OSINT:
Hidden In Plain Sight

SecPro
SECURITY PROFESSIONALS

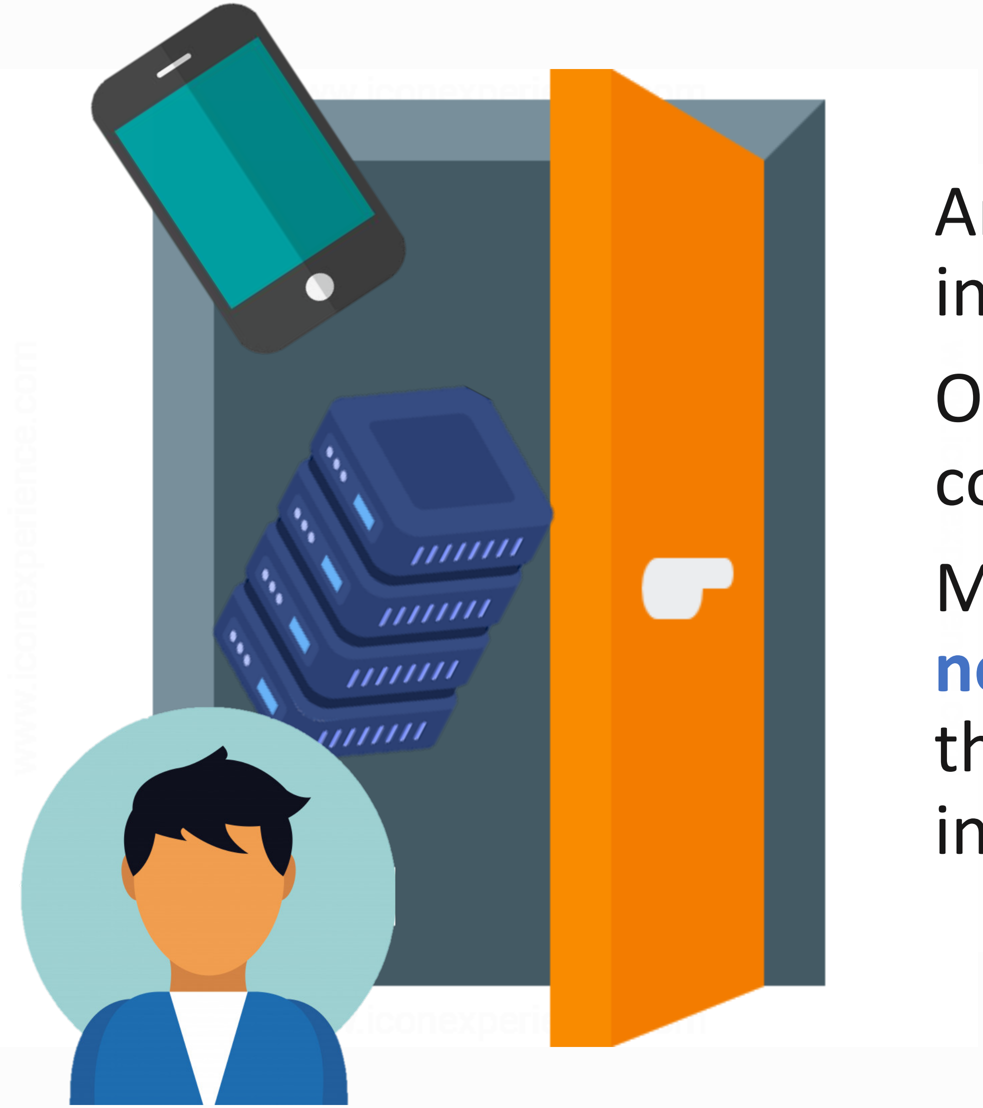
Whoami: David F. Pereira

25+ Years Experience in:

- Ethical Hacking
- Pentesting
- DFIR
- Incident Response

- International Instructor for Different Law Enforcement Agencies and Companies around the world
- Colombian Army Veteran

What is OSINT Used for?



An OSINT collection is conducted to **gather** information on a specific **target or topic.**

OSINT investigations can be conducted on companies or individuals.

Many companies, employees or individuals **are not aware** of the sensitivity and importance of the **information they publish** and how this information can be used by various actors.

What can we Obtain?

Corporative Information:

- Infrastructure (IP Addresses, Existing Networks, Open Ports, Domains, Subdomains)
- WEB Presence
- Business Partners
- Service Providers
- Operating Systems and their respective versions
- Software Used and their respective Versions
- Antivirus / Antimalware
- Geographical Locations
- Employees
- Former Employees
- Email Addresses
- Telephone Numbers
- Etc......

What can we Obtain?

SecPro
SECURITY PROFESSIONALS

Personal Information:

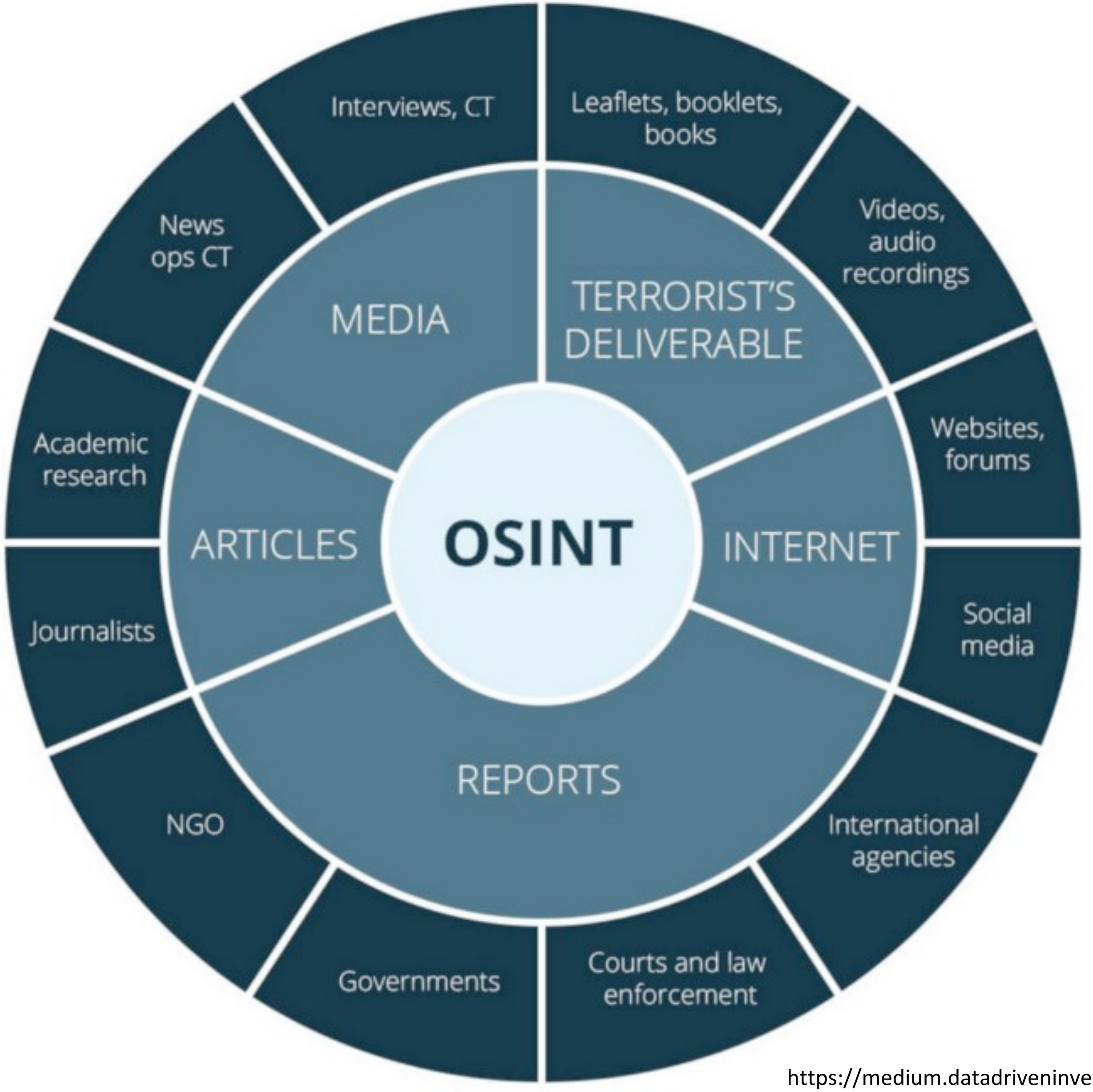
- Social Networking
- Email addresses
- Full name(s)
- Company or entity where you work (Position, Area, Influence, etc.)
- Related companies or entities
- Telephone numbers
- Addresses, relatives, friends
- Passwords
- Sports you practice
- Properties
- Vehicles
- Likes, habits, hobbies
- Etc.....

OSINT Planning

How to get Started?

Where to start the process is one of the first hurdles we will face, but using a simple methodology such as Gap Analysis, we can easily orient ourselves;

OSINT Sources



https://medium.datadriveninvestor.com/open-source-intelligence-osint-101-d96f47ff2ff1

What areas exist on the Internet?

