to search for information on social media platforms in real time and offers detailed analytical data. Malicious users utilize this tool to monitor a target individual across different social networking sites and gather details such as full URLs to their profiles, their posts, and other personal details.

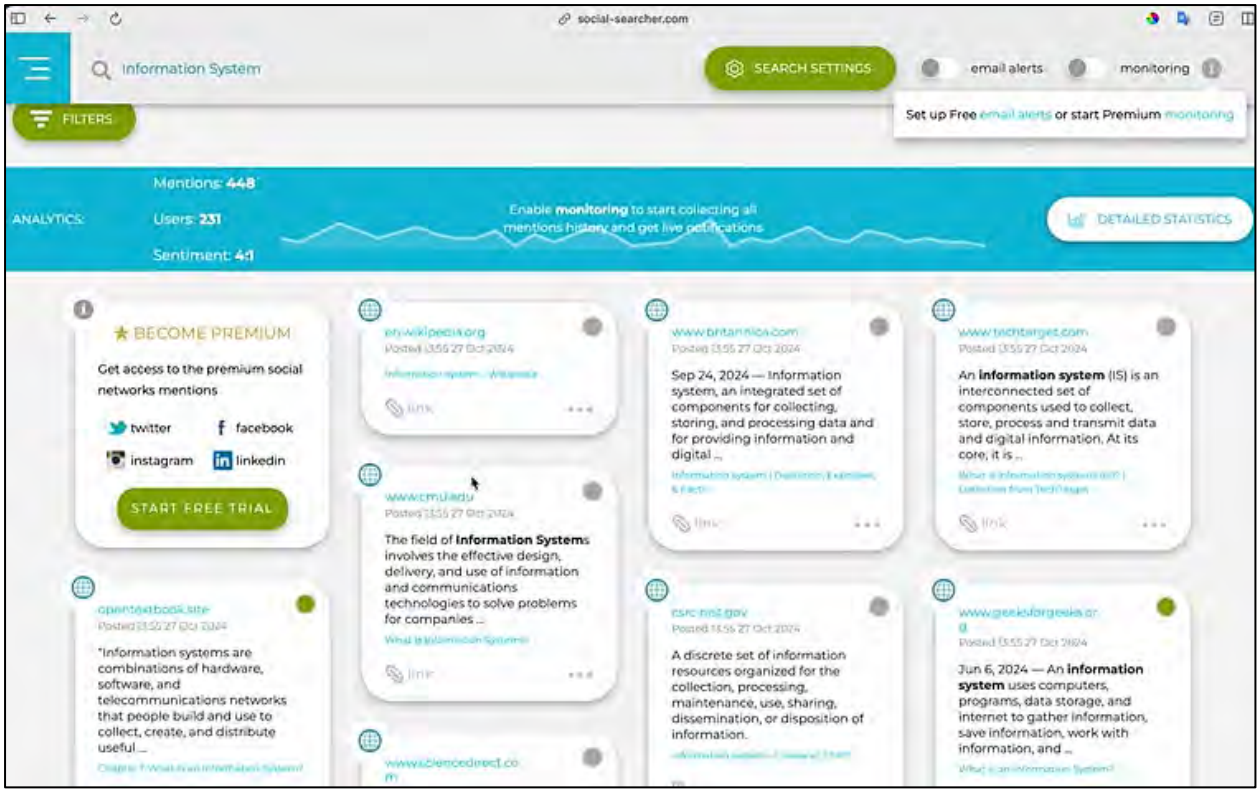

Figure 2-57: Screenshot of Social Searcher

Footprinting through Social Networking with AI

With the help of AI, attackers can easily leverage social media platforms to gather important information about their victims.

For instance, a perpetrator could utilize ChatGPT to accomplish this by crafting a suitable prompt like:

> "Use Sherlock to gather personal information about Sundar Pichai and save the result in recon2.txt"
>
> To search for online accounts linked to the name "Sundar Pichai" using Sherlock and save the results in a file called "recon2", you can execute the following command:
>
> sherlock SundarPichai --output recon2

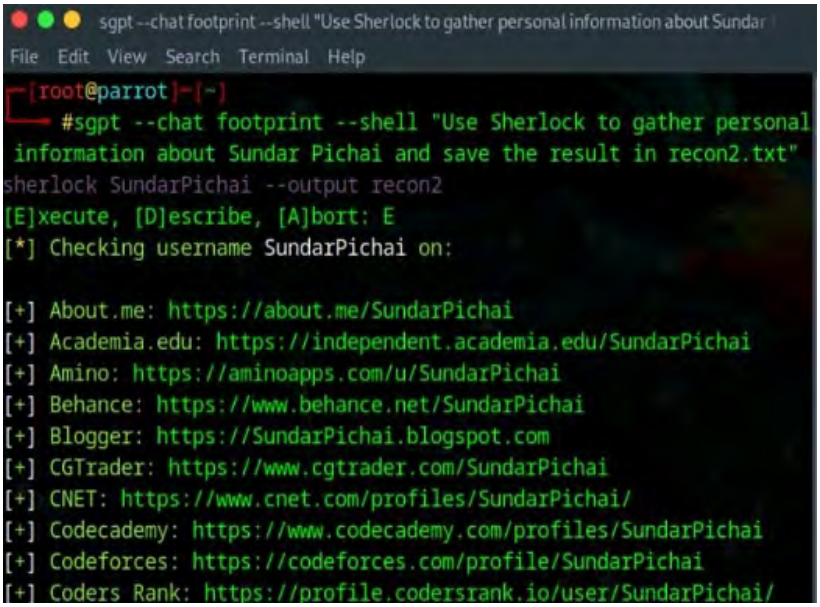

Figure 2-58: Search for Online Accounts Associated with the Name "Sundar Pichai"

Table 2-11 describes each option used in the command.

Command Breakdown	Description
sherlock	Runs Sherlock (a tool for searching for online accounts across various platforms).
SundarPichai	Name for which we want to search online accounts.
--output recon2	Specifies the output file where the results will be saved (recon2).

Table 2-11: sherlock Command Description

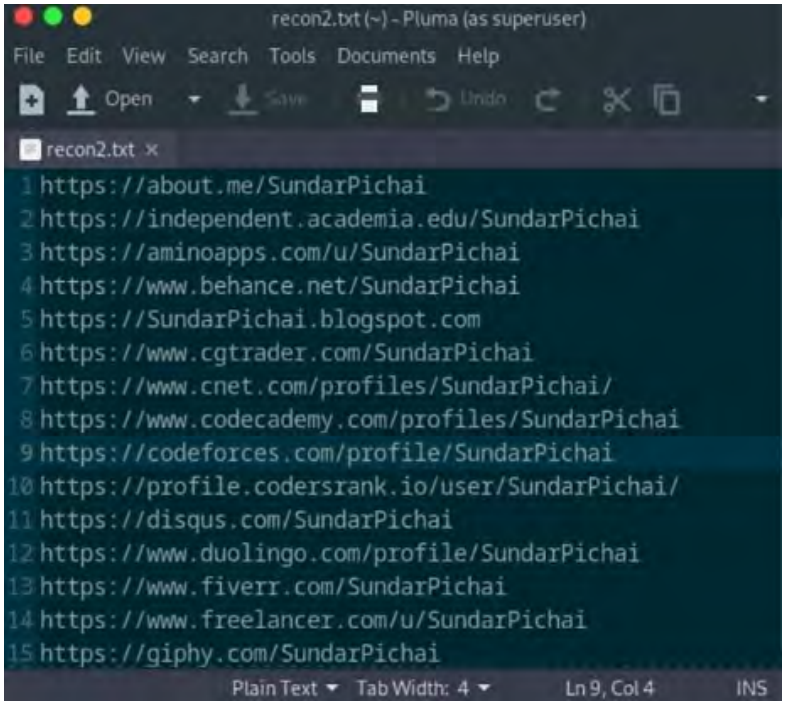

Figure 2-59: Output File

Whois Footprinting

Gathering network information, such as Whois data, is crucial for planning an attack. This section covers Whois footprinting, which involves collecting domain details about an organization's owner, registrar, registration, name server, and contact info. We will look at how to perform Whois lookups, analyze results, find IP geolocation, and the tools for gathering this information.

Whois Lookup

Whois is a protocol used for querying databases that contain information about the registered users or assignees of an Internet resource, such as domain names, IP address blocks, or autonomous systems. This protocol operates on TCP port 43 to listen for requests. Regional Internet Registries (RIRs) oversee Whois databases that hold personal details of domain owners. For each resource, the Whois database provides text records that include information about the resource itself as well as relevant data concerning assignees, registrants, and administrative details (such as creation and expiration dates).

There are three types of data models used to store and retrieve Whois information:

1. **Thick Whois (Distributed Model):** Retains complete Whois information from all registrars for a specific set of data.
2. **Thin Whois (Centralized Model):** Keeps only the name of the Whois server for a domain's registrar, which contains the full details of the data being queried.
3. **Decentralized Whois:** Maintains complete Whois information with various independent entities managing the Whois database.

The Whois query yields the following details:

- Domain name details
- Domain registrar
- Contact details of the domain owner

- Domain name servers NetRange
- When a domain has been created
- Expiry records Records last updated
- Domain status (available, registered, or suspended)
- IP address information

An attacker can send a query to a Whois database server to gather information about the target domain, and the Whois server answers the request with the sought-after details. With this information, an attacker can construct a diagram of the organization's network, deceive domain owners through social engineering tactics, and acquire internal information about the network.

Regional Internet Registries (RIRs)

The Regional Internet Registries (RIRs) consist of the following:

- American Registry for Internet Numbers (ARIN) (https://www.arin.net)
- African Network Information Center (AFRINIC) (https://www.afrinic.net)
- Asia Pacific Network Information Center (APNIC) (https://www.apnic.net)
- Réseaux IP Européens Network Coordination Centre (RIPE) (https://www.ripe.net)
- Latin American and Caribbean Network Information Center (LACNIC) (https://www.lacnic.net)

Whois Lookup Result

Whois services like https://whois.domaintools.com and https://www.tamos.com can assist in conducting Whois lookups. The image below displays the outcome of a Whois lookup performed using both of the previously mentioned Whois services. These services execute a Whois lookup by inputting the desired domain or IP address. Batch IP Converter, found at http://www.sabsoft.com, offers details about an IP address, hostname, or domain, including the country, state or province, city, phone number, fax number, and the names of the network provider, administrator, and technical support contact. It accommodates Internationalized Domain Names (IDNs), meaning users can search for domain names that contain non-English characters. Additionally, it is compatible with IPv6 addresses.

Whois Record for CertifiedHacker.com

− Domain Profile

Registrant	PERFECT PRIVACY, LLC
Registrant Country	us
Registrar	Network Solutions, LLC IANA ID: 2 URL: http://networksolutions.com Whois Server: whois.networksolutions.com domain.operations@web.com (p) 18777228662
Registrar Status	clientTransferProhibited
Dates	7,164 days old Created on 2002-07-29 Expires on 2022-07-29 Updated on 2021-08-22
Name Servers	NS1.BLUEHOST.COM (has 2,681,575 domains) NS2.BLUEHOST.COM (has 2,681,575 domains)
Tech Contact	PERFECT PRIVACY, LLC 5335 Gate Parkway care of Network Solutions PO Box 459, Jacksonville, FL, 32256, us kq9t994x73e@networksolutionsprivateregistration.com (p) 15707088622
IP Address	162.241.216.11 - 1,745 other sites hosted on this server
IP Location	- Utah - Provo - Unified Layer
ASN	AS26337 OIS1, US (registered Oct 09, 2013)
Domain Status	Registered And Active Website
IP History	13 changes on 13 unique IP addresses over 16 years
Registrar History	3 registrars with 2 drops
Hosting History	6 changes on 4 unique name servers over 19 years

Figure 2-60: Screenshot of Whois

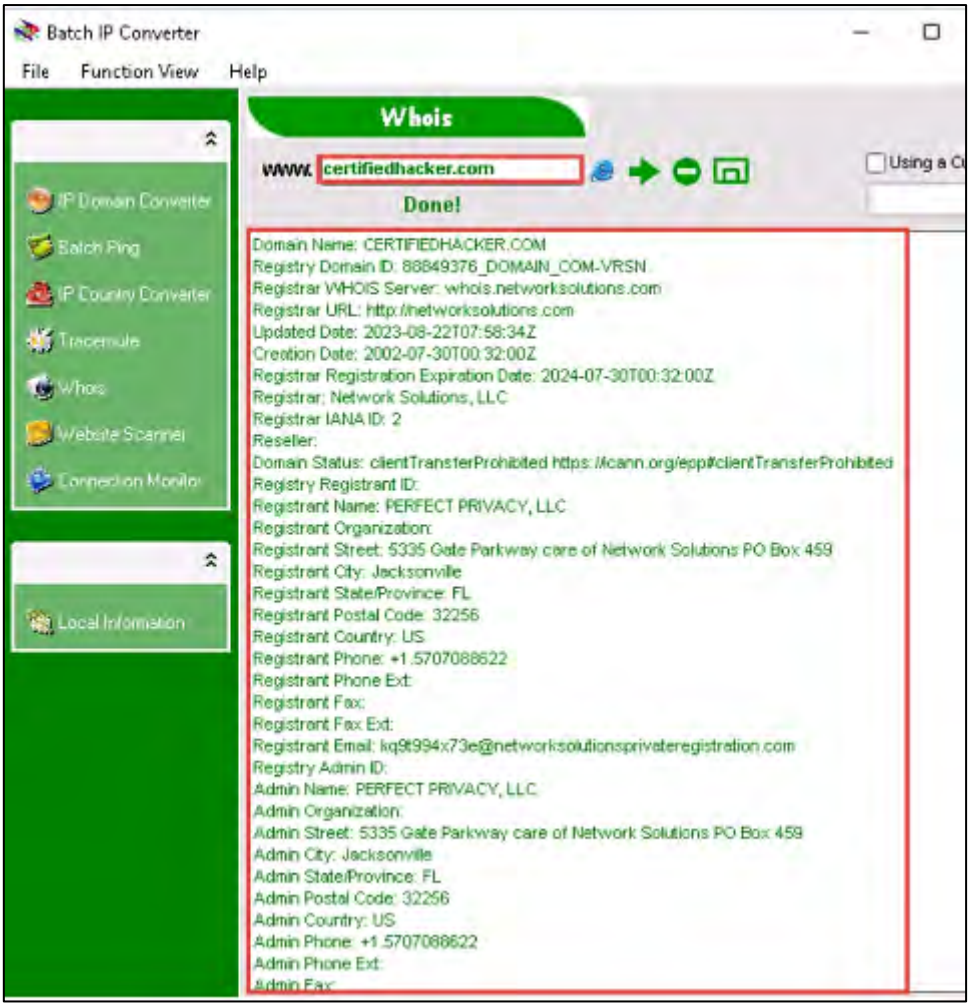

Figure 2-61: Screenshot of Batch IP Converter

EXAM TIP: Attackers also utilize Whois lookup tools like WHOIS Domain Lookup and Active Whois to carry out a Whois inquiry on the targeted domain.

Finding IP Geolocation Information

IP geolocation allows for the gathering of details about a target, including their country, state/region, city, latitude and longitude coordinates, ZIP/postal code, time zone, internet connection speed, ISP (the service provider), domain name, international dialing country code, area code, weather station code and name, mobile carrier, and elevation.

By utilizing the data acquired through IP geolocation, an attacker may seek to collect additional information about a target through methods such as social engineering, surveillance, and non-technical tactics like dumpster diving, hoaxing, or impersonating a technical expert. With the insights gained, an attacker could also establish a compromised server in close proximity to the victim's location. If the victim's precise location is identified, the attacker can engage in malicious acts, infect the victim with malware tailored for that region, attempt unauthorized access to the target device or launch an attack on it.

EXAM TIP: Attackers can launch social engineering attacks such as spamming and phishing by using IP geolocation lookup tools such as IP2Location, IP Location Finder, and IP Address Geographical Location Finder.

IP Geolocation Lookup Tools

IP2Location

As illustrated in Figure 2-62, attackers utilize the IP2Location tool to determine a visitor's geographic location, including details such as country, region, city, latitude and longitude of the city, ZIP code, time zone, connection speed, ISP, domain name, IDD country code, area code, weather station code and name, mobile carrier, elevation, and usage type data through a proprietary IP address lookup database and technology.

Figure 2-62: Screenshot of IP2Location

DNS Footprinting

Once the Whois records of the target have been gathered, the subsequent step in the footprinting process is Domain Name System (DNS) footprinting. Attackers engage in DNS footprinting to collect details about DNS servers, DNS records, and the kinds of servers utilized by the target organization. This data allows attackers to pinpoint the hosts linked to the target network and potentially exploit the target organization further. This section outlines how to

retrieve DNS information and conduct reverse DNS lookups using a variety of DNS interrogation tools.

Extracting DNS Information

DNS footprinting provides insights into DNS zone information. This DNS zone information consists of DNS domain names, hostnames, IP addresses, and various other details concerning a network. Attackers utilize DNS data to identify crucial hosts within the network, subsequently conducting social engineering attacks to collect additional information.

DNS footprinting assists in uncovering the following records related to the targeted DNS:

Record Type	Description
A	Points to a host's IP address
AAAA	Points to a host's IPv6 address
MX	Points to domain's mail server
NS	Points to host's name server
CNAME	Canonical naming allows aliases to a host
SOA	Indicate authority for a domain
SRV	Service records
PTR	Maps IP address to a hostname
RP	Responsible person
HINFO	Host information record includes CPU type and OS
TXT	Unstructured text records

Table 2-12: DNS Records with Description

DNS Interrogation Tools

Attackers utilize DNS interrogation tools like SecurityTrails, Fierce, DNSChecker, zdns, and DNSdumpster.com to conduct DNS footprinting. These tools can retrieve various IP addresses through IP routing lookups. If the target network permits unknown and unauthorized users to transfer DNS zone data, it becomes simple for an attacker to gather DNS information using a DNS interrogation tool. When an attacker makes a query to a DNS server with a DNS interrogation tool, the server replies with a record structure that holds details about the target DNS. DNS records deliver crucial information regarding the locations and types of servers.

SecurityTrails

SecurityTrails is a sophisticated DNS enumeration tool that can generate a DNS map of the target domain's network. It can gather both current and historical DNS records, including A, AAAA, NS, MX, SOA, and TXT, aiding in the construction of the DNS architecture. Additionally, it discovers all the existing subdomains of the target domain utilizing brute-force methods.

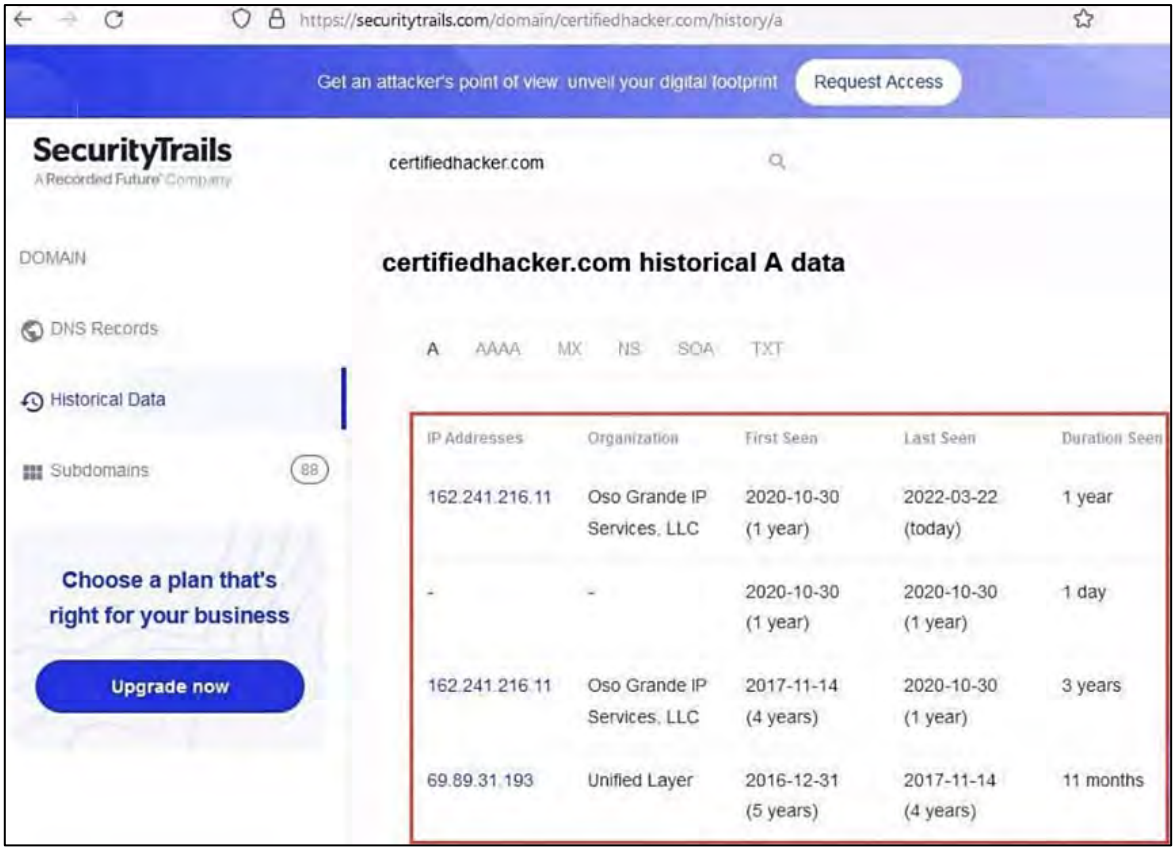

Figure 2-63: Screenshot of SecurityTrails

Fierce

Fierce is a DNS reconnaissance utility used for probing and gathering essential details about a target domain. This tool allows attackers to enumerate subdomains associated with the target domain. It also helps them pinpoint non-contiguous IP ranges and hostnames related to certain domains or subdomains. By collecting this data, attackers can construct a network landscape and identify potential targets for exploitation.

