OSINT Resources for 2019

Steve Micallef Follow
Dec 27, 2018 · 9 min read

Whether you are new to OSINT (Open Source Intelligence) or use it regularly in your professional life for reconnaissance, threat intelligence or investigations, the recent speed of growth in the field means constant development in terms of tooling, data, content and community. In this post I aim to highlight some essentials that everyone relying on OSINT should know, plus newer resources that might provide additional insights.

Photo by Tom Sodoge on Unsplash

First, the Essentials

If you are new to OSINT or come from a less technical background, there are some foundational resources you should gain a solid grasp of first because they'll really help you get better use out of the other tools mentioned later in this post, in addition to gaining a deeper understanding the data they present:

- **DNS:** With tools like host, dig and nslookup you can look-up different types of DNS records (A, CNAME, NS, MX, TXT, etc), use alternate name servers and more. For instance, did you know that

Quad9's DNS will always resolve any malicious host to 127.0.0.1? This means by using their name server to perform your lookup, you can quickly check if it's potentially malicious:

```
$ nslookup m-tesla.pw 9.9.9.9
Server:       9.9.9.9
Address:      9.9.9.9#53

Non-authoritative answer:
Name:   m-tesla.pw
Address: 127.0.0.1
```

- **Whois:** Probably everyone knows about performing Whois lookups on domains, but did you know you can also perform Whois on IP addresses, networks and ASNs? Let's see who owns 8.8.8.8 — yes, Google, but Whois shows us that the broader network range is owned by Level 3 Communications (now CenturyLink), who have sub-allocated 8.8.8.0/24 to Google:

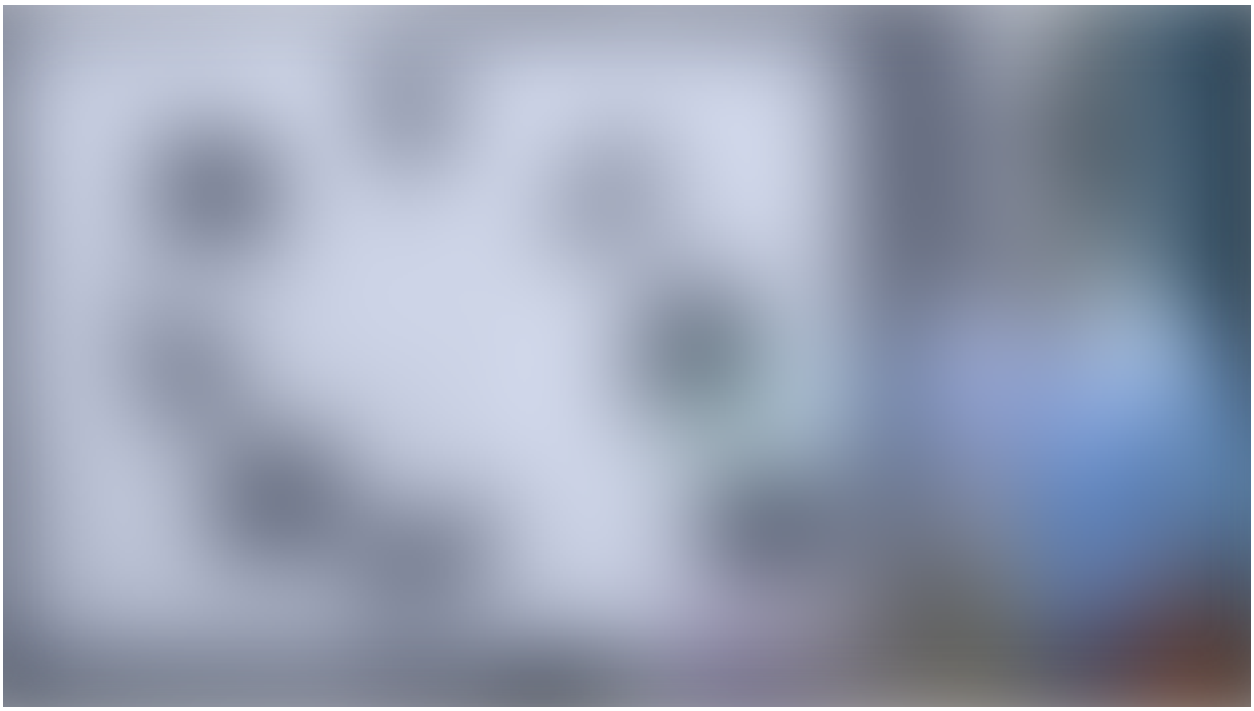
```
$ whois 8.8.8.8
...

NetRange:       8.0.0.0 - 8.127.255.255
CIDR:           8.0.0.0/9
NetName:        LVLT-ORG-8-8
NetHandle:      NET-8-0-0-0-1
Parent:         NET8 (NET-8-0-0-0-0)
NetType:        Direct Allocation
OriginAS:
Organization:   Level 3 Parent, LLC (LPL-141)
RegDate:        1992-12-01
```

```
Updated:        2018-04-23
Ref:            https://rdap.arin.net/registry/ip/8.0.0.0

OrgName:        Level 3 Parent, LLC
OrgId:          LPL-141
Address:        100 CenturyLink Drive
City:           Monroe
StateProv:      LA
PostalCode:     71203
Country:        US
RegDate:        2018-02-06
Updated:        2018-02-22
...

NetRange:       8.8.8.0 - 8.8.8.255
CIDR:           8.8.8.0/24
NetName:        LVLT-GOGL-8-8-8
NetHandle:      NET-8-8-8-0-1
Parent:         LVLT-ORG-8-8 (NET-8-0-0-0-1)
NetType:        Reallocated
OriginAS:
Organization:   Google LLC (GOGL)
RegDate:        2014-03-14
Updated:        2014-03-14
Ref:            https://rdap.arin.net/registry/ip/8.8.8.0

...
```

