🕵️ OSINT for a Website: A Beginner's Guide

This guide breaks down the process of gathering information about a website using publicly available tools and techniques. This process, known as **OSINT**, is a crucial first step in cybersecurity, allowing you to understand a website's infrastructure, technology, and potential vulnerabilities without ever directly interacting with it in a malicious way.

---

1) WHOIS: The Domain's ID Card

What it is:

WHOIS is a public directory that contains information about who owns a domain name. It's like looking up a home address in a public phone book. The information can include the owner's contact information, when the domain was registered, and when it's set to expire.

Why we do it:

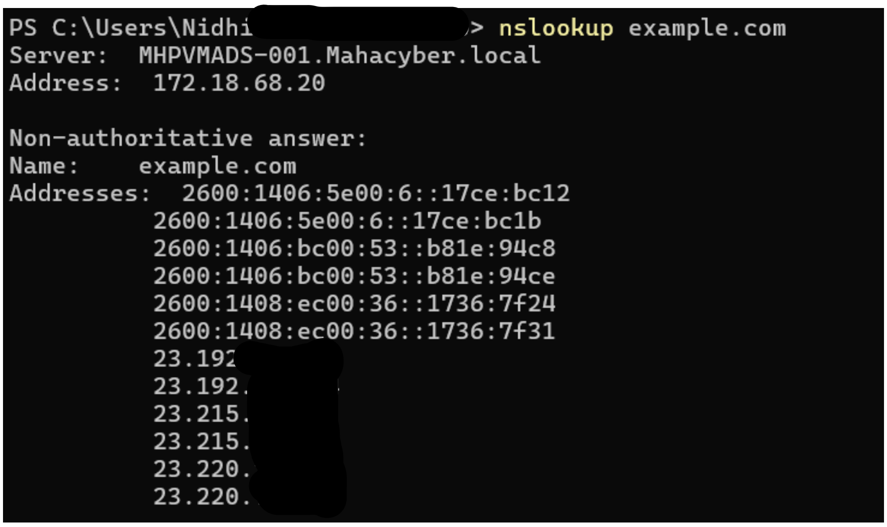
This step helps us understand the domain's **ownership**, **age**, and **registration status**. This information can be useful for identifying the owner (if their details aren't private) and assessing the domain's legitimacy. A very new domain might be a sign of a new business or, in some cases, a phishing site.

Observations:

While your screenshot doesn't show a direct WHOIS output, the typical information you'd find would include the **Registrar name** (e.g., GoDaddy, Namecheap), the **Creation Date**, the **Expiration Date**, and the **Nameservers**.

Key Takeaway:

WHOIS gives you the **foundation** of a domain's identity.



```
PS C:\Users\Nidhi[redacted]> nslookup example.com
Server:  MHPVMADS-001.Mahacyber.local
Address:  172.18.68.20

Non-authoritative answer:
Name:    example.com
Addresses:  2600:1406:5e00:6::17ce:bc12
          2600:1406:5e00:6::17ce:bc1b
          2600:1406:bc00:53::b81e:94c8
          2600:1406:bc00:53::b81e:94ce
          2600:1408:ec00:36::1736:7f24
          2600:1408:ec00:36::1736:7f31
          23.192[redacted]
          23.192.
          23.215.
          23.215.
          23.220.
          23.220.
```

2) DNS Records: The Internet's Address Book

What it is:

The Domain Name System (DNS) is the internet's phone book. It translates human-readable domain names (like example.com) into machine-readable IP addresses (like [redacted]). We're looking at different types of records that act as different entries in this address book.

- **A Records:** The most fundamental record. It maps a domain name to an **IPv4 address**.
- **AAAA Records:** Maps a domain name to an **IPv6 address**.
- **MX Records:** Specifies the **mail servers** responsible for accepting email messages on behalf of the domain.
- **NS Records:** Lists the **nameservers** that are authoritative for the domain.
- **TXT Records:** Can hold arbitrary text, often used for things like **email authentication** (SPF, DKIM) or domain verification.

Why we do it:

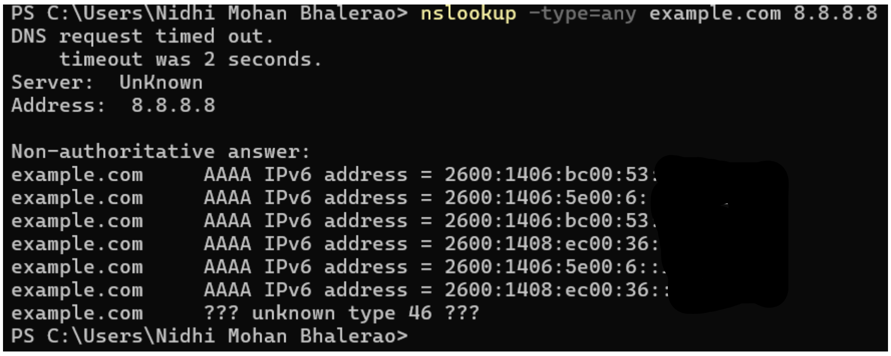
Understanding a website's DNS records reveals its **core infrastructure**. You can see where it's hosted (A records), what service handles its email (MX records), and its official nameservers. This gives you a clear picture of its digital footprint.

