Reconnaissance Paper

-Samriddhi

Executive Summary

This paper provides a thorough overview of the digital footprint of the Broad Institute of MIT and Harvard, identifying numerous subdomains, hosts, email addresses, IP addresses, DNS records, internal documents, and cloud storage buckets linked to the company using a variety of passive reconnaissance techniques. Malicious actors might use these resources to undertake focused attacks on the Broad Institute's infrastructure.

The results demonstrated that the institute's website runs on a NGINX HTTP server and makes use of HTML5. Additional information on network components, hosting, IP addresses, SSL/TLS versions, and geolocation was obtained using tools such as Netcraft. These facts provide useful information that threat actors can use to exploit security weaknesses.

Additionally, Metagoofil helped discover various public documents containing sensitive information on subjects like genetic mutations and healthcare data, which may contain proprietary research or sensitive PII. Such data poses a security risk as it could be misused for identity theft or data leaks.

The Broad Institute should implement several preventative security measures to reduce these threats. Regular vulnerability assessments and penetration tests, secure coding techniques, timely security patch and update applications, careful website and network infrastructure monitoring, limiting access to sensitive data, and frequent security and awareness training for staff are a few of these.

In conclusion, the results underscore the need for a proactive cybersecurity strategy, stressing the necessity of ongoing observation, assessment, and adjustment to protect the organization's security posture from possible dangers.

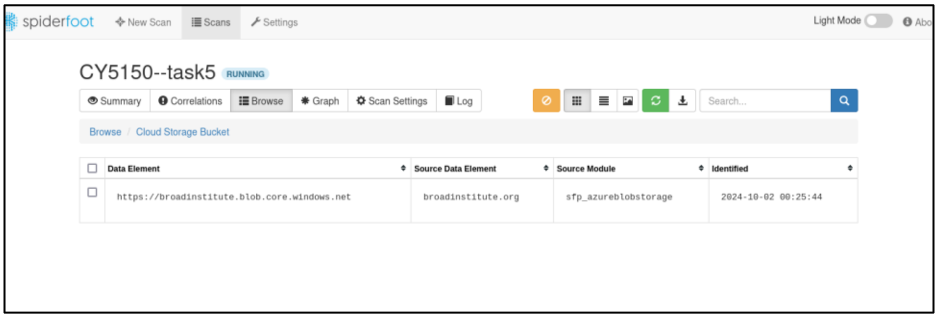
Introduction

The Broad Institute of MIT and Harvard is a world-renowned biomedical and genomic research institute founded in 2004. The institute focuses on advancing knowledge in areas such as genomics, personalized medicine, chemical biology, cancer research, and infectious diseases. With a collaborative approach, the Broad Institute brings together scientists from MIT, Harvard, and affiliated hospitals to tackle complex scientific challenges through interdisciplinary research. The institute is dedicated to improving human health by driving innovation in biomedicine, making groundbreaking discoveries, and developing transformative technologies. Its contributions span various sectors, including healthcare, pharmaceuticals, biotechnology, and academic research, positioning the Broad Institute as a leading force in the life sciences.

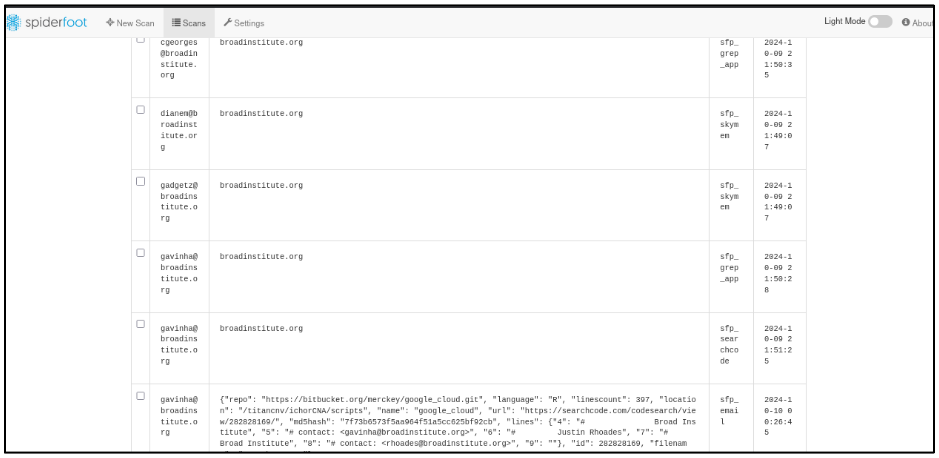
As a repository of large-scale genomic data, Broad Institute would be an interesting case study for understanding how high-value organizations protect sensitive information, manage cyber risks, and, by exploring their digital footprint, can reveal valuable networking or partnership data.

Passive Reconnaissance Tools

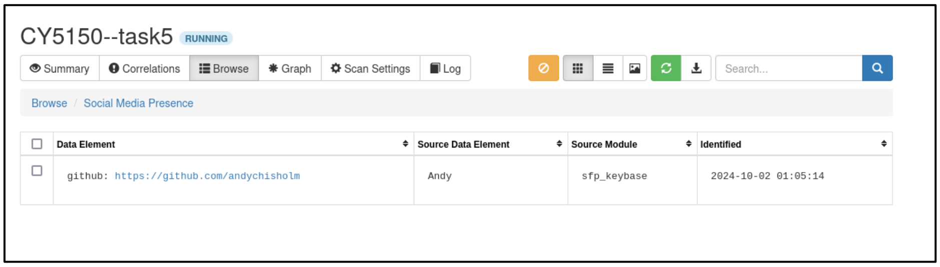
1. **Spiderfoot**: It is a comprehensive OSINT tool that is used for domain discovery. It collects data from multiple sources and gives a detailed description of emails, domains, IP address information, certificate logs, etc. for a target website.
    - Summary of Output: The Spiderfoot scan showed information about cloud storage, email addresses of actual employees of Broad Institute, and social media presence.



☐	Data Element	Source Data Element	Source Module	Identified
☐	https://broadinstitute.blob.core.windows.net	broadinstitute.org	sfp_azureblobstorage	2024-10-02 00:25:44



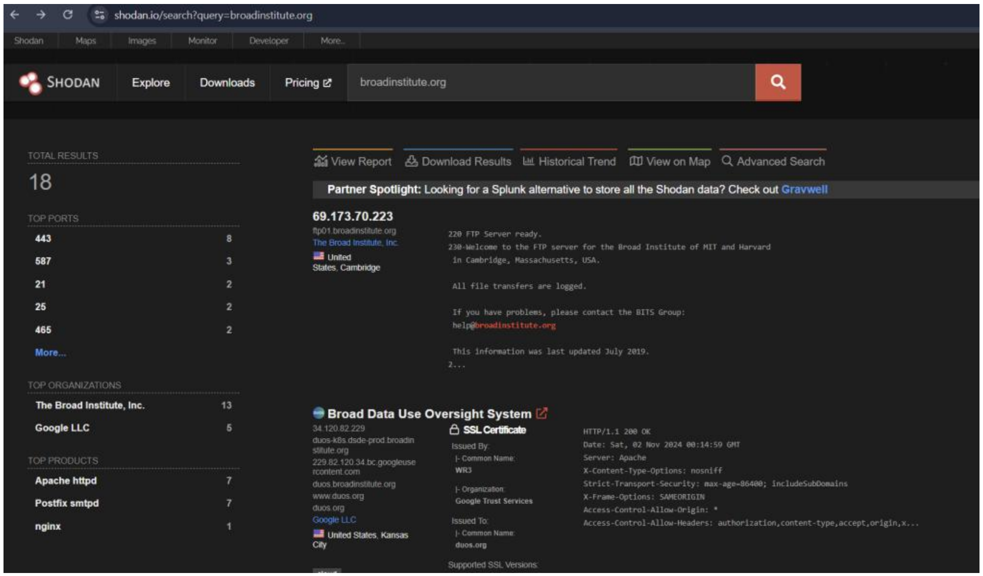
☐	cgeorges@broadinstitute.org	broadinstitute.org	sfp_grep_app	2024-10-09 21:50:35
☐	dianem@broadinstitute.org	broadinstitute.org	sfp_skymem	2024-10-09 21:49:07
☐	gadgetz@broadinstitute.org	broadinstitute.org	sfp_skymem	2024-10-09 21:49:07
☐	gavinha@broadinstitute.org	broadinstitute.org	sfp_grep_app	2024-10-09 21:50:28
☐	gavinha@broadinstitute.org	broadinstitute.org	sfp_searchcode	2024-10-09 21:51:25
☐	gavinha@broadinstitute.org	{"repo": "https://bitbucket.org/merckey/google_cloud.git", "language": "R", "linescount": 397, "location": "/titancnv/ichorCNA/scripts", "name": "google_cloud", "url": "https://searchcode.com/codesearch/view/282828169/", "md5hash": "7f73b6573f5aa964f51a5cc625bf92cb", "lines": {"4": "# Broad Institute", "5": "# contact: <gavinha@broadinstitute.org>", "6": "# Justin Rhoades", "7": "# Broad Institute", "8": "# contact: <rhoades@broadinstitute.org>", "9": ""}, "id": 282828169, "filenam	sfp_email	2024-10-10 00:26:45