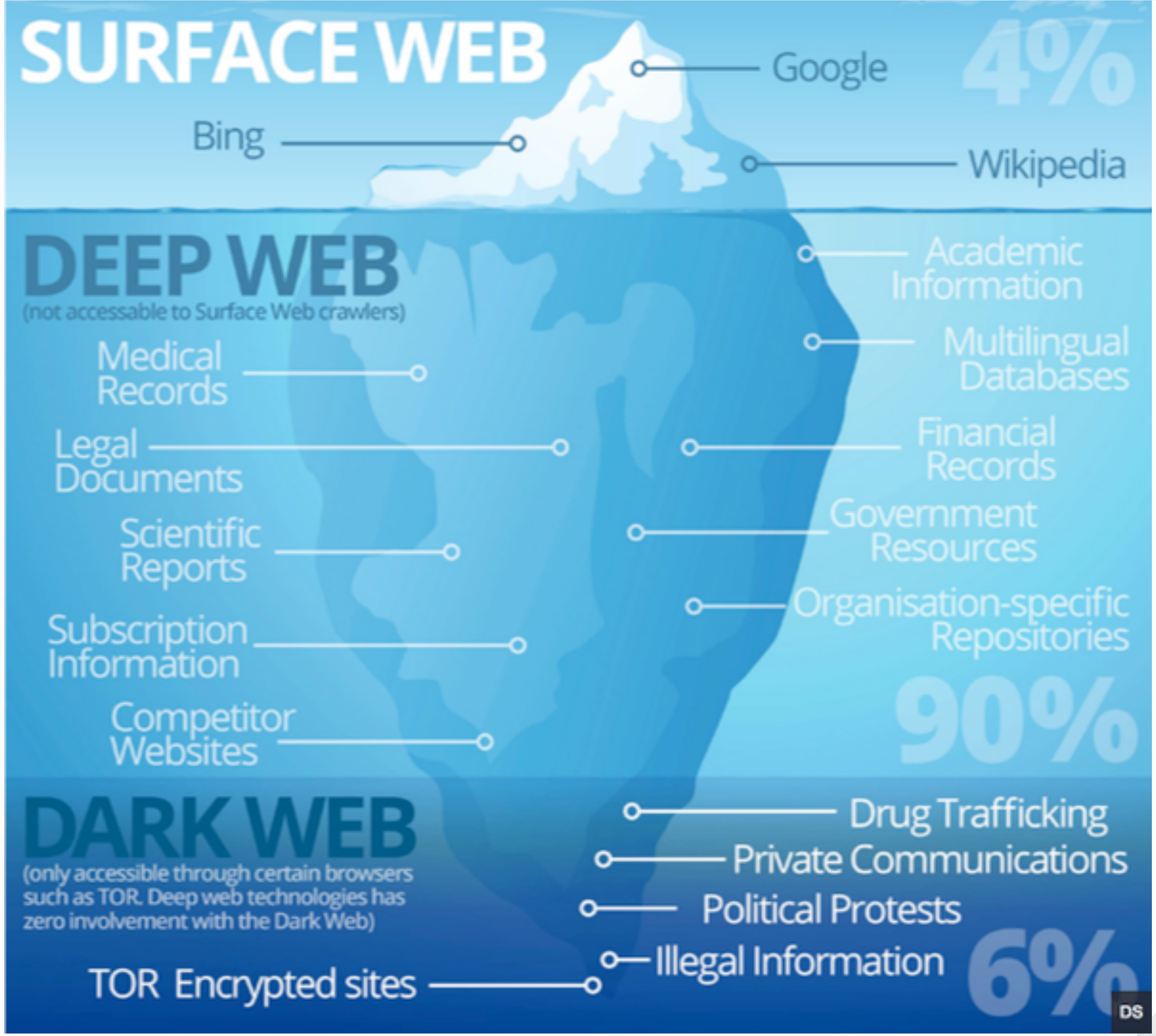


Image Credit : elixirofknowledge
https://www.elixirofknowledge.com/2016/07/dark-web.html

Surface WEB and Search Engines

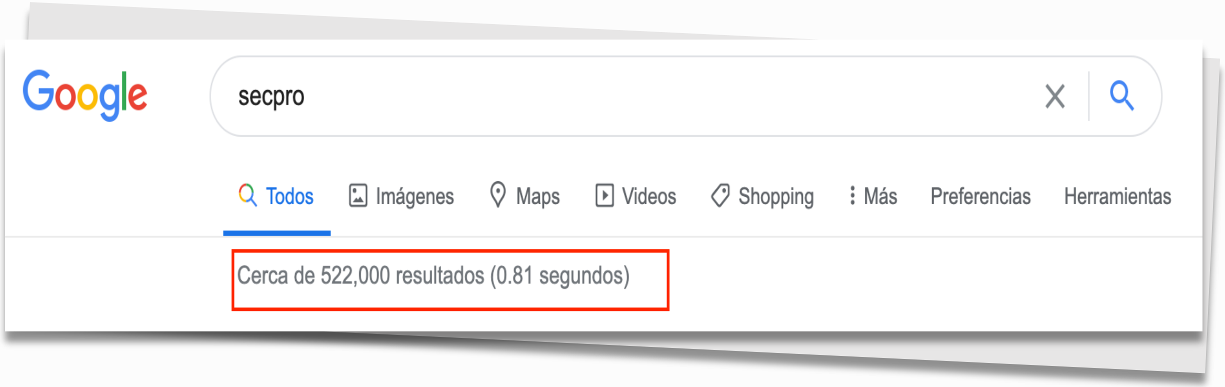
In our daily life we interact with the Surface WEB, which is also called Clear WEB; in order to find information, we can use traditional search engines such as:

- https://www.google.com
- https://www.bing.com
- https://www.yahoo.com

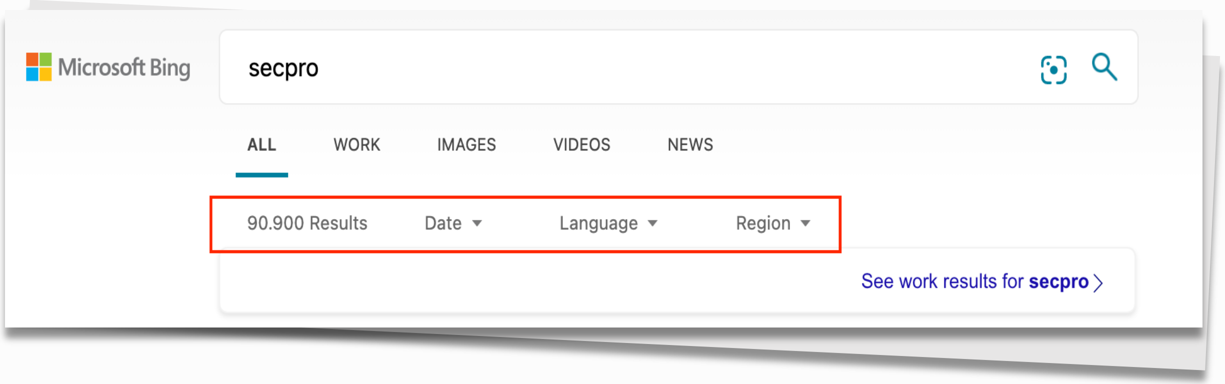
Search Engines

SecPro
SECURITY PROFESSIONALS

Traditional **search engines** have **different indexing algorithms** oriented to the Surface WEB, which offer different results.

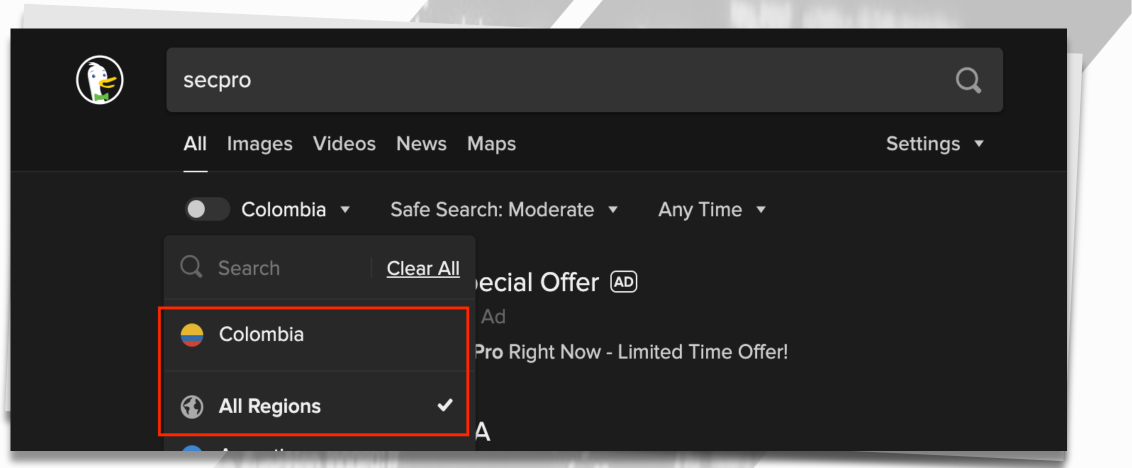


Cerca de 522,000 resultados (0.81 segundos)



90.900 Results   Date ▾   Language ▾   Region ▾

See work results for **secpro** ›

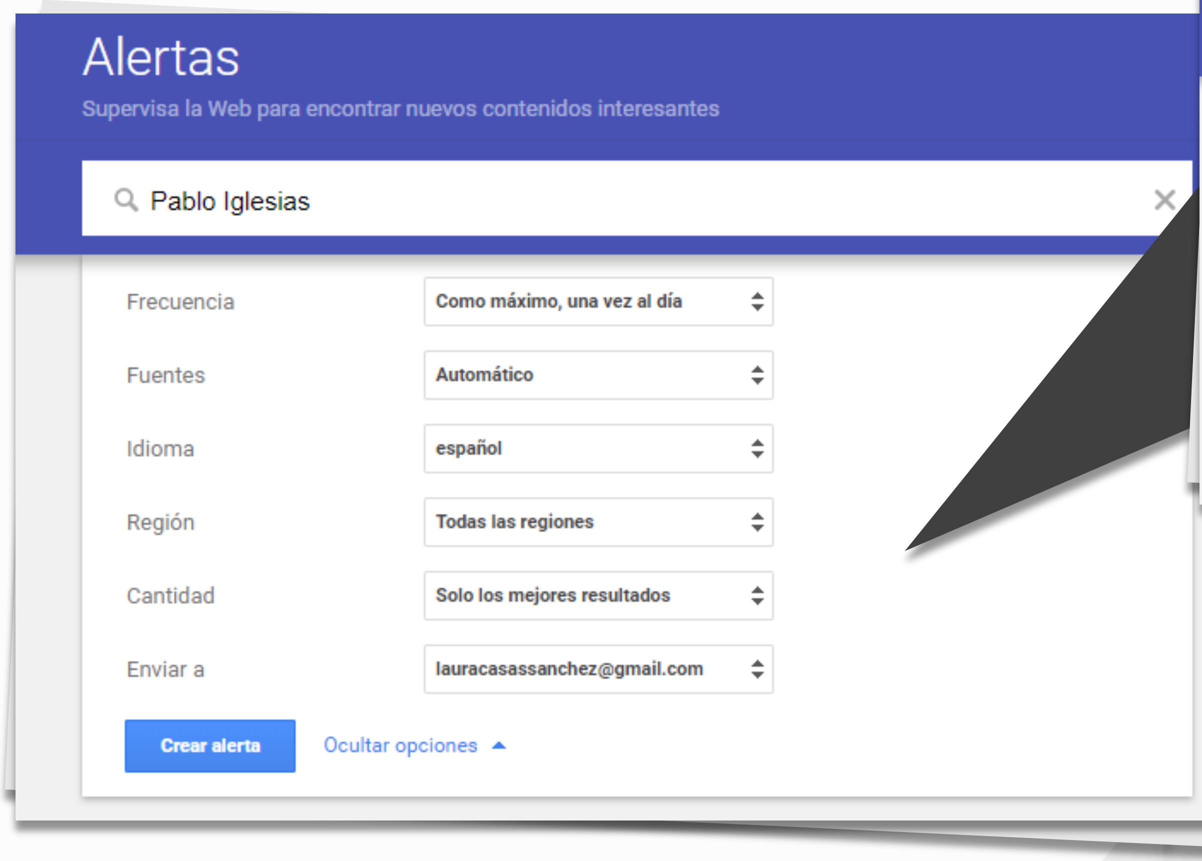


Always use different engines and compare results.