Observations:

Your `nslookup` output shows multiple **IP addresses** (both IPv4 and IPv6) for `example.com`, indicating that the site likely uses a Content Delivery Network (CDN) or has a distributed setup for load balancing. This is a common practice for large websites to improve performance and reliability. Your `nslookup -type=TXT` output also shows an SPF record, which helps prevent email spoofing.

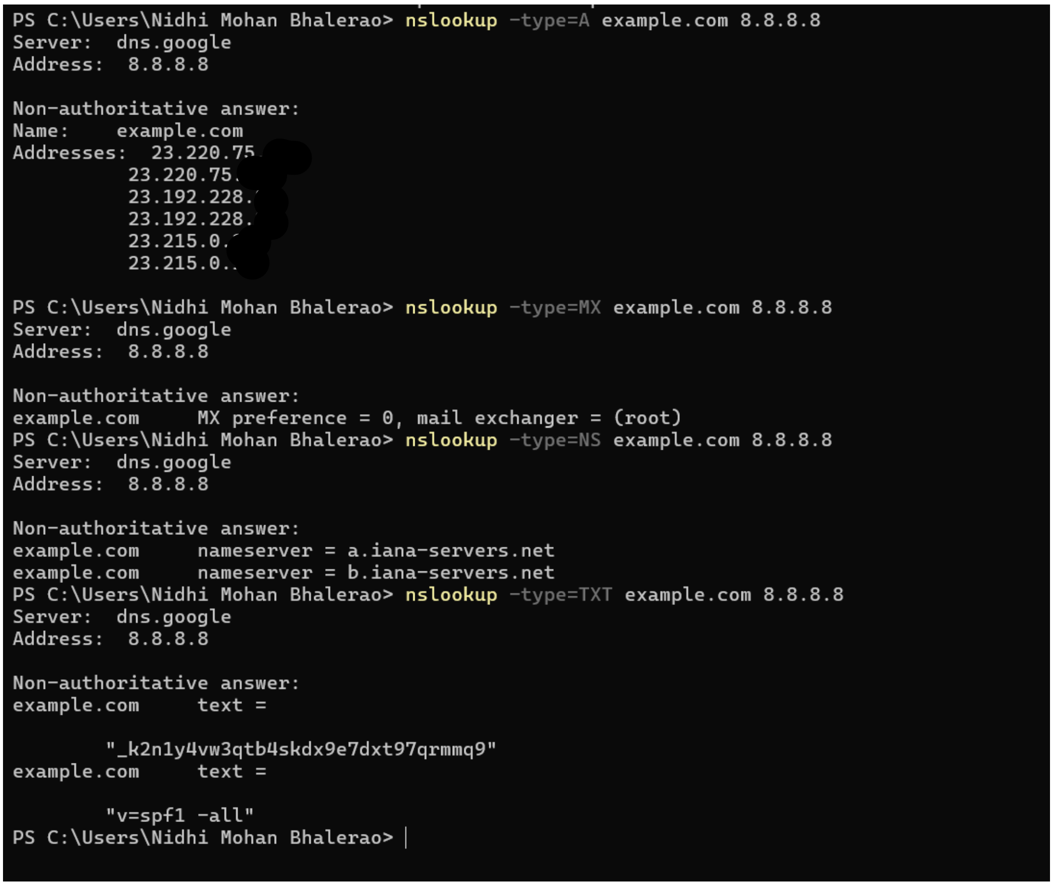
Key Takeaway:

DNS records provide a **technical blueprint** of the website's digital infrastructure.



```
PS C:\Users\Nidhi Mohan Bhalerao> nslookup -type=any example.com 8.8.8.8
DNS request timed out.
    timeout was 2 seconds.
Server:  UnKnown
Address:  8.8.8.8

Non-authoritative answer:
example.com     AAAA IPv6 address = 2600:1406:bc00:53:[redacted]
example.com     AAAA IPv6 address = 2600:1406:5e00:6:
example.com     AAAA IPv6 address = 2600:1406:bc00:53
example.com     AAAA IPv6 address = 2600:1408:ec00:36:
example.com     AAAA IPv6 address = 2600:1406:5e00:6::
example.com     AAAA IPv6 address = 2600:1408:ec00:36::
example.com     ??? unknown type 46 ???
PS C:\Users\Nidhi Mohan Bhalerao>
```



```
PS C:\Users\Nidhi Mohan Bhalerao> nslookup -type=A example.com 8.8.8.8
Server:  dns.google
Address:  8.8.8.8

Non-authoritative answer:
Name:    example.com
Addresses:  23.220.75[redacted]
          23.220.75.
          23.192.228.
          23.192.228.
          23.215.0.
          23.215.0.

PS C:\Users\Nidhi Mohan Bhalerao> nslookup -type=MX example.com 8.8.8.8
Server:  dns.google
Address:  8.8.8.8

Non-authoritative answer:
example.com     MX preference = 0, mail exchanger = (root)
PS C:\Users\Nidhi Mohan Bhalerao> nslookup -type=NS example.com 8.8.8.8
Server:  dns.google
Address:  8.8.8.8

Non-authoritative answer:
example.com     nameserver = a.iana-servers.net
example.com     nameserver = b.iana-servers.net
PS C:\Users\Nidhi Mohan Bhalerao> nslookup -type=TXT example.com 8.8.8.8
Server:  dns.google
Address:  8.8.8.8

Non-authoritative answer:
example.com     text =

        "_k2n1y4vw3qtb4skdx9e7dxt97qrmmq9"
example.com     text =

        "v=spf1 -all"
PS C:\Users\Nidhi Mohan Bhalerao> |
```

3) Certificate Transparency & Subdomain Discovery

What it is:

Certificate Transparency (CT) is a public log of all SSL/TLS certificates issued by Certificate Authorities (CAs). When a CA issues a certificate for a domain (e.g., `example.com`), it's logged publicly. This log often includes certificates for **subdomains** like `blog.example.com` or `dev.example.com`. Tools like **crt.sh** query this log.