☐	Data Element	Source Data Element	Source Module	Identified
☐	github: https://github.com/andychisholm	Andy	sfp_keybase	2024-10-02 01:05:14

    - How information can be used by an attacker: Attackers can exploit the email addresses, cloud storage information from the Spiderfoot's scan to social engineer, do phishing attacks that could lead to data breaches and unauthorized access to the organization. Additionally, the social media presence of an employee can be used for malicious purposes.
    - Suggested Controls: Broad Institute can use email filtering and monitoring and train employees in phishing and social engineering attacks. They should also limit public exposure to email addresses by using generic contact forms on websites.

2. **Shodan**: A search engine used for ports and devices on the internet. It gives information about operating systems, hosts, open ports, etc.
    - Summary of Output: The shodan search showed that Broad Institute is using some of the unsecured ports like port 25,21,80. It also gave information about the SSL certificate used by the organization.



Partner Spotlight: Looking for a Splunk alternative to store all the Shodan data? Check out Gravwell

TOTAL RESULTS

18

TOP PORTS	
443	8
587	3
21	2
25	2
465	2
More...	

TOP ORGANIZATIONS	
The Broad Institute, Inc.	13
Google LLC	5

TOP PRODUCTS	
Apache httpd	7
Postfix smtpd	7
nginx	1

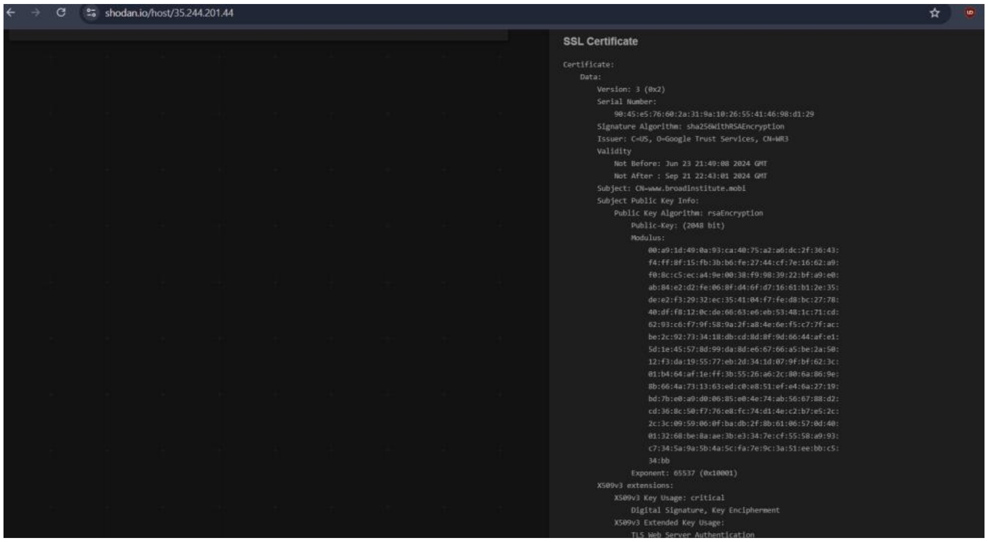
69.173.70.223
ftp01.broadinstitute.org
The Broad Institute, Inc.
United States, Cambridge

```
220 FTP Server ready.
230-Welcome to the FTP server for the Broad Institute of MIT and Harvard
 in Cambridge, Massachusetts, USA.

 All file transfers are logged.

 If you have problems, please contact the BITS Group:
 help@broadinstitute.org

 This information was last updated July 2019.
2...
```

Broad Data Use Oversight System
34.120.82.229
duos-k8s.dsde-prod.broadinstitute.org
229.82.120.34.bc.googleusercontent.com
duos.broadinstitute.org
www.duos.org
duos.org
Google LLC
United States, Kansas City

SSL Certificate
Issued By:
|- Common Name:
WR3
|- Organization:
Google Trust Services
Issued To:
|- Common Name:
duos.org
Supported SSL Versions:

```
HTTP/1.1 200 OK
Date: Sat, 02 Nov 2024 00:14:59 GMT
Server: Apache
X-Content-Type-Options: nosniff
Strict-Transport-Security: max-age=86400; includeSubDomains
X-Frame-Options: SAMEORIGIN
Access-Control-Allow-Origin: *
Access-Control-Allow-Headers: authorization,content-type,accept,origin,x...
```