- **Port scanners (Nmap, masscan, …):** Open ports indicate what *services* a server exposes. Port 443 is usually HTTPS; port 22 is usually SSH and so on. Port scanners will automate the process of identifying all the services a given IP/host has open, will potentially reveal the software being used, its version and can sometimes also identify the operating system of the host.
- **Google search syntax:** Many books cover this extensively (even one fully dedicated to the topic) so I won't even attempt to

consolidate it into a paragraph, but check out the Google Hacking Database (GHDB) and you'll get the idea.

- **Python:** Eventually you'll want to do something unique which brings together different tools and APIs for your specific use case. I've specifically mentioned Python because it's an approachable language, usually easy to read and the standard library is rich with functionality. And if the standard library doesn't have what you need you can bet there is already a module out there for it. Most OSINT tools I've come across are written in Python for this reason so if you're new to Python and are looking to apply it specifically for OSINT, take a look at the online courses Justin Seitz (the author of Hunchly) has put together, or simply the Python Tutorial — it's surprisingly readable.
- **Think creatively!** One of the things I love about OSINT is that it's often a big puzzle of loosely connected (or often disconnected) pieces. You get one piece of information which leads to another and then another. You hit a dead end, have to pivot and meanwhile the information you are gathering along the way is building up a more coherent picture. The key is to not give up, think creatively and be resourceful. Hopefully at the end what you have is this:



Sorry, I just had to sneak in a Brooklyn 99 reference here somewhere.

With these essentials covered, you are ready to stand on the shoulders of giants by utilising more sophisticated tools and platforms others have created, many of which build upon the basics above but at a much larger scale.

. . .

Internet Scanners

The great thing about Internet Scanners is that when you want to find out what services your target has exposed (open ports, protocols, applications, content), they do most of the heavy lifting for you so you just need to query their database instead of performing port

scans yourself. The other major benefit: it's completely passive so your target won't know anything about your search.

I performed a simple search of 'spiderfoot' in SHODAN which reveals some people running it openly exposed on the Internet. 👀

- SHODAN: Probably the undisputed king of Internet scanners, provides a rich query language, API and most importantly a ton of data to sift through.
- Censys: This platform is growing fast; high quality data and a nice interface and API to work with.
- BinaryEdge: These guys have uncovered a number of leaks through their platform and have recently opened up free access to

the public. They also include BitTorrent data and have an API.

Passive DNS

Passive DNS services tap into Internet DNS traffic to build up a history of DNS resolutions. We know that with DNS we can resolve a name to an IP, or an IP to a name. But what if we want to find all the names that resolve to a given IP? Or see what IPs a name has historically resolved to? Passive DNS solves this problem. No single service will ever have a complete picture and the freshness of data will vary, so it's often best to query multiple Passive DNS services to get the most comprehensive result.

Searching for my personal website binarypool.com on SecurityTrails reveals DNS records going back to 2008!

- SecurityTrails: Putting SecurityTrails under the Passive DNS

category is pigeon-holing a bit as they do much more than passive DNS. They have a ton of data, a rich API and a nice UI. They are also growing fast as they build out their platform and provide a free usage tier.

- Robtex: Robtex was my first introduction to passive DNS and got me hooked. It's used by pretty much all OSINT tools because it's been around so long, is free and has a lot of high quality data going back years. Don't be fooled by the plain UI — it's rich with data, and they provide an API.
- HackerTarget: Another rich source of free passive DNS data and available through an API. Worth also checking out all the other tools they freely provide.