Why we do it:

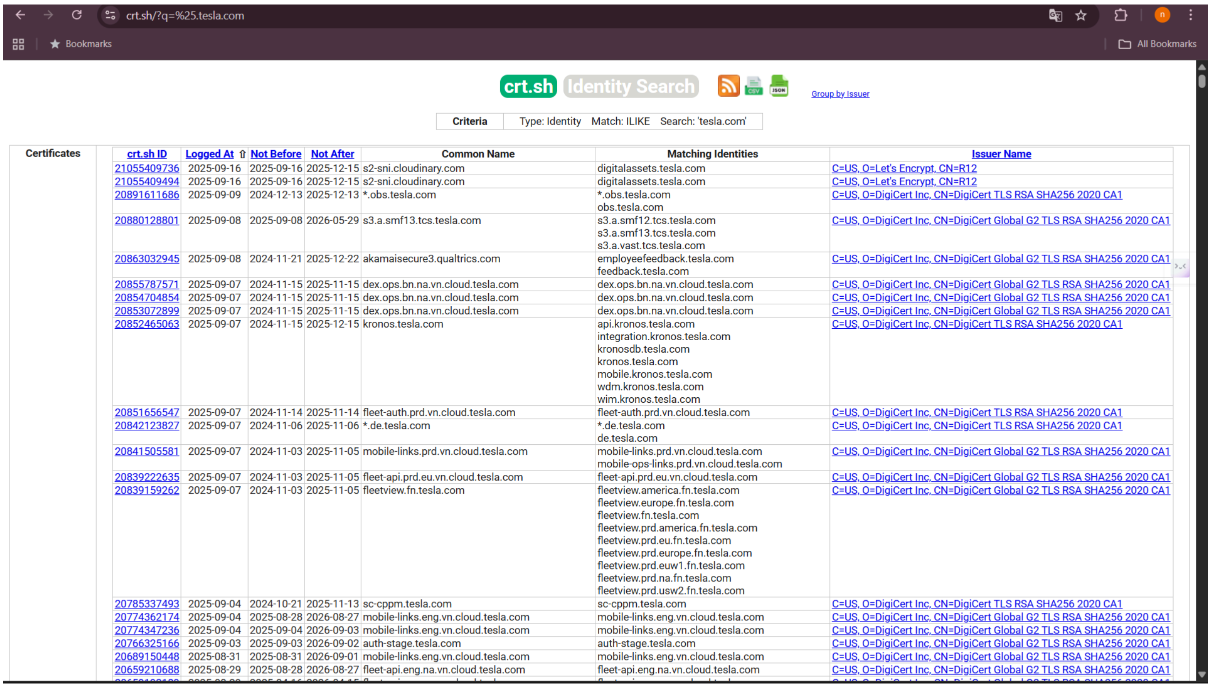
This step is a goldmine for discovering **hidden or forgotten subdomains**. Attackers often target these subdomains because they might be less secure or running older software than the main website. Finding them is a critical part of the reconnaissance process.

Observations:

Your `crt.sh` screenshot for `example.com` shows numerous certificates, revealing multiple subdomains like `example.com` and `*.example.com`. The `*` indicates a **wildcard certificate**, which can cover any subdomain.

Key Takeaway:

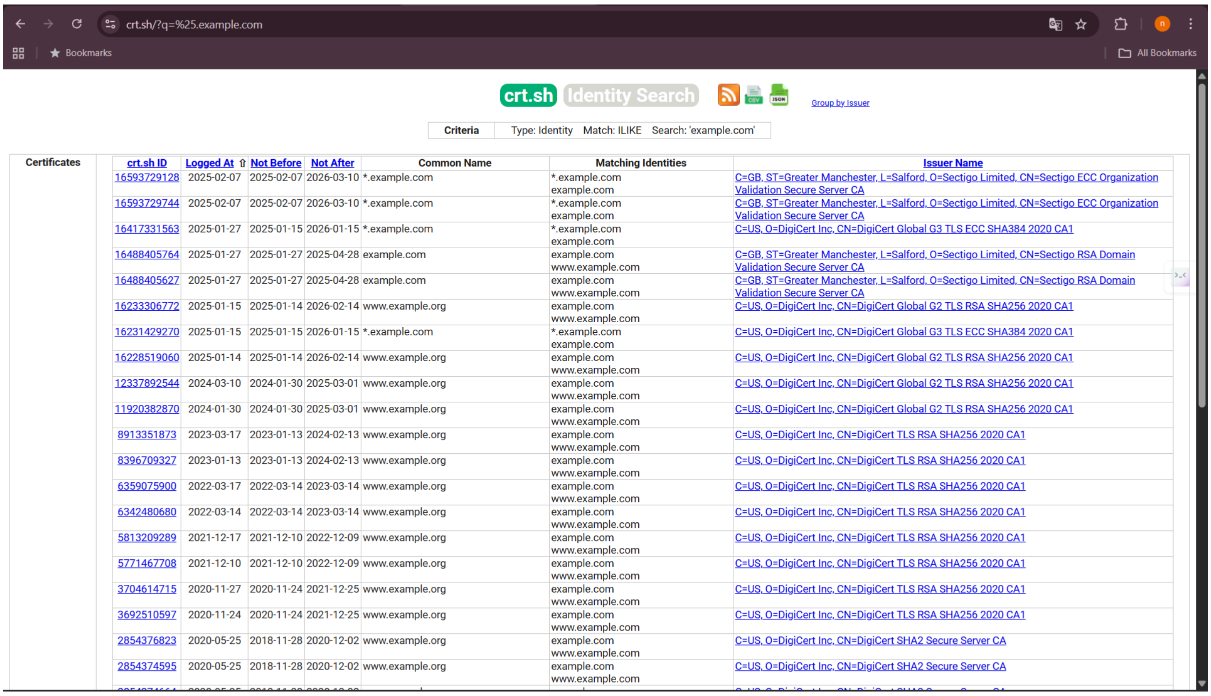
crt.sh is a passive but powerful way to find a site's **forgotten corners**.



Criteria | Type: Identity Match: ILIKE Search: 'tesla.com'

Certificates

crt.sh ID	Logged At ⇧	Not Before	Not After	Common Name	Matching Identities	Issuer Name
21055409736	2025-09-16	2025-09-16	2025-12-15	s2-sni.cloudinary.com	digitalassets.tesla.com	C=US, O=Let's Encrypt, CN=R12
21055409494	2025-09-16	2025-09-16	2025-12-15	s2-sni.cloudinary.com	digitalassets.tesla.com	C=US, O=Let's Encrypt, CN=R12
20891611686	2025-09-09	2024-12-13	2025-12-13	*.obs.tesla.com	*.obs.tesla.com obs.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
20880128801	2025-09-08	2025-09-08	2026-05-29	s3.a.smf13.tcs.tesla.com	s3.a.smf12.tcs.tesla.com s3.a.smf13.tcs.tesla.com s3.a.vast.tcs.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20863032945	2025-09-08	2024-11-21	2025-12-22	akamaisecure3.qualtrics.com	employeefeedback.tesla.com feedback.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20855787571	2025-09-07	2024-11-15	2025-11-15	dex.ops.bn.na.vn.cloud.tesla.com	dex.ops.bn.na.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20854704854	2025-09-07	2024-11-15	2025-11-15	dex.ops.bn.na.vn.cloud.tesla.com	dex.ops.bn.na.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20853072899	2025-09-07	2024-11-15	2025-11-15	dex.ops.bn.na.vn.cloud.tesla.com	dex.ops.bn.na.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20852465063	2025-09-07	2024-11-15	2025-12-15	kronos.tesla.com	api.kronos.tesla.com integration.kronos.tesla.com kronosdb.tesla.com kronos.tesla.com mobile.kronos.tesla.com wdm.kronos.tesla.com wim.kronos.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
20851656547	2025-09-07	2024-11-14	2025-11-14	fleet-auth.prd.vn.cloud.tesla.com	fleet-auth.prd.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
20842123827	2025-09-07	2024-11-06	2025-11-06	*.de.tesla.com	*.de.tesla.com de.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
20841505581	2025-09-07	2024-11-03	2025-11-05	mobile-links.prd.vn.cloud.tesla.com	mobile-links.prd.vn.cloud.tesla.com mobile-ops-links.prd.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20839222635	2025-09-07	2024-11-03	2025-11-05	fleet-api.prd.eu.vn.cloud.tesla.com	fleet-api.prd.eu.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20839159262	2025-09-07	2024-11-03	2025-11-05	fleetview.fn.tesla.com	fleetview.america.fn.tesla.com fleetview.europe.fn.tesla.com fleetview.fn.tesla.com fleetview.prd.america.fn.tesla.com fleetview.prd.eu.fn.tesla.com fleetview.prd.europe.fn.tesla.com fleetview.prd.euw1.fn.tesla.com fleetview.prd.na.fn.tesla.com fleetview.prd.usw2.fn.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20785337493	2025-09-04	2024-10-21	2025-11-13	sc-cppm.tesla.com	sc-cppm.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
20774362174	2025-09-04	2025-08-28	2026-08-27	mobile-links.eng.vn.cloud.tesla.com	mobile-links.eng.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20774347236	2025-09-04	2025-09-04	2026-09-03	mobile-links.eng.vn.cloud.tesla.com	mobile-links.eng.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20766325166	2025-09-03	2025-09-03	2026-09-02	auth-stage.tesla.com	auth-stage.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20689150448	2025-08-31	2025-08-31	2026-09-01	mobile-links.eng.vn.cloud.tesla.com	mobile-links.eng.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
20659210688	2025-08-29	2025-08-28	2026-08-27	fleet-api.eng.na.vn.cloud.tesla.com	fleet-api.eng.na.vn.cloud.tesla.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1