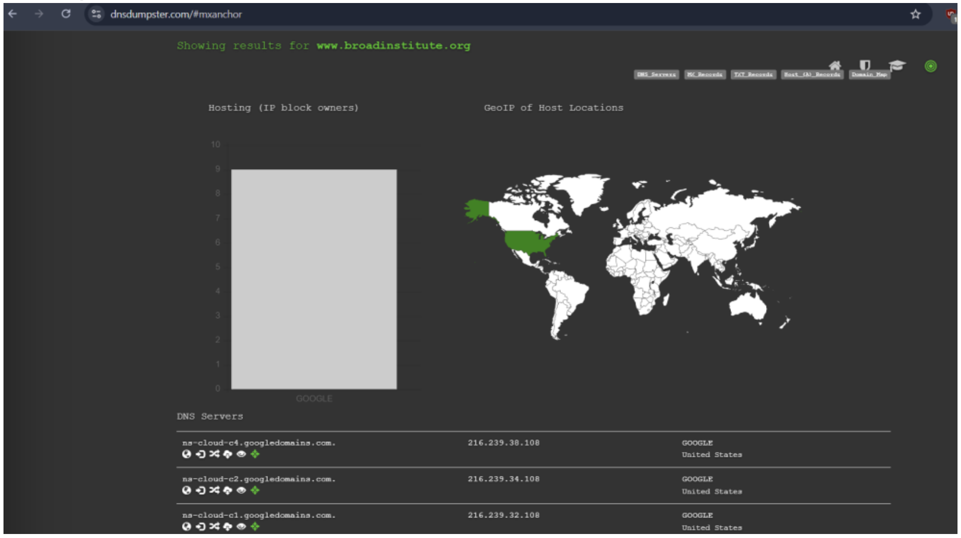
SSL Certificate

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            90:45:e5:76:60:2a:31:9a:10:26:55:41:46:98:d1:29
        Signature Algorithm: sha256WithRSAEncryption
        Issuer: C=US, O=Google Trust Services, CN=WR3
        Validity
            Not Before: Jun 23 21:40:08 2024 GMT
            Not After : Sep 21 22:43:01 2024 GMT
        Subject: CN=www.broadinstitute.mobi
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (2048 bit)
                Modulus:
                    00:a9:1d:49:0a:93:ca:40:75:a2:a6:dc:2f:36:43:
                    f4:ff:8f:15:fb:3b:b6:fe:27:44:cf:7e:16:62:a9:
                    f0:8c:c5:ec:a4:9e:00:38:f9:98:39:22:bf:a9:e0:
                    ab:84:e2:d2:fe:06:8f:d4:6f:d7:16:61:b1:2e:35:
                    de:e2:f3:29:32:ec:35:41:04:f7:fe:d8:bc:27:78:
                    40:df:f8:12:0c:de:66:63:e6:eb:53:48:1c:71:cd:
                    62:93:c6:f7:9f:58:9a:2f:a8:4e:6e:f5:c7:7f:ac:
                    be:2c:92:73:34:18:db:cd:8d:8f:9d:66:44:af:e1:
                    5d:1e:45:57:8d:99:da:8d:e6:67:66:a5:be:2a:50:
                    12:f3:da:19:55:77:eb:2d:34:1d:07:9f:bf:62:3c:
                    01:b4:64:af:1e:ff:3b:55:26:a6:2c:80:6a:86:9e:
                    8b:66:4a:73:13:63:ed:c0:e8:51:ef:e4:6a:27:19:
                    bd:7b:e0:a9:d0:06:85:e0:4e:74:ab:56:67:88:d2:
                    cd:36:8c:50:f7:76:e8:fc:74:d1:4e:c2:b7:e5:2c:
                    2c:3c:09:59:06:0f:ba:db:2f:8b:61:06:57:0d:40:
                    01:32:68:be:8a:ae:3b:e3:34:7e:cf:55:58:a9:93:
                    c7:34:5a:9a:5b:4a:5c:fa:7e:9c:3a:51:ee:bb:c5:
                    34:bb
                Exponent: 65537 (0x10001)
        X509v3 extensions:
            X509v3 Key Usage: critical
                Digital Signature, Key Encipherment
            X509v3 Extended Key Usage:
                TLS Web Server Authentication
```

    - How information can be used by an attacker: Attackers could use open ports or outdated operating systems as entry points for attacks.
    - Suggested Controls: Regularly scan and close unused or unnecessary ports, patching and updating all internet-facing systems, and employing firewalls and intrusion detection/prevention systems (IDS/IPS) can be implemented.

3. **DNS Dumpster**: a DNS enumeration tool that is used to discover DNS records, subdomains, and other infrastructure related data for a target domain. A passive tool which does not interact directly with target's network.

- Summary of Output: The DNS Dumpster showed DNS server locations, MX records, TXT records and the infrastructure of Broad like the mail servers used by the organization.



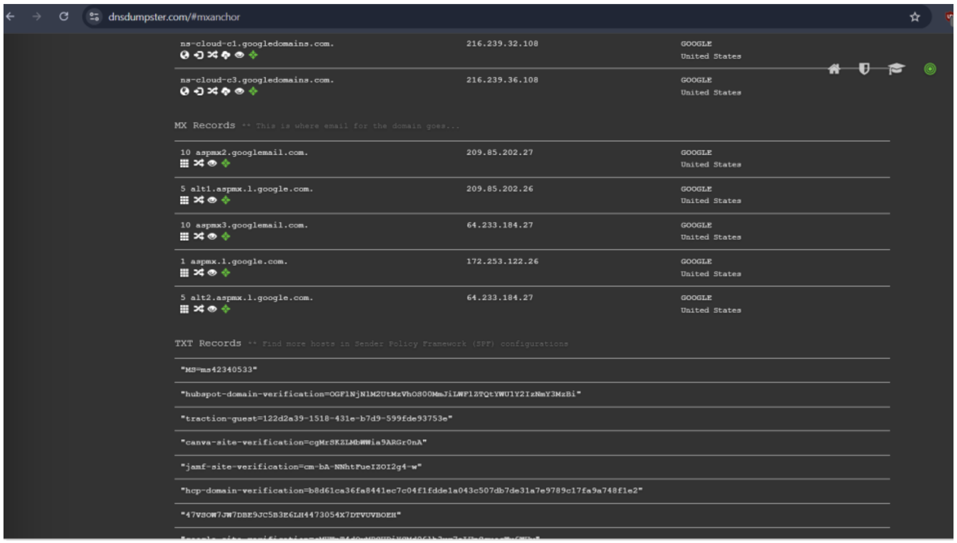
Showing results for **www.broadinstitute.org**

Hosting (IP block owners)

GeoIP of Host Locations

DNS Servers

ns-cloud-c4.googledomains.com.	216.239.38.108	GOOGLE United States
ns-cloud-c2.googledomains.com.	216.239.34.108	GOOGLE United States
ns-cloud-c1.googledomains.com.	216.239.32.108	GOOGLE United States



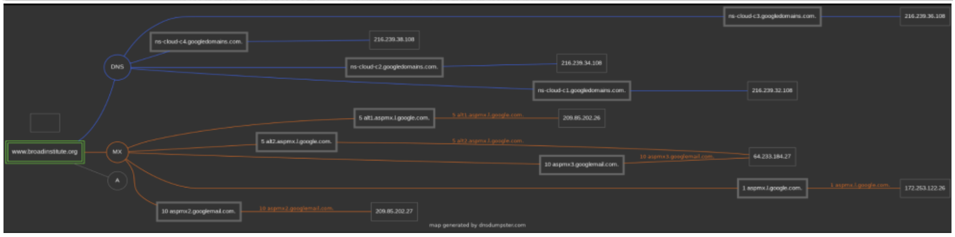
ns-cloud-c1.googledomains.com.	216.239.32.108	GOOGLE United States
ns-cloud-c3.googledomains.com.	216.239.36.108	GOOGLE United States

MX Records ** This is where email for the domain goes...

10 aspmx2.googlemail.com.	209.85.202.27	GOOGLE United States
5 alt1.aspmx.l.google.com.	209.85.202.26	GOOGLE United States
10 aspmx3.googlemail.com.	64.233.184.27	GOOGLE United States
1 aspmx.l.google.com.	172.253.122.26	GOOGLE United States
5 alt2.aspmx.l.google.com.	64.233.184.27	GOOGLE United States

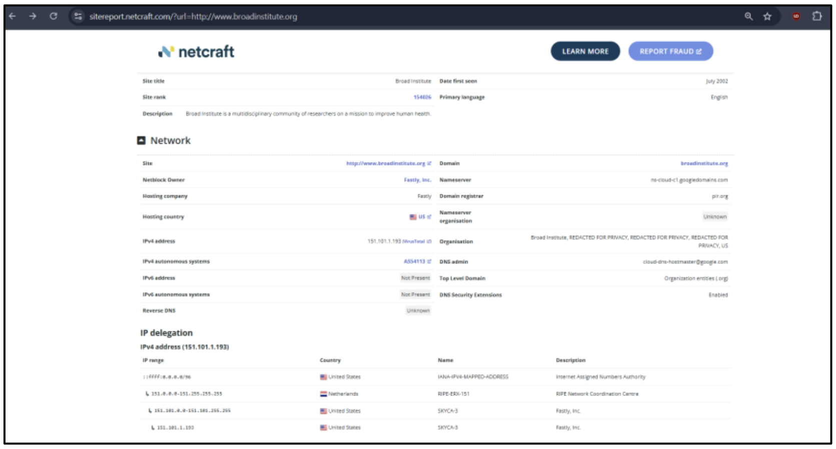
TXT Records ** Find more hosts in Sender Policy Framework (SPF) configurations

