# Automating OSINT, OPSEC, and Dark Web Research

## Putting Information in Everyone's Hands

## Introduction

Digital presence is how we appear online as individuals, organizations, products, and services motivated by fulfilling requirements or managing relationships for personal or professional reasons. Its landscape spans the open web, deep web, dark web, information repositories, archives, directories, social networks, applications, cloud infrastructure, domains, subdomains, news stories, and more