Criteria | Type: Identity Match: ILIKE Search: 'example.com'

Certificates

crt.sh ID	Logged At ⇧	Not Before	Not After	Common Name	Matching Identities	Issuer Name
16593729128	2025-02-07	2025-02-07	2026-03-10	*.example.com	*.example.com example.com	C=GB, ST=Greater Manchester, L=Salford, O=Sectigo Limited, CN=Sectigo ECC Organization Validation Secure Server CA
16593729744	2025-02-07	2025-02-07	2026-03-10	*.example.com	*.example.com example.com	C=GB, ST=Greater Manchester, L=Salford, O=Sectigo Limited, CN=Sectigo ECC Organization Validation Secure Server CA
16417331563	2025-01-27	2025-01-15	2026-01-15	*.example.com	*.example.com example.com	C=US, O=DigiCert Inc, CN=DigiCert Global G3 TLS ECC SHA384 2020 CA1
16488405764	2025-01-27	2025-01-27	2025-04-28	example.com	example.com www.example.com	C=GB, ST=Greater Manchester, L=Salford, O=Sectigo Limited, CN=Sectigo RSA Domain Validation Secure Server CA
16488405627	2025-01-27	2025-01-27	2025-04-28	example.com	example.com www.example.com	C=GB, ST=Greater Manchester, L=Salford, O=Sectigo Limited, CN=Sectigo RSA Domain Validation Secure Server CA
16233306772	2025-01-15	2025-01-14	2026-02-14	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
16231429270	2025-01-15	2025-01-15	2026-01-15	*.example.com	*.example.com example.com	C=US, O=DigiCert Inc, CN=DigiCert Global G3 TLS ECC SHA384 2020 CA1
16228519060	2025-01-14	2025-01-14	2026-02-14	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
12337892544	2024-03-10	2024-01-30	2025-03-01	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
11920382870	2024-01-30	2024-01-30	2025-03-01	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert Global G2 TLS RSA SHA256 2020 CA1
8913351873	2023-03-17	2023-01-13	2024-02-13	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
8396709327	2023-01-13	2023-01-13	2024-02-13	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
6359075900	2022-03-17	2022-03-14	2023-03-14	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
6342480680	2022-03-14	2022-03-14	2023-03-14	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
5813209289	2021-12-17	2021-12-10	2022-12-09	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
5771467708	2021-12-10	2021-12-10	2022-12-09	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
3704614715	2020-11-27	2020-11-24	2021-12-25	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
3692510597	2020-11-24	2020-11-24	2021-12-25	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert TLS RSA SHA256 2020 CA1
2854376823	2020-05-25	2018-11-28	2020-12-02	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert SHA2 Secure Server CA
2854374595	2020-05-25	2018-11-28	2020-12-02	www.example.org	example.com www.example.com	C=US, O=DigiCert Inc, CN=DigiCert SHA2 Secure Server CA

4) Passive DNS: Mapping the Network

What it is:

Passive DNS is a technique that involves collecting and storing DNS query responses from around the internet. Tools like **DNSDumpster** use this data to create a map of a domain's network, showing its subdomains, IP addresses, and the organizations that host them.

Why we do it:

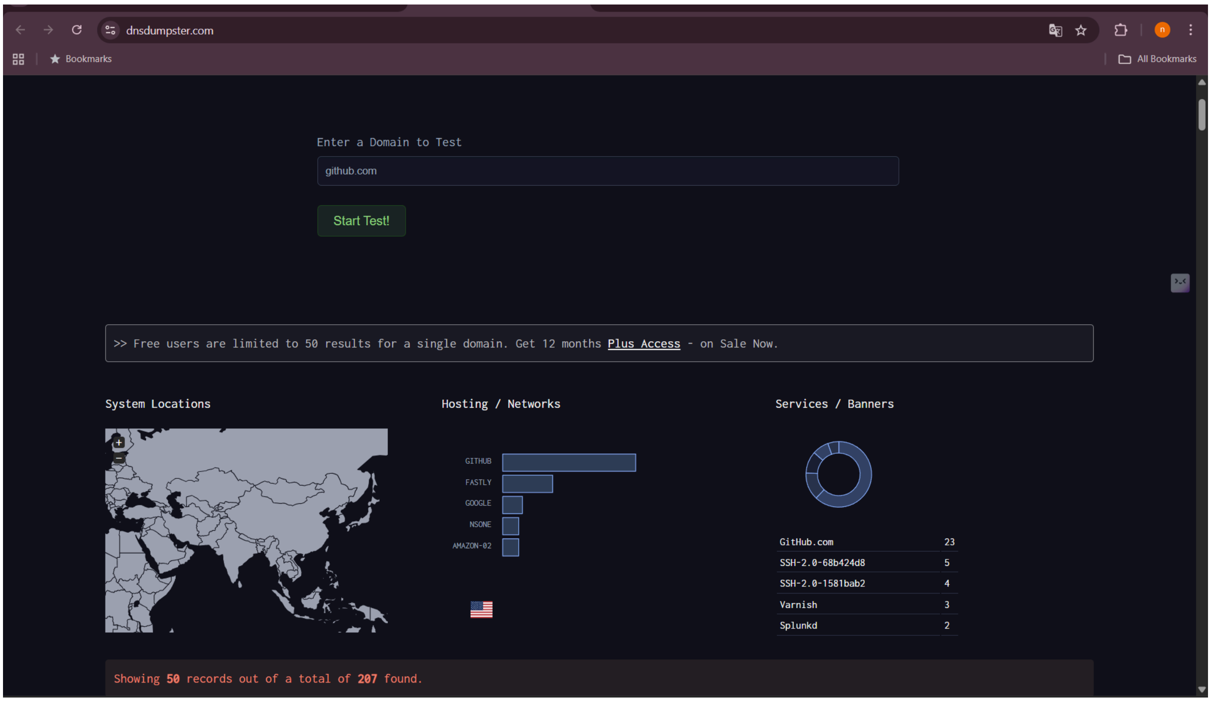
This step helps you **visualize the domain's entire network footprint**. You can see not only the domain's main IPs but also those of its subdomains, and which companies (like Fastly, Google, or GitHub) are hosting different parts of the network. This can reveal dependencies and relationships.