```
"MS=ms42340533"
"hubspot-domain-verification=OGFlNjNlM2UtMzVhOS00MmJiLWFlZTQtYWU1Y2IzNmY3MzBi"
"traction-guest=122d2a39-1518-431e-b7d9-599fde93753e"
"canva-site-verification=cgMrSKZLMbWWia9ARGr0nA"
"jamf-site-verification=cm-bA-NNhtFueIZOI2g4-w"
"hcp-domain-verification=b8d61ca36fa8441ec7c04f1fdde1a043c507db7de31a7e9789c17fa9a748f1e2"
"47VSOW7JW7DBE9JC5B3E6LH4473054X7DTVUVBOEH"
```


www.broadinstitute.org
DNS
MX
A
ns-cloud-c3.googledomains.com.
216.239.36.108
ns-cloud-c4.googledomains.com.
216.239.38.108
ns-cloud-c2.googledomains.com.
216.239.34.108
ns-cloud-c1.googledomains.com.
216.239.32.108
5 alt1.aspmx.l.google.com.
5 alt1.aspmx.l.google.com.
209.85.202.26
5 alt2.aspmx.l.google.com.
5 alt2.aspmx.l.google.com.
64.233.184.27
10 aspmx3.googlemail.com.
10 aspmx3.googlemail.com.
1 aspmx.l.google.com.
1 aspmx.l.google.com.
172.253.122.26
10 aspmx2.googlemail.com.
10 aspmx2.googlemail.com.
209.85.202.27
map generated by dnsdumpster.com

- How information can be used by an attacker: Information on subdomains or DNS records can enable domain spoofing, phishing, or targeted attacks on specific server of the infrastructure.
- Suggested Controls: The use of DNS security measures like DNSSEC, monitoring of DNS records for unauthorized changes and implementing Split DNS to keep internal records separate from public ones can be done.

4. **Netcraft**: Based on publicly available data, Netcraft gathers data and gives information like hosting information, phishing or malware data, technologies used by the target's website.
   - Summary of Output: The Netcraft results showed network details, Web trackers/technologies, hosting details, IP addresses, etc.



sitereport.netcraft.com/?url=http://www.broadinstitute.org

netcraft LEARN MORE REPORT FRAUD

Site title	Broad Institute	Date first seen	July 2002
Site rank	154026	Primary language	English
Description	Broad Institute is a multidisciplinary community of researchers on a mission to improve human health.		

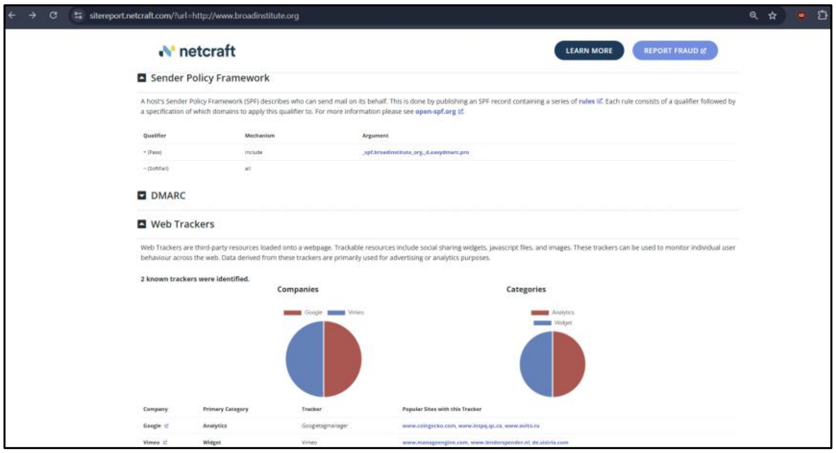
Network

Site	http://www.broadinstitute.org	Domain	broadinstitute.org
Netblock Owner	Fastly, Inc.	Nameserver	ns-cloud-c1.googledomains.com
Hosting company	Fastly	Domain registrar	pir.org
Hosting country	US	Nameserver organisation	Unknown
IPv4 address	151.101.1.193 (VirusTotal)	Organisation	Broad Institute, REDACTED FOR PRIVACY, REDACTED FOR PRIVACY, REDACTED FOR PRIVACY, US
IPv4 autonomous systems	AS54113	DNS admin	cloud-dns-hostmaster@google.com
IPv6 address	Not Present	Top Level Domain	Organization entities (.org)
IPv6 autonomous systems	Not Present	DNS Security Extensions	Enabled
Reverse DNS	Unknown		

IP delegation

IPv4 address (151.101.1.193)

IP range	Country	Name	Description
::ffff:0.0.0.0/96	United States	IANA-IPV4-MAPPED-ADDRESS	Internet Assigned Numbers Authority
↳ 151.0.0.0-151.255.255.255	Netherlands	RIPE-ERX-151	RIPE Network Coordination Centre
↳ 151.101.0.0-151.101.255.255	United States	SKYCA-3	Fastly, Inc.
↳ 151.101.1.193	United States	SKYCA-3	Fastly, Inc.



sitereport.netcraft.com/?url=http://www.broadinstitute.org

netcraft LEARN MORE REPORT FRAUD

Sender Policy Framework

A host's Sender Policy Framework (SPF) describes who can send mail on its behalf. This is done by publishing an SPF record containing a series of rules. Each rule consists of a qualifier followed by a specification of which domains to apply this qualifier to. For more information please see open-spf.org.

Qualifier	Mechanism	Argument
+ (Pass)	include	_spf.broadinstitute_org._d.easydmarc.pro
~ (SoftFail)	all	

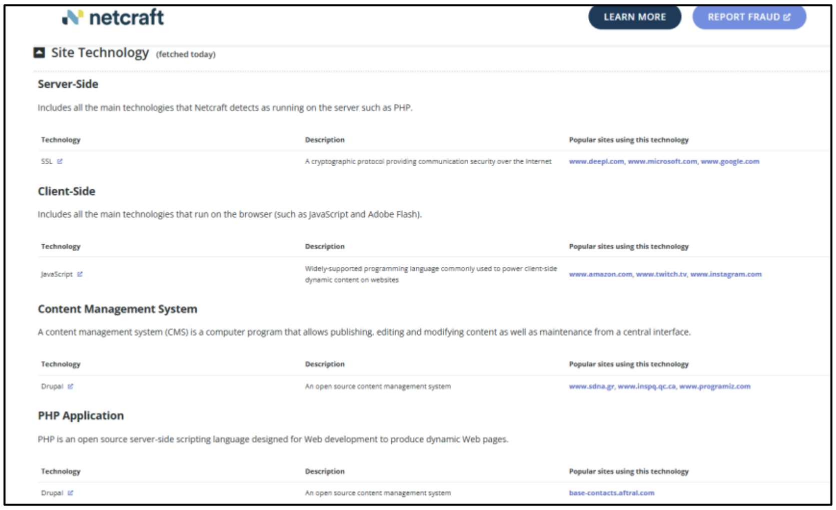
DMARC

Web Trackers

Web Trackers are third-party resources loaded onto a webpage. Trackable resources include social sharing widgets, javascript files, and images. These trackers can be used to monitor individual user behaviour across the web. Data derived from these trackers are primarily used for advertising or analytics purposes.

2 known trackers were identified.

Company	Primary Category	Tracker	Popular Sites with this Tracker
Google	Analytics	Googletagmanager	www.coingecko.com, www.inspq.qc.ca, www.avito.ru
Vimeo	Widget	Vimeo	www.manageengine.com, www.lenderspender.nl, de.sistrix.com



netcraft LEARN MORE REPORT FRAUD

Site Technology (fetched today)

Server-Side

Includes all the main technologies that Netcraft detects as running on the server such as PHP.

Technology	Description	Popular sites using this technology
SSL	A cryptographic protocol providing communication security over the Internet	www.deepl.com, www.microsoft.com, www.google.com

Client-Side

Includes all the main technologies that run on the browser (such as JavaScript and Adobe Flash).

Technology	Description	Popular sites using this technology
JavaScript	Widely-supported programming language commonly used to power client-side dynamic content on websites	www.amazon.com, www.twitch.tv, www.instagram.com

Content Management System

A content management system (CMS) is a computer program that allows publishing, editing and modifying content as well as maintenance from a central interface.

Technology	Description	Popular sites using this technology
Drupal	An open source content management system	www.sdna.gr, www.inspq.qc.ca, www.programiz.com

PHP Application

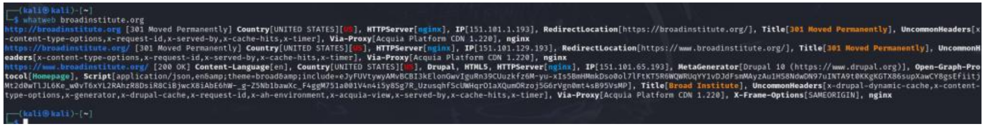
PHP is an open source server-side scripting language designed for Web development to produce dynamic Web pages.

Technology	Description	Popular sites using this technology
Drupal	An open source content management system	base-contacts.aftral.com

   - How information can be used by an attacker: Attackers could identify known vulnerabilities in web technologies and attempt to exploit them.

- Suggested Controls: Using secure web server configurations and disable unnecessary services, patching CMS, frameworks, and other web technologies and continuous monitoring of website activity for any indicators of compromise.

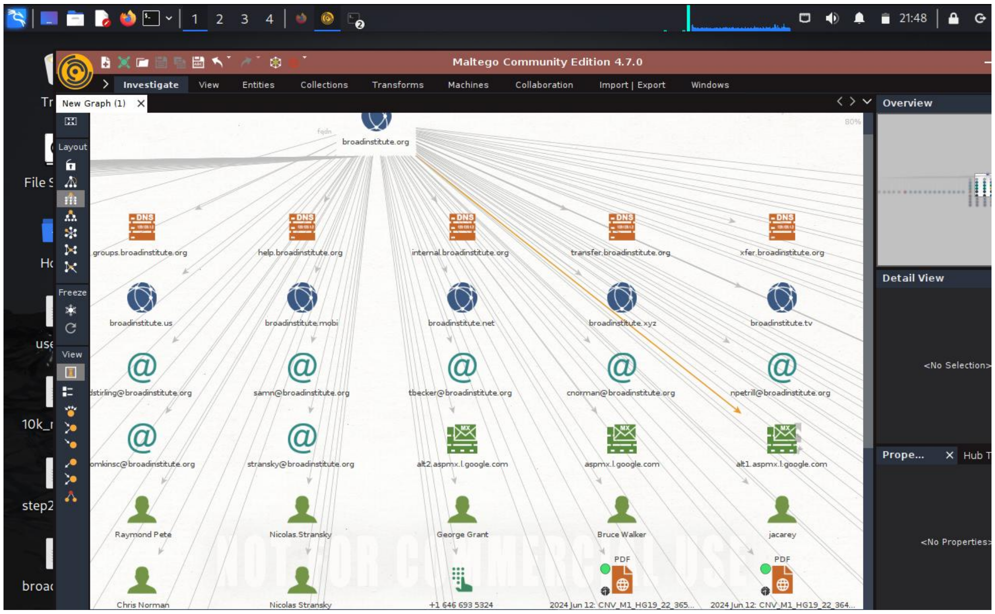
5. **Whatweb**: A fingerprinting web application used to identify technologies used by a website, including the CMS systems, server type, and plugins/frameworks used.
   - Summary of Output: The results of the Whatweb tool for Broad identified the website was built using HTML5, the HTTP server used(nginx), IP address, location (US).

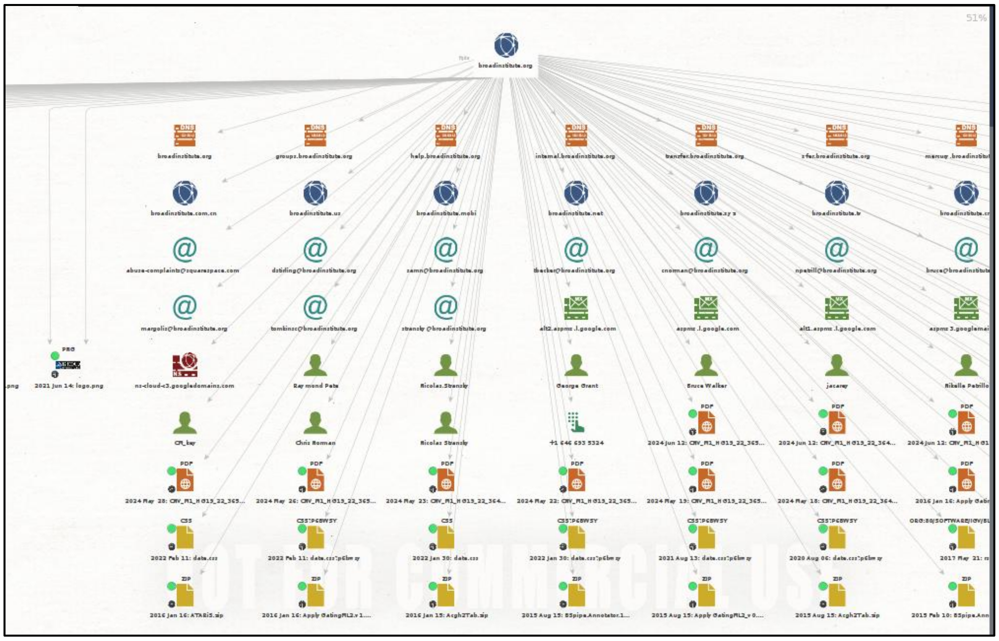