Reputation Systems

If one of your main goals with OSINT is threat intelligence, you're in luck because the number and quality of sources is huge, so I've cheated a bit here and listed my two favourites plus a link to a service which aggregates all of them and compares their originality for you.

Looking up a suspicious domain name in VirusTotal reveals reputation information plus much more.

- VirusTotal: Google-owned at Google-scale — a huge platform providing reputation data, passive DNS and more. Access is free but query volumes tightly controlled with throttling. Shouldn't be an issue if you're only using the UI.
- Greynoise: A newcomer focused more on identifying Internet scanners (like those mentioned in the section above). If you're investigating suspicious IPs then this is a great resource to eliminate false positives ("anti-threat intelligence").

- FireHOL IP Lists: As stated on the site, " The objective is to create a blacklist that can be safe enough to be used on all systems, with a firewall, to block access entirely, from and to its listed IPs." They also maintain historic data, analyse for data uniqueness and more.

Reverse Whois

Reverse Whois is one of my favorite OSINT resources because it's so powerful and often yields surprising (and funny) results. While regular Whois only provides searching by the domain name, Reverse Whois resources enable you to search current and historic Whois records by fields such as name, phone number and email address. More concretely, if abc@xyz.com is the contact for the domain you're investigating, you can find all the other domains registered under that e-mail address. This is incredibly useful for when attempting to identify shadow IT issues or discovering the full perimeter of your target beyond the primary domain name(s).

Searching for generalcounsel@trumporg.com reveals other domains registered under that address.

- ViewDNS.info: Many free tools are provided on their website, and pricing for the API is quite reasonable. Without the API you can still query quite a bit from the website and the data goes very far back. You can also download data in bulk.
- WhoXY: Their data covers over 2000 TLDs, is available via an API and in bulk. You can also pay for query volumes in chunks at a rate of $2 per 1000 queries. Not bad at all.

- SecurityTrails: I had to mention SecurityTrails again because yes, they provide reverse Whois data too.

Automation Tooling

With the ever increasing number of OSINT data sources you will eventually need some automation so that you're not spending all your time switching between browser windows, copying and pasting and increasing the chance of human error. Tools will enable you collect, structure, correlate and visualise OSINT data and even monitor for changes over time.

There are a growing number of tools entering the OSINT arena, covering sub-domain enumeration, social media correlation and so on. Sometimes a general broad-coverage tool is good to give you access to a lot of data points at once, and other times you want a very narrow and specific tool for just one data point. One critical theme however is freshness — try and use tools that are actively maintained so that you can be confident they are using the latest APIs of the data sources they integrate with.

SpiderFoot is available as open source, or a cloud-hosted version (SpiderFoot HX, above) which is currently in Private Beta.

- SpiderFoot: Being the author of SpiderFoot, it's impossible for me to be impartial so I'll just say that it provides a web-based UI, CLI, is actively maintained and has over 150 modules for integrating with and analysing OSINT data sources. Also, the SpiderFoot HX Private Beta is still open at the time of writing, which has much more functionality.
- Maltego: Available as Community Edition (free) or commercial. Maltego has impressive visualisation capabilities and works with the model of OSINT "transforms" which transform one type of

data (e.g. e-mail address) to another (e.g. a person's name).

- theHarvester: Very popular open source pure CLI OSINT tool which integrates with a number of data sources including several mentioned in this article.

Community

One of the best things to happen in the OSINT space over the past couple of years is the growth in the community. Blog posts, chat groups, aggregated resource lists and even Podcasts are now available:

A podcast 100% dedicated to OSINT.

- The OSINT Podcast: Jake Creps does a good job of highlighting recent topics, covering new OSINT resources and tools, and interviews of key people in the OSINT space.

- Michael Bazzell's website (and newsletter!) is regular reading for me, and I frequently use it to find ideas for new SpiderFoot modules. It's more on the investigative side and less technical, but still very useful for picking up new sources of OSINT.
- OSINTcurio.us: A blog that went live only recently and one to keep an eye on. Entertaining reads about how some people got into OSINT, with increasingly more instructional material coming.
- Awesome OSINT: The motherload of links to OSINT resources. It probably has every OSINT source known to man and is hosted on Github so if you want to contribute, it's just a pull request away.
- OSINT Rocket Chat: With hundreds (thousands?) of members and growing, you'll find a helpful and active community of investigators, InfoSec people, researchers and hobbyists.

. . .

Conclusion

As is typical with any post of this nature, I've really just scratched the surface of OSINT resources available today. Nonetheless I hope you found it useful and in future posts I'll cover some other aspects of the growing world of OSINT. If you want to see some of the resources above (and others) in action you can check out my previous post about analysing a Bitcoin scam.