Advanced search engine tools

Some search engines allow you to **program** and **filter searches** so that they automatically inform you about new results that are found following certain guidelines. For example, Google Alerts.

https://www.google.com/alerts

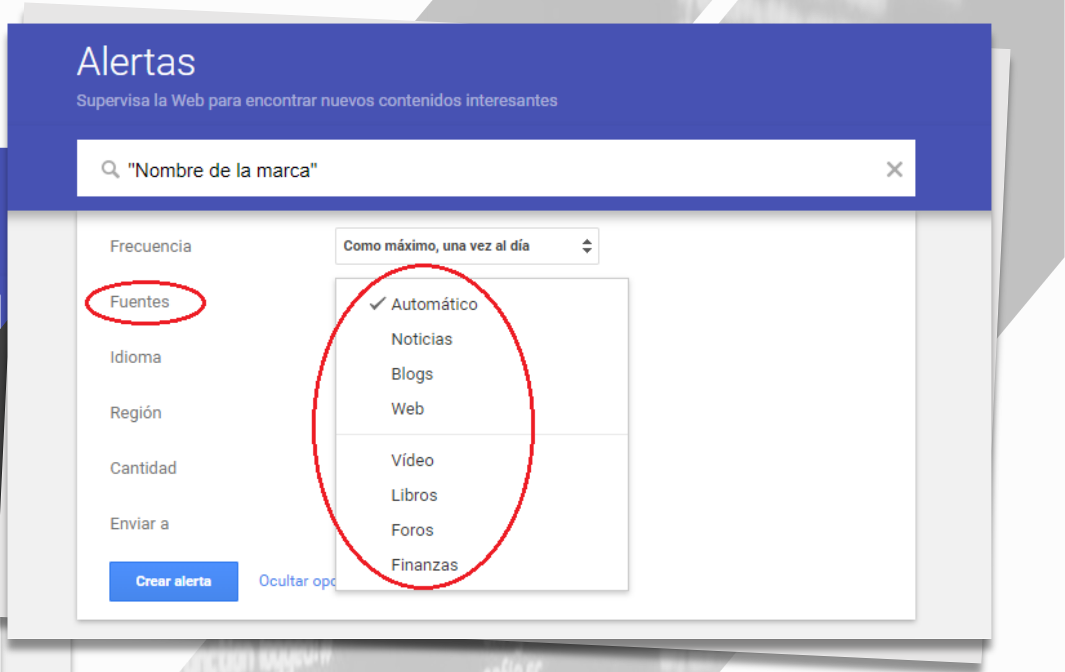


Alertas

Supervisa la Web para encontrar nuevos contenidos interesantes

Pablo Iglesias

Frecuencia	Como máximo, una vez al día
Fuentes	Automático
Idioma	español
Región	Todas las regiones
Cantidad	Solo los mejores resultados
Enviar a	lauracasassanchez@gmail.com



Alertas

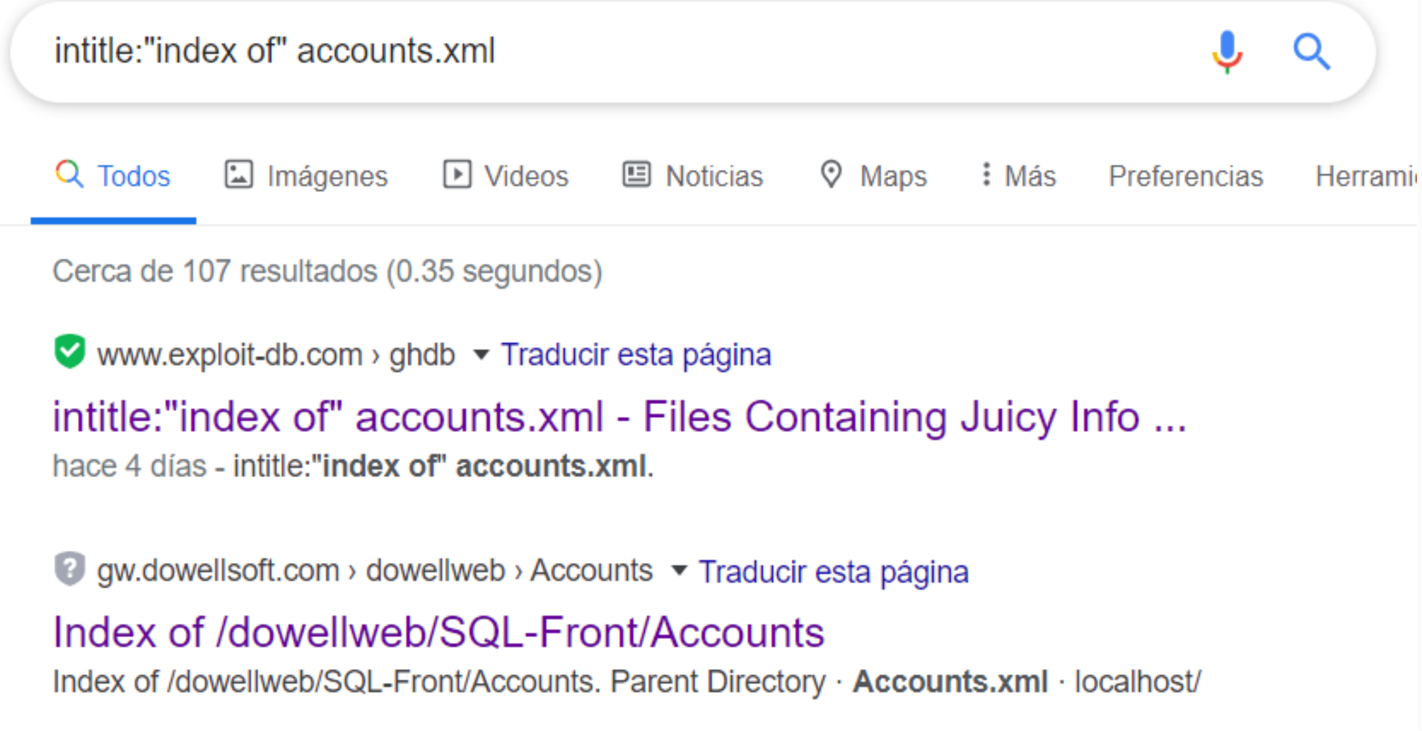
Supervisa la Web para encontrar nuevos contenidos interesantes

"Nombre de la marca"

Frecuencia: Como máximo, una vez al día

Google Operators

Example: intitle:"index of" accounts.xml



intitle:"index of" accounts.xml

Cerca de 107 resultados (0.35 segundos)

www.exploit-db.com › ghdb ▾ Traducir esta página

intitle:"index of" accounts.xml - Files Containing Juicy Info ...

hace 4 días - intitle:**"index of" accounts.xml**.

gw.dowellsoft.com › dowellweb › Accounts ▾ Traducir esta página

Index of /dowellweb/SQL-Front/Accounts

Index of /dowellweb/SQL-Front/Accounts. Parent Directory · **Accounts.xml** · localhost/

This query generates a list of pages that have the accounts.xml directory exposed in a listable directory.

Additional refinement parameters

Parameters	Explanation
Spaces indicate "and/or"	"secpro cybersecurity" = "secpro and cybersecurity" and "secpro or cybersecurity"
AND	“Secpro and cybersecurity”
OR	“Secpro or cybersecurity”
NOT or "-” to erase results	Used to eliminate results "Not Secpro Cybersecurity" or "-Secpro Cybersecurity" = No results for Secpro but results for Cybersecurity.
"*”	Look for anything that comes after an initial parameter; "Cyber*", it would search for: cybernetics, cyberwarfare, cyberdefense, etc.
quotation marks ""	Search for the exact phrase "david pereira hacker".
To give more relevance to a term: ^	secpro^4 Cybersecurity
$	Search Price of an item , for example: playstation5 $500

Dorks useful for different Search Engines

Dorks Listing for Different Search Engines

Parameter	Description	Google	Duck D. Go	Yahoo	Bing
cache:[url]	Displays the Cached version of the selected page.	√			
define:	Definition of the requested based on Google Dictionary.	√			
related:[url]	Search for web pages similar to the selected one.	√			
site:[url]	Limit the search to a specific page.	√	√	√	√
intitle:[text] or allintitle:[text]	Find pages with a certain word (or words) in the title. In Bing you have to put a space between the (:) and the word to be searched. Allintitle is used to search for several words at the same time.	√	√	√	√
inurl:[text] or allinurl:[text]	Find pages with a certain word (or words) in the URL.		√		
filetype:[extension] o ext:[extension]	Search for specific file types; e.g. filetype:doc, filetype:pdf - site:dian.gov.co filetype:xls	√	√		√
intext:[text], allintext:[text], inbody:[text]	Search for specific text on a page; for Bing and Yahoo the query is: inbody:[text]. For DuckDuckGo the query is intext:[text]. For Google the query is intext:[text] or allintext:[text].	√	√		√
inanchor:[text]	Search for web pages with links to the specified words.	√			

Dorks useful for different Search Engines

Dorks Listing for Different Search Engines

Parameter	Description	Google	Duck D. Go	Yahoo	Bing
location:[ISO code] o loc:[ISO Code], region:[Region code]	Search for information about a specific location by word or iso code: https://en.wikipedia.org/wiki/ISO_3166-1		✓		✓
contains:[text]	Search for sites that contain links to specific file types e.g. contains:pdf				✓
altloc:[ISO Code]	Search by location and language (e.g. pt-us or en-us).				✓
feed:[Example: rss]	Find RSS feeds related to a specific topic.		✓	✓	✓
ip:[IP address]	Find sites hosted on a specific IP.			✓	
language:[language code]	Delivers sites created in the specified language.		✓	✓	
book:[title]	Search for books related to the words consulted.	✓			
weather:	Displays the weather in a specific city: weather:bogotá	✓			
map:[Location]	Displays the geographic map of the selected location	✓			
linkfromdomain:[url]	Displays sites whose links are mentioned in the specified URL (May Fail)				✓

Google Hacking Data Base (GHDB)

Google hacking database (GHDB) is a project in charge of storing and keeping updated a series of queries in Google which provide interesting information about potential targets; this type of queries are called Google dorks.

The term dorks initially referred to a foolish or inept person disclosed by Google (Johny Long, DEFCON13).