```
┌──(kali㉿kali)-[~]
└─$ whatweb broadinstitute.org
http://broadinstitute.org [301 Moved Permanently] Country[UNITED STATES][US], HTTPServer[nginx], IP[151.101.1.193], RedirectLocation[https://broadinstitute.org/], Title[301 Moved Permanently], UncommonHeaders[x
-content-type-options,x-request-id,x-served-by,x-cache-hits,x-timer], Via-Proxy[Acquia Platform CDN 1.220], nginx
https://broadinstitute.org/ [301 Moved Permanently] Country[UNITED STATES][US], HTTPServer[nginx], IP[151.101.129.193], RedirectLocation[https://www.broadinstitute.org/], Title[301 Moved Permanently], UncommonH
eaders[x-content-type-options,x-request-id,x-served-by,x-cache-hits,x-timer], Via-Proxy[Acquia Platform CDN 1.220], nginx
https://www.broadinstitute.org/ [200 OK] Content-Language[en], Country[UNITED STATES][US], Drupal, HTML5, HTTPServer[nginx], IP[151.101.65.193], MetaGenerator[Drupal 10 (https://www.drupal.org)], Open-Graph-Pro
tocol[Homepage], Script[application/json,en&amp;theme=broad&amp;include=eJyFUVtywyAMvBCBI3kElonGwvIguRn39CUuzkfz6M-yu-xIs5BmHMmkDso0ol7lFtKT5R6WQWRUqYY1vDJdFsmMAyzAu1H58NdwDN97uINTA9t0KKgKGTX86supXawCY8gsEfiitj
Mt2d0wTlJL6Ke_w0vT6xYL2RAhzR8DsiR8CiBjwcXBiAbE6hW-_g-Z5Nb1bawXc_F4ggM751a001V4n4i5y8Sg7R_UzusqhfScUWHqrO1aXQumORzoj5G6rVgn0mt4sB95VsMP], Title[Broad Institute], UncommonHeaders[x-drupal-dynamic-cache,x-content-
type-options,x-generator,x-drupal-cache,x-request-id,x-ah-environment,x-acquia-view,x-served-by,x-cache-hits,x-timer], Via-Proxy[Acquia Platform CDN 1.220], X-Frame-Options[SAMEORIGIN], nginx