Observations:

Your `DNSDumpster` screenshot for `github.com` shows its IPs and which hosting providers (Fastly, Google, etc.) are associated with them. The map view helps visualize these connections.

Key Takeaway:

Passive DNS gives you a **top-down map** of the domain's connected assets.



Enter a Domain to Test

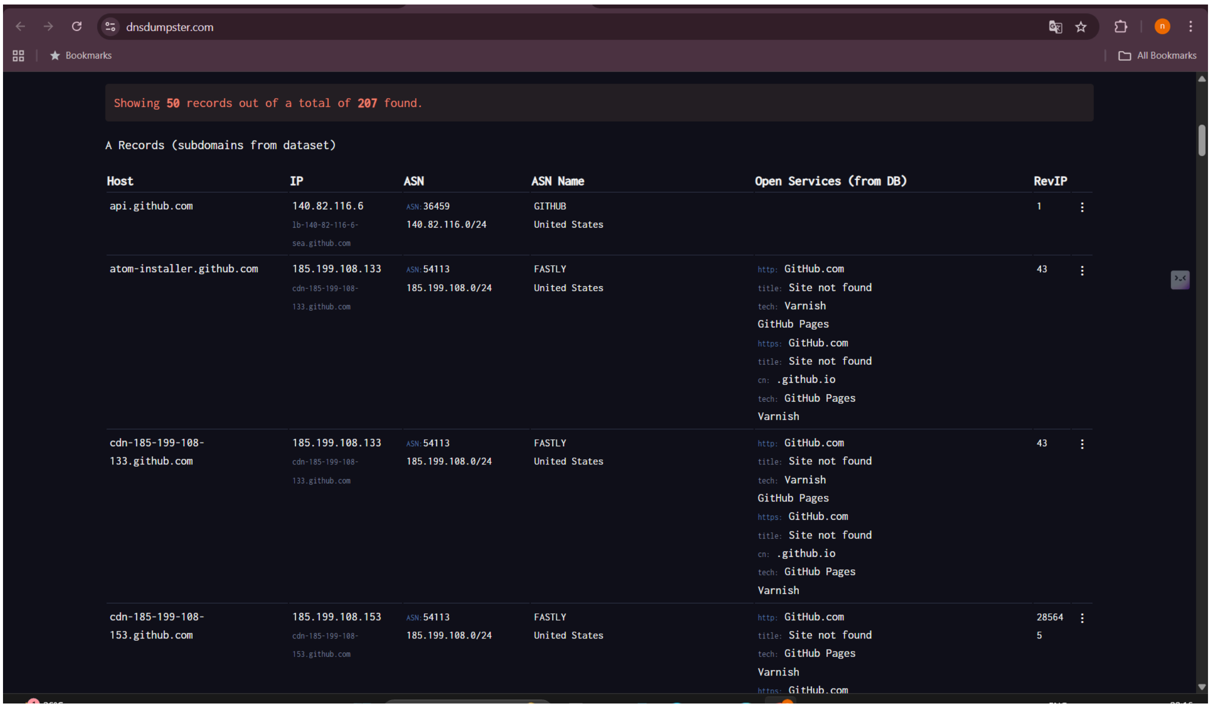
github.com

Start Test!

>> Free users are limited to 50 results for a single domain. Get 12 months Plus Access - on Sale Now.

GitHub.com	23
SSH-2.0-68b424d8	5
SSH-2.0-1581bab2	4
Varnish	3
Splunkd	2

Showing **50** records out of a total of **207** found.



Showing **50** records out of a total of **207** found.