Link to GHDB: https://www.exploit-db.com/google-hacking-database

GOOGLE
HACKING-DATABASE

Meta search engines

Meta search is a web-based service that aggregates data from various search engines. A meta search engine does not have a database of indexed pages of its own. Instead, it **submits** a user's queries to **several search engines** and **combines the top results** from each into an overall list.

Some important features are:

- Grouping of search results
- Suggestions of related terms
- Can be focused on specific topics such as images, business, etc.

Meta search engines

The key to obtaining good results is to have enough keywords to be able to narrow down the search.

Example of differences in results when searching for keywords:

- David Pereira
- David Pereira Cybersecurity

Meta search engines:

HTML Tools: Metasearch.html

http://www.etools.ch/search.do
https://www.startpage.com/
https://www.metacrawler.com/
https://www.alltheinternet.com/

OSINT to IT Infrastructure / Enterprises

SecPro
SECURITY PROFESSIONALS

Download Web Sites

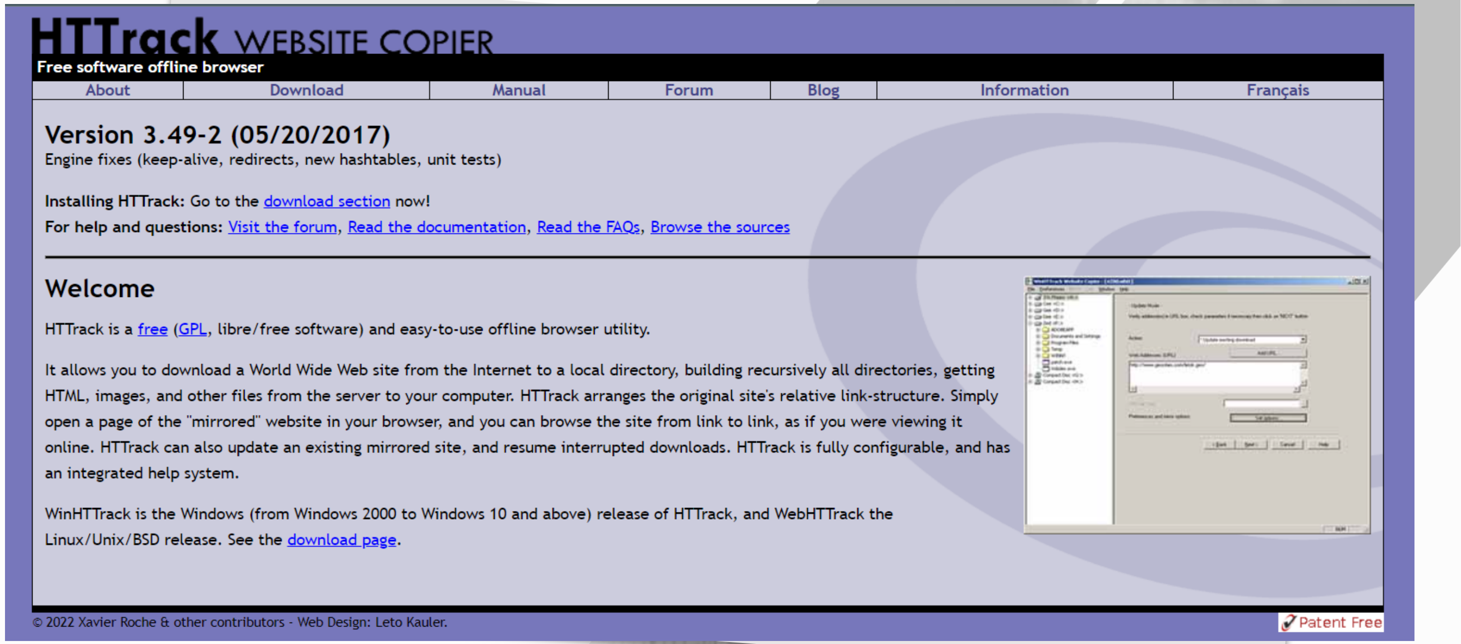
SecPro
SECURITY PROFESSIONALS

HTTRACK

https://www.httrack.com/

Tool that allows you to completely download a WEB site;

$ webhttrack &



Version 3.49-2 (05/20/2017)

Engine fixes (keep-alive, redirects, new hashtables, unit tests)

Installing HTTrack: Go to the download section now!

For help and questions: Visit the forum, Read the documentation, Read the FAQs, Browse the sources

Welcome

HTTrack is a free (GPL, libre/free software) and easy-to-use offline browser utility.

It allows you to download a World Wide Web site from the Internet to a local directory, building recursively all directories, getting HTML, images, and other files from the server to your computer. HTTrack arranges the original site's relative link-structure. Simply open a page of the "mirrored" website in your browser, and you can browse the site from link to link, as if you were viewing it online. HTTrack can also update an existing mirrored site, and resume interrupted downloads. HTTrack is fully configurable, and has an integrated help system.

WinHTTrack is the Windows (from Windows 2000 to Windows 10 and above) release of HTTrack, and WebHTTrack the Linux/Unix/BSD release. See the download page.

© 2022 Xavier Roche & other contributors - Web Design: Leto Kauler.

HASH of Files and Folders

In some investigations it is necessary to be able to demonstrate the integrity of the evidence and the non-alteration of the data collected; therefore we must use HASH generation tools for this purpose;

As a Hash Algorithm we recommend SHA256.

Since the level of collision is minimal.

Tools:

https://www.nirsoft.net/utils/hash_my_files.html

- sha256sum

Metadata

Every digital file in the world contains in its interior, information related to:

- File Name
- Description
- Creation Date
- Modification Date
- Date of last access;

In some file types we add:

- User who created it
- Last user who accessed it
- Printing Date
- Editing Tool
- Geolocation
- Device used to create it
- Among others

This information is called **metadata**.

OSINT of Metadata in OLE Files

In OLE (Object Linked and Embedded) document files, i.e. Microsoft documents (doc, xls, ppt).

Mainly we will find depending on the type of file:

Possible Metadata in an OLE File
Title
Topic
Type
Author
Business / Company
Status
Date and time of creation
Last modification date and time
Number of pages
Last recorded by:
Last printed on:

Metadata Search

In Kali Linux:

```
$ metagoofil.py -d alibaba.com -t xls –f
```

https://github.com/decalage2/oletools

```
$ olemeta *.xls |more
```

Metadata

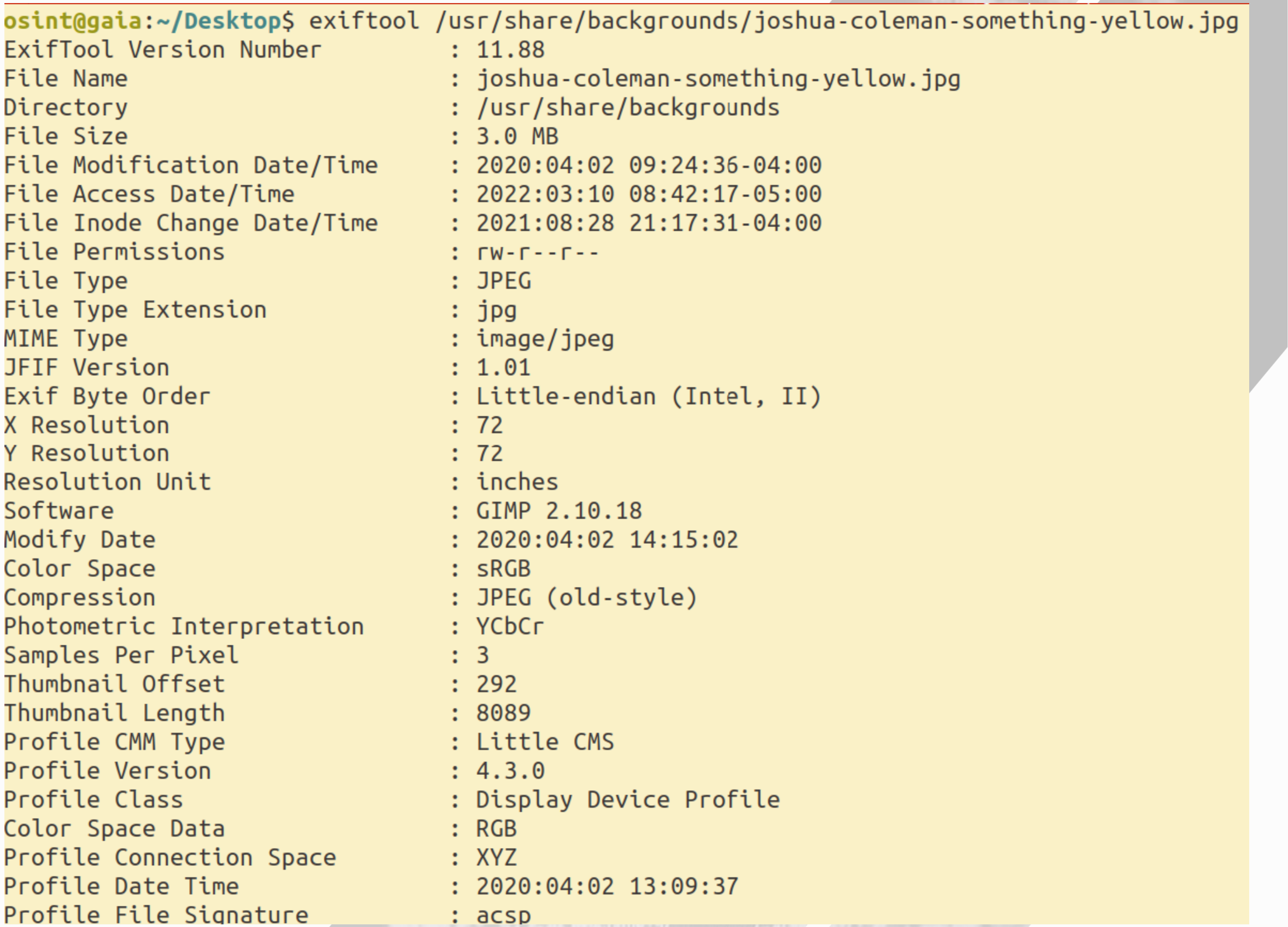
Exiftool

This tool can also be used in command line like this:

Exiftool file.jpeg

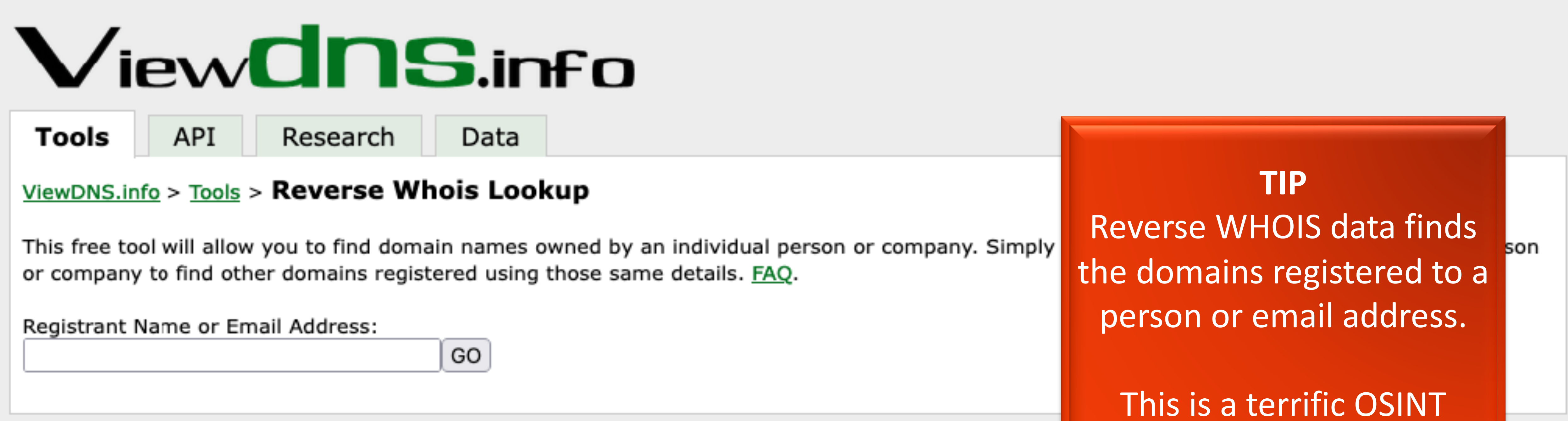
Examples:

https://exiftool.org/examples.html



```
osint@gaia:~/Desktop$ exiftool /usr/share/backgrounds/joshua-coleman-something-yellow.jpg
ExifTool Version Number         : 11.88
File Name                       : joshua-coleman-something-yellow.jpg
Directory                       : /usr/share/backgrounds
File Size                       : 3.0 MB
File Modification Date/Time     : 2020:04:02 09:24:36-04:00
File Access Date/Time           : 2022:03:10 08:42:17-05:00
File Inode Change Date/Time     : 2021:08:28 21:17:31-04:00
File Permissions                : rw-r--r--
File Type                       : JPEG
File Type Extension             : jpg
MIME Type                       : image/jpeg
JFIF Version                    : 1.01
Exif Byte Order                 : Little-endian (Intel, II)
X Resolution                    : 72
Y Resolution                    : 72
Resolution Unit                 : inches
Software                        : GIMP 2.10.18
Modify Date                     : 2020:04:02 14:15:02
Color Space                     : sRGB
Compression                     : JPEG (old-style)
Photometric Interpretation      : YCbCr
Samples Per Pixel               : 3
Thumbnail Offset                : 292
Thumbnail Length                : 8089
Profile CMM Type                : Little CMS
Profile Version                 : 4.3.0
Profile Class                   : Display Device Profile
Color Space Data                : RGB
Profile Connection Space        : XYZ
Profile Date Time               : 2020:04:02 13:09:37
Profile File Signature          : acsp
```

Reverse DNS Data



ViewDNS.info > Tools > **Reverse Whois Lookup**

This free tool will allow you to find domain names owned by an individual person or company. Simply son or company to find other domains registered using those same details. FAQ.

Registrant Name or Email Address:

TIP
Reverse WHOIS data finds the domains registered to a person or email address.

This is a terrific OSINT technique to identifify other less- well- knwn targets to attack.

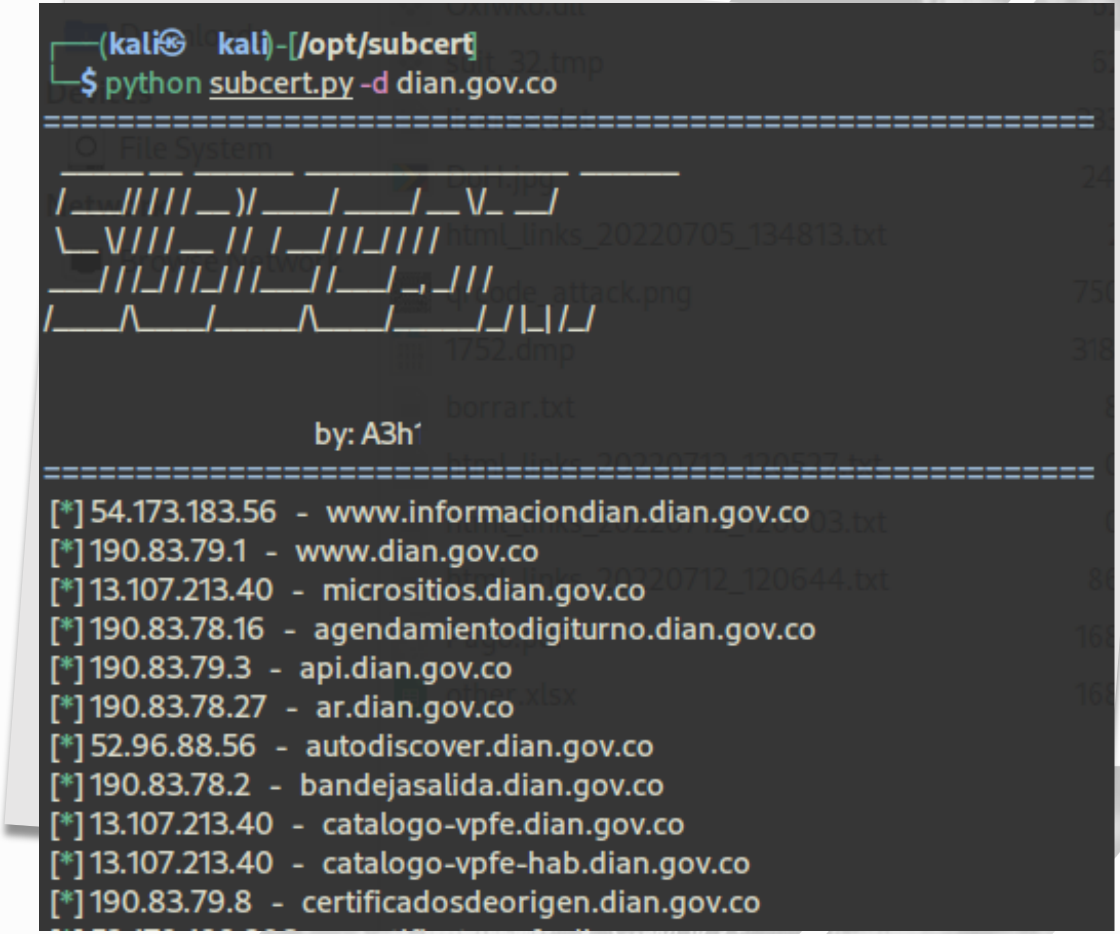
https://viewdns.info/

Subdomains Enumeration

SSL Certificates Enumeration

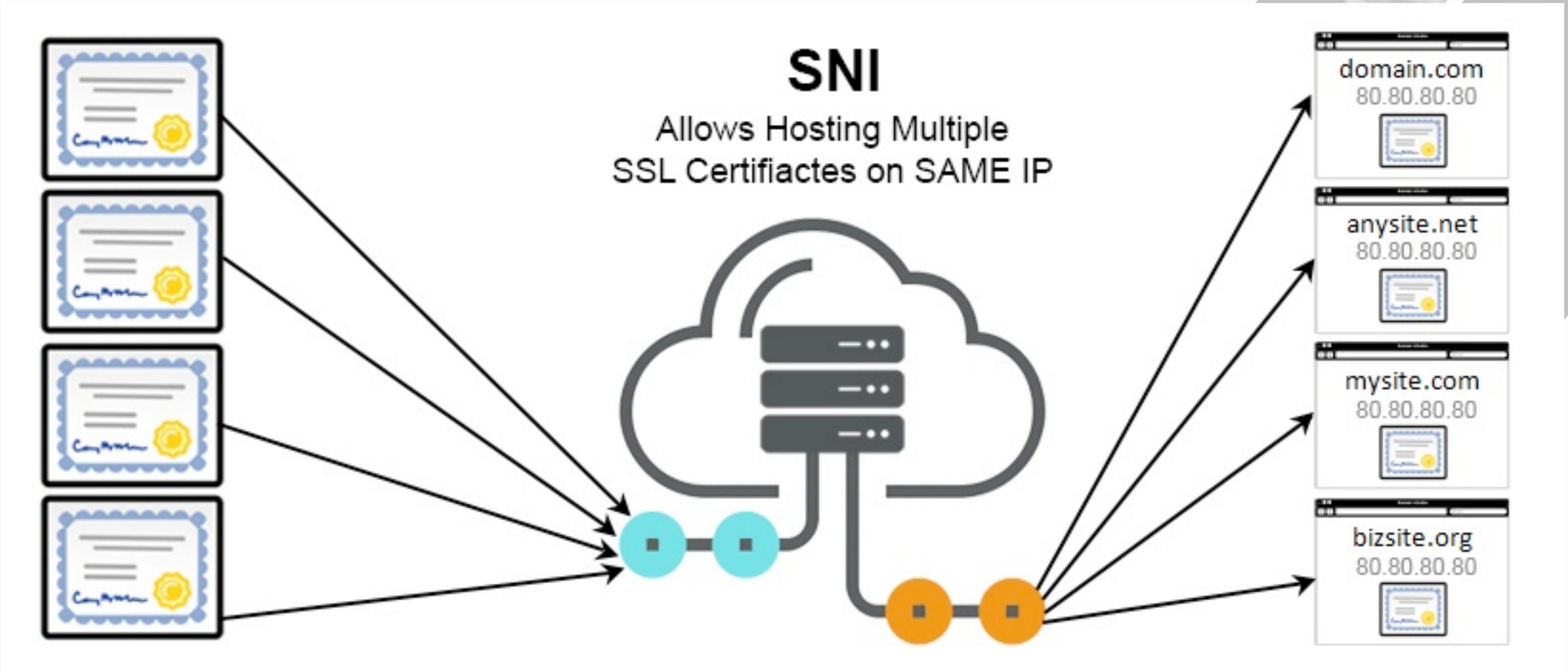
Certificate Transparency Logs

- Subcert: https://github.com/A3h1nt/Subcert
  - $/opt/subcert/subcert.py
- https://crt.sh/



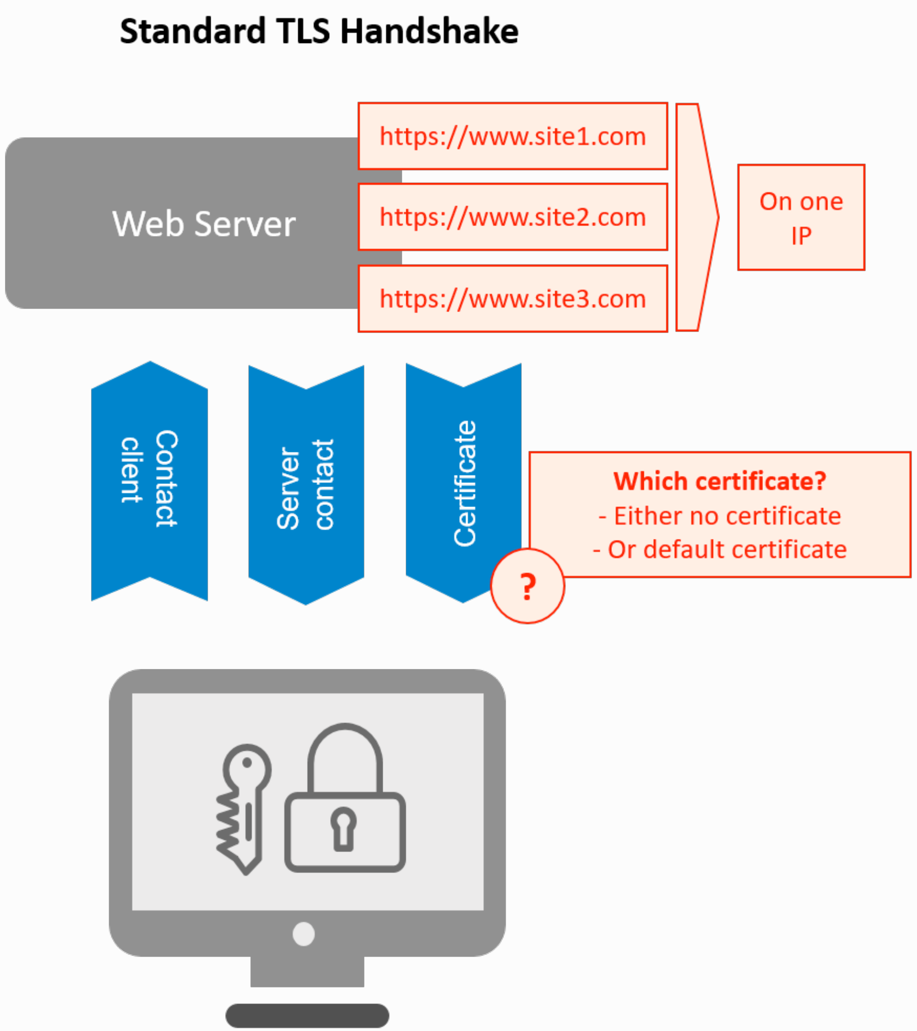
```
┌──(kali㉿kali)-[/opt/subcert]
└─$ python subcert.py -d dian.gov.co
==================================================================
  _____ __  ______  ___________________  _______
 / ___// / / / __ )/ ____/ ____/ __ \/_  __/
 \__ \/ / / / __  / /   / __/ / /_/ / / /
 ___/ / /_/ / /_/ / /___/ /___/ _, _/ / /
/____/\____/_____/\____/_____/_/ |_| /_/


                    by: A3h1
==================================================================
[*] 54.173.183.56  -  www.informaciondian.dian.gov.co
[*] 190.83.79.1  -  www.dian.gov.co
[*] 13.107.213.40  -  micrositios.dian.gov.co
[*] 190.83.78.16  -  agendamientodigiturno.dian.gov.co
[*] 190.83.79.3  -  api.dian.gov.co
[*] 190.83.78.27  -  ar.dian.gov.co
[*] 52.96.88.56  -  autodiscover.dian.gov.co
[*] 190.83.78.2  -  bandejasalida.dian.gov.co
[*] 13.107.213.40  -  catalogo-vpfe.dian.gov.co
[*] 13.107.213.40  -  catalogo-vpfe-hab.dian.gov.co
[*] 190.83.79.8  -  certificadosdeorigen.dian.gov.co
```

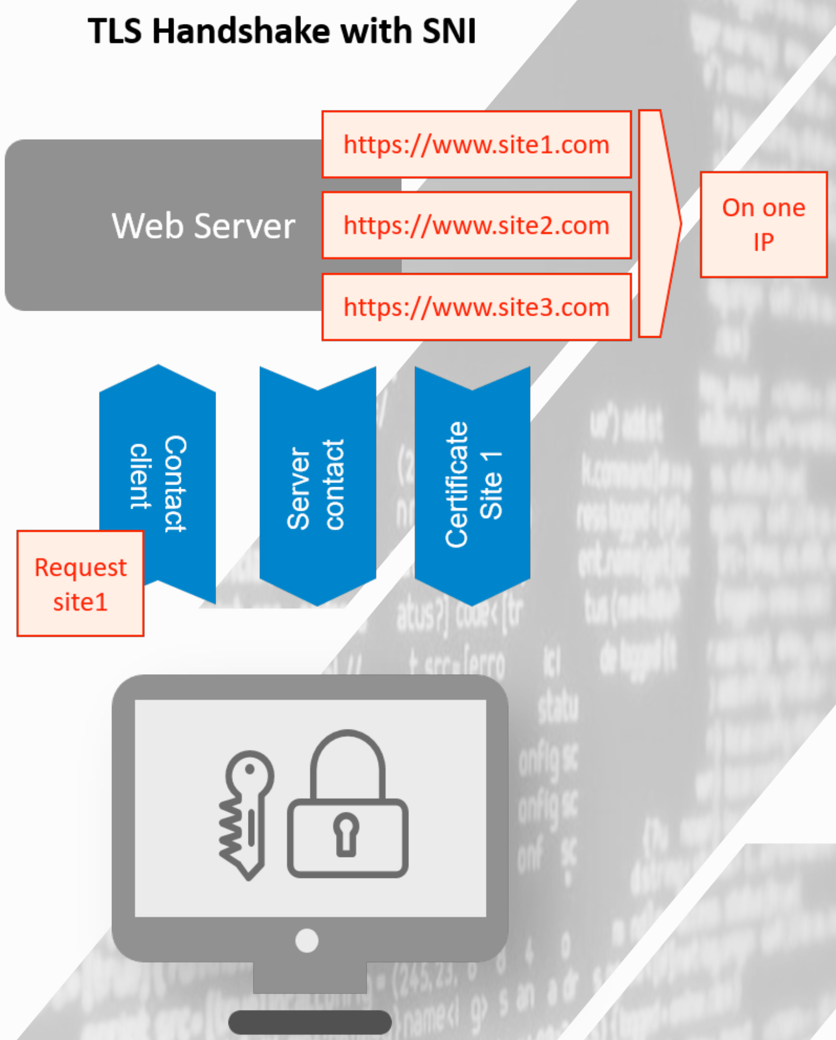
SNI - (Server Name Indication)



https://www.ssl2buy.com/wiki/server-name-indication-sni-use-multiple-ssl-on-a-single-ip

SNI - (Server Name Indication)





https://www.globalsign.com/en/blog/what-is-server-name-indication

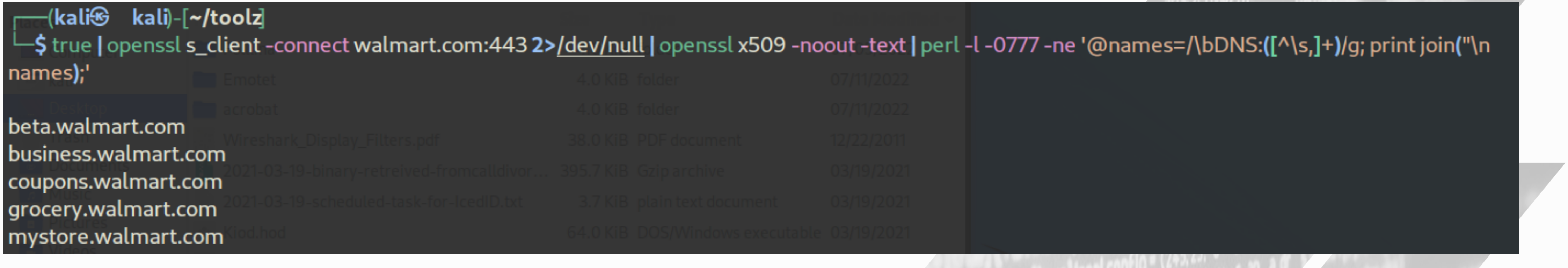
Subdomains Enumeration

SecPro
SECURITY PROFESSIONALS

SSL Certificates Enumeration

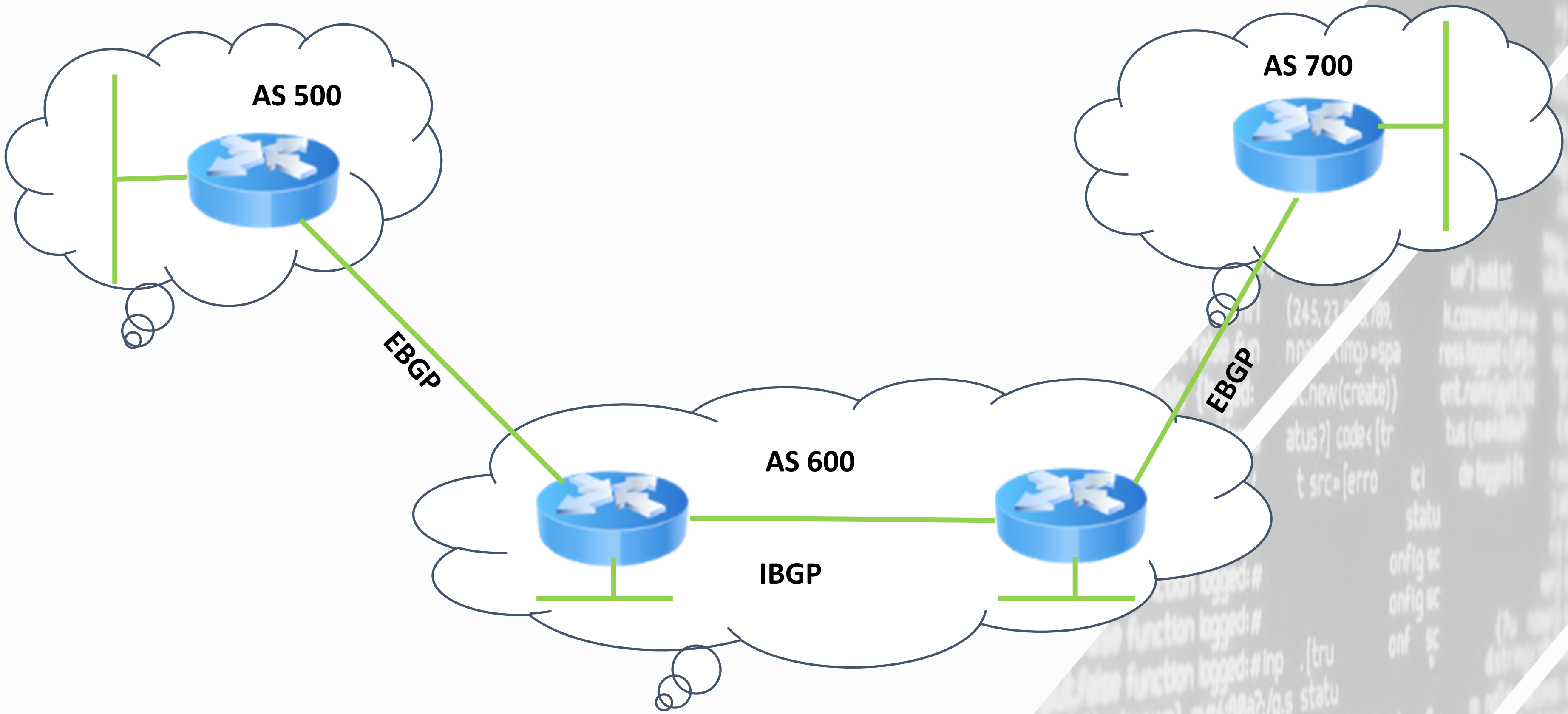
Subject Alternative Name (SAN)

- Structured list that includes all the Domain Names and IP addresses protected by a Certificate
  - $ true | openssl s_client -connect example.com:443 2>/dev/null | openssl x509 -noout -text | perl -l -0777 -ne '@names=/\bDNS:([^\s,]+)/g; print join("\n", sort @names);'



```
┌──(kali㉿ kali)-[~/toolz]
└─$ true | openssl s_client -connect walmart.com:443 2>/dev/null | openssl x509 -noout -text | perl -l -0777 -ne '@names=/\bDNS:([^\s,]+)/g; print join("\n
names);'

beta.walmart.com
business.walmart.com
coupons.walmart.com
grocery.walmart.com
mystore.walmart.com
```

Border Gateway protocol – Autonomous System



Subdomain Enumeration

- ASN Enumeration
  - Autonomous System Numbers are given to big networks for large Companies;
  - These ASN’s allow an attacker to detect IT Infrastructure
    - Tools:
      - https://www.robtex.com/
      - https://bgp.he.net/
      - Amass https://github.com/OWASP/Amass
        - $ amass intel -asn 1705
      - https://nmap.org/nsedoc/scripts/targets-asn.html
        - nmap --script targets-asn --script-args targets-asn.asn=1100

Search for devices and services -> AS

Robtex

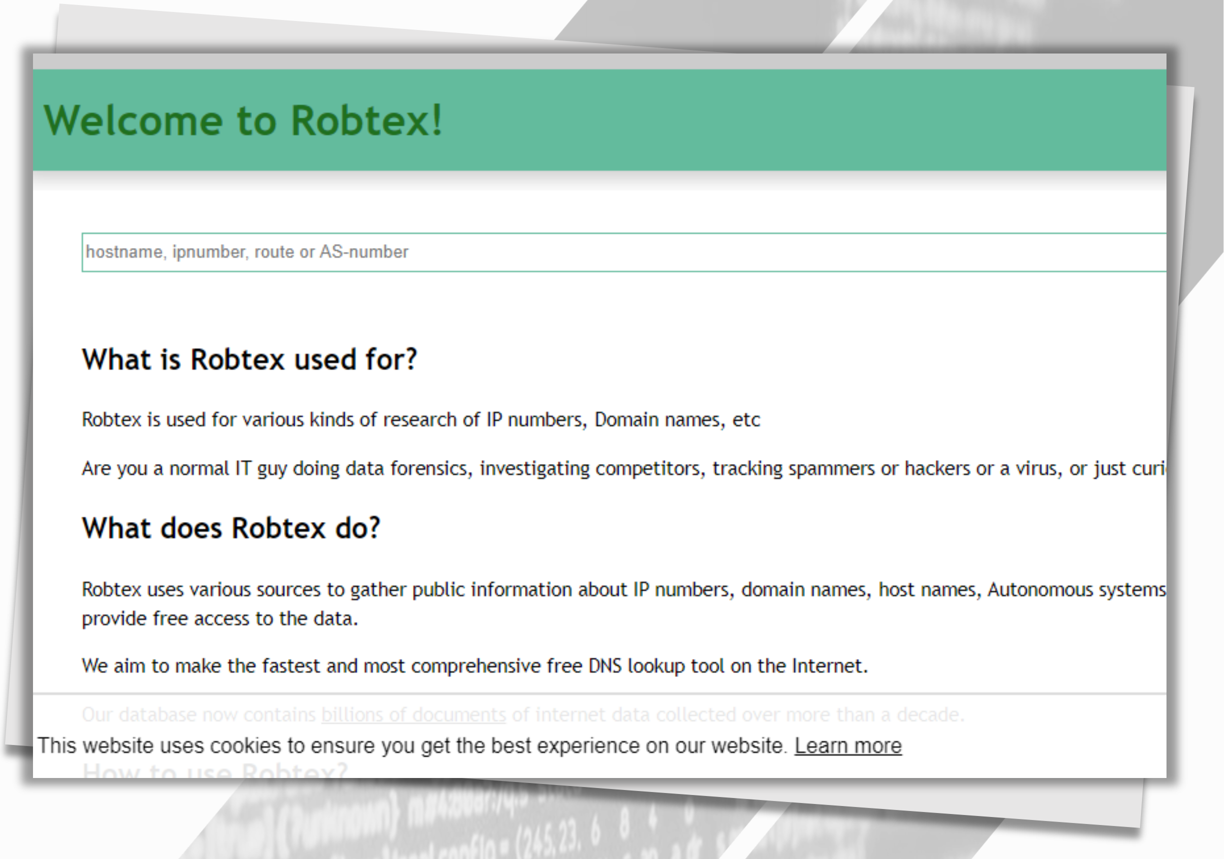
Robtex uses various sources to gather public information about IP addresses, domain names, hostnames, autonomous systems, routes, etc.

It then indexes the information in a large database and provides free access to them.

It aims to be the fastest and most comprehensive free DNS lookup tool on the Internet.

https://www.robtex.com/

https://www.robtex.com/dns-lookup/secpro.co#dnsrel



Welcome to Robtex!

What is Robtex used for?

Robtex is used for various kinds of research of IP numbers, Domain names, etc

Are you a normal IT guy doing data forensics, investigating competitors, tracking spammers or hackers or a virus, or just curi

What does Robtex do?

Robtex uses various sources to gather public information about IP numbers, domain names, host names, Autonomous systems
provide free access to the data.

We aim to make the fastest and most comprehensive free DNS lookup tool on the Internet.

Our database now contains billions of documents of internet data collected over more than a decade.

This website uses cookies to ensure you get the best experience on our website. Learn more

How to use Robtex?

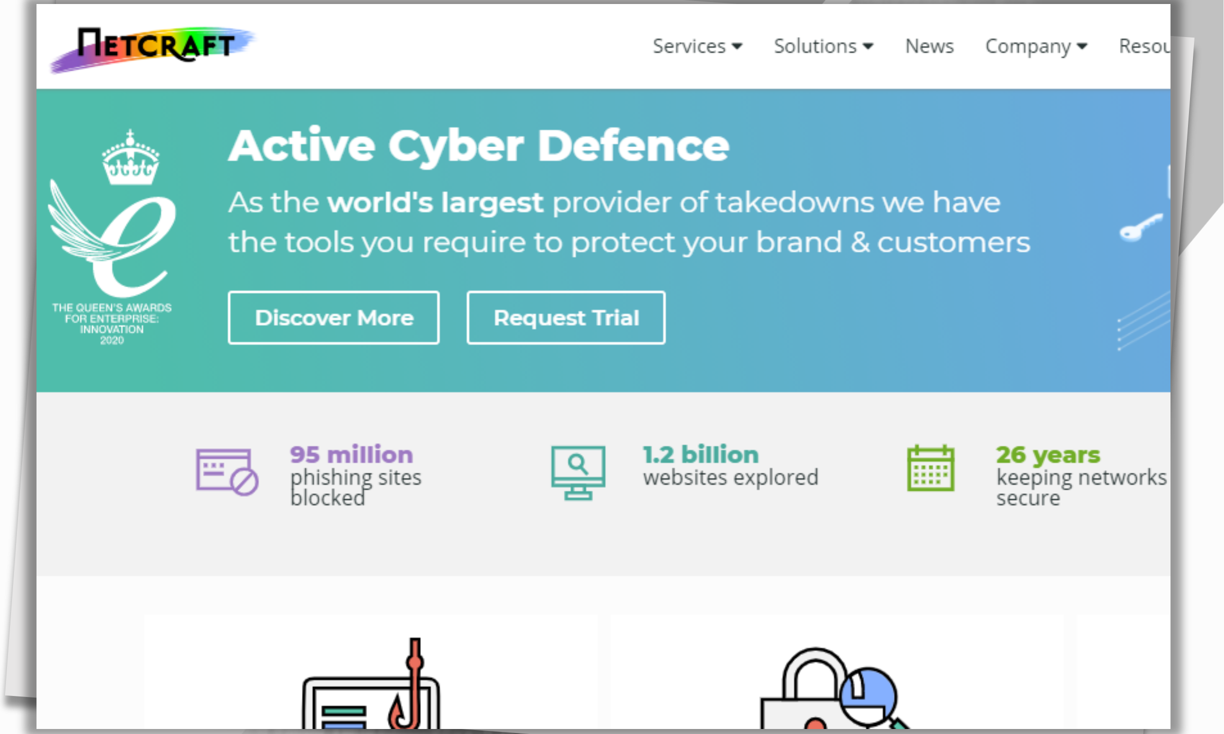
Search for Infrastructure Information

SecPro SECURITY PROFESSIONALS

Netcraft Site Report

Netcraft is a UK-based Internet services company that provides Internet security services, including cybercrime detection, application security testing and automated vulnerability scanning.

https://sitereport.netcraft.com/



NETCRAFT

Services ▾ Solutions ▾ News Company ▾ Resou

Active Cyber Defence

As the **world's largest** provider of takedowns we have the tools you require to protect your brand & customers

Discover More

Request Trial

THE QUEEN'S AWARDS FOR ENTERPRISE: INNOVATION 2020

95 million phishing sites blocked

1.2 billion websites explored

26 years keeping networks secure

Search for Technologies Information

Built With

https://builtwith.com/

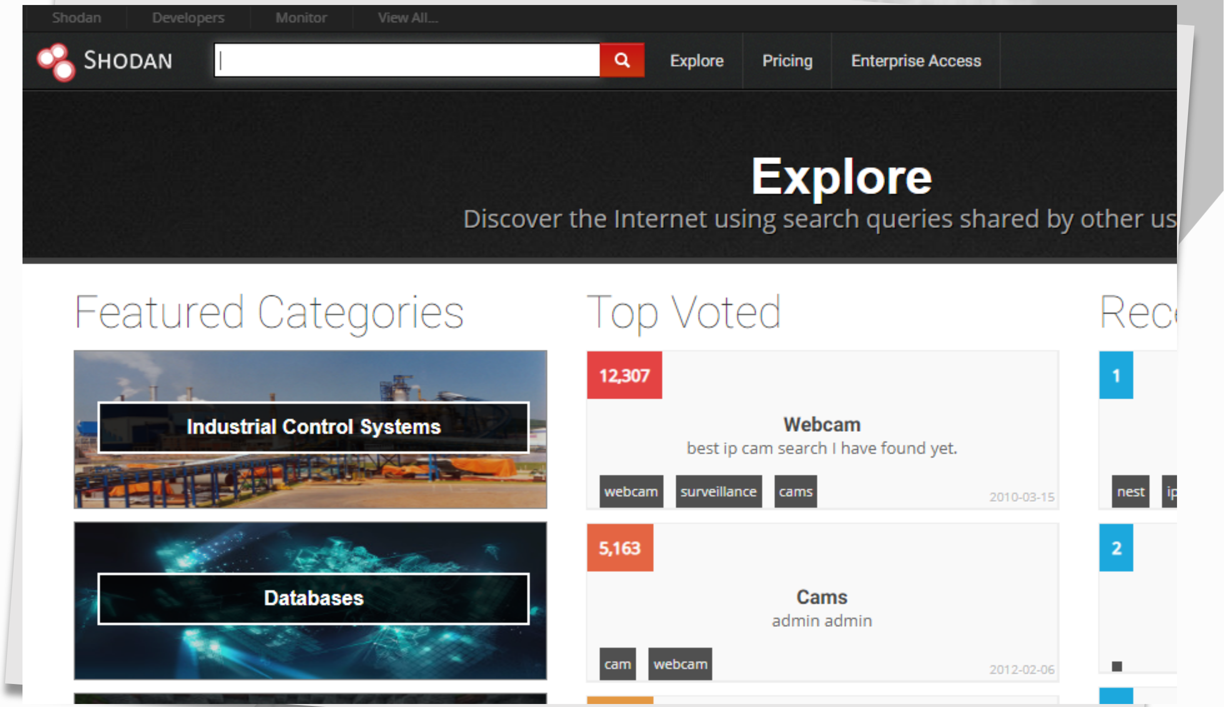
Search for devices and services

Shodan

Shodan is the search engine for everything on the Internet. While Google and other search engines only index the web, Shodan indexes almost everything else: webcams, water treatment facilities, yachts, medical devices, traffic lights, wind turbines, license plate readers, smart TVs, refrigerators, everything connected to the Internet https://www.shodan.io/.

Examples of searches:

- port:”5060" modbus
- product:”checkpoint"
- product:"kubernetes"
- hostname:”dian.gov.co"
- host:”180.10.10.10"



Explore

Discover the Internet using search queries shared by other us

Featured Categories

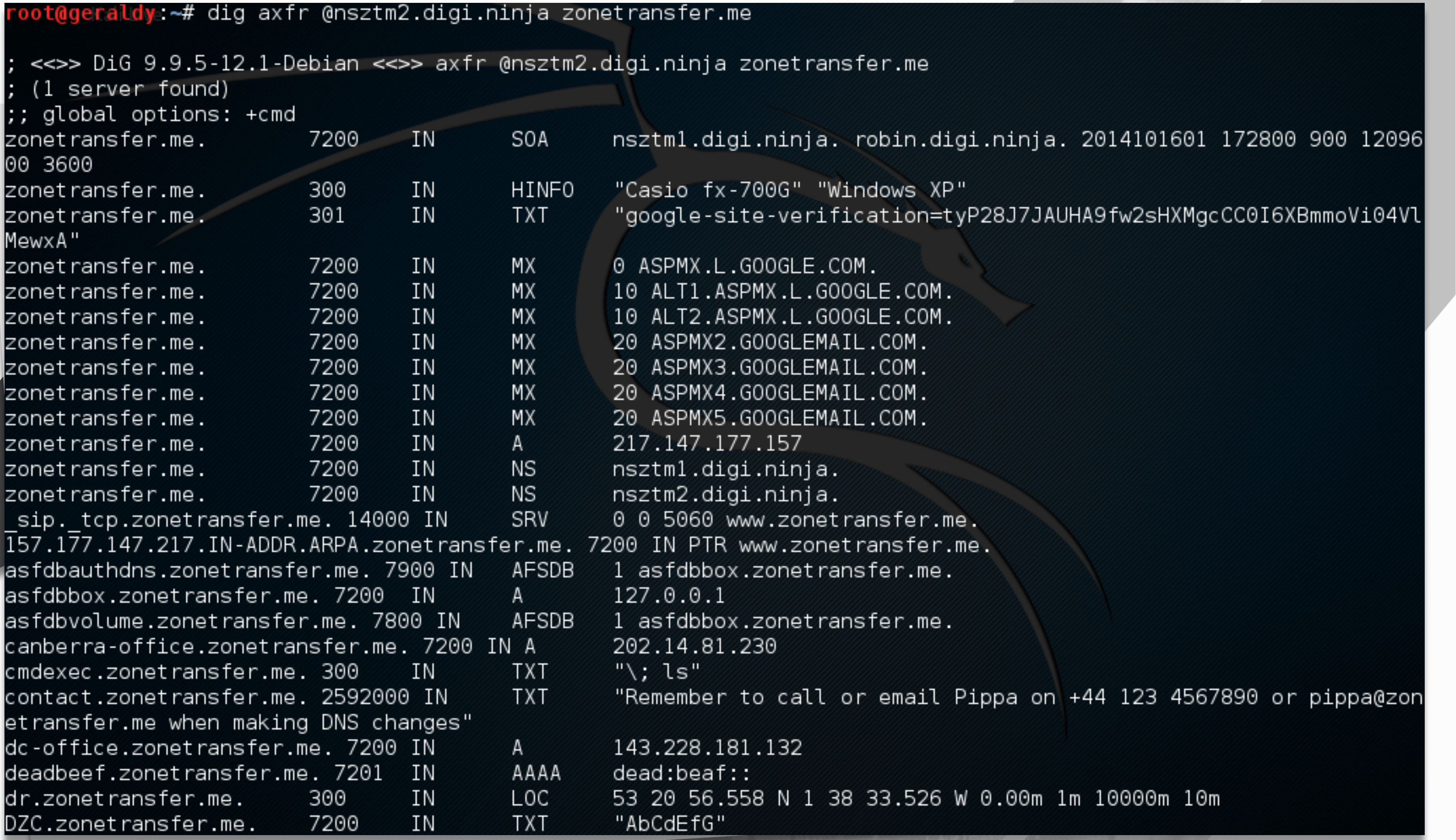
Top Voted

12,307 Webcam best ip cam search I have found yet. webcam surveillance cams 2010-03-15

5,163 Cams admin admin cam webcam 2012-02-06

Zone transfer

Zone transfer is a technique that exploits errors in the configuration of a DNS server to impersonate a new server and extract all records from a vulnerable DNS server.



```
root@geraldy:~# dig axfr @nsztm2.digi.ninja zonetransfer.me

; <<>> DiG 9.9.5-12.1-Debian <<>> axfr @nsztm2.digi.ninja zonetransfer.me
; (1 server found)
;; global options: +cmd
zonetransfer.me.        7200    IN      SOA     nsztm1.digi.ninja. robin.digi.ninja. 2014101601 172800 900 12096
00 3600
zonetransfer.me.        300     IN      HINFO   "Casio fx-700G" "Windows XP"
zonetransfer.me.        301     IN      TXT     "google-site-verification=tyP28J7JAUHA9fw2sHXMgcCC0I6XBmmoVi04Vl
MewxA"
zonetransfer.me.        7200    IN      MX      0 ASPMX.L.GOOGLE.COM.
zonetransfer.me.        7200    IN      MX      10 ALT1.ASPMX.L.GOOGLE.COM.
zonetransfer.me.        7200    IN      MX      10 ALT2.ASPMX.L.GOOGLE.COM.
zonetransfer.me.        7200    IN      MX      20 ASPMX2.GOOGLEMAIL.COM.
zonetransfer.me.        7200    IN      MX      20 ASPMX3.GOOGLEMAIL.COM.
zonetransfer.me.        7200    IN      MX      20 ASPMX4.GOOGLEMAIL.COM.
zonetransfer.me.        7200    IN      MX      20 ASPMX5.GOOGLEMAIL.COM.
zonetransfer.me.        7200    IN      A       217.147.177.157
zonetransfer.me.        7200    IN      NS      nsztm1.digi.ninja.
zonetransfer.me.        7200    IN      NS      nsztm2.digi.ninja.
_sip._tcp.zonetransfer.me. 14000 IN     SRV     0 0 5060 www.zonetransfer.me.
157.177.147.217.IN-ADDR.ARPA.zonetransfer.me. 7200 IN PTR www.zonetransfer.me.
asfdbauthdns.zonetransfer.me. 7900 IN   AFSDB   1 asfdbbox.zonetransfer.me.
asfdbbox.zonetransfer.me. 7200  IN      A       127.0.0.1
asfdbvolume.zonetransfer.me. 7800 IN    AFSDB   1 asfdbbox.zonetransfer.me.
canberra-office.zonetransfer.me. 7200 IN A      202.14.81.230
cmdexec.zonetransfer.me. 300    IN      TXT     "\; ls"
contact.zonetransfer.me. 2592000 IN     TXT     "Remember to call or email Pippa on +44 123 4567890 or pippa@zon
etransfer.me when making DNS changes"
dc-office.zonetransfer.me. 7200 IN      A       143.228.181.132
deadbeef.zonetransfer.me. 7201  IN      AAAA    dead:beaf::
dr.zonetransfer.me.     300     IN      LOC     53 20 56.558 N 1 38 33.526 W 0.00m 1m 10000m 10m
DZC.zonetransfer.me.    7200    IN      TXT     "AbCdEfG"
```

Thank You!!!

David Pereira

@davidpereiracib
https://www.youtube.com/user/dfpluc2
https://www.linkedin.com/in/davidfpereira/

SecPro
SECURITY PROFESSIONALS