┌──(kali㉿kali)-[~]
└─$
```

   - How information can be used by an attacker: A specific vulnerability in software and plugin versions can be easily exploited.
   - Suggested Controls: The web servers should be configured in a way that they mask software version details. These plugins and software should be regularly updated.

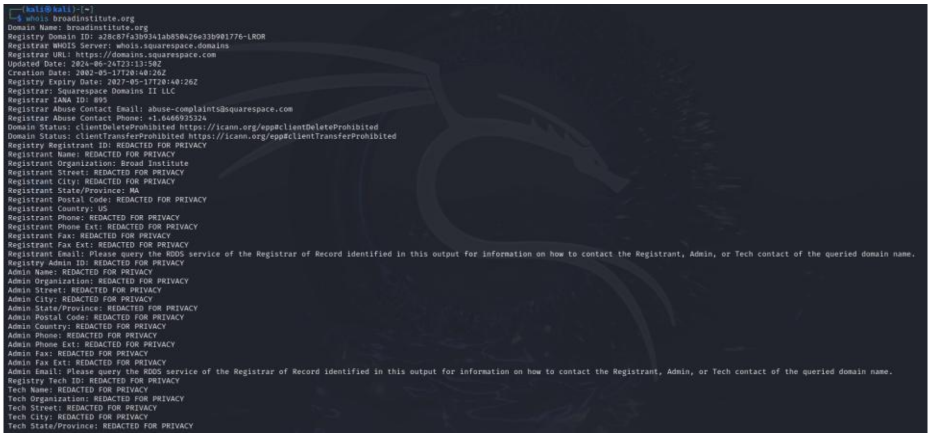
6. **Maltego**: A data linkage tool that gives out target's social media profiles, emails, and domains into a graph for analysis.
   - Summary of Output: The output from Maltego showed a detailed map of domains, emails, and names from the organization. It was even able to identify a phone for someone named "George Grant".




51%
broadinstitute.org
broadinstitute.org
groups.broadinstitute.org
help.broadinstitute.org
internal.broadinstitute.org
transfer.broadinstitute.org
xfer.broadinstitute.org
mercury.broadinstitut
broadinstitute.com.cn
broadinstitute.us
broadinstitute.mobi
broadinstitute.net
broadinstitute.sy s
broadinstitute.tv
broadinstitute.cr
abuse-complaints@squarespace.com
dstirling@broadinstitute.org
samn@broadinstitute.org
tbecker@broadinstitute.org
cnorman@broadinstitute.org
npetrill@broadinstitute.org
bruce@broadinstitute
margolis@broadinstitute.org
tomkinsc@broadinstitute.org
stransky@broadinstitute.org
alt2.aspmx.l.google.com
aspmx.l.google.com
alt1.aspmx.l.google.com
aspmx3.googlemai
PNG
.png
2021 Jun 14: logo.png
ns-cloud-c3.googledomains.com
Raymond Pete
Nicolas.Stransky
George Grant
Bruce Walker
jacarey
Nikelle Petrillo
CR_key
Chris Norman
Nicolas Stransky
+1 646 693 5324
PDF
2024 Jun 12: CNV_M1_H G13_22_365...
2024 Jun 12: CNV_M1_H G13_22_364...
2024 Jun 12: CNV_M1_H G1
2024 May 28: CNV_M1_H G13_22_365...
2024 May 26: CNV_M1_H G13_22_365...
2024 May 23: CNV_M1_H G13_22_364...
2024 May 22: CNV_M1_H G13_22_365...
2024 May 19: CNV_M1_H G13_22_365...
2024 May 18: CNV_M1_H G13_22_364...
2016 Jan 16: Apply Gatir
CSS
CSS?P6BW5Y
ORG:80/SOFTWARE/IGV/BL
2022 Feb 11: data.css
2022 Feb 11: data.css?p6bw sy
2022 Jan 30: data.css
2022 Jan 30: data.css?p6bw sy
2021 Aug 13: data.css?p6bw sy
2020 Aug 06: data.css?p6bw sy
2017 May 21:
ZIP
2016 Jan 16: ATARiS.zip
2016 Jan 16: Apply GatingML2.v 1....
2016 Jan 15: Acgh2Tab.zip
2015 Aug 15: BSpipe.Annotator.1...
2015 Aug 15: Apply GatingML2_v 0....
2015 Aug 15: Acgh2Tab.zip
2015 Feb 10: BSpipe.Ann
NOT FOR COMMERCIAL USE

- <u>How information can be used by an attacker:</u> Attackers can perform social engineering or impersonation attacks using social media or email accounts.
- <u>Suggested Controls:</u> Employees should be educated on privacy settings on their social media accounts, restricting themselves from sharing their contact details publicly.

7. **Whois**: A Whois scan for a target domain can provide important information about the owner of the domain, registration details, contact information, Admin name, location, etc.
   - <u>Summary of Output:</u> The Whois scan revealed that the organization is registered in MA(US), IANA ID and the Registrar URL, server.

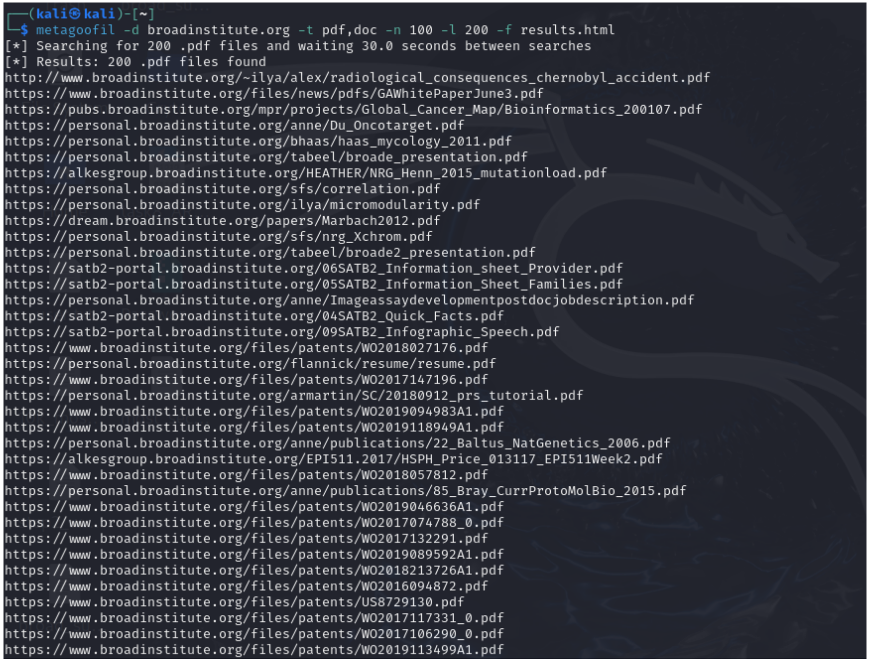