A Records (subdomains from dataset)

Host	IP	ASN	ASN Name	Open Services (from DB)	RevIP
api.github.com	140.82.116.6 lb-140-82-116-6-sea.github.com	ASN: 36459 140.82.116.0/24	GITHUB United States		1
atom-installer.github.com	185.199.108.133 cdn-185-199-108-133.github.com	ASN: 54113 185.199.108.0/24	FASTLY United States	http: GitHub.com title: Site not found tech: Varnish GitHub Pages https: GitHub.com title: Site not found cn: .github.io tech: GitHub Pages Varnish	43
cdn-185-199-108-133.github.com	185.199.108.133 cdn-185-199-108-133.github.com	ASN: 54113 185.199.108.0/24	FASTLY United States	http: GitHub.com title: Site not found tech: Varnish GitHub Pages https: GitHub.com title: Site not found cn: .github.io tech: GitHub Pages Varnish	43
cdn-185-199-108-153.github.com	185.199.108.153 cdn-185-199-108-153.github.com	ASN: 54113 185.199.108.0/24	FASTLY United States	http: GitHub.com title: Site not found tech: GitHub Pages Varnish https: GitHub.com	285645

5) Technology Stack Fingerprinting

What it is:

Technology fingerprinting involves identifying the software and services a website uses. Tools like **BuiltWith** or Wappalyzer analyze a website's code and headers to reveal what it's built with, such as the Content Management System (CMS), JavaScript libraries, advertising platforms, and analytics tools.

Why we do it:

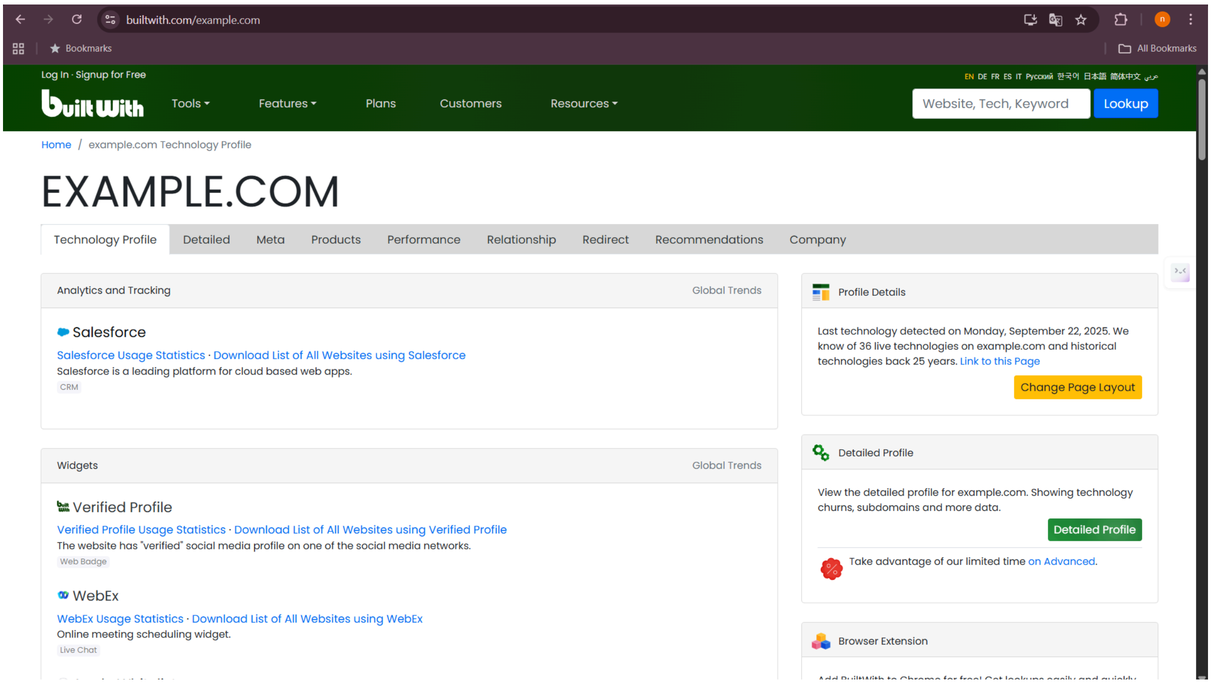
Knowing the technology stack is crucial because it helps identify **known vulnerabilities** associated with specific software versions. For example, if you find a website is running an old version of WordPress, you can search for public exploits for that version.

Observations:

Your **BuiltWith** screenshot for `example.com` shows technologies like **Salesforce** and **WebEx**, which indicates that this domain is likely used by a company for business operations, rather than being a simple informational website.

Key Takeaway:

BuiltWith is like a **tech-savvy detective** that tells you what tools a website is using.



EXAMPLE.COM

Analytics and Tracking

Salesforce

Salesforce Usage Statistics · Download List of All Websites using Salesforce

Salesforce is a leading platform for cloud based web apps.

CRM

Widgets

Verified Profile

Verified Profile Usage Statistics · Download List of All Websites using Verified Profile

The website has "verified" social media profile on one of the social media networks.

Web Badge

WebEx

WebEx Usage Statistics · Download List of All Websites using WebEx

Online meeting scheduling widget.

Live Chat

Profile Details

Last technology detected on Monday, September 22, 2025. We know of 36 live technologies on example.com and historical technologies back 25 years. Link to this Page

Detailed Profile

View the detailed profile for example.com. Showing technology churns, subdomains and more data.

Take advantage of our limited time on Advanced.

Browser Extension

Add BuiltWith to Chrome for free! Get lookups easily and quickly

---

6) Shodan: The Search Engine for Devices

What it is:

Shodan is a search engine for internet-connected devices. Instead of searching for web pages, it searches for servers, routers, webcams, and other devices, providing information about their open ports and service banners. A **service banner** is a response from a service that often includes its name and version (e.g., `Apache/2.4.41`).