Attackers may utilize the commands listed below to conduct DNS reconnaissance with the Fierce tool:

1. Execute this command to initiate a basic scan on the specified domain (certifiedhacker.com) without any extra parameters:

```
fierce --domain certifiedhacker.com
```

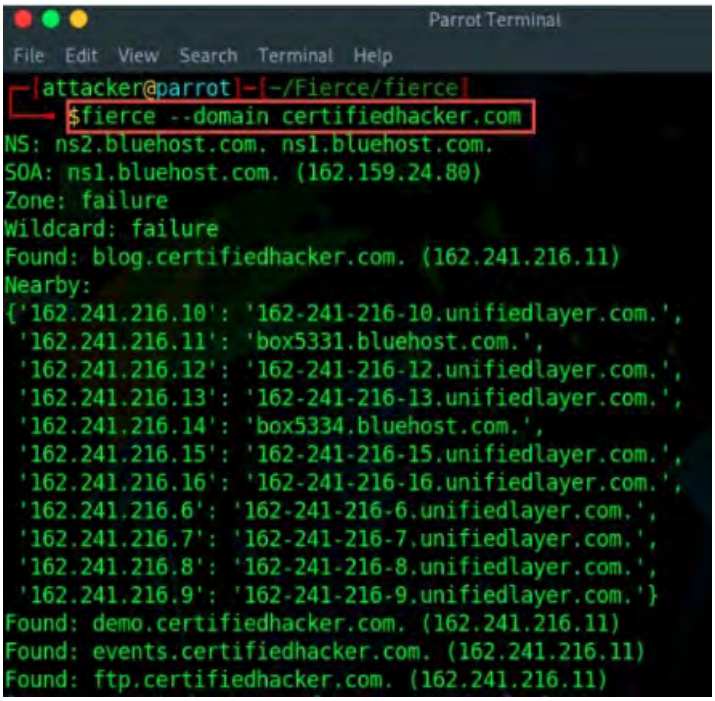

Figure 2-64: Screenshot showing Reconnaissance of the Targeted Domain

2. Execute the command below to search the designated domain for particular subdomains (in this case, subdomains including terms like write, admin, and mail):

```
fierce –domain certifiedhacker.com –subdomains write admin mail
```

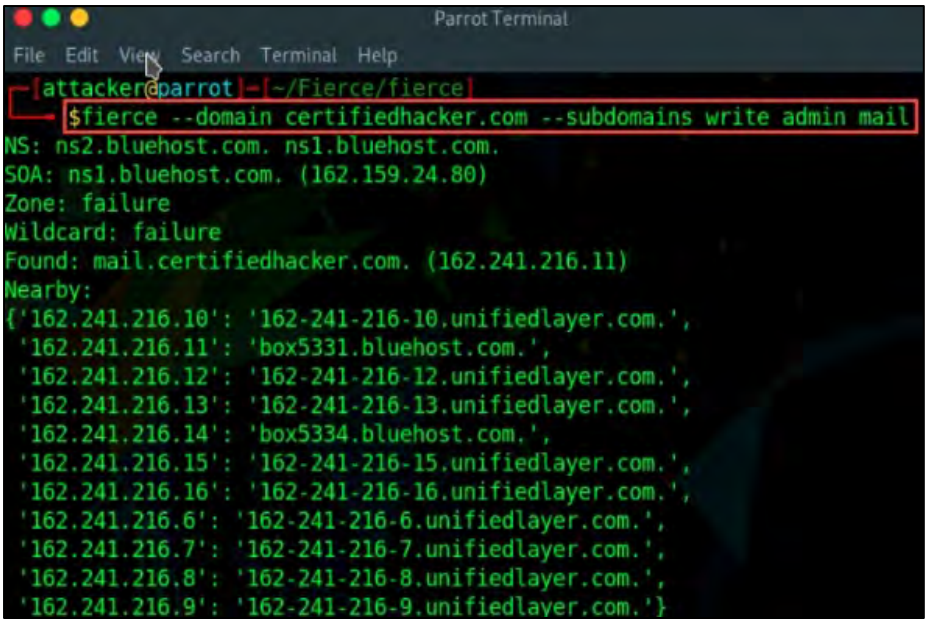

Figure 2-65: Screenshot showing a Simple Scan on Targeted Domain

3. Execute the following command to examine domains close to the identified records of the target domain:

```
fierce –domain certifiedhacker.com –subdomains mail –traverse 10
```

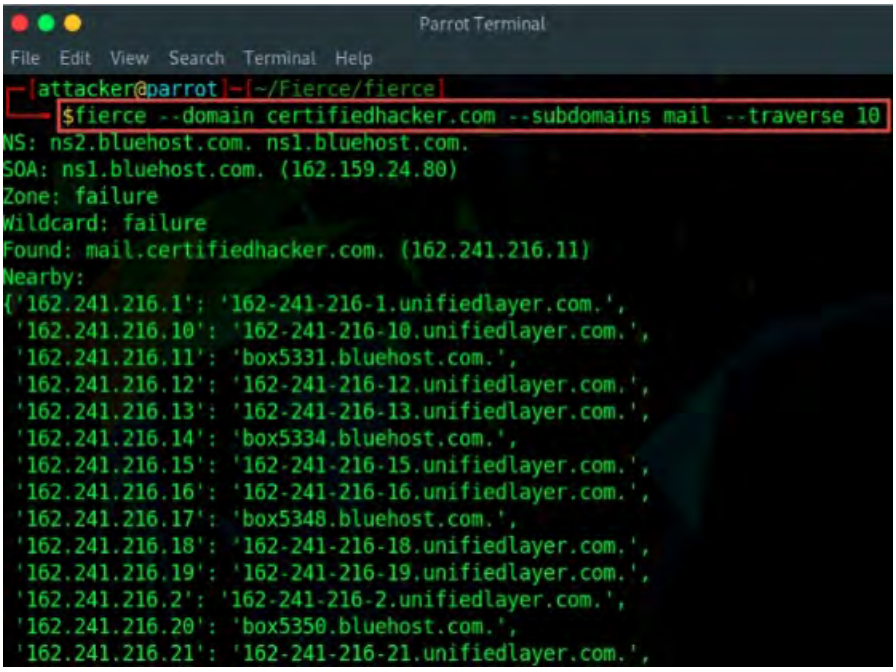

Figure 2-66: Screenshot showing the Scanning of Domains near Discovered Records

In the command provided, the option **--traverse 10** directs Fierce to look for co\ntiguous blocks of IP addresses within a range of 10.

4. Execute the command below to try establishing an HTTP connection on the identified domains of the target:

```
fierce --domain certifiedhacker.com --subdomains mail --connect
```

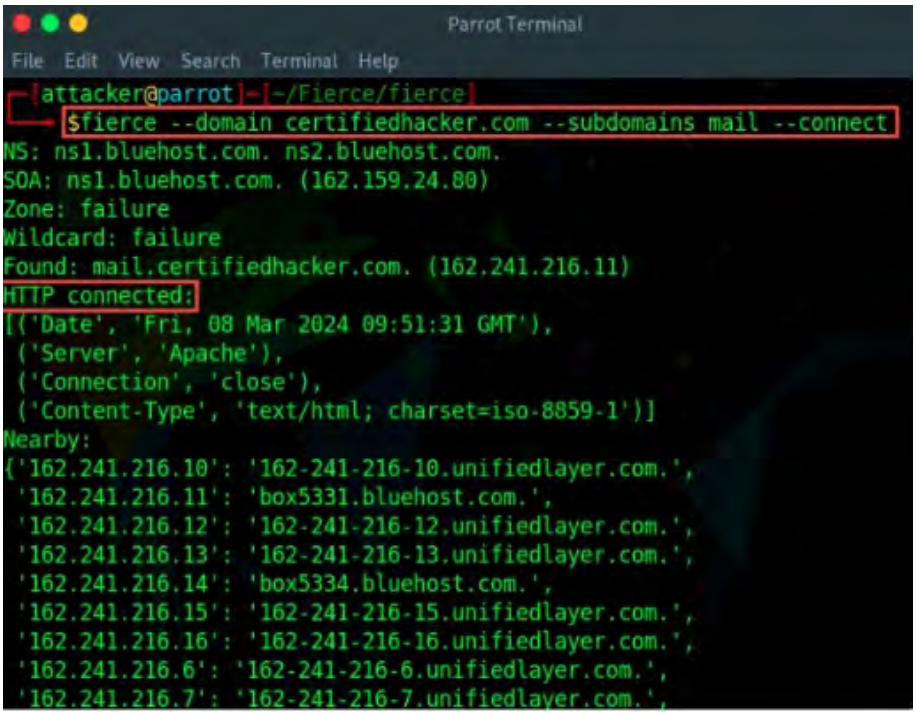

Figure 2-67: Screenshot showing the HTTP connection on discovered domains

5. Execute the command below to perform a comprehensive scan of all identified records for the target domain, specifically a complete detailed examination:

```
fierce --domain certifiedhacker.com --wide
```

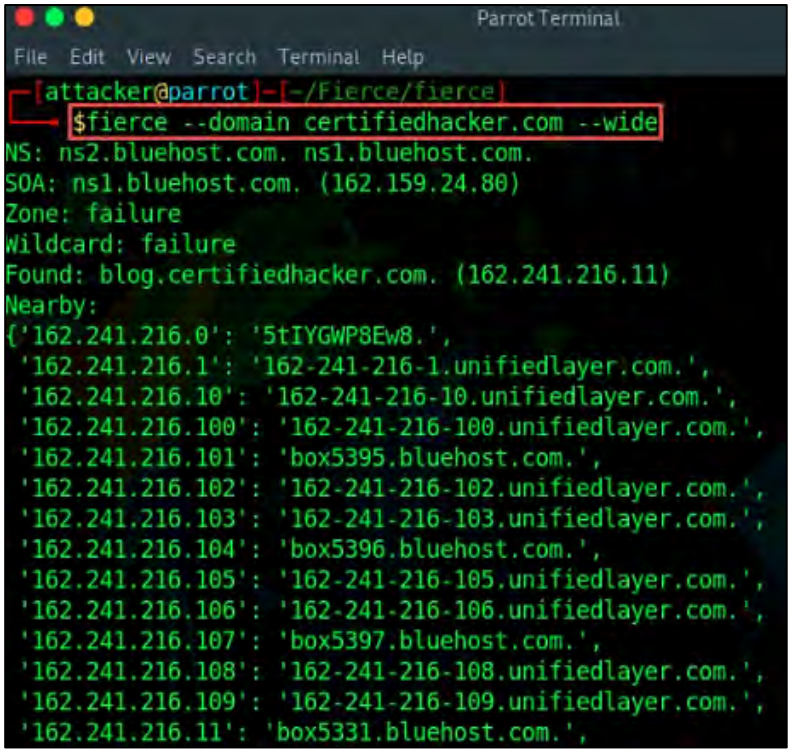

Figure 2-68: Screenshot showing the Scanning of Discovered Records of all the Classes

DNS Lookup with AI

Attackers can utilize AI-driven technologies to improve and streamline their footprinting activities. By leveraging AI, these individuals can easily conduct reverse DNS lookups on a target and obtain useful information.

For instance, attackers could employ ChatGPT to carry out this task by utilizing a suitable prompt like:

```
Install and use DNSRecon to perform DNS enumeration on the target domain www.certifiedhacker.com.
```

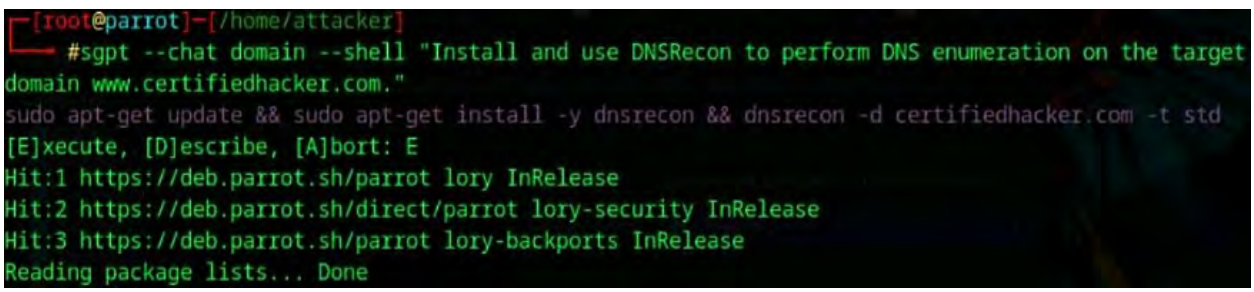

Figure 2-69: Prompt for Installing and performing DNSRecon with AI

The shell command outlined below is intended for conducting DNS enumeration utilizing the "dnsrecon" tool on the www.certifiedhacker.com domain:

```
sudo apt-get update && sudo apt-get install -y dnsrecon && dnsrecon -d certifiedhacker.com -t std
```

Table 2-13 describes each option used in the command.

Command Breakdown	Description
sudo apt-get update	Refreshes the package lists for upgrades and installations of new packages.
&&	Links commands to run them in succession.
sudo apt-get install -y dnsrecon	Installs the dnsrecon tool with an automatic "yes" response to all prompts.
dnsrecon -d certifiedhacker.com -t std	Calls the dnsrecon tool to carry out DNS enumeration on the certifiedhacker.com domain using standard enumeration methods.

Table 2-13: DNS Enumeration Command Description

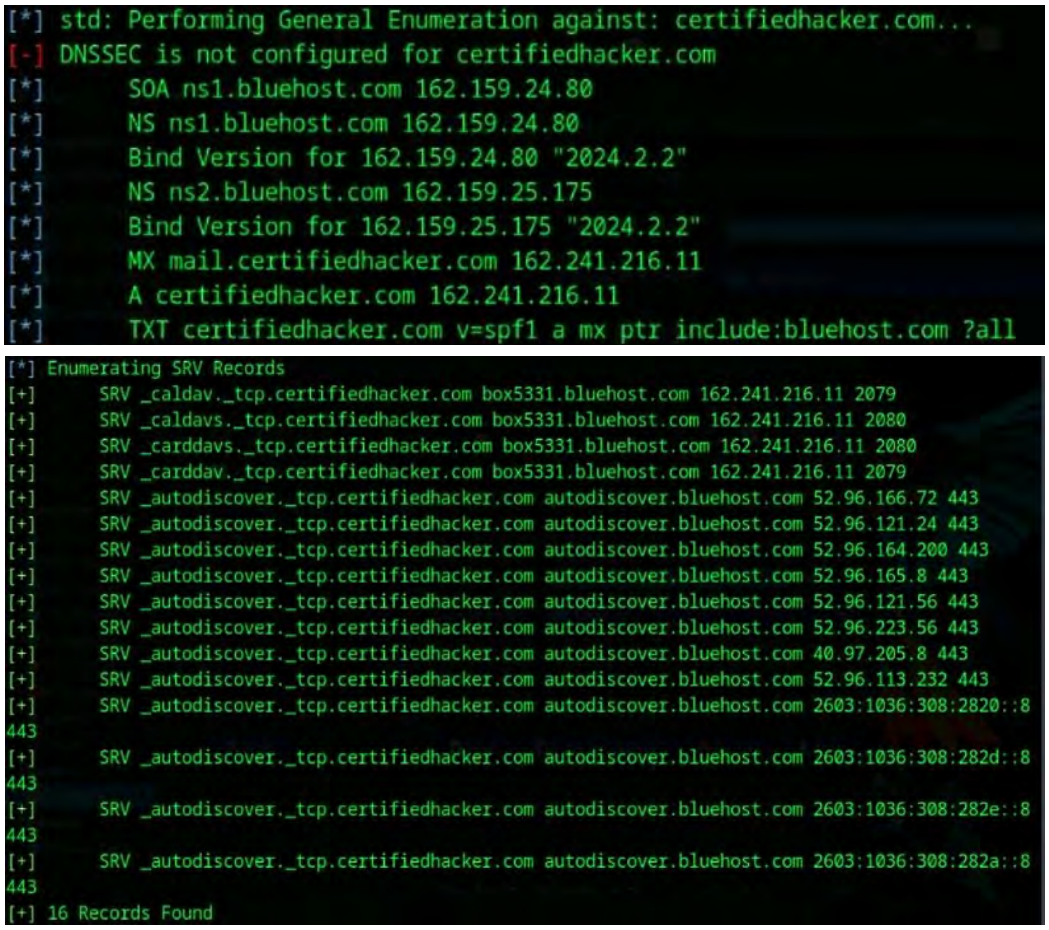

Figure 2-70: Outputs for Performing DNSRecon with AI

Reverse DNS Lookup

DNS lookup is utilized to determine the IP addresses associated with a specific domain name, while a reverse DNS operation is conducted to retrieve the domain name corresponding to a particular IP address. When you type a domain name into a web browser, the DNS translates that domain name into an IP address and sends the request for further handling. This process

of converting a domain name into an IP address occurs through the use of a record. Attackers often carry out a reverse DNS search on an IP range to identify a DNS PTR record for those IP addresses.

Various tools such as DNSRecon, Reverse Lookup, puredns, Reverse IP Domain Check, and Reverse IP Lookup are employed by attackers to execute a reverse DNS search on the target system. Once you have an IP address or a block of IP addresses, these tools can be utilized to find the corresponding domain name.

DNSRecon

As shown in Figure 2-71, attackers utilize the command below to carry out a reverse DNS lookup on the target host:

```
dnsrecon -r 162.241.216.0-162.241.216.255
```

In this command, the **-r** option indicates the range of IP addresses (from first to last) for a brute force reverse lookup.

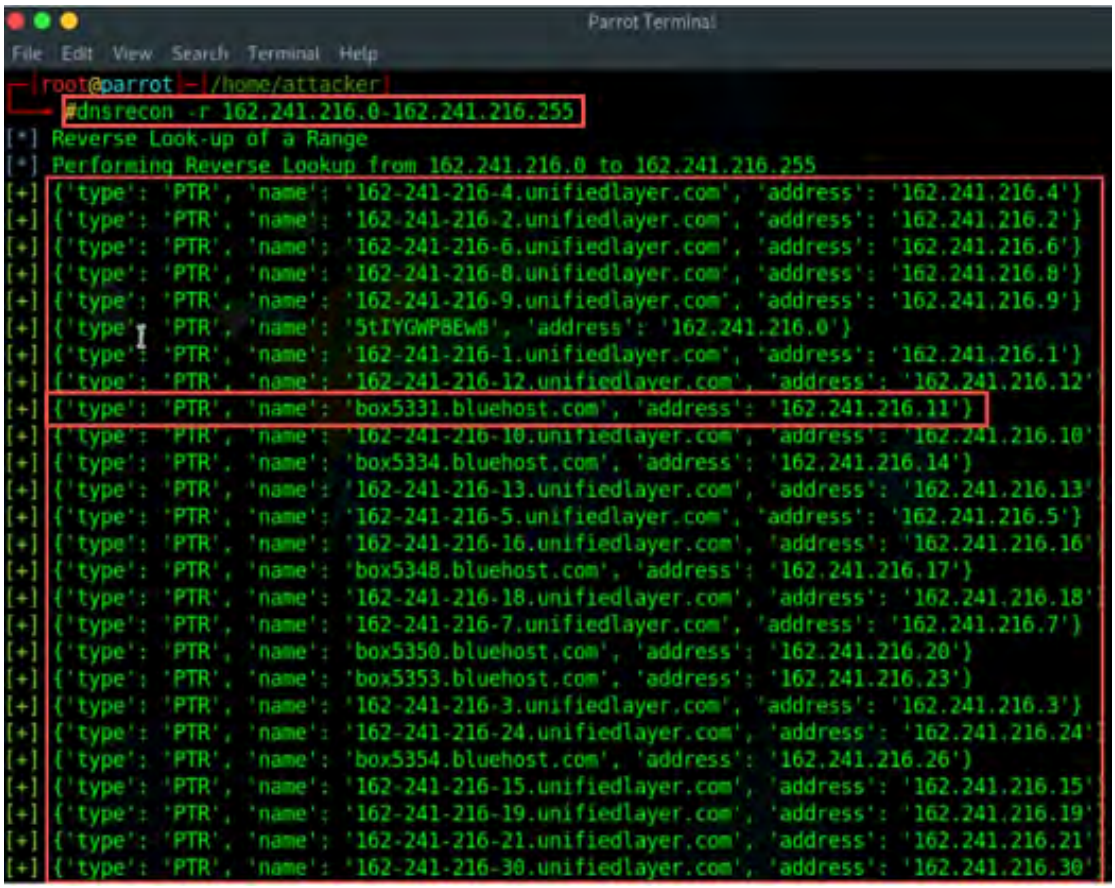

Figure 2-71: Screenshot of DNSRecon showing Reverse DNS Lookup Information

Reverse Lookup

The Reverse Lookup tool, as depicted in Figure 2-72, conducts a reverse IP lookup by using an IP address to find the corresponding DNS PTR record associated with that address.

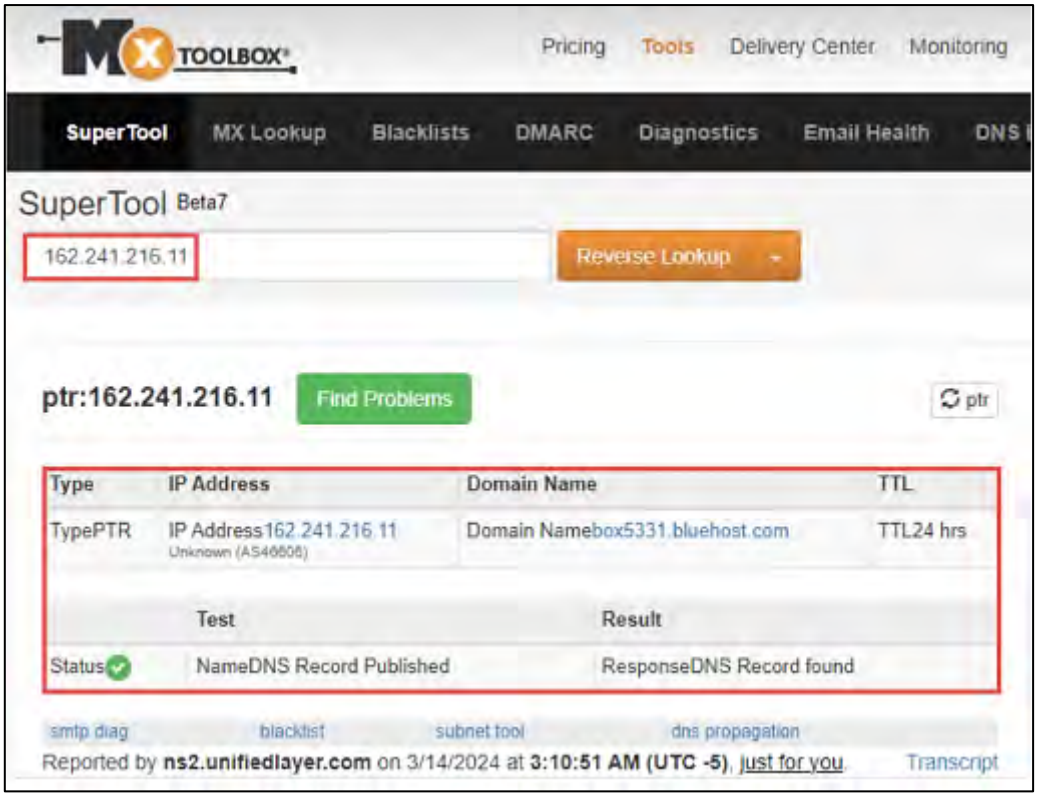

Figure 2-72: Screenshot of the Reverse Lookup Tool

Network and Email Footprinting

After retrieving DNS information, the next step is to gather network-related data and track email communications. This section explains how to identify the network range, perform traceroute analysis, and utilize various traceroute tools. Additionally, it outlines the process for tracking email communications, collecting information from email headers, and using email tracking tools.

Locate the Network Range

To conduct network footprinting, it is essential to collect fundamental and significant details about the targeted organization, such as its activities, its employees, and the nature of its work. The responses to these inquiries yield insights that assist in understanding the internal architecture of the target network.

Once this information is compiled, an attacker can ascertain the network range of the target system. Detailed data concerning IP allocation and the specifics of this allocation can be found in the relevant regional registry database. An attacker can also uncover the subnet mask of the domain and trace the path between their system and the target system. Commonly utilized traceroute tools include NetScanTools Pro and PingPlotter.

Gaining access to private IP addresses can be advantageous for attackers. The Internet Assigned Numbers Authority (IANA) has allocated the following three segments of IP address space for private networks: 10.0.0.0–10.255.255.255 (10/8 prefix), 172.16.0.0–172.31.255.255 (172.16/12 prefix), and 192.168.0.0–192.168.255.255 (192.168/16 prefix).

By using the network range, an attacker can glean information regarding the network's structure and identify which machines within the network are active. The network range also assists in determining the network topology, access control devices, and the operating system utilized in

the target network. To discover the network range of the target network, it is necessary to input the server IP address (acquired through Whois footprinting) into the ARIN Whois database search tool. A user may also navigate to the ARIN website (https://www.arin.net/about/welcome/region) and enter the server IP into the SEARCH Site or Whois text box. This will provide the network range of the target network. Poorly configured DNS servers present attackers with a favorable opportunity to acquire a list of internal machines within the network. Furthermore, if an attacker traces the route to a machine, it may be feasible to obtain the internal IP address of the gateway, which could be beneficial.

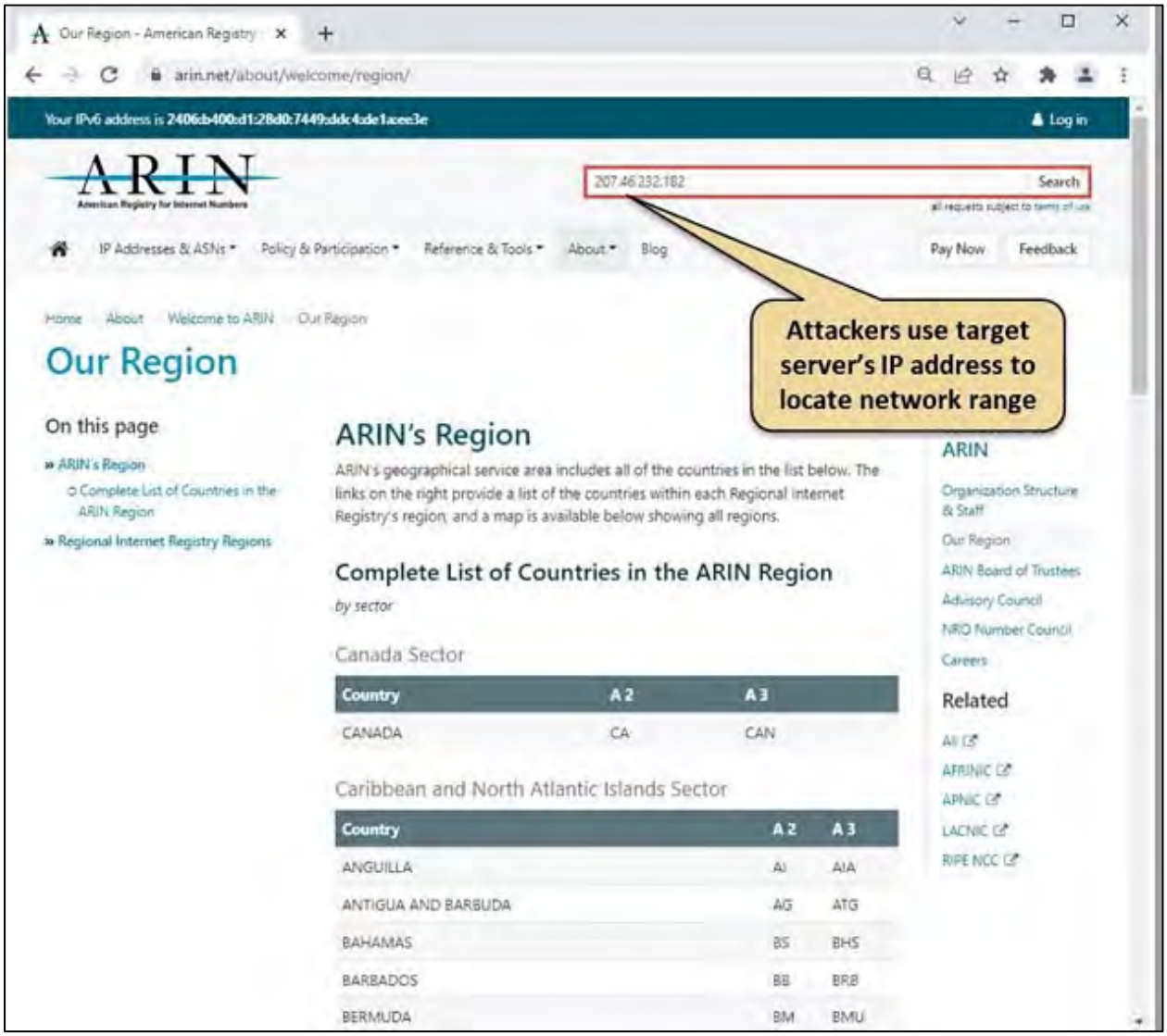

Figure 2-73: Screenshot of ARIN's Region

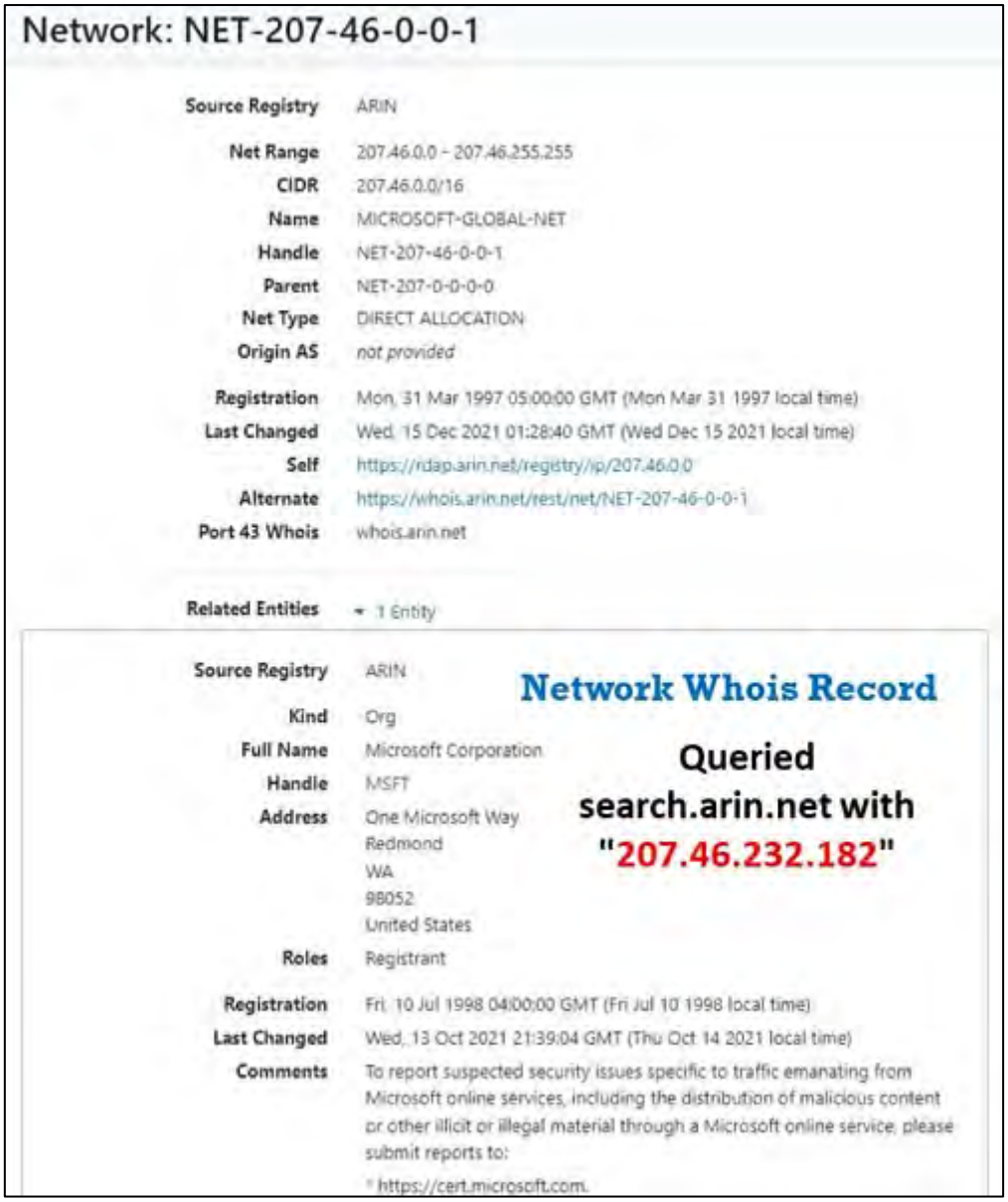

Figure 2-74: Screenshot showing the Result of ARIN Whois Database Search Result

EXAM TIP: Attackers usually utilize multiple tools to gather network information, as a single tool often cannot provide all the needed data.

Traceroute

Identifying the path to the target host on the network is essential for testing against man-in-the-middle attacks and similar threats. Most operating systems include a Traceroute tool to accomplish this objective. It traces the route that packets directed towards the target host follow across the network.

Traceroute operates using the ICMP protocol and the Time to Live (TTL) field within the IP header to determine the target host's path on the network. The Traceroute tool can provide insights into the trajectory taken by IP packets between two devices. It is capable of counting the number of routers the packets pass through, measuring the round-trip time (the duration it takes to transit between two routers), and, if available, identifying the routers' names and their

network affiliations through DNS entries. Additionally, it can trace geographical locations. This functionality takes advantage of a characteristic of the Internet Protocol known as TTL. The TTL field signifies the maximum number of routers a packet is permitted to traverse. Each router processing a packet decreases the TTL count in the ICMP header by one. Once the count reaches zero, the router discards the packet and sends an ICMP error message back to the original sender.

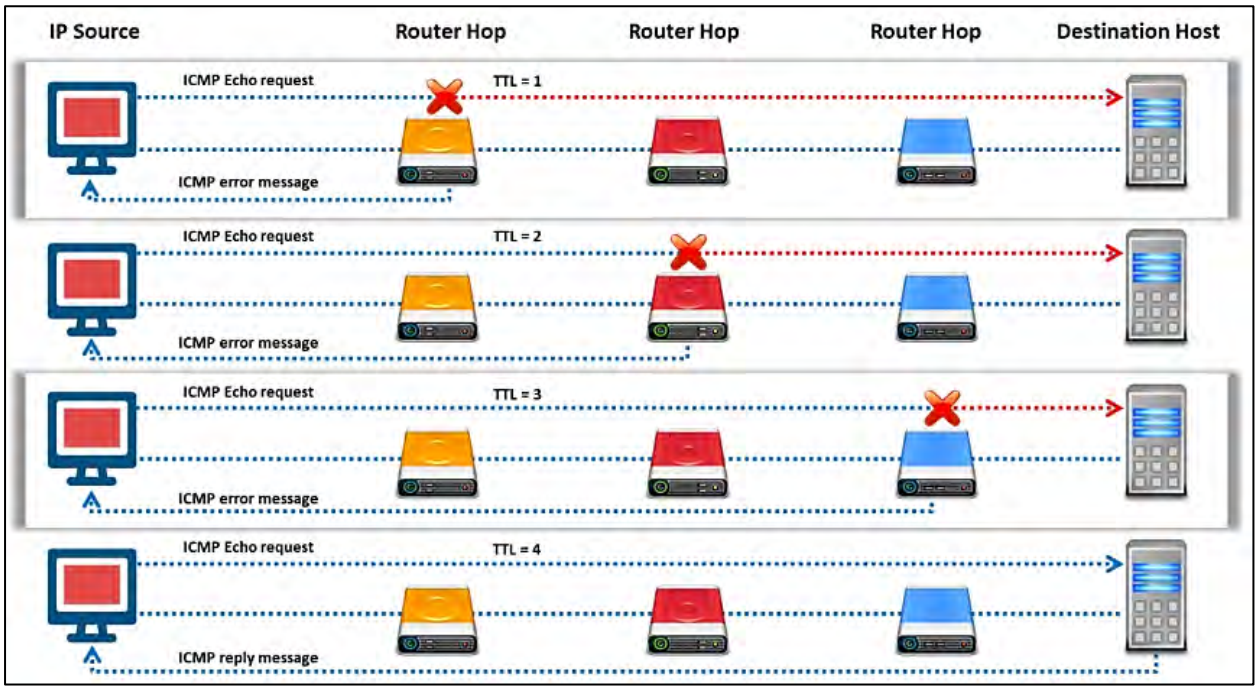

Figure 2-75: Illustration of Traceroute

The utility captures the router's IP address and DNS name and transmits another packet with a TTL value set to two. This packet successfully traverses the first router but times out at the subsequent router in the sequence. The second router also returns an error message to the originating host. Traceroute proceeds in this manner, documenting the IP address and name of every router until either the packet successfully reaches the target host or it concludes that the host is unreachable. Throughout this process, it measures the time taken for each packet to complete a round trip to every router. In the end, when it arrives at the destination, a standard ICMP ping response is sent back to the sender. This utility provides insights into the IP addresses of the intermediate hops along the path from the source to the target host.

ICMP Traceroute

The Windows operating system, by default, utilizes ICMP for traceroute. Open the command prompt and enter the **tracert** command followed by the destination IP address or domain name like this:

```
C:\>tracert 216.239.36.10
```

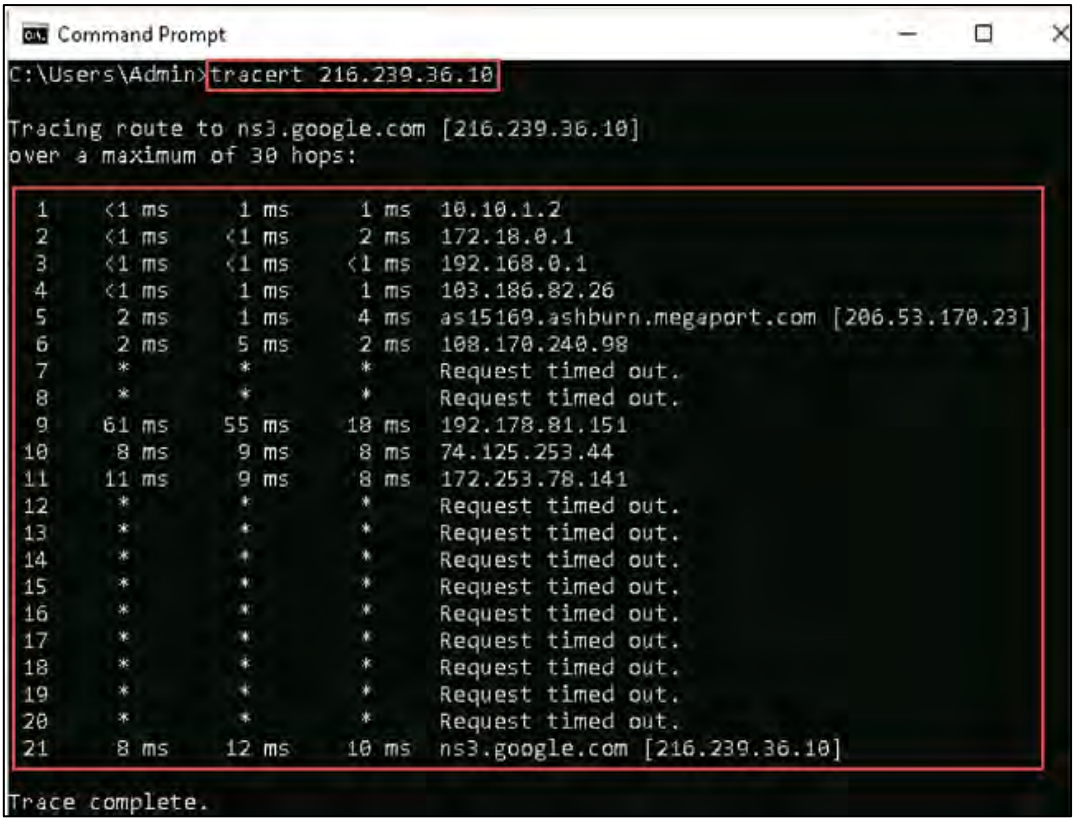

Figure 2-76: Output of tracert

TCP Traceroute

Numerous devices within a network are typically set up to prevent ICMP traceroute messages. In such cases, an attacker may utilize TCP or UDP traceroute, commonly referred to as Layer 4 traceroute. To execute this on a Linux operating system, open the terminal and enter the tcptraceroute command along with the target IP address or domain name as follows:

```
sudo tcptraceroute www.google.com
```

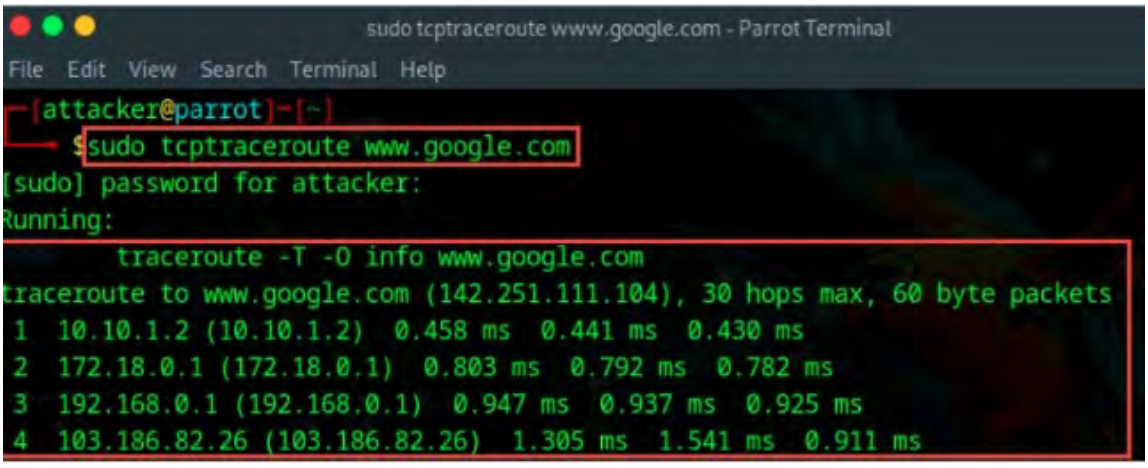

Figure 2-77: Output of TCP Traceroute

UDP Traceroute

Similar to Windows, Linux features its own traceroute tool that employs the UDP protocol to trace the path to a target. To use it, open the terminal in the Linux OS and enter the **traceroute** command followed by the destination IP address or domain name, like:

```
traceroute www.google.com
```

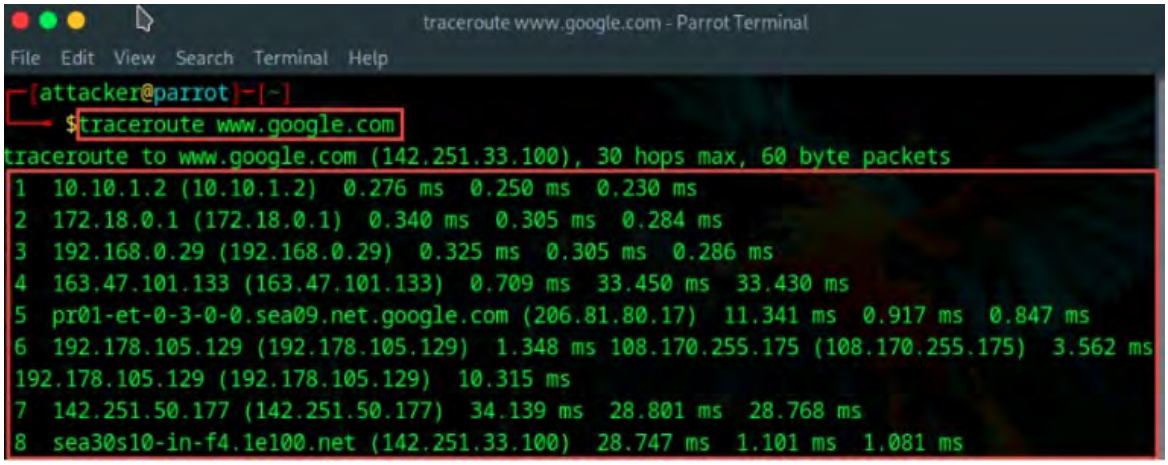

Figure 2-78: Output of UDP Traceroute

Traceroute with AI

Attackers can utilize AI-driven technologies to improve and automate their footprinting activities. With AI's assistance, attackers can easily conduct tracerouting on a target.