```
┌──(kali㉿kali)-[~]
└─$ whois broadinstitute.org
Domain Name: broadinstitute.org
Registry Domain ID: a28c87fa3b9341ab850426e33b901776-LROR
Registrar WHOIS Server: whois.squarespace.domains
Registrar URL: https://domains.squarespace.com
Updated Date: 2024-06-24T23:13:50Z
Creation Date: 2002-05-17T20:40:26Z
Registry Expiry Date: 2027-05-17T20:40:26Z
Registrar: Squarespace Domains II LLC
Registrar IANA ID: 895
Registrar Abuse Contact Email: abuse-complaints@squarespace.com
Registrar Abuse Contact Phone: +1.6466935324
Domain Status: clientDeleteProhibited https://icann.org/epp#clientDeleteProhibited
Domain Status: clientTransferProhibited https://icann.org/epp#clientTransferProhibited
Registry Registrant ID: REDACTED FOR PRIVACY
Registrant Name: REDACTED FOR PRIVACY
Registrant Organization: Broad Institute
Registrant Street: REDACTED FOR PRIVACY
Registrant City: REDACTED FOR PRIVACY
Registrant State/Province: MA
Registrant Postal Code: REDACTED FOR PRIVACY
Registrant Country: US
Registrant Phone: REDACTED FOR PRIVACY
Registrant Phone Ext: REDACTED FOR PRIVACY
Registrant Fax: REDACTED FOR PRIVACY
Registrant Fax Ext: REDACTED FOR PRIVACY
Registrant Email: Please query the RDDS service of the Registrar of Record identified in this output for information on how to contact the Registrant, Admin, or Tech contact of the queried domain name.
Registry Admin ID: REDACTED FOR PRIVACY
Admin Name: REDACTED FOR PRIVACY
Admin Organization: REDACTED FOR PRIVACY
Admin Street: REDACTED FOR PRIVACY
Admin City: REDACTED FOR PRIVACY
Admin State/Province: REDACTED FOR PRIVACY
Admin Postal Code: REDACTED FOR PRIVACY
Admin Country: REDACTED FOR PRIVACY
Admin Phone: REDACTED FOR PRIVACY
Admin Phone Ext: REDACTED FOR PRIVACY
Admin Fax: REDACTED FOR PRIVACY
Admin Fax Ext: REDACTED FOR PRIVACY
Admin Email: Please query the RDDS service of the Registrar of Record identified in this output for information on how to contact the Registrant, Admin, or Tech contact of the queried domain name.
Registry Tech ID: REDACTED FOR PRIVACY
Tech Name: REDACTED FOR PRIVACY
Tech Organization: REDACTED FOR PRIVACY
Tech Street: REDACTED FOR PRIVACY
Tech City: REDACTED FOR PRIVACY
Tech State/Province: REDACTED FOR PRIVACY
```

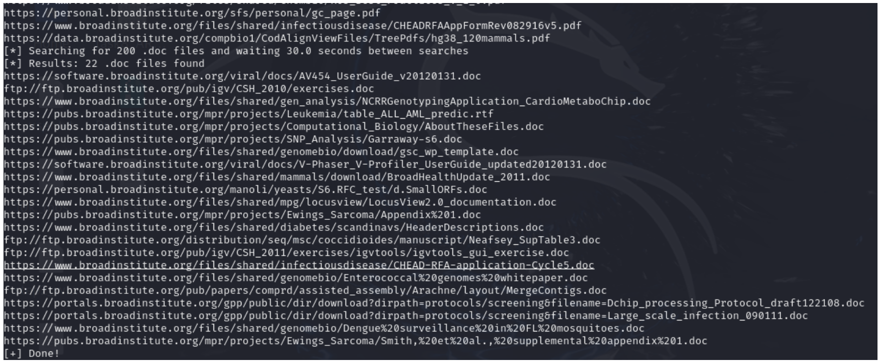
   - <u>How information can be used by an attacker:</u> The data obtained from Whois provides detailed information about the domain, registration details, contact information, owner details, etc. (not found for Broad Institute)

- Suggested Controls: Use domain privacy protection to mask Whois information, which was done by Broad Institute. (Redacted For Privacy as shown in the screenshot above)

8. **Metagoofil**: a metadata gathering tool designed to extract data from public documents that are associated with a particular domain. These documents may include PDFs, Word files, and other commonly used formats.
    - Summary of Output: The Metagoofil output showed documents related to mutations, genetics, and other healthcare related files.



```
┌──(kali㉿kali)-[~]
└─$ metagoofil -d broadinstitute.org -t pdf,doc -n 100 -l 200 -f results.html
[*] Searching for 200 .pdf files and waiting 30.0 seconds between searches
[*] Results: 200 .pdf files found
http://www.broadinstitute.org/~ilya/alex/radiological_consequences_chernobyl_accident.pdf
https://www.broadinstitute.org/files/news/pdfs/GAWhitePaperJune3.pdf
https://pubs.broadinstitute.org/mpr/projects/Global_Cancer_Map/Bioinformatics_200107.pdf
https://personal.broadinstitute.org/anne/Du_Oncotarget.pdf
https://personal.broadinstitute.org/bhaas/haas_mycology_2011.pdf
https://personal.broadinstitute.org/tabeel/broade_presentation.pdf
https://alkesgroup.broadinstitute.org/HEATHER/NRG_Henn_2015_mutationload.pdf
https://personal.broadinstitute.org/sfs/correlation.pdf
https://personal.broadinstitute.org/ilya/micromodularity.pdf
https://dream.broadinstitute.org/papers/Marbach2012.pdf
https://personal.broadinstitute.org/sfs/nrg_Xchrom.pdf
https://personal.broadinstitute.org/tabeel/broade2_presentation.pdf
https://satb2-portal.broadinstitute.org/06SATB2_Information_sheet_Provider.pdf
https://satb2-portal.broadinstitute.org/05SATB2_Information_Sheet_Families.pdf
https://personal.broadinstitute.org/anne/Imageassaydevelopmentpostdocjobdescription.pdf
https://satb2-portal.broadinstitute.org/04SATB2_Quick_Facts.pdf
https://satb2-portal.broadinstitute.org/09SATB2_Infographic_Speech.pdf
https://www.broadinstitute.org/files/patents/WO2018027176.pdf
https://personal.broadinstitute.org/flannick/resume/resume.pdf
https://www.broadinstitute.org/files/patents/WO2017147196.pdf
https://personal.broadinstitute.org/armartin/SC/20180912_prs_tutorial.pdf
https://www.broadinstitute.org/files/patents/WO2019094983A1.pdf
https://www.broadinstitute.org/files/patents/WO2019118949A1.pdf
https://personal.broadinstitute.org/anne/publications/22_Baltus_NatGenetics_2006.pdf
https://alkesgroup.broadinstitute.org/EPI511.2017/HSPH_Price_013117_EPI511Week2.pdf
https://www.broadinstitute.org/files/patents/WO2018057812.pdf
https://personal.broadinstitute.org/anne/publications/85_Bray_CurrProtoMolBio_2015.pdf
https://www.broadinstitute.org/files/patents/WO2019046636A1.pdf
https://www.broadinstitute.org/files/patents/WO2017074788_0.pdf
https://www.broadinstitute.org/files/patents/WO2017132291.pdf
https://www.broadinstitute.org/files/patents/WO2019089592A1.pdf
https://www.broadinstitute.org/files/patents/WO2018213726A1.pdf
https://www.broadinstitute.org/files/patents/WO2016094872.pdf
https://www.broadinstitute.org/files/patents/US8729130.pdf
https://www.broadinstitute.org/files/patents/WO2017117331_0.pdf
https://www.broadinstitute.org/files/patents/WO2017106290_0.pdf
https://www.broadinstitute.org/files/patents/WO2019113499A1.pdf
```

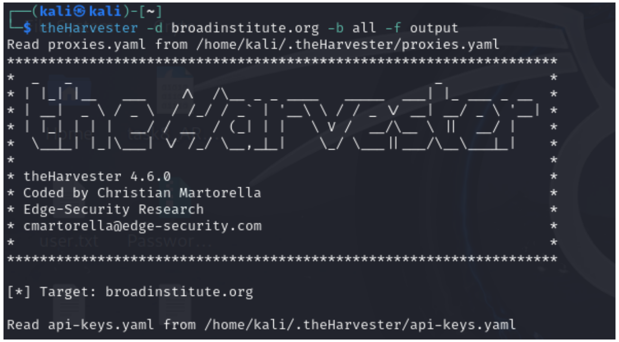