Why we do it:

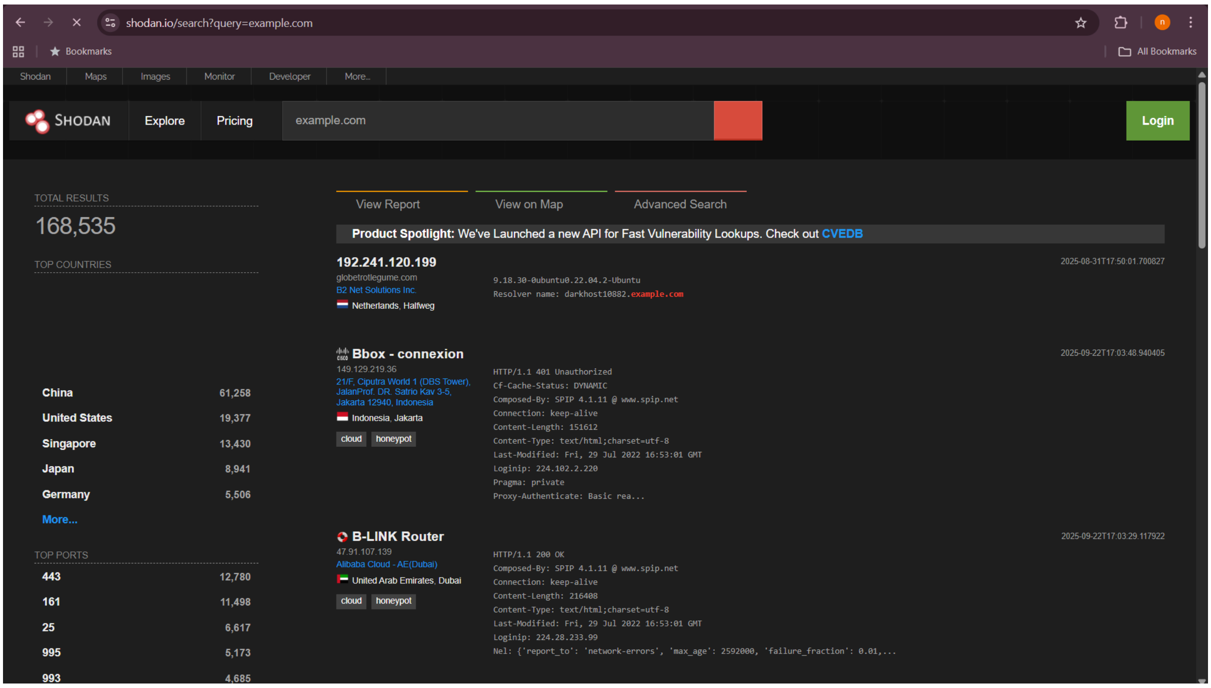
By searching for the **IP addresses** you found earlier, you can see what ports are open on the server and what services are running on them. This is the closest you get to the actual machine. Open ports and outdated service banners can signal security weaknesses.

Observations:

Your `Shodan` screenshot shows details for the IP `192.241.120.199`, including open ports and the service banner `nginx/1.18.0`. This tells you a web server is running Nginx and provides its version number, which can be useful for further research.

Key Takeaway:

Shodan gives you a **live peek** at what's running on a server's external ports.



TOTAL RESULTS

168,535

TOP COUNTRIES

China	61,258
United States	19,377
Singapore	13,430
Japan	8,941
Germany	5,506

More...

TOP PORTS

443	12,780
161	11,498
25	6,617
995	5,173
993	4,685

Product Spotlight: We've Launched a new API for Fast Vulnerability Lookups. Check out CVEDB

192.241.120.199 2025-08-31T17:50:01.700827
globetrotlegume.com
B2 Net Solutions Inc.
Netherlands, Halfweg

```
9.18.30-0ubuntu0.22.04.2-Ubuntu
Resolver name: darkhost10882.example.com
```

Bbox - connexion 2025-09-22T17:03:48.940405
149.129.219.36
21/F, Ciputra World 1 (DBS Tower), JalanProf. DR. Satrio Kav 3-5, Jakarta 12940, Indonesia
Indonesia, Jakarta
cloud honeypot

```
HTTP/1.1 401 Unauthorized
Cf-Cache-Status: DYNAMIC
Composed-By: SPIP 4.1.11 @ www.spip.net
Connection: keep-alive
Content-Length: 151612
Content-Type: text/html;charset=utf-8
Last-Modified: Fri, 29 Jul 2022 16:53:01 GMT
Loginip: 224.102.2.220
Pragma: private
Proxy-Authenticate: Basic rea...
```

B-LINK Router 2025-09-22T17:03:29.117922
47.91.107.139
Alibaba Cloud - AE(Dubai)
United Arab Emirates, Dubai
cloud honeypot

```
HTTP/1.1 200 OK
Composed-By: SPIP 4.1.11 @ www.spip.net
Connection: keep-alive
Content-Length: 216408
Content-Type: text/html;charset=utf-8
Last-Modified: Fri, 29 Jul 2022 16:53:01 GMT
Loginip: 224.28.233.99
Nel: {'report_to': 'network-errors', 'max_age': 2592000, 'failure_fraction': 0.01,...
```