For instance, they can employ ChatGPT to carry out this task by using a suitable prompt like:

```
Perform network tracerouting to discover the routers on the path to a target host www.certifiedhacker.com.
```

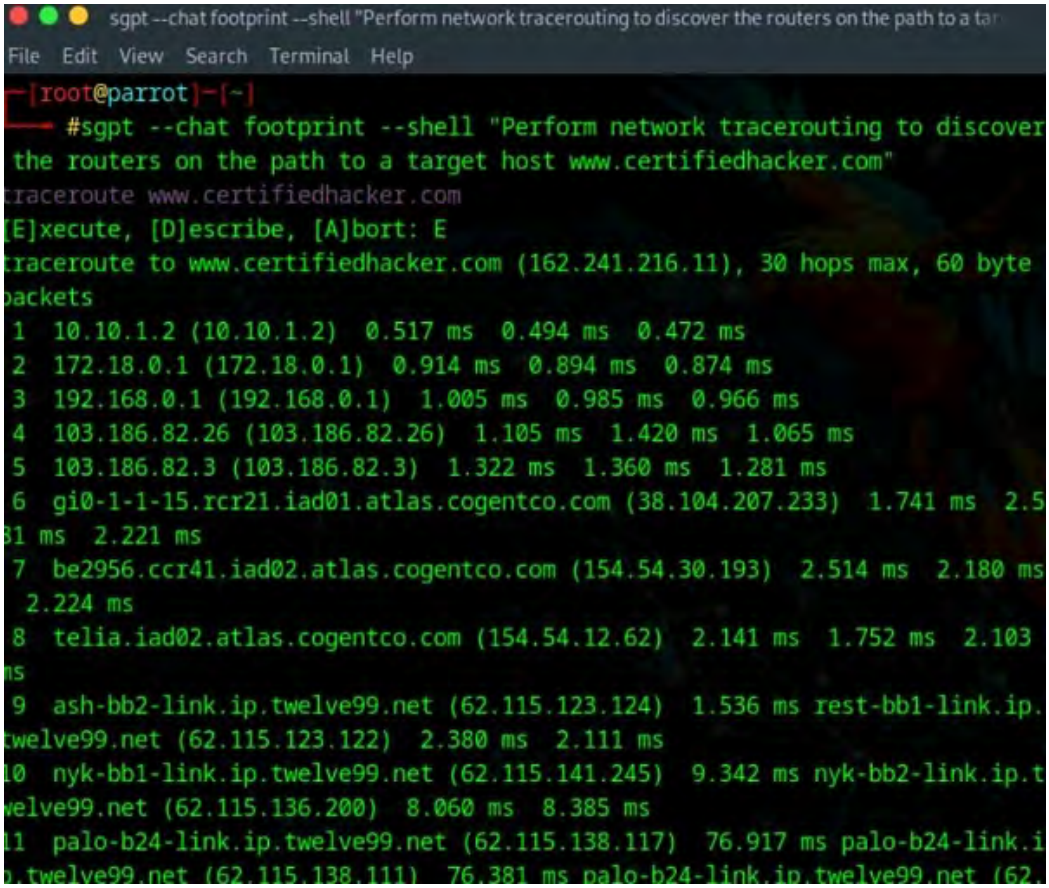

Figure 2-79: Prompt and Output for performing Traceroute with AI

The given shell command is intended to execute network tracerouting using the "traceroute" utility to identify the routers along the path to the www.certifiedhacker.com host:

```
traceroute www.certifiedhacker.com
```

This command starts the traceroute tool to locate the routers on the route to the www.certifiedhacker.com host by dispatching packets to the destination with progressively increasing Time To Live (TTL) values and examining the responses obtained from the intermediate routers.

 **EXAM TIP:** Traceroute utility helps find the IP addresses of intermediate devices such as routers and firewalls present between a source and its destination.

Traceroute Analysis

After executing multiple traceroutes, an attacker can determine the location of a hop within the target network. Consider the following results from the traceroutes:

- traceroute 1.10.10.20, the penultimate hop is 1.10.10.1
- traceroute 1.10.20.10, the third-to-last hop is 1.10.10.1
- traceroute 1.10.20.10, the penultimate hop is 1.10.10.50
- traceroute 1.10.20.15, the third-to-last hop is 1.10.10.1
- traceroute 1.10.20.15, the penultimate hop is 1.10.10.50

By examining these results, an attacker can pinpoint the intermediate devices or hosts along the route to the target network, as illustrated in Figure 2-80.

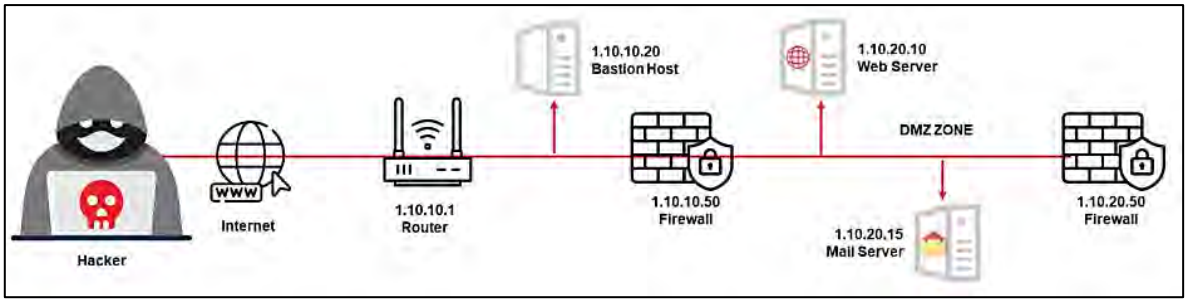

Figure 2-80: Traceroute Analysis

Traceroute Tools

Traceroute tools like NetScanTools Pro, PingPlotter, Traceroute NG, and tracert are valuable for gathering information about the geographical positions of routers, servers, and IP devices within a network. These tools aid in tracing, identifying, and monitoring network activity on a global map. Some features of these tools include:

- Hop-by-hop traceroutes
- Reverse tracing
- Historical analysis
- Packet loss reporting
- Reverse DNS
- Ping plotting
- Port probing
- Network problem detection
- Performance metrics analysis
- Monitoring network performance

NetScanTools Pro

Individuals can utilize NetScanTools Pro to track the path packets take from their computer to the intended device, whether on a local area network or over the Internet. The software provides ICMP, UDP, or TCP traceroute options, enabling users to pinpoint the intermediary devices along the way. Additionally, it assists users in determining the country associated with each IPv4 address at each hop.

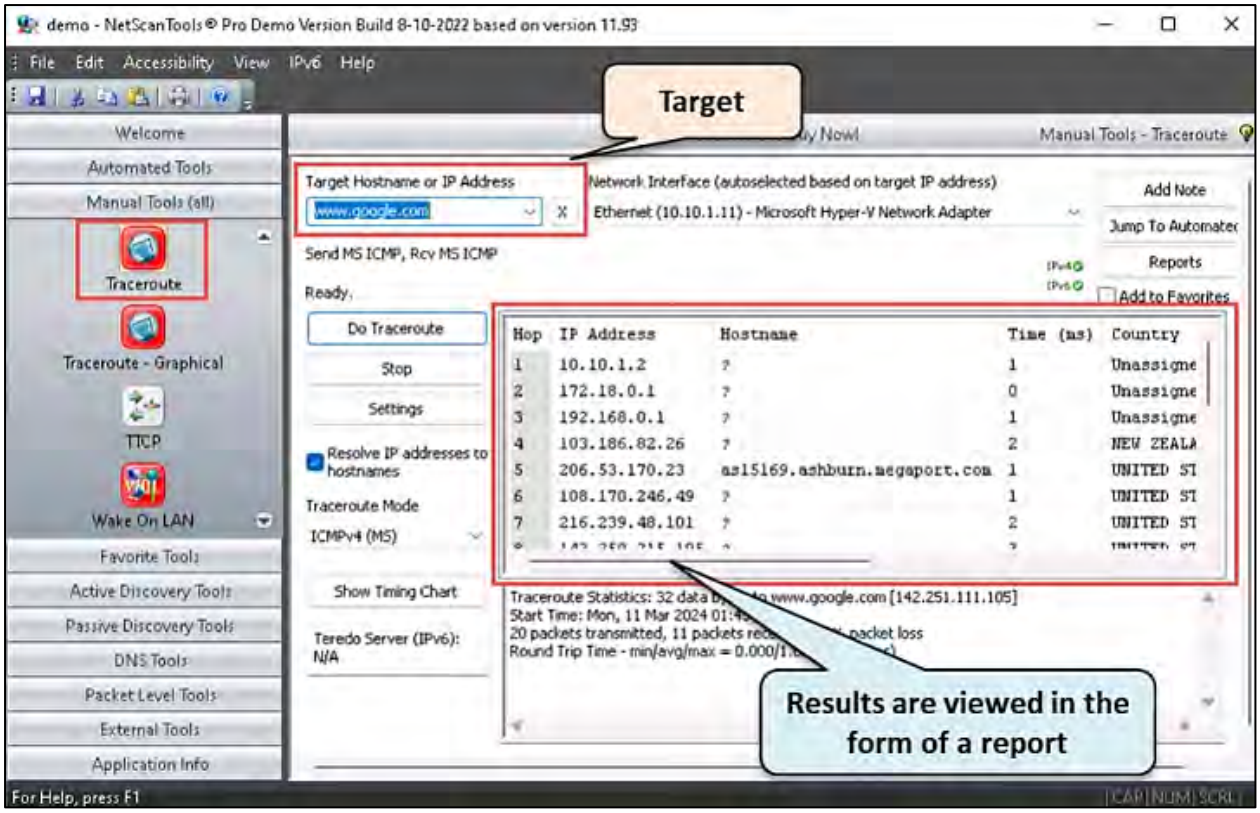

Figure 2-81: Screenshot of NetScanTools Pro

PingPlotter

PingPlotter enables users to gather traceroute information for specific hosts through the use of ICMP, UDP, and TCP packets. It autonomously identifies network hops and monitors latency as well as packet loss over a period. With this tool, users can represent the traceroute information in easily interpretable graphs. This application assists users in pinpointing bandwidth issues, WiFi disruptions, or hardware malfunctions within the target network.

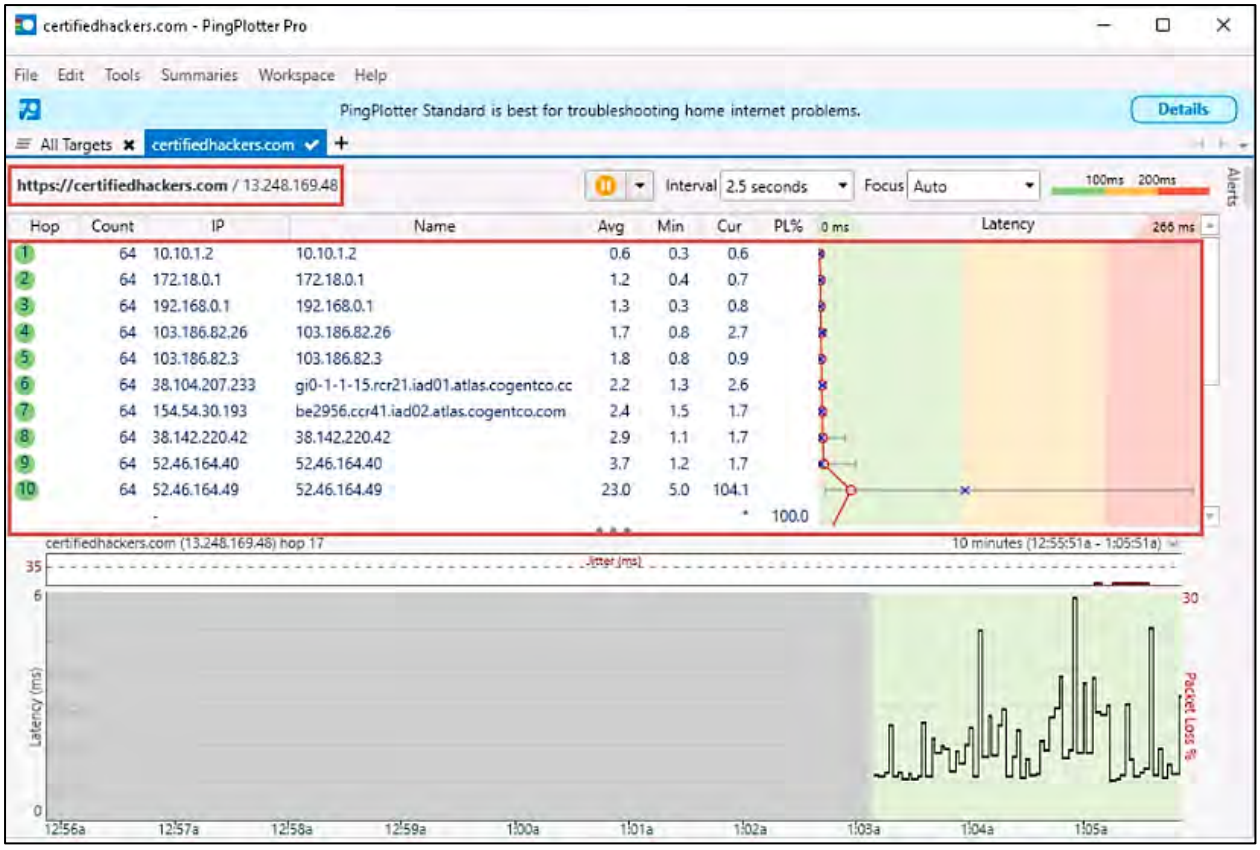

Figure 2-82: Screenshot of PingPlotter

Tracking Email Communications

Email tracking observes the email communications of a specific user. This type of surveillance is feasible through digitally time-stamped records that indicate when the target opens and receives a certain email. Email tracking tools enable an attacker to gather details such as IP addresses, mail servers, and the service providers associated with sending the email. This information can be utilized by attackers to create a hacking plan and to execute social engineering and other forms of cyberattacks. Notable examples of email tracking tools include IP2LOCATION's Email Header Tracer, MxToolbox, DNS Checker Email Header Analyzer, and Social Catfish.

Data about the victim acquired using email tracking tools includes:

- **Recipient's System IP address:** Permits the tracking of the recipient's IP address.
- **Geolocation:** Provides an estimation and displays the recipient's location on a map and may even calculate the distance from the attacker's position.
- **Email Received and Read:** Alerts the attacker when the recipient has received and opened the email.
- **Read Duration:** Indicates the amount of time the recipient spends reading the email sent by the sender.
- **Proxy Detection**: Supplies data regarding the type of server utilized by the recipient.
- **Links:** Monitors whether the links sent to the recipient via email have been accessed.
- **Operating System and Browser Information:** Discloses details about the operating system and browser employed by the recipient. The attacker can leverage this

information to identify vulnerabilities in that specific version of the operating system and browser to initiate further attacks.

- **Forward Email:** Indicates if the email received by the user is sent to someone else.
- **Device Type:** Offers details regarding the kind of device utilized to access and view the email, such as a desktop computer, mobile device, or laptop.
- **Path Travelled:** Monitors the journey the email took through email transfer agents from the source to the destination system.

Collecting Information from Email Header

An email header includes information about the sender, routing details, addressing format, date, subject line, and recipient. Email headers also assist intruders in tracking the route an email takes before reaching the recipient. Every email header serves as a valuable resource for an attacker to initiate assaults on the target. The method of accessing the email header differs among various email applications.

Commonly used email programs include:

- eM Client
- Mailbird
- Hiri
- Mozilla Thunderbird
- Spike
- Claws Mail
- SmarterMail Webmail
- Outlook
- Apple Mail
- ProtonMail
- AOL Mail
- Tuta

The email header includes the following details:

- Sender's mail server
- Date and time of receipt by the originator's email servers
- Authentication system used by the sender's mail server
- Data and time of sending the message
- A unique number assigned by mx.google.com to identify the message
- Sender's full name
- The sender's IP address and the address from which the message was sent

An attacker can gather and trace all this information by conducting a thorough examination of the complete email header.

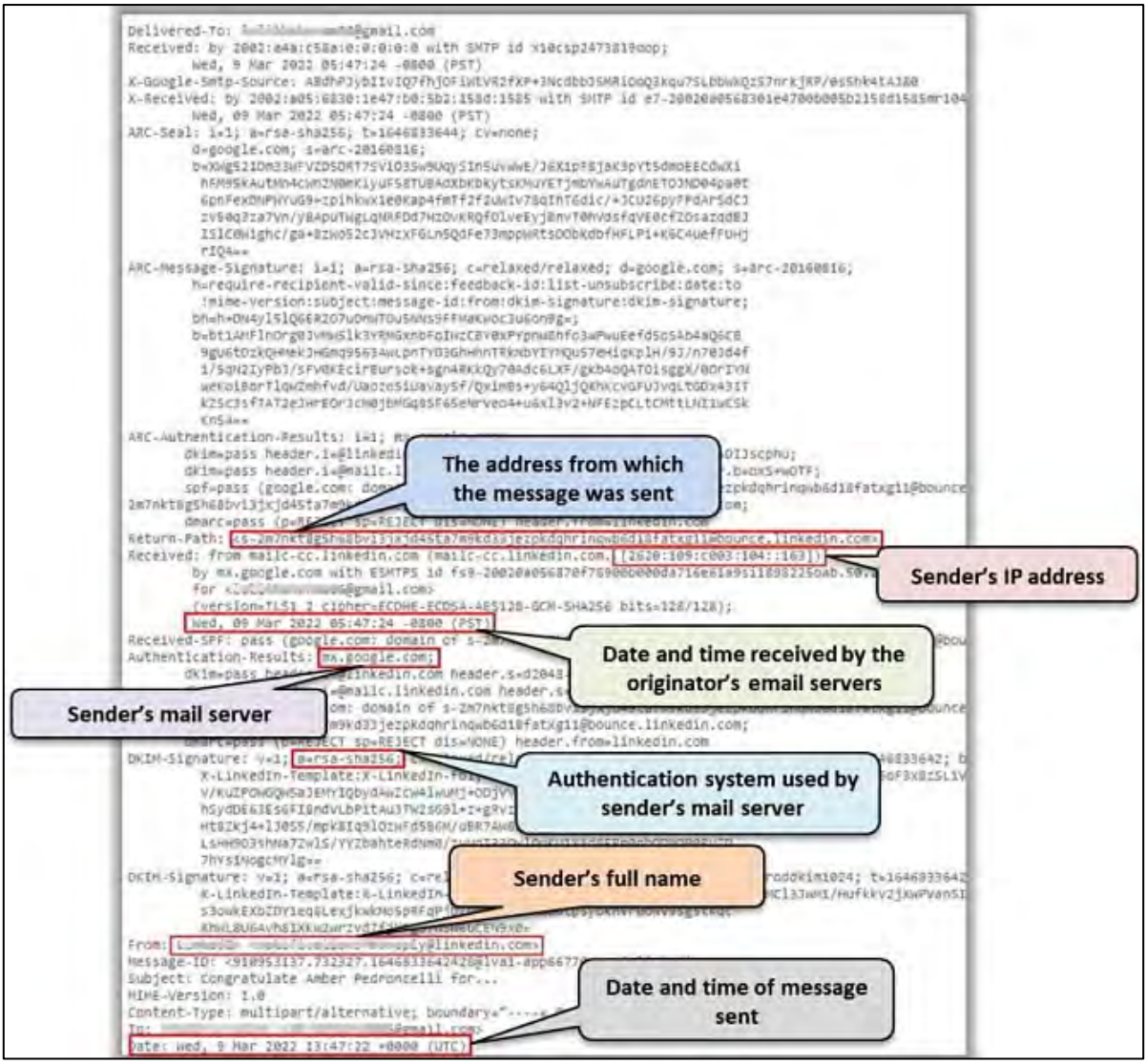

Figure 2-83: Screenshot showing Detailed Analysis of the Email Header

Email Tracking Tools

Email tracking tools like IP2LOCATION's Email Header Tracer, MxToolbox, eMailTrackerPro, Holehe, DNS Checker Email Header Analyzer, and Social Catfish enable an attacker to monitor an email and obtain details such as the sender's identity, mail server, sender's IP address, location, and more. Attackers exploit the gathered information to trace the email route from their position to the intended mail server by utilizing IP addresses found in the email header.

eMailTrackerPro

As demonstrated in Figure 2-84, attackers utilize eMailTrackerPro to examine email headers and retrieve details like the sender's geographical location, IP address, and more. This tool enables an attacker to revisit the traces later by storing previous records.

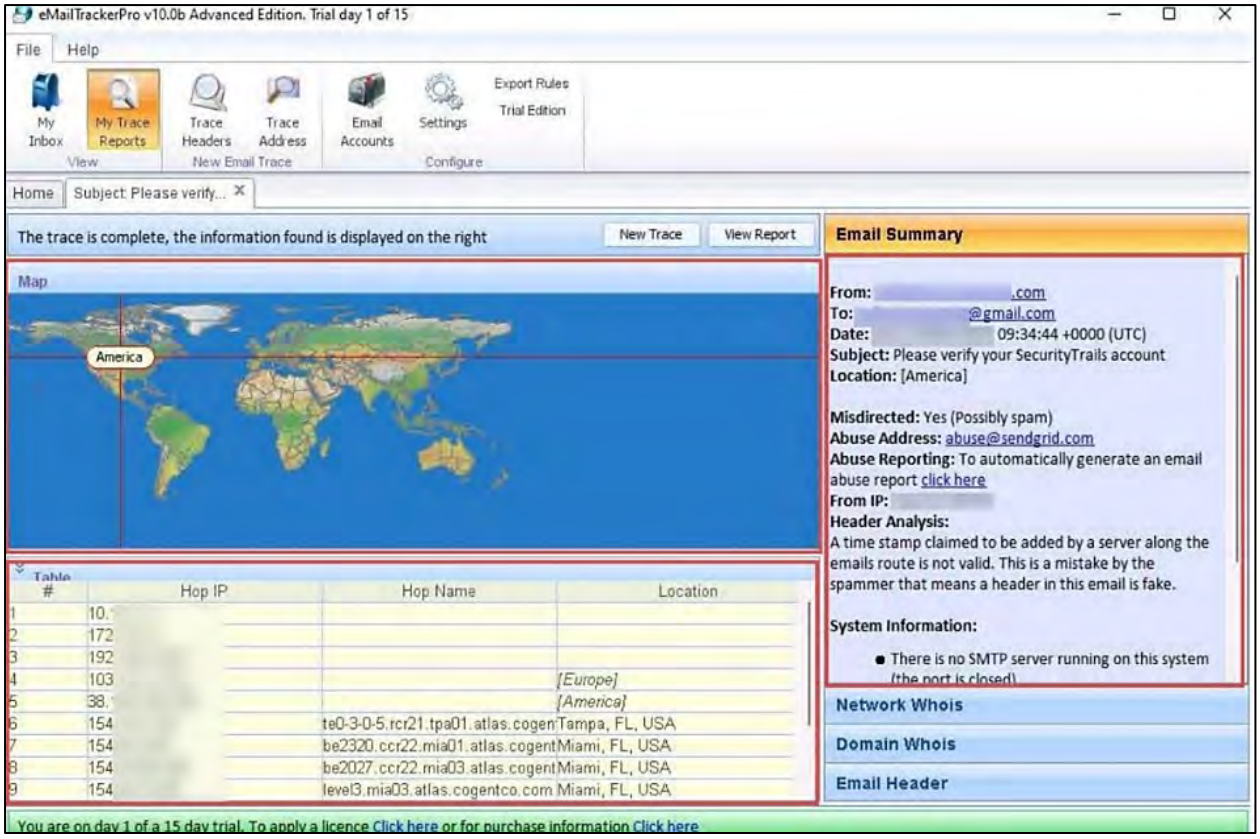

Figure 2-84: Screenshot of eMailTrackerPro

IP2LOCATION's Email Header Tracer

IP2LOCATION's Email Header Tracer is a publicly available tool that attackers can utilize to examine and trace the journey of emails through their headers. It allows attackers to identify the original location of the target and the mail servers that the email has traveled through by analyzing the IP addresses present in the email header.

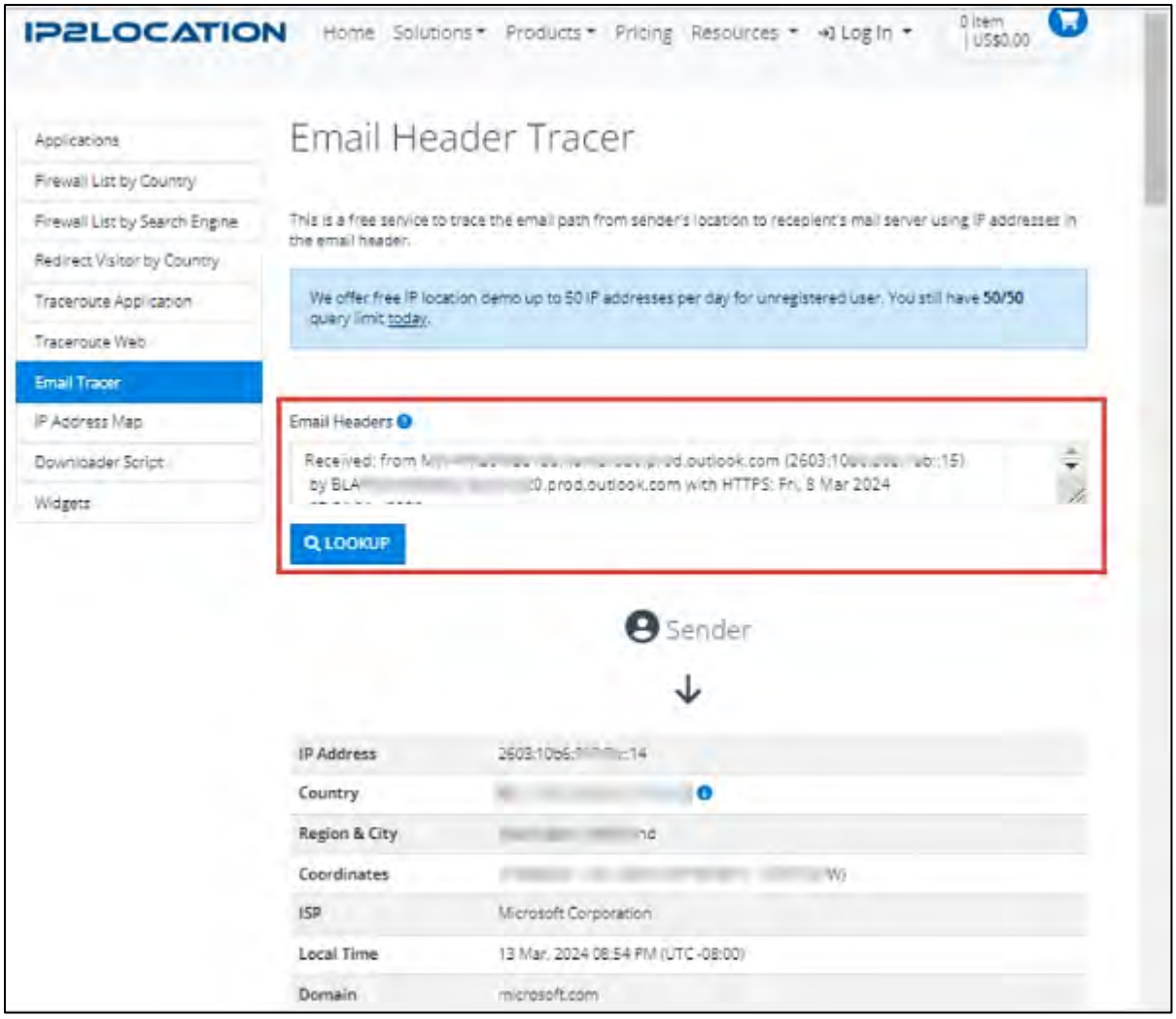

Figure 2-85: Screenshot of IP2LOCATION's Email Header Tracer-1

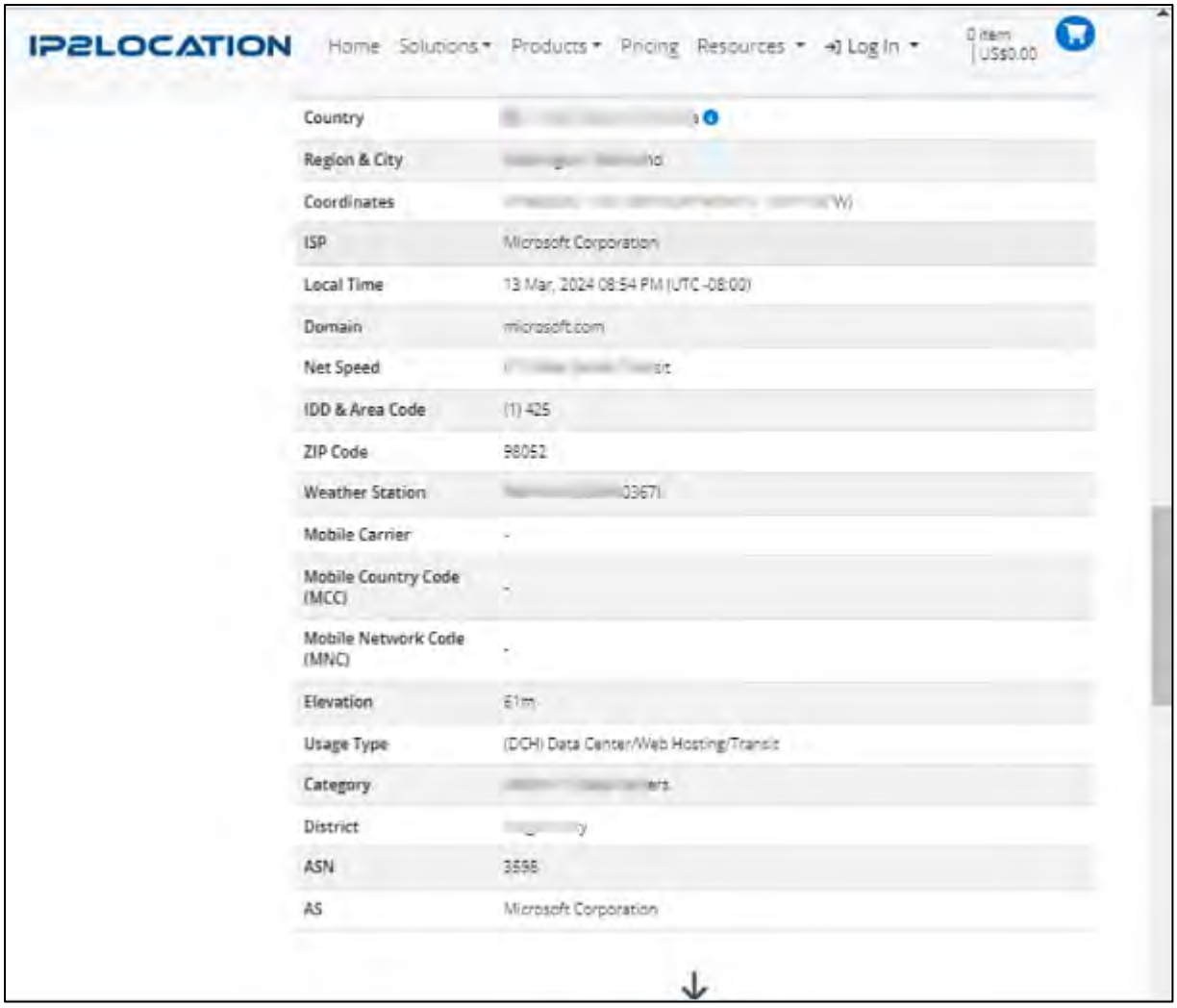

Figure 2-86: Screenshot of IP2LOCATION's Email Header Tracer-2

Footprinting through Social Engineering

Footprinting through social engineering refers to the practice of gathering information from individuals by taking advantage of their vulnerabilities. This section discusses both the idea and the methods employed to collect information via social engineering.

Social engineering is a process that does not rely on technical means, wherein an attacker deceives an individual into unintentionally revealing confidential information. In simpler terms, the target is unaware that their sensitive information is being extracted. The attacker exploits the trusting nature of people and their propensity to share confidential data. To execute social engineering, an attacker must initially earn the trust of an authorized user and then mislead that user into disclosing sensitive information. The primary objective of social engineering is to acquire the necessary confidential data and utilize it for malicious activities, including unauthorized system access, identity theft, corporate espionage, network breaches, fraud, and more.

The types of information that can be obtained through social engineering may encompass credit card information, social security numbers, usernames and passwords, other personal data, details about security products in use, operating system and software versions, IP addresses, server names, network structure details, and others. Various techniques can be employed for

social engineering, including eavesdropping, shoulder surfing, dumpster diving, impersonation, tailgating, third-party authorization, piggybacking, and reverse social engineering, among others.

Collecting Information through Social Engineering on Social Networking Sites

Social networking platforms are digital services, platforms, or other websites that enable individuals to connect and form interpersonal relationships. The utilization of social networking platforms is growing rapidly. Instances of such platforms include LinkedIn, Facebook, Instagram, Twitter, Pinterest, YouTube, and others. Each social networking platform serves its unique purpose and offers distinct features. One platform may link friends and family, while another focuses on enabling users to share their professional profiles. Social networking platforms are accessible to everyone. Attackers may exploit this openness to collect sensitive information from users, either by examining their public profiles or by establishing a deceptive account to impersonate a legitimate user. Users often share personal details on social networking platforms, including their date of birth, educational background, job history, names of spouses, and more. Organizations frequently share information about potential partners, websites, and upcoming company news.

For an attacker, social networking platforms can provide valuable insights about a targeted individual or organization. The attacker can only collect information that users have publicly shared. There are no restrictions for attackers to view public pages associated with accounts on social networking platforms. To gain additional information about the target, attackers might create fraudulent accounts and employ social engineering tactics to entice the victim into disclosing more personal details. For instance, an attacker may send a friend request to the targeted individual from a bogus account; if the victim accepts the request, the attacker can then view even the restricted content of the target's profile on that platform.

Information Available on Social Networking Sites

Individuals often keep profiles on social media platforms to share basic details about themselves and to foster and maintain connections with others. A profile typically includes personal details like a person's name, contact information (such as cell phone numbers and email addresses), data about friends, information regarding family members, as well as interests and activities. People commonly connect with friends and communicate with them. Malicious individuals can extract sensitive information through these conversations. Social media sites also enable users to share images and videos. If users do not configure their privacy settings correctly for their photo albums, then malicious individuals can access the pictures and videos they have shared. Users might participate in groups to engage in games or discuss their opinions and interests. Attackers can obtain details about a victim's preferences by monitoring the groups they join, potentially tricking the victim into disclosing additional information. Users can

create events to inform others about forthcoming occasions, which allows attackers to learn about the user's activities.

The user activities on social networking platforms and the corresponding information that an attacker may gather are outlined in Table 2-14.

What Users Do	What Attacker Gets
Maintain profile	Contact info, location, and related information
Connect to friends, chat	Friends list, friends' info, and related information
Share photos and videos	Identity of family members, interests, and related information
Play games, join groups	Interests
Create events	Activities

Table 2-14: Activities of Users on the Social Networking Sites and the Respective Information

Similar to individuals, companies utilize social networking platforms to engage with people, advertise their products, and obtain feedback regarding their offerings and services. The actions taken by an organization on social networking platforms and the relevant information an attacker might gather are encapsulated in Table 2-15.

What Organizations Do	What Attacker Gets
User surveys	Business strategies
Promote products	Product profile
User support	Social engineering
Recruitment	Platform/technology information
Background check to hire employees	Type of business

Table 2-15: Activities of the Organization on the Social Networking Sites and the Respective Information

Collecting Information Using Eavesdropping, Shoulder Surfing, Dumpster Diving, and Impersonation

Eavesdropping, shoulder surfing, dumpster diving, and impersonation are commonly utilized social engineering methods for gathering information from individuals.

Eavesdropping

Eavesdropping refers to the interception of any form of communication—be it audio, video, or text—without the knowledge or consent of the involved parties. This also involves accessing confidential messages from communication platforms, such as instant messaging or fax transmissions. An attacker may obtain information by tapping into phone conversations or capturing audio, video, or written exchanges.

Shoulder Surfing

Shoulder surfing is a strategy where attackers discreetly watch a targeted individual to obtain sensitive information. In this technique, an attacker positions themselves behind the victim and quietly observes their on-screen actions, such as typing usernames, passwords, and other confidential details. This method is particularly effective for acquiring passwords, personal identification numbers, security codes, account details, credit card digits, and similar

information. Attackers can carry out shoulder surfing easily in busy locations, as it is simple to stand behind and monitor the victim without their awareness.

Dumpster Diving

This inappropriate method, also referred to as trashing, involves the attacker searching through garbage bins for information. The attacker can acquire crucial details like phone bills, contact lists, financial data, operational documents, printouts of source code, and other sensitive materials from the waste bins of the target company's trash, printer refuse, or even sticky notes left on employees' desks. Additionally, attackers might collect account information from ATM waste bins. The gathered information can assist the attacker in executing their attacks.

Impersonation

Impersonation is a tactic where an attacker acts as a legitimate or authorized individual. Attackers can conduct impersonation attacks in person or utilize phones and other communication methods to deceive targets into disclosing their information. The attacker may pose as a delivery person, janitor, business associate, client, technician, or may even act as a visitor. By employing this approach, an attacker can acquire sensitive information by inspecting terminals for passwords, looking for important documents on desks, rummaging through bins, and more. The attacker might also attempt to eavesdrop on confidential discussions or engage in "shoulder surfing" to collect sensitive data.

Footprinting Tasks using Advanced Tools and AI

Numerous organizations provide tools that assist in gathering information. This section outlines the tools utilized to gather data from different sources. Footprinting tools are employed to gather fundamental details about target systems for exploitation purposes. The information obtained from footprinting tools may include the target's IP location, routing data, business details, physical address, phone number, social security number, information regarding the source of an email and a file, as well as DNS and domain information.

Advanced Footprinting Tools

Numerous advanced footprinting tools help attackers to accomplish their goals effectively and easily. Some of the tools are discussed here.

Maltego

Maltego is an automated tool that can help identify the connections and real-world associations between individuals, groups, organizations, websites, internet infrastructure, documents, and more. Malicious actors can utilize the various entities available in the tool to gather information such as email addresses, a collection of phone numbers, and details about a target's internet infrastructure (including domains, DNS names, Netblocks, and IP address information).

As illustrated in Figure 2-87, attackers can add a Website entity, rename it to the target’s domain, and retrieve the associated email addresses and phone numbers for that target.

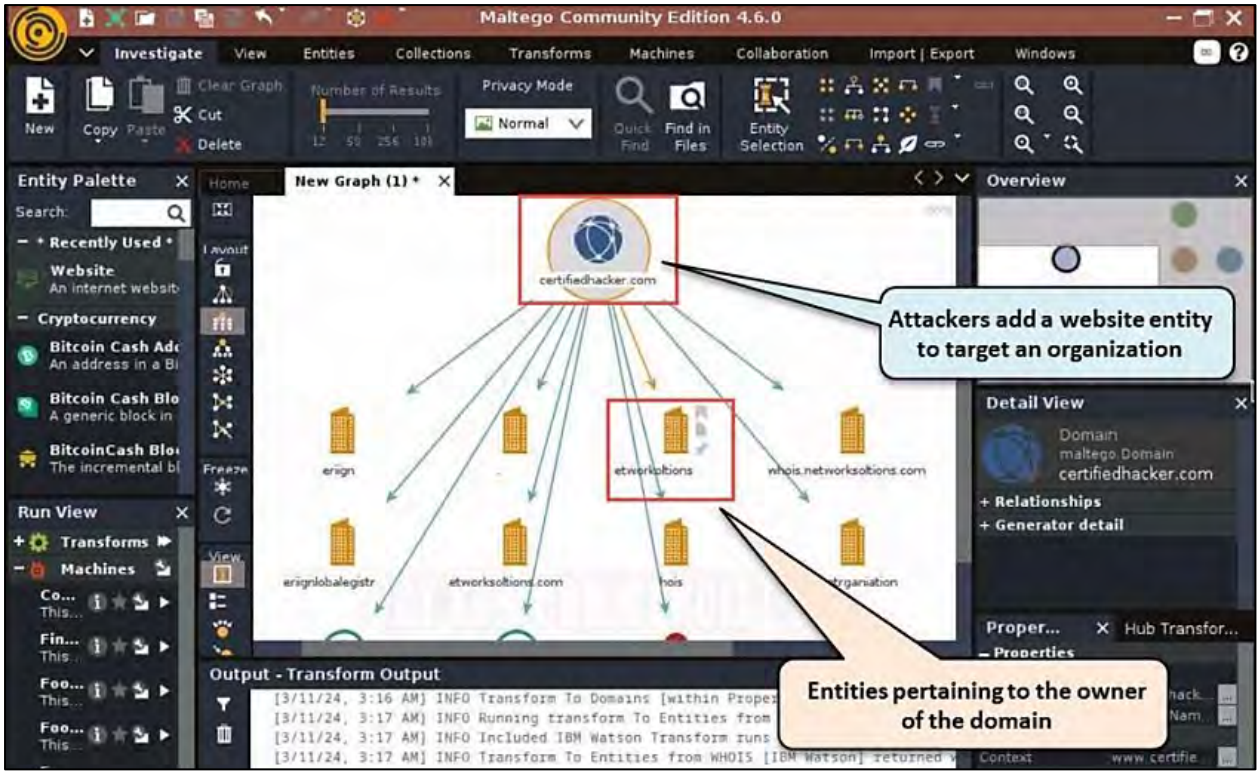

Figure 2-87: Screenshot of Maltego

Recon-ng

Recon-ng is a framework for web reconnaissance that consists of separate modules for interacting with databases, creating a setting where open-source web-based reconnaissance can take place.

In Figure 2-88, it can be seen that attackers utilize the module "recon/domains-hosts/brute_hosts" to gather a list of hosts connected to the target URL.

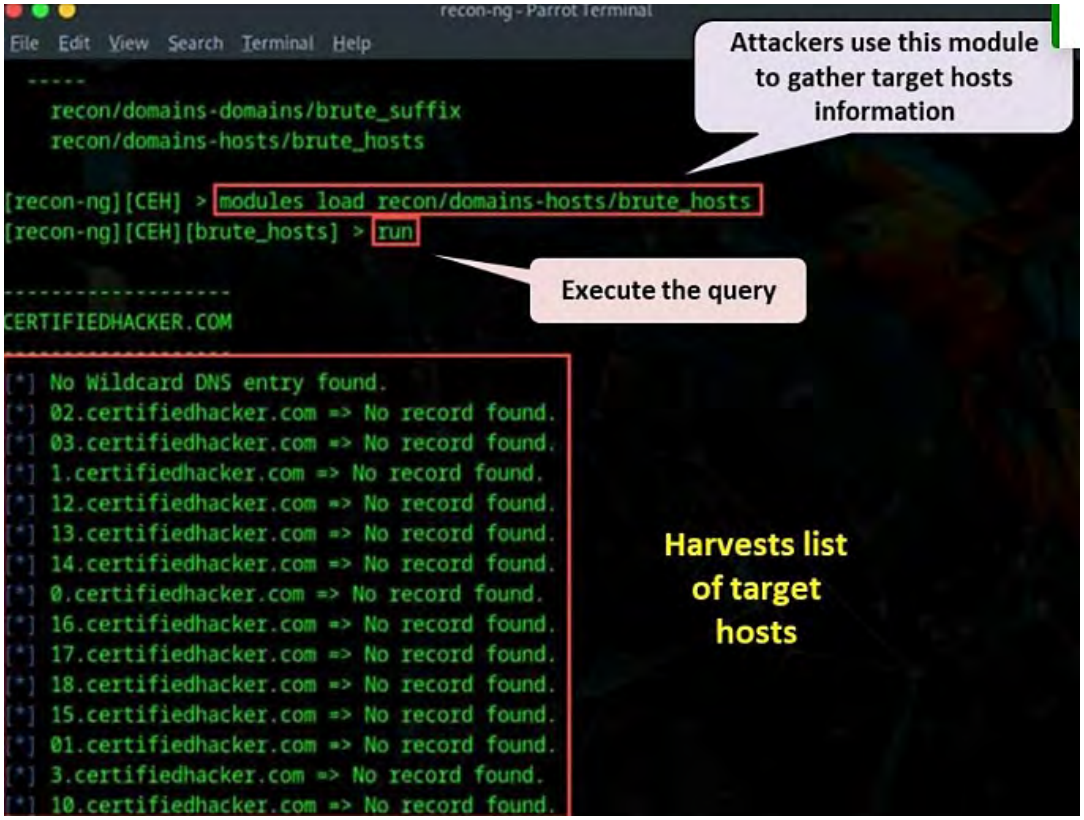

Figure 2-88: Screenshot of Recon-ng

FOCA

Fingerprinting Organizations with Collected Archives (FOCA) is a utility primarily designed to uncover metadata and concealed information within the documents it analyzes. FOCA can scan and evaluate a diverse range of documents, with Microsoft Office, Open Office, and PDF files being the most frequently processed. Features include:

- **Web Search:** Identifies hosts and domain names by exploring URLs linked to the main domain. Each link is reviewed to extract data from its corresponding host and domain names.
- **DNS Search:** Inspects each domain to identify the host names configured in NS, MX, and SPF servers to uncover new host and domain names.
- **IP Resolution:** Determines the IP address associated with each hostname by comparing against the DNS for accurate results. This process involves analyzing the organization's internal DNS.
- **PTR Scanning:** Discovers additional servers within the same segment of a specified address; the IP FOCA executes a PTR log scan.
- **Bing IP:** Initiates a search process through FOCA for new domain names that are linked to each discovered IP address.
- **Common Names:** Conducts dictionary attacks targeting the DNS.

As illustrated in Figure 2-89, attackers investigate the target domain to gather information about the files stored within it. The retrieved files can be accessed through a web browser. Additionally, the attackers can access further details, including network domains, roles, vulnerabilities, and metadata associated with the target domain.

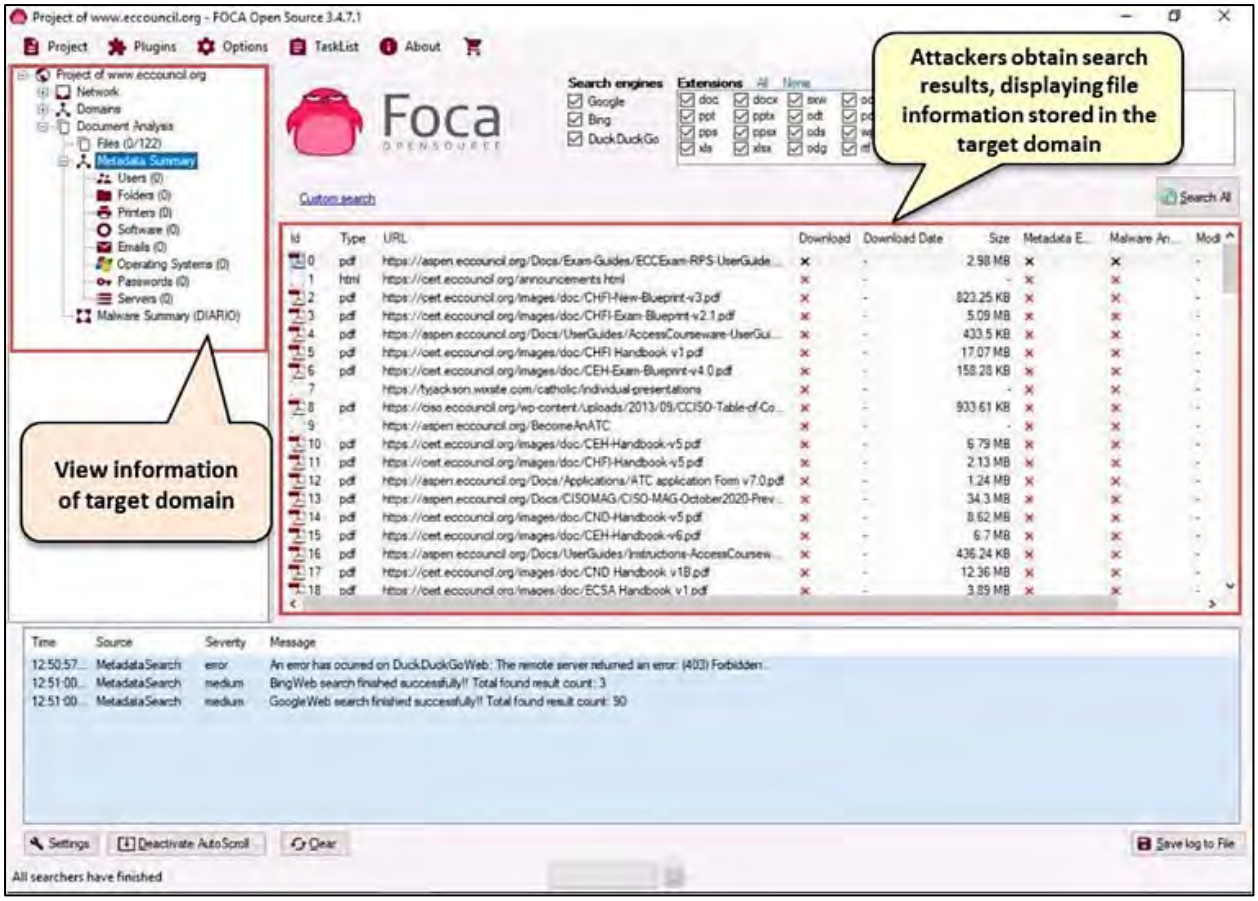

Figure 2-89: Screenshot of FOCA

subfinder

Subfinder is a tool designed for discovering subdomains that assist attackers in locating valid subdomains for websites by utilizing passive online resources. It offers support for several output formats, including JSON, file, and standard output.

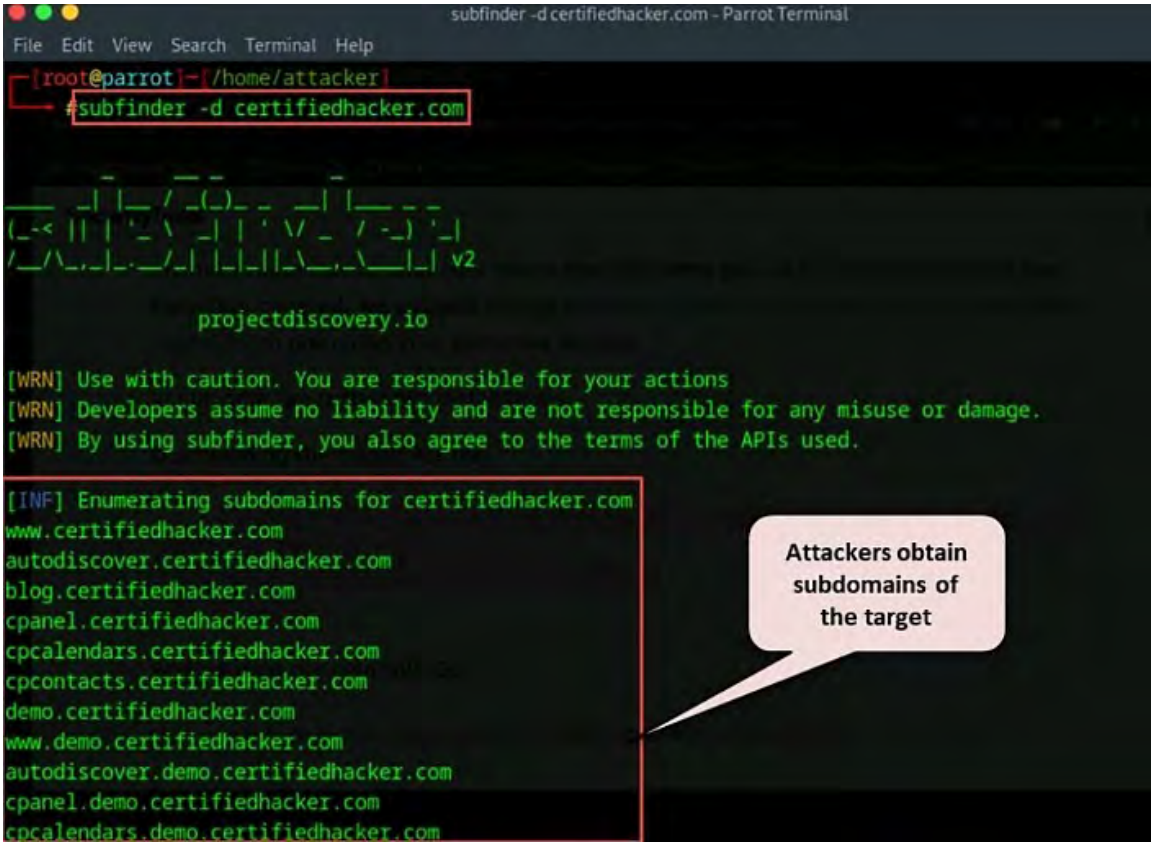

Figure 2-90: Screenshot of subfinder

OSINT Framework

The OSINT Framework is a framework for gathering open source intelligence that assists security professionals in conducting automated footprinting and reconnaissance, as well as OSINT research and intelligence collection. Its primary focus is to collect information using free tools or resources. This framework features a straightforward web interface that organizes various OSINT tools by category, displayed as an OSINT tree structure on the website.

Figure 2-91 illustrates the tools listed, which include the following indicators:

- (T): Signifies a tool that needs to be installed and operated locally
- (D): Google dork
- (R): Requires user registration
- (M): Denotes a URL containing the search term that must be edited manually to include the term

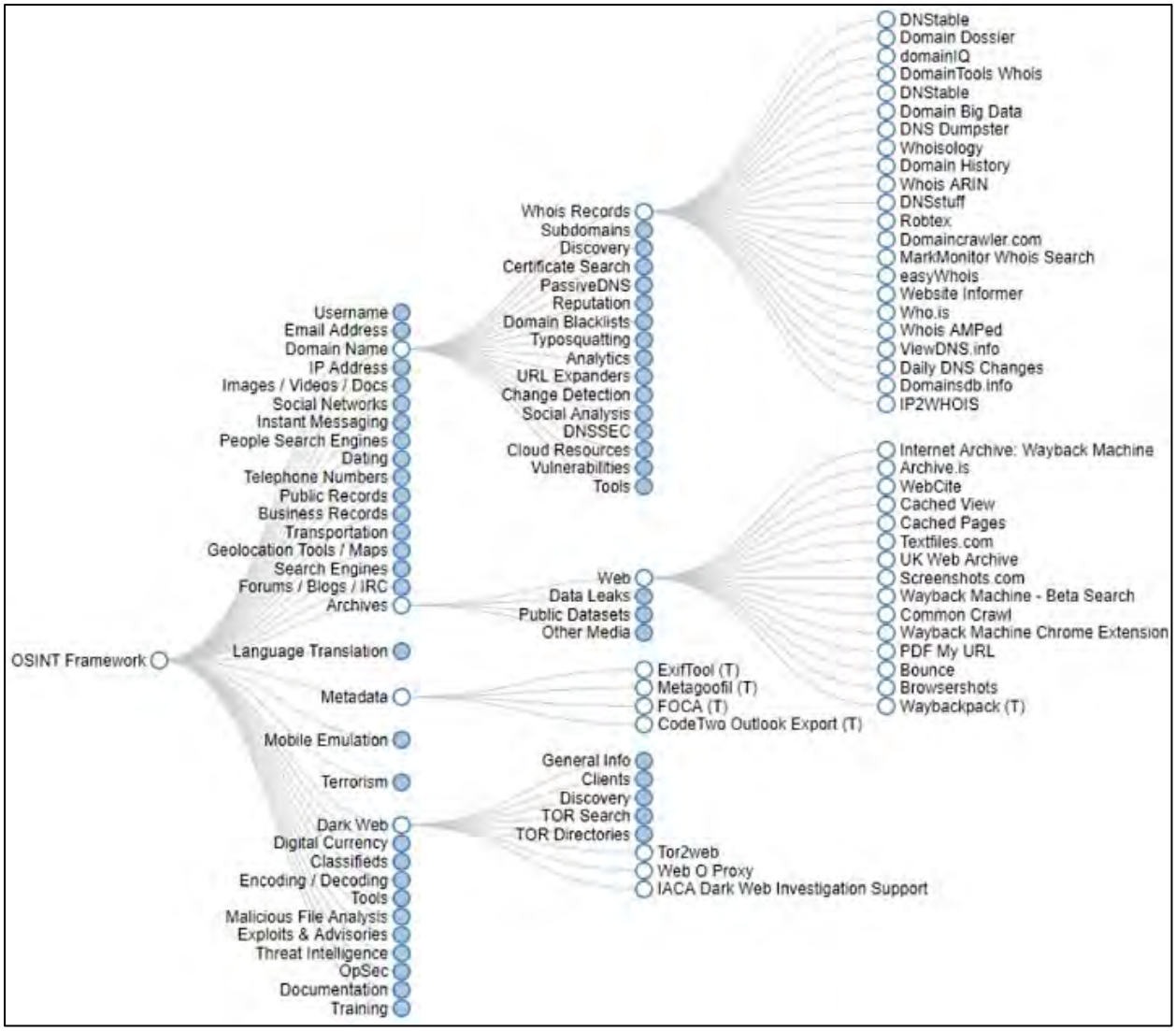

Figure 2-91: Screenshot of OSINT Framework

Recon-Dog

Recon-Dog serves as a comprehensive tool for all fundamental information collection requirements. It utilizes APIs to obtain details regarding the target system. Features include:

- **Censys:** Employs censys.io to amass extensive information about an IP address
- **NS lookup:** Executes name server queries
- **Port scan:** Checks the most frequently used TCP ports
- **Detect CMS:** Capable of identifying over 400 content management systems
- **Whois lookup:** Conducts a Whois search
- **Detect honeypot:** Leverages shodan.io to determine if the target is a honeypot
- **Find subdomains:** Uses findsubdomains.com to discover subdomains
- **Reverse IP lookup:** Performs a reverse IP search to identify domains linked to an IP address
- **Detect technologies:** Utilizes wappalyzer.com to recognize more than 1000 technologies
- **All:** Executes all utilities against the target

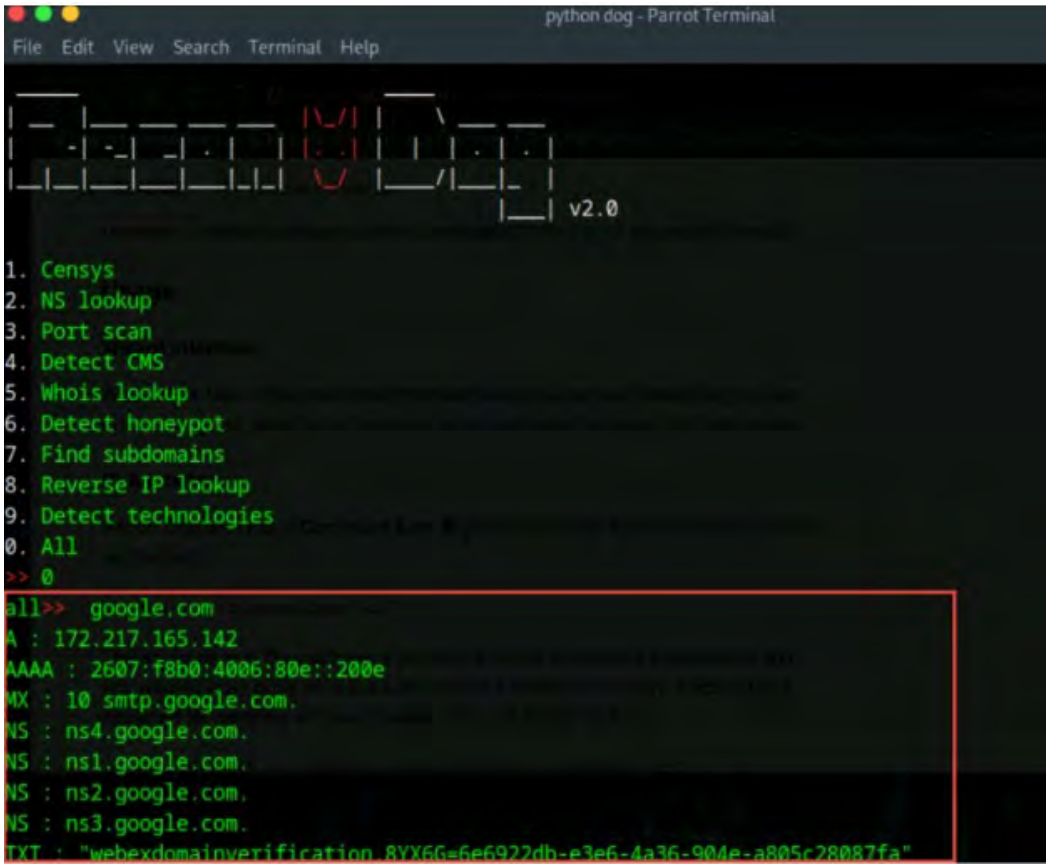

Figure 2-92: Screenshot of Recon-Dog

BillCipher

BillCipher serves as a tool for collecting information about a website or IP address. It is compatible with any operating system that runs Python 2, Python 3, and Ruby. This tool offers multiple features, including DNS lookup, Whois lookup, port scanning, zone transfer, host finder, and reverse IP lookup, all of which assist in gathering essential information.

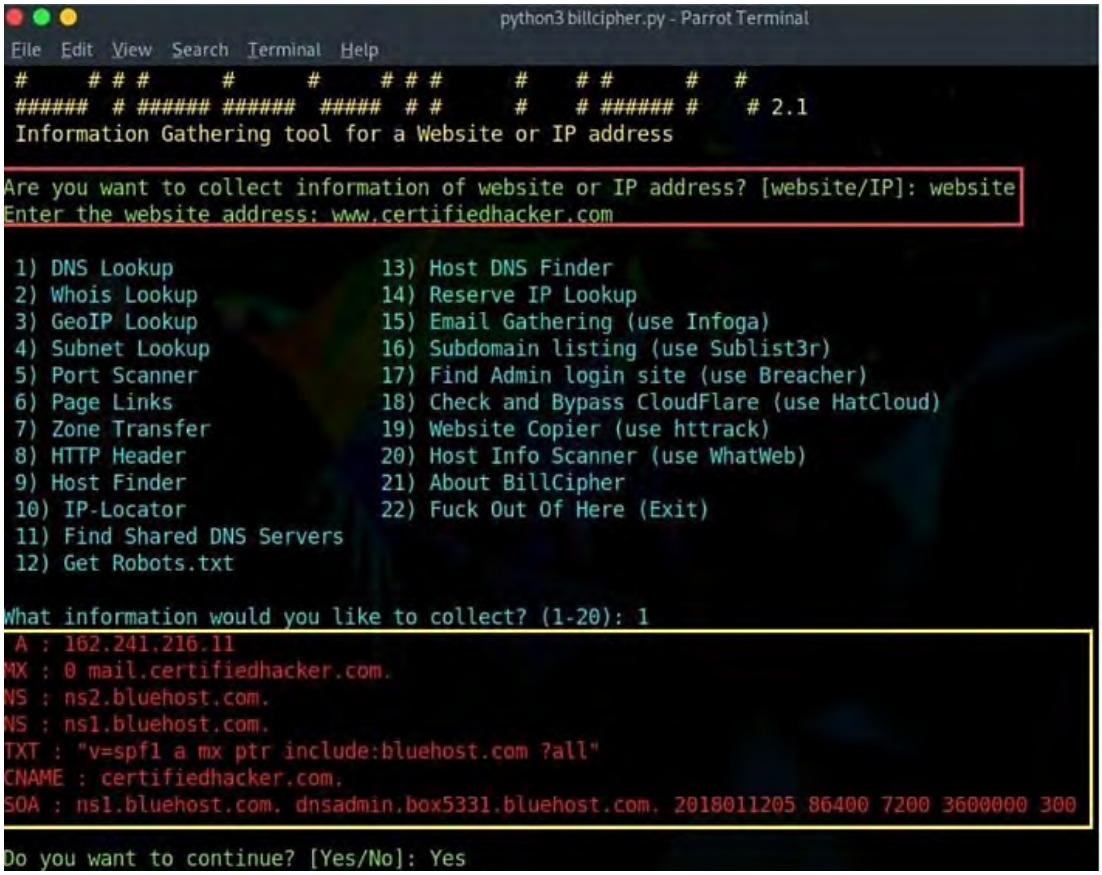

Figure 2-93: Screenshot of BillCipher

AI-Powered OSINT Tools

AI has transformed Open-Source Intelligence (OSINT) by greatly improving investigative skills through sophisticated data gathering, analysis, and forecasting. AI streamlines data processing, extracts pertinent insights, and provides actionable intelligence more effectively than conventional techniques while also augmenting OSINT tools.

AI-driven tools present many benefits for OSINT. Below are several key scenarios where AI can offer substantial support to OSINT researchers:

- **Web Scraping:** AI methodologies harness online information from sources like social media, blogs, forums, and deep web archives. This information allows for the tracking of entities over time or the observation of public behavior. Machine-learning algorithms can automate the retrieval of specific data, such as social media comments and responses.
- **Pattern Recognition:** Machine learning (ML) approaches can pinpoint entities within extensive datasets and examine files to uncover the connections among various entities, including names, company information, addresses, emails, phone numbers, and pertinent details.
- **Content Summarization**: NLP techniques can condense vast amounts of data into concise summaries. OSINT collectors can leverage this feature to extract relevant information from large datasets. For instance, an AI summarization tool can pull out company names from numerous PDF documents totaling hundreds of pages.
- **Sentiment Analysis:** AI systems can analyze text to gauge human emotions, which is especially beneficial for understanding public sentiment. OSINT researchers can utilize

AI to evaluate the emotional tone of users based on their social media posts and comments or to forecast consumer behavior from reviews.

- **Image Recognition**: A branch of AI known as computer vision can scrutinize digital media such as images and videos. In OSINT investigations, computer vision can aid in:
    - Face Recognition: Spotting and tracking individuals across various media.
    - Metadata Analysis: Extracting metadata from digital assets.
    - Reverse Image Search: Improving reverse image search functions and identifying deepfake images.
- **AI Detection:** AI can also recognize content produced by other AI systems, which is essential for spotting malicious activities enabled by AI.

Benefits of Integrating AI in OSINT

1. Enhanced Efficiency: AI boosts the efficiency of OSINT by streamlining tasks like web scraping and data extraction, which speeds up the process of data collection and analysis. This enables investigators to concentrate on more complex analysis and decision-making, ultimately accelerating investigations and providing timely insights.

2. Broader Reach: AI broadens the reach of OSINT by examining extensive data from the surface web, deep web, and dark web, ensuring a thorough intelligence gathering. Its capability to process large datasets and identify connections helps investigators uncover hidden patterns and relationships that are challenging to find manually.

3. Improved Visibility: AI improves the visibility of intelligence data by linking billions of seemingly unrelated data points into coherent networks of information, allowing investigators to swiftly recognize suspicious activities and establish connections between threat actors or events. AI-driven tools display these networks through user-friendly graphical interfaces, simplifying the task for investigators to identify and understand complex relationships and trends within the data.

4. Enhanced Investigator Safety: AI increases the safety of investigators by facilitating anonymized and automated investigations. This minimizes the risk of revealing an investigator's identity or jeopardizing sensitive information. AI tools are capable of conducting thorough investigations without requiring direct human involvement in potentially hazardous environments like the dark web.

Taranis AI

Taranis AI is a sophisticated OSINT tool that employs AI to boost information collection and situational analysis. It leverages NLP and AI to enhance the quality of data obtained from various sources, such as websites, for the purpose of gathering unstructured news articles. Analysts subsequently convert these AI-enhanced articles into structured reports that serve as the foundation for deliverables like PDF files that are ultimately published.

Features of Taranis AI include:

- **Enhanced OSINT Functionality:** Taranis AI taps into various data sources, encompassing websites, to gather unstructured news articles and delivers a thorough and enriched intelligence feed.
- **AI-Driven Analysis:** Taranis improves the quality and relevance of the collected articles through AI and NLP techniques.

- **Diverse Output Formats:** Taranis AI produces a range of end products, such as structured reports and PDF documents, customized to fulfill specific informational needs and criteria.
- **Effortless Publication:** The platform facilitates the straightforward publication of completed intelligence products, guaranteeing the prompt distribution of vital information to stakeholders.

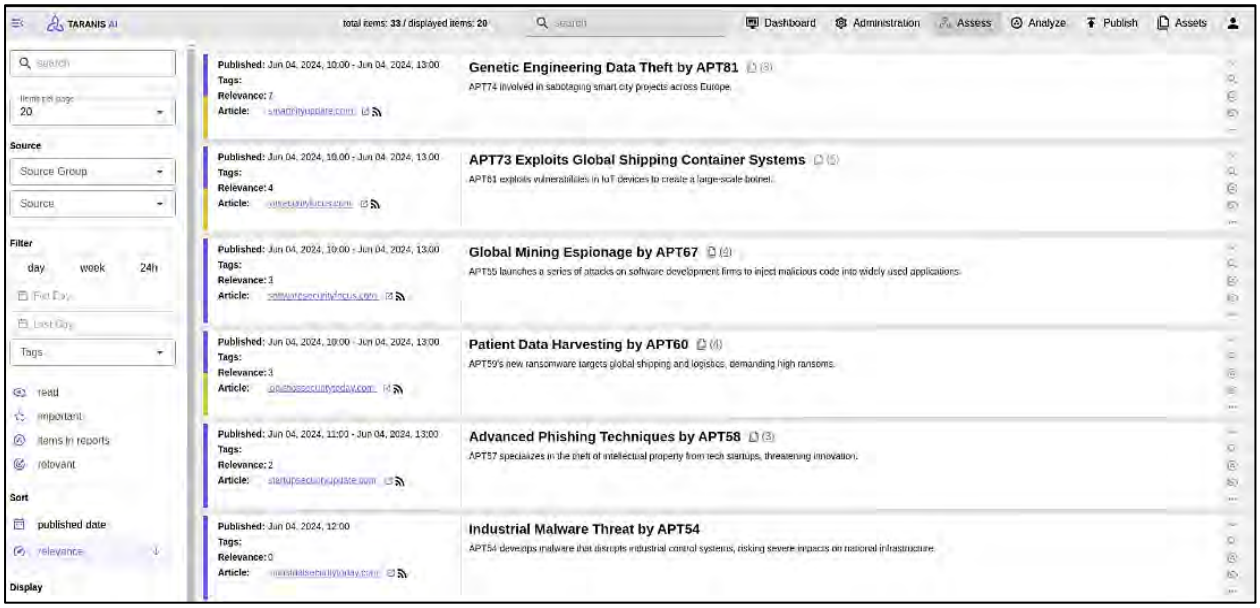

Figure 2-94: Screenshot of Taranis AI

OSS Insight

OSS Insight utilizes AI to explore the GitHub ecosystem by examining a vast dataset of over five billion GitHub events. This feature allows it to provide in-depth insights and tools that improve the understanding and navigation of the open-source landscape. From comprehensive repository analytics that include metrics like stars, forks, and commits to insights regarding developer productivity and collaboration trends, OSS Insight is equipped with robust resources for making informed decisions and strategic planning in open-source software development.

Key Features of OSS Insight include:

- **GPT-Powered Data Exploration:** This feature allows users to query GitHub data using natural language, create SQL queries, and visually present the findings. It aids ethical hackers in gathering intelligence on repositories, developer activities, and trends without needing advanced SQL knowledge, thus improving investigative efficiency.
- **Technical Fields Analytics:** The platform organizes GitHub collections within specific technical areas and offers analytics in fields like web frameworks, AI, and Web3. Ethical hackers can utilize this information to pinpoint emerging technologies' potential vulnerabilities and gain insights into specialized security environments.
- **Developer Analytics:** The tool tracks developer productivity indicators such as commits, pull requests, and code contributions while also examining collaboration patterns and engagement levels. Ethical hackers can leverage this to evaluate contributor activity and reliability, identify potential security practice weaknesses, and spot high-risk or low-activity contributors.

- **Repository Analytics:** This tool evaluates metrics related to GitHub repositories, including popularity (measured by stars and forks), update frequency, and community involvement. It provides historical trends and comparative insights that assist in benchmarking and strategic decision-making. Ethical hackers can assess the health and security posture of repositories, monitor how issues and pull requests are managed, and identify repositories that might be vulnerable to attacks due to low engagement or outdated code.
- **Compare Projects:** This feature allows for straightforward comparisons of metrics across various GitHub projects, including levels of activity, contributor demographics, issue resolution efficiency, and technical metrics. This assists ethical hackers in identifying projects that exhibit better security practices, active communities, or potential vulnerabilities for more strategic targeting and assessment.

The significance of OSS Insight in OSINT for ethical hacking includes:

- **Comprehensive Data Analysis:** By examining more than five billion events on GitHub, OSS Insight offers a vast array of data that ethical hackers can leverage to gather intelligence about software vulnerabilities, trending frameworks, and new developments. This information can enhance vulnerability evaluations and assist in recognizing potential attack points.
- **Real-time and Historical Data:** The combination of real-time data updates along with historical records from the GHArchive guarantees that ethical hackers have access to the latest information regarding ongoing changes and previous incidents within the GitHub environment.
- **AI-Powered Querying:** The AI-driven GitHub Data Explorer streamlines the process of querying intricate datasets using natural language, which makes it easier for ethical hackers to pull out specific information pertinent to their investigations.

Figure 2-95: Screenshot of OSS Insight

Additional AI-Enhanced OSINT Tools

AI tools for OSINT that utilize artificial intelligence to improve the effectiveness and precision of gathering open-source intelligence include:

DorkGPT

DorkGPT is an AI-driven tool intended to aid in Google Dorking, a method used to discover information that isn't readily accessible via ordinary search queries. It utilizes the functionality of Generative Pre-trained Transformer (GPT) models to create and refine search queries, assisting users in revealing sensitive details, hidden pages, and other pertinent data for cybersecurity, ethical hacking, or research objectives.

DorkGenius

DorkGenius is an AI-driven application that automates the Google Dorking process and assists users in formulating advanced search queries to locate specific information online. It is particularly beneficial for identifying hidden files, directories, sensitive details, and security weaknesses, especially for ethical hackers.

Google Word Sniper

Google Word Sniper aids in refining search queries for more effective results on Google. It targets specific keywords and phrases, making it simpler to locate particular information, concealed content, and specialized data. This tool is advantageous for researchers, marketers, and cybersecurity experts, as it boosts their capability to discover valuable information buried within search results.

Cylect.io

Cylect.io is a sophisticated OSINT tool powered by AI that consolidates numerous databases into an intuitive interface, offering a wide array of resources for ethical hackers and facilitating effective and confident OSINT investigations. Created to overcome the shortcomings of conventional search engines, Cylect.io streamlines the search process and improves both the speed and precision of data gathering in investigative scenarios.

ChatPDF

ChatPDF serves as an OSINT tool that utilizes AI to analyze and extract information from PDF documents through a conversational interface. Users can upload PDF files and engage with the tool to swiftly obtain specific data, summaries, and insights, making it a highly beneficial resource for ethical hacking.

Bardeen.ai

Bardeen.ai functions as an automation tool that can be utilized for OSINT by allowing users to simplify and automate the processes of data collection and analysis from various online sources. This increases the speed and precision of OSINT activities, making it a valuable asset for cybersecurity professionals, researchers, and investigators.

DarkGPT

DarkGPT is an AI assistant that employs GPT-4-200K to query leaked databases, facilitating efficient and targeted searches within compromised data sources. This allows users to obtain crucial information and insights, enhancing the OSINT abilities of cybersecurity analysts and researchers.

PenLink Cobwebs

PenLink Cobwebs is a cutting-edge OSINT tool powered by AI that specializes in collecting and analyzing data from various online sources. It provides extensive capabilities for gathering, processing, and visualizing information to aid cybersecurity investigations.

Explore AI

Explore AI is an AI-driven YouTube search engine that employs artificial intelligence to search for and extract information from YouTube videos, simplifying access to information for ethical hacking objectives.

AnyPicker

AnyPicker is a robust visual web scraping and AI OSINT tool that enables users to collect data from websites without the need for programming knowledge. This tool allows for the simultaneous scraping of multiple pages and offers a live preview of the extracted results, enhancing flexibility and efficiency in gathering web data.

Create and Run Custom Python Script to Automate Footprinting Tasks with Al

Attackers can utilize AI-driven technologies to improve and automate their footprinting tasks. With AI's support, these individuals can easily develop and execute personalized footprinting scripts to gather crucial information about their targets. By creating specialized scripts powered by Artificial Intelligence (AI), they can effectively perform a sequence of website footprinting commands to collect data regarding a target domain for cybersecurity evaluations.

For instance, attackers may employ ChatGPT to assist in crafting such a script by using a suitable prompt like:

"Develop a Python script which will accept the domain name www.microsoft.com as input and execute a series of website footprinting commands, including DNS lookups, WHOIS records retrieval, email enumeration, and more, to gather information about the target domain"

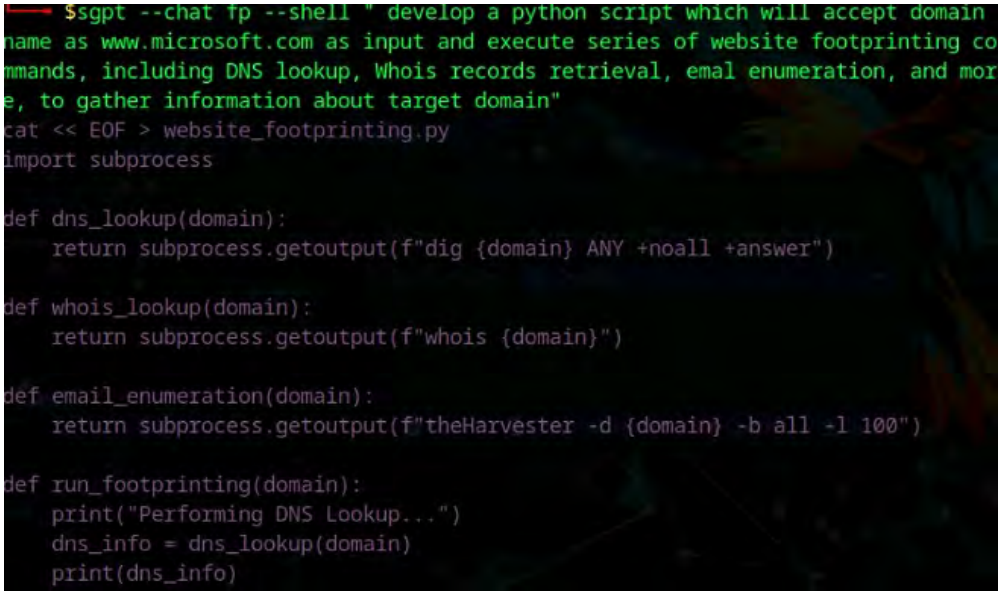

Figure 2-96: Prompt and Output for performing Traceroute with AI – Screenshot-1

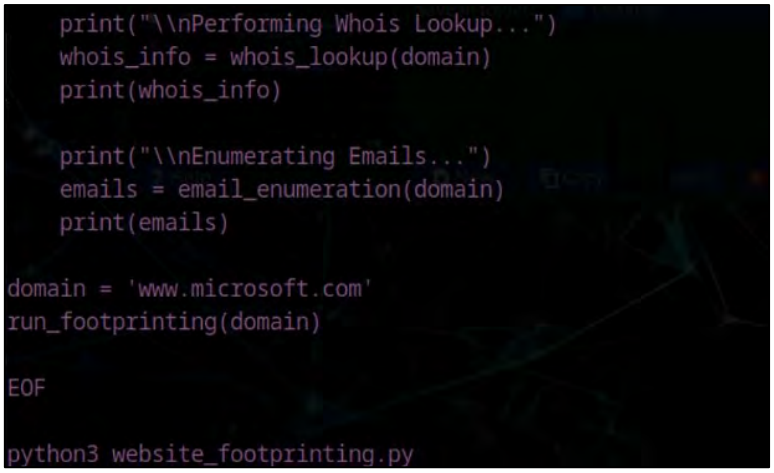

Figure 2-97: Prompt and Output for performing Traceroute with AI – Screenshot-2

The following shell script is created to conduct website footprinting, which includes DNS lookups, retrieval of WHOIS records, email enumeration, and additional methods to collect information about the specified domain: www.microsoft.com.

```
cat <<< EOF > website_footprinting.py
import subprocess

def dns_lookup(domain):
    return subprocess.getoutput(f"dig {domain} ANY +noall +answer")
def whois_lookup(domain):
    return subprocess.getoutput(f"whois {domain}")

def email_enumeration(domain):
    return subprocess.getoutput(f"theHarvester -d {domain} -b all -l 100")

def run_footprinting(domain):
    print("Performing DNS Lookup...")
    dns_info = dns_lookup(domain)
    print(dns_info)

    print("\nPerforming Whois Lookup...")
    whois_info = whois_lookup(domain)
    print(whois_info)

    print("\nEnumerating Emails...")
    emails = email_enumeration(domain)
    print(emails)

domain = 'www.microsoft.com'
run_footprinting(domain)
EOF
python3 website_footprinting.py
```

Figure 2-98: Shell Script to perform Website Footprinting

This Python script includes four functions: dns_lookup, whois_lookup, email_enumeration, and run_footprinting.

Table 2-16 describes the four functions used in the Python Script.

Functions	Description
dns_lookup (domain)	Conducts a DNS lookup for the provided domain utilizing the dig command.
whois_lookup(domain)	Fetches WHOIS records for the given domain by means of the whois command.
email_enumeration(domain)	Gathers email addresses linked to the specified domain using theHarvester tool.
run_footprinting(domain)	Carries out a series of web footprinting commands (DNS lookup, WHOIS lookup, email enumeration) for the specified domain and displays the results.

Table 2-16: Python Script Functions with Descriptions

This script can be executed with Python3 to carry out website footprinting on the given domain (www.microsoft.com).

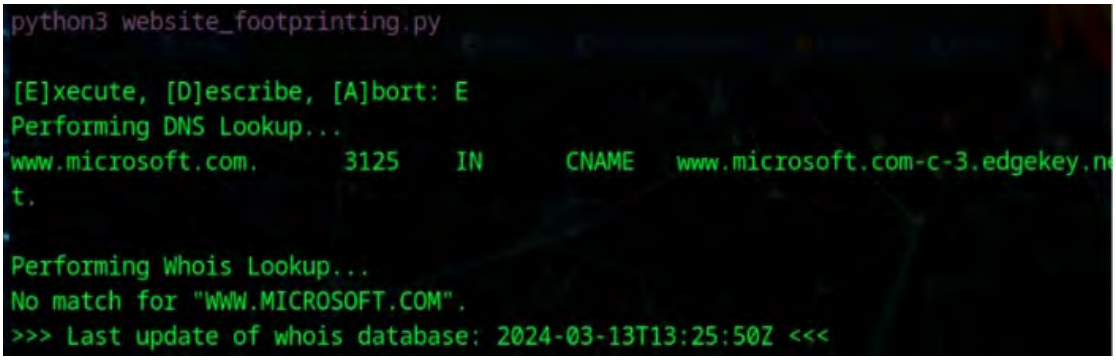

Figure 2-99: Output file for script with AI – Screenshot-1

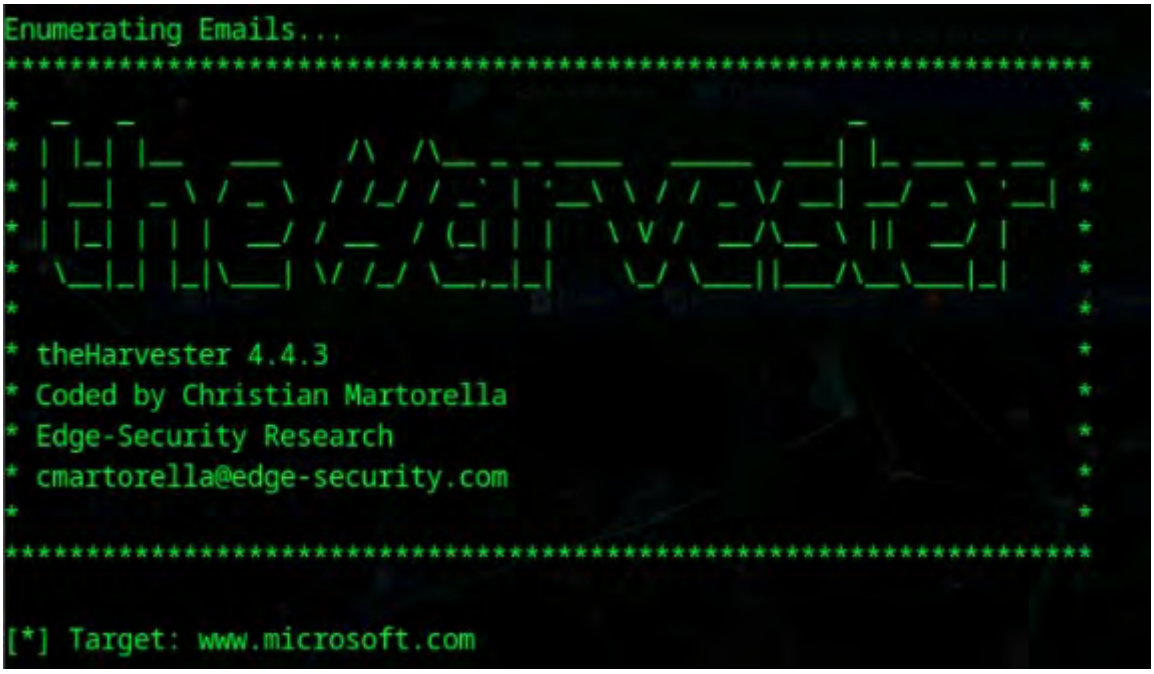

Figure 2-100: Output file for script with AI – Screenshot-2

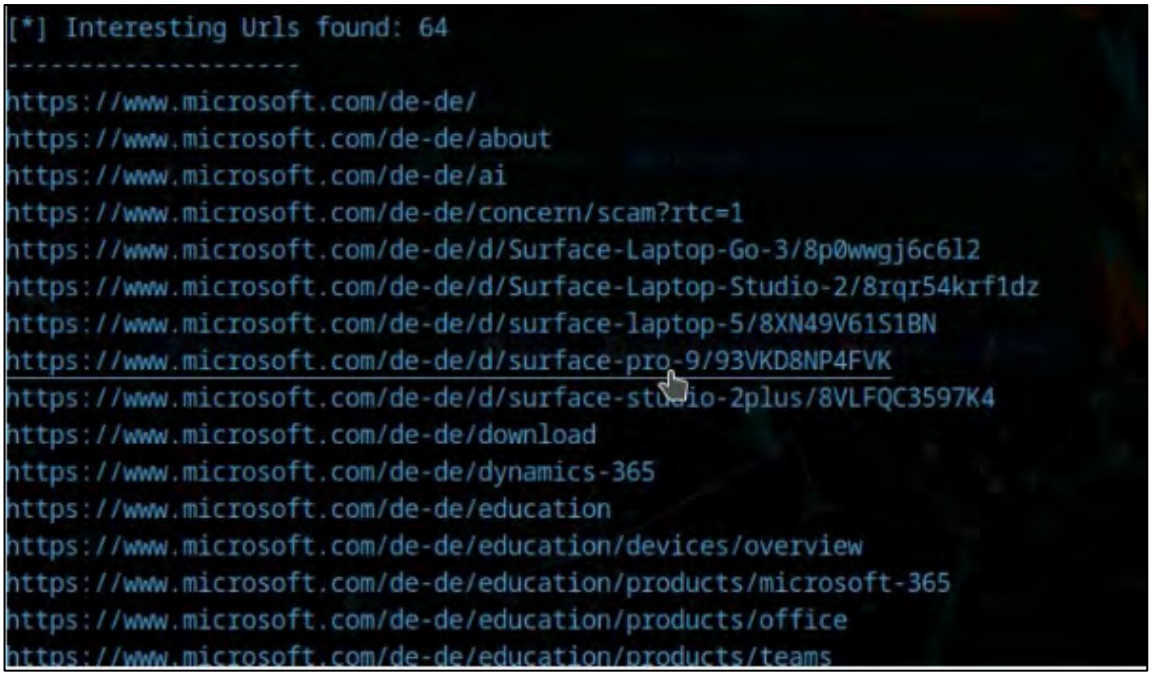

Figure 2-101: Output file for script with AI – Screenshot-3

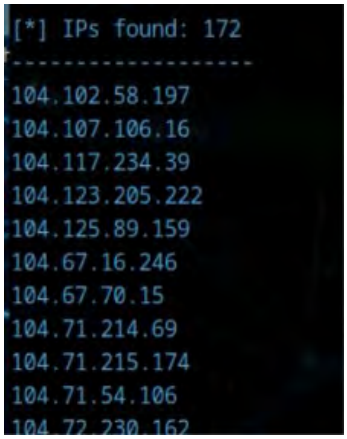

Figure 2-102: Output file for script with AI – Screenshot-4

Footprinting Countermeasures

This section discusses measures to counter footprinting, specifically actions taken to prevent or mitigate the disclosure of information. Some of the countermeasures against footprinting include:

- Limit employees' access to social media platforms from the organization's network
- Set up web servers to prevent any information leakage
- Train employees to use false identities on blogs, forums, and groups
- Avoid disclosing sensitive information in press releases, annual reports, and product catalogs
- Minimize the volume of information shared on a website or the Internet
- Employ footprinting techniques to identify and eliminate any sensitive data that is publicly accessible
- Stop search engines from caching web pages and utilize anonymous registration services

- Create and enforce security protocols like information security and password policies to control the information employees can disclose to external parties
- Use multi-factor authentication methods to strengthen the security of the organization's systems and assets
- Separate internal and external DNS or implement split DNS and limit zone transfers to authorized servers
- Disable directory listings on web servers
- Conduct ongoing security awareness training to inform employees about various social engineering tactics and threats
- Choose privacy services for a Whois lookup database
- Refrain from domain-level cross-linking for sensitive assets
- Secure and encrypt sensitive information with password protection
- Use captchas and rate limiting on public-facing services to curb automated tools from rapidly gathering information
- Disable any unnecessary protocols
- Always utilize TCP/IP and IPsec filters for additional layers of defense
- Configure Internet Information Services (IIS) to prevent information disclosure through banner grabbing
- Conceal the IP address and associated information by using a VPN or placing the server behind a secure proxy
- Request archive.org to erase the website's history from the archive database
- Keep the domain name profile confidential
- Store critical documents like business plans and proprietary material offline to avoid exploitation
- Educate employees to counter social engineering methods and attacks
- Cleanse the information shared with Internet registrars to obscure the organization's direct contact details
- Turn off geo-tagging features on cameras to avoid geolocation tracking
- Refrain from disclosing personal location or travel plans on social media platforms
- Disable geolocation access on all mobile devices when not needed
- Ensure that no sensitive information, such as strategic plans, product details, or sales forecasts, is visible on notice boards or walls
- Deactivate or delete the accounts of former employees
- Set mail servers to ignore messages from anonymous senders
- Deploy honeypots or honeynets within the network to attract and identify attackers, diverting potential footprinters away from critical systems

Summary

Footprinting involves the passive collection of publicly available information about a target system, network, or organization to identify vulnerabilities while minimizing detection. It includes passive methods that rely on public data and active methods that involve direct interaction.

Search engines play a vital role in identifying sensitive information, such as exposed directories and configuration files, using techniques like Google hacking and dorking to uncover unintended data leaks. Specialized platforms like Shodan, combined with social media analysis,

enable attackers to collect detailed organizational insights, including hierarchies and physical locations.

Whois databases offer access to domain registration details, while DNS records reveal critical information about network infrastructure. Network and email footprinting techniques involve mapping network ranges and examining email headers. Advanced tools, machine learning, and OSINT frameworks further streamline and enhance the efficiency of information-gathering processes.

The chapter concludes with strategies to mitigate risks, such as limiting public information and training employees to resist social engineering.

Mind Map

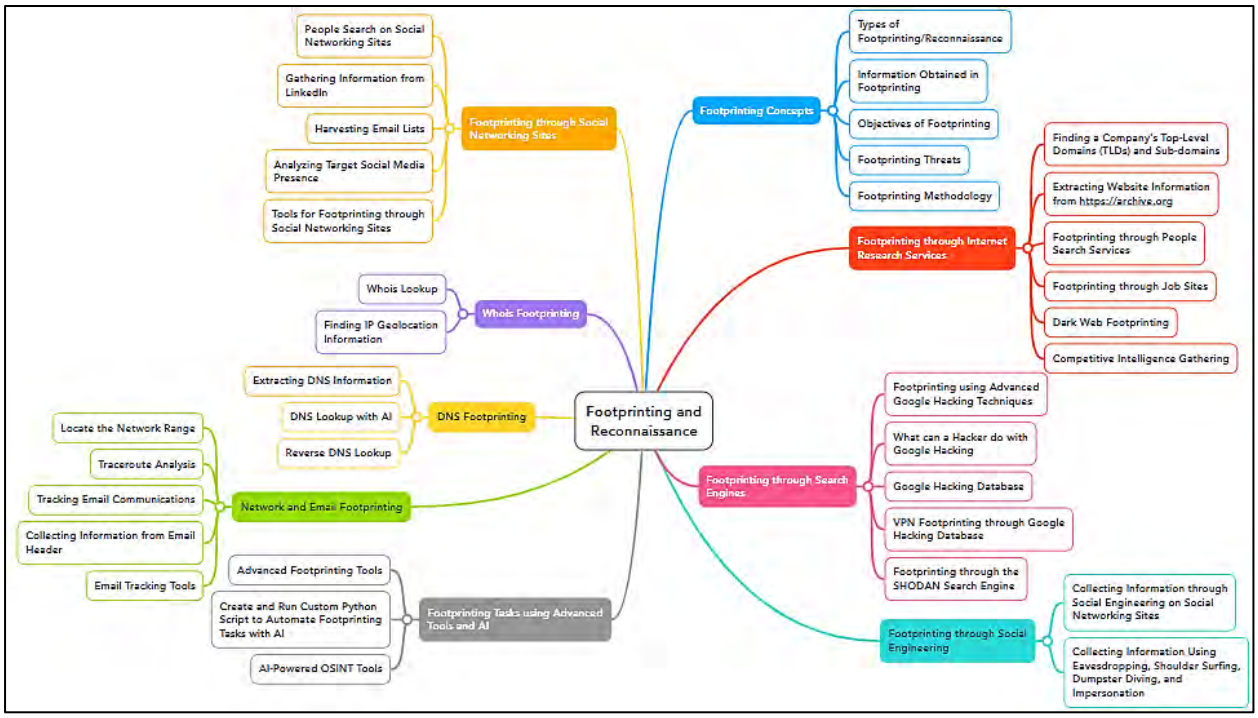

Figure 2-103: Mind Map

Practice Questions

1. Footprinting is the process of gathering _____ about a target organization to identify potential vulnerabilities.

A. public

B. private

C. encrypted

D. unauthorized

2. Which type of footprinting involves using pretexting or impersonation?

A. Network footprinting

B. Social engineering footprinting

C. Email footprinting

D. DNS footprinting

3. What information can a Whois query reveal?

A. DNS records

B. Domain registration details

C. IP geolocation data

D. Encrypted user credentials

4. Google Dorking is a technique used for footprinting through search engines.

A. True

B. False

5. What is the process of collecting information about a target system or network without interacting with it directly?

A. Active Footprinting

B. Aggressive Footprinting

C. Passive Footprinting

D. Intrusive Footprinting

6. _____ search engine is often used for advanced technical searches and finding specific file types.

A. Google

B. Bing

C. DuckDuckGo

D. Shodan

7. Which of the following represents the most effective sequence for a reconnaissance plan during footprinting?

A. Identify network ranges → Perform social engineering → Query Whois database → Enumerate DNS records

B. Query Whois database → Enumerate DNS records → Identify network ranges → Use social engineering

C. Enumerate DNS records → Query Whois database → Use social engineering → Identify network ranges

D. Perform social engineering → Identify network ranges → Query Whois database → Enumerate DNS records

8. What is the primary purpose of MX records retrieved during DNS footprinting?

A. To identify mail exchange servers for a domain

B. To provide a backup for DNS records

C. To map IP addresses to hostnames

D. To redirect traffic to an external CDN

9. How can AI-driven tools significantly enhance footprinting processes?

A. By automating data collection and analysis in real time

B. By bypassing firewalls for more extensive reconnaissance

C. By decrypting secured DNS traffic

D. By enabling advanced payload delivery

10. Setting up a Virtual Private Network (VPN) is a viable countermeasure to hide your real IP address during active footprinting attempts.

A. True

B. False

11. theHarvester query _____ platform to discover employees and their job roles.

A. GitHub

B. LinkedIn

C. Facebook

D. Twitter

12. Which type of information can traceroute reveal during network footprinting?

A. Email spoofing attempts

B. Geolocation of the target's DNS server

C. Encrypted traffic flows

D. IP addresses of intermediate routers

13. How can attackers misuse tools like Google Earth and Wikimapia?

A. To deploy malware on a target's network

B. To identify building entrances, security cameras, and vulnerabilities

C. To brute force an organization’s Wi-Fi network

D. To generate phishing emails targeting employees

14. _____ browser is required to access the dark web.

A. Chrome

B. Tor

C. Safari

D. Microsoft Edge

15. What is Shodan primarily used for in footprinting?

A. Searching for leaked credentials

B. Exploiting web application vulnerabilities

C. Analyzing encrypted network traffic

D. Identifying and enumerating internet-connected devices

16. Which tactic can attackers use to gather information from users on social networking sites without directly contacting them?

A. Sending friend requests to obtain private information

B. Analyzing publicly available profile information to deduce potential passwords

C. Phishing users through fake login pages

D. Using malware to infect the target's device

17. Which of the following is a common countermeasure against DNS footprinting?

A. Disabling zone transfers for DNS servers

B. Enabling ICMP requests to the public

C. Publishing internal IP addresses in DNS records

D. Using DNS over HTTPS (DoH)

18. Which of the following tools is commonly used to visually display the route taken by packets across a network and its geographical locations?

A. tracert

B. NetScanTools Pro

C. PingPlotter

D. Traceroute NG

19. _____ tool can be used to trace the path between the local system and a target system, displaying the hops along the way and their geographical locations.

A. Nmap

B. Metasploit

C. Tracert

D. Burp Suite

20. In which phase of a penetration test is theHarvester most commonly used?

A. Exploitation phase

B. Initial reconnaissance phase

C. Reporting phase

D. Post-exploitation phase

21. _____ protocol is most commonly used by traceroute for its operations in IPv4 networks.

A. TCP

B. ICMP

C. UDP

D. HTTP

22. What type of information can attackers commonly find on the dark web about a target organization?

A. Corporate social media passwords

B. Employee activity logs from their office computers

C. Leaked credentials, database dumps, and intellectual property

D. Internal network configurations in real time

23. Which feature of DNSRecon supports reverse DNS lookups?

A. Zone transfer analysis

B. PTR record enumeration

C. DNS spoofing detection

D. Real-time DNS query interception

24. What is the primary purpose of competitive intelligence gathering in cybersecurity?

A. To collect information about the competitor's business strategies and vulnerabilities

B. To exploit weaknesses in a competitor's software products

C. To disrupt a competitor's network operations

D. To monitor competitor's marketing efforts

25. Which of the following Google Dork queries is commonly used to locate exposed VPN login portals?

A. inurl:"/vpn/login" filetype:php

B. intitle:"VPN admin" inurl:login

C. inurl:"/login" site:*.vpn

D. inurl:/sslvpn/Login?login

Answers

1. **Answer: A**

Explanation: Footprinting involves gathering public information about a target organization from accessible sources, such as websites, social media, WHOIS databases, and other publicly available data. This information is used to identify potential vulnerabilities or weak points in the organization's security posture.

2. Answer: B

Explanation: Social engineering footprinting involves using pretexting or impersonation to gather sensitive information from individuals within the target organization. Attackers may pose as a trusted entity (e.g., a vendor or employee) to deceive people into disclosing valuable data, such as login credentials or confidential information.

3. Answer: B

Explanation: A Whois query reveals domain registration details, such as the domain owner's name, contact information, registration and expiration dates, and the domain's associated nameservers. This information can be useful for identifying the organization or individual behind a domain.

4. Answer: A

Explanation: Google Dorking is a technique used for footprinting through search engines. It involves using advanced search operators (also known as "Google Dorks") to find specific information on websites, such as vulnerable files, exposed sensitive data, or misconfigured systems.

5. Answer: C

Explanation: Passive Footprinting involves collecting information about a target system or network without directly interacting with it. This is done by gathering publicly available data, such as WHOIS records, social media posts, and domain information, without sending probes or making direct queries to the target system.

6. Answer: D

Explanation: Shodan is a search engine specifically designed for advanced technical searches. It allows users to find internet-connected devices, including servers, routers, and IoT devices. Shodan is particularly useful for identifying specific file types, devices, and vulnerabilities based on their banners, ports, and other technical information.

7. Answer: B

Explanation: Starting with the Whois database provides foundational details like domain ownership. Then, enumerating DNS records reveals critical information about the target's infrastructure. Identifying network ranges allows for further mapping of the target's network. Finally, social engineering can leverage all of this information to manipulate individuals within the organization for more sensitive details.

8. Answer: A

Explanation: Mail Exchange (MX) records in DNS footprinting are used to identify the mail servers responsible for receiving email for a domain. By retrieving MX records, an attacker or security analyst can identify the mail exchange servers and assess their configuration or potential vulnerabilities.

9. Answer: A

Explanation: AI-driven tools can significantly enhance footprinting processes by automating data collection and analyzing vast amounts of information in real time. These tools can process large volumes of open-source data, identify patterns, and extract relevant details faster than manual methods.

10. Answer: A

Explanation: Setting up a Virtual Private Network (VPN) is a viable countermeasure to hide your real IP address during active footprinting attempts. A VPN routes your internet traffic through a remote server, masking your original IP address and making it more difficult for attackers or network administrators to trace your activities.

11. Answer: B

Explanation: theHarvester is a tool used for gathering Open-Source Intelligence (OSINT) during the footprinting phase of a penetration test. It can query platforms like LinkedIn to discover employees of a target organization along with their job roles.

12. Answer: D

Explanation: Traceroute is a network diagnostic tool that reveals the IP addresses of intermediate routers along the path taken by packets from the source to the destination. By tracing the route, it helps map out the network topology, showing each hop between routers and servers.

13. Answer: B

Explanation: Tools like Google Earth and Wikimapia can be misused by attackers to visually assess physical locations. They can help attackers identify building entrances, security cameras, gates, vulnerabilities in perimeter fences, and other aspects of the physical infrastructure of a target organization.

14. Answer: B

Explanation: To access the dark web, a specialized browser like Tor (The Onion Router) is required. Tor allows users to browse the internet anonymously by routing traffic through multiple encrypted layers, which helps conceal their identity and location. Websites on the dark web typically have ".onion" domains, which can only be accessed using the Tor browser.

15. Answer: D

Explanation: Shodan is primarily used for identifying and enumerating internet-connected devices. It allows users to search for devices such as servers, routers, webcams, industrial control systems, and other IoT devices that are exposed to the internet.

16. Answer: B

Explanation: Attackers can gather information from users on social networking sites by analyzing publicly available profile information. This includes details like birthdates, names of family members, pets, favorite places, or interests, which can be used to guess security questions or deduce potential passwords.

17. Answer: A

Explanation: A common countermeasure against DNS footprinting is disabling zone transfers for DNS servers. Zone transfers allow a DNS server to share its zone file (which contains domain information) with another server. If not properly secured, attackers can exploit this to gather a lot of sensitive information about the domain and its associated IP addresses.

18. Answer: C

Explanation: PingPlotter is commonly used to visually display the route taken by packets across a network, along with its geographical locations. It provides real-time visualization of the network path, helping users track latency, packet loss, and other network performance metrics.

19. Answer: C

Explanation: Tracert (or Traceroute) is a network diagnostic tool that can be used to trace the path between the local system and a target system. It displays the intermediate hops along the way, showing the IP addresses of routers or devices that the packets pass through.

20. Answer: B

Explanation: theHarvester is most commonly used in the initial reconnaissance phase of a penetration test. During this phase, the tool helps gather Open-Source Intelligence (OSINT) by collecting information such as email addresses, domain names, subdomains, and employee details from various public sources like search engines, social media platforms, and WHOIS records.

21. Answer: B

Explanation: ICMP (Internet Control Message Protocol) is most commonly used by traceroute for its operations in IPv4 networks. Traceroute sends ICMP Echo Request messages to each hop along the route to the destination and waits for ICMP Time Exceeded messages to return. These messages help identify the path taken by packets from the source to the destination.

22. Answer: C

Explanation: On the dark web, attackers can commonly find leaked credentials, database dumps, and intellectual property related to a target organization. This information may be sold or shared in underground marketplaces, making it accessible to malicious actors.

23. Answer: B

Explanation: PTR (Pointer) record enumeration is the feature of DNSRecon that supports reverse DNS lookups. PTR records map IP addresses to domain names, which is the reverse of the typical DNS query that maps domain names to IP addresses.

24. Answer: A

Explanation: The primary purpose of competitive intelligence gathering in cybersecurity is to collect information about a competitor's business strategies and vulnerabilities. This includes analyzing publicly available information to assess potential weaknesses in their products, services, and security practices. The goal is to gain insights into how a competitor operates and identify areas where they may be vulnerable to cyberattacks.

25. Answer: D

Explanation: The Google Dork query inurl:/sslvpn/Login?login is commonly used to locate exposed SSL VPN login portals. The query specifically targets URLs that contain the /sslvpn/Login?login path, which is often used by SSL VPN login pages.