```
https://personal.broadinstitute.org/sfs/personal/gc_page.pdf
https://www.broadinstitute.org/files/shared/infectiousdisease/CHEADRFAAppFormRev082916v5.pdf
https://data.broadinstitute.org/compbio1/CodAlignViewFiles/TreePdfs/hg38_120mammals.pdf
[*] Searching for 200 .doc files and waiting 30.0 seconds between searches
[*] Results: 22 .doc files found
https://software.broadinstitute.org/viral/docs/AV454_UserGuide_v20120131.doc
ftp://ftp.broadinstitute.org/pub/igv/CSH_2010/exercises.doc
https://www.broadinstitute.org/files/shared/gen_analysis/NCRRGenotypingApplication_CardioMetaboChip.doc
https://pubs.broadinstitute.org/mpr/projects/Leukemia/table_ALL_AML_predic.rtf
https://pubs.broadinstitute.org/mpr/projects/Computational_Biology/AboutTheseFiles.doc
https://pubs.broadinstitute.org/mpr/projects/SNP_Analysis/Garraway-s6.doc
https://www.broadinstitute.org/files/shared/genomebio/download/gsc_wp_template.doc
https://software.broadinstitute.org/viral/docs/V-Phaser_V-Profiler_UserGuide_updated20120131.doc
https://www.broadinstitute.org/files/shared/mammals/download/BroadHealthUpdate_2011.doc
https://personal.broadinstitute.org/manoli/yeasts/S6.RFC_test/d.SmallORFs.doc
https://www.broadinstitute.org/files/shared/mpg/locusview/LocusView2.0_documentation.doc
https://pubs.broadinstitute.org/mpr/projects/Ewings_Sarcoma/Appendix%201.doc
https://www.broadinstitute.org/files/shared/diabetes/scandinavs/HeaderDescriptions.doc
ftp://ftp.broadinstitute.org/distribution/seq/msc/coccidioides/manuscript/Neafsey_SupTable3.doc
ftp://ftp.broadinstitute.org/pub/igv/CSH_2011/exercises/igvtools/igvtools_gui_exercise.doc
https://www.broadinstitute.org/files/shared/infectiousdisease/CHEAD-RFA-application-Cycle5.doc
https://www.broadinstitute.org/files/shared/genomebio/Enterococcal%20genomes%20whitepaper.doc
ftp://ftp.broadinstitute.org/pub/papers/comprd/assisted_assembly/Arachne/layout/MergeContigs.doc
https://portals.broadinstitute.org/gpp/public/dir/download?dirpath=protocols/screening&filename=Dchip_processing_Protocol_draft122108.doc
https://portals.broadinstitute.org/gpp/public/dir/download?dirpath=protocols/screening&filename=Large_scale_infection_090111.doc
https://www.broadinstitute.org/files/shared/genomebio/Dengue%20surveillance%20in%20FL%20mosquitoes.doc
https://pubs.broadinstitute.org/mpr/projects/Ewings_Sarcoma/Smith,%20et%20al.,%20supplemental%20appendix%201.doc
[+] Done!
```

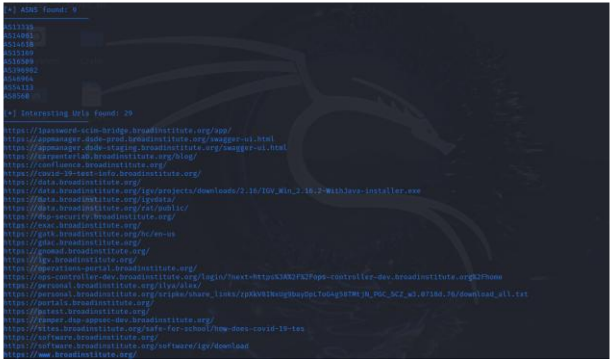
- How information can be used by an attacker: Metadata from internal documents can reveal sensitive details, including ongoing research which can be exploited.
- Suggested Controls: The use of metadata cleaning tools before publishing documents on the internet, educating employees on metadata privacy practices.

9. **theHarvester**: is an open-source reconnaissance tool that is used to collect information from public sources and search engines. The information usually consists of email addresses, IP addresses, ASNS, hostnames, etc.

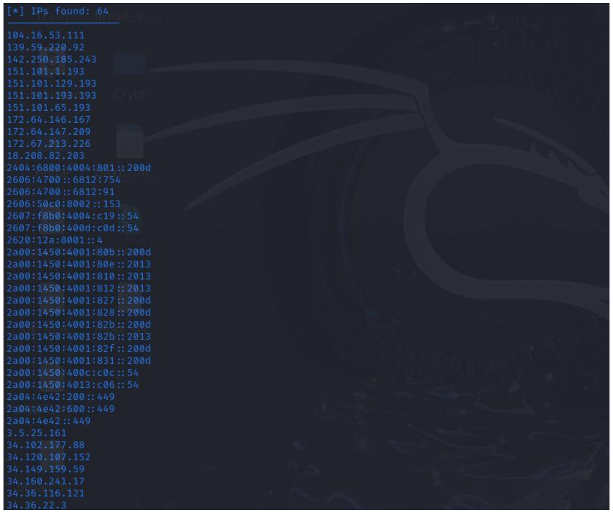
- Summary of Output: The results of the theHarvester on Broad provided information about ASNS, Interesting URLs, and IP addresses.



```
┌──(kali㉿kali)-[~]
└─$ theHarvester -d broadinstitute.org -b all -f output
Read proxies.yaml from /home/kali/.theHarvester/proxies.yaml
*******************************************************************
*  _   _                                            _             *
* | |_| |__   ___    /\  /\__ _ _ ____   _____  ___| |_ ___ _ __  *
* | __|  _ \ / _ \  / /_/ / _` | '__\ \ / / _ \/ __| __/ _ \ '__| *
* | |_| | | |  __/ / __  / (_| | |   \ V /  __/\__ \ ||  __/ |    *
*  \__|_| |_|\___| \/ /_/ \__,_|_|    \_/ \___||___/\__\___|_|    *
*                                                                 *
* theHarvester 4.6.0                                              *
* Coded by Christian Martorella                                   *
* Edge-Security Research                                          *
* cmartorella@edge-security.com                                   *
*                                                                 *
*******************************************************************

[*] Target: broadinstitute.org

Read api-keys.yaml from /home/kali/.theHarvester/api-keys.yaml
```



```
[*] ASNS found: 9
--------------------
AS13335
AS14061
AS14618
AS15169
AS16509
AS396982
AS46964
AS54113
AS8560

[*] Interesting Urls found: 29
--------------------
https://1password-scim-bridge.broadinstitute.org/app/
https://appmanager.dsde-prod.broadinstitute.org/swagger-ui.html
https://appmanager.dsde-staging.broadinstitute.org/swagger-ui.html
https://carpenterlab.broadinstitute.org/blog/
https://confluence.broadinstitute.org/
https://covid-19-test-info.broadinstitute.org/
https://data.broadinstitute.org/
https://data.broadinstitute.org/igv/projects/downloads/2.16/IGV_Win_2.16.2-WithJava-installer.exe
https://data.broadinstitute.org/igvdata/
https://data.broadinstitute.org/rat/public/
https://dsp-security.broadinstitute.org/
https://exac.broadinstitute.org/
https://gatk.broadinstitute.org/hc/en-us
https://gdac.broadinstitute.org/
https://gnomad.broadinstitute.org/
https://igv.broadinstitute.org/
https://operations-portal.broadinstitute.org/
https://ops-controller-dev.broadinstitute.org/login/?next=https%3A%2F%2Fops-controller-dev.broadinstitute.org%2Fhome
https://personal.broadinstitute.org/ilya/alex/
https://personal.broadinstitute.org/sripke/share_links/zpXkV8INxUg9bayOpLToG4g58TMtjN_PGC_SCZ_w3.0718d.76/download_all.txt
https://portals.broadinstitute.org/
https://pstest.broadinstitute.org/
https://ramper.dsp-appsec-dev.broadinstitute.org/
https://sites.broadinstitute.org/safe-for-school/how-does-covid-19-tes
https://software.broadinstitute.org/
https://software.broadinstitute.org/software/igv/download
https://www.broadinstitute.org/
```



```
[*] IPs found: 64
-------------------
104.16.53.111
139.59.220.92
142.250.185.243
151.101.1.193
151.101.129.193
151.101.193.193
151.101.65.193
172.64.146.167
172.64.147.209
172.67.213.226
18.208.82.203
2404:6800:4004:801::200d
2606:4700::6812:754
2606:4700::6812:91
2606:50c0:8002::153
2607:f8b0:4004:c19::54
2607:f8b0:400d:c0d::54
2620:12a:8001::4
2a00:1450:4001:80b::200d
2a00:1450:4001:80e::2013
2a00:1450:4001:810::2013
2a00:1450:4001:812::2013
2a00:1450:4001:827::200d
2a00:1450:4001:828::200d
2a00:1450:4001:82b::200d
2a00:1450:4001:82b::2013
2a00:1450:4001:82f::200d
2a00:1450:4001:831::200d
2a00:1450:400c:c0c::54
2a00:1450:4013:c06::54
2a04:4e42:200::449
2a04:4e42:600::449
2a04:4e42::449
3.5.25.161
34.102.177.88
34.120.107.152
34.149.159.59
34.160.241.17
34.36.116.121
34.36.22.3
```

- How information can be used by an attacker: The public information can be used for phishing, network mapping, and other attacks. Attackers can gain unauthorized access to organization internal systems and tamper sensitive data.
- Suggested Controls: The use of strong access control and network segmentation should be done, conducting regular security assessments of exposed assets, and monitor network traffic for unauthorized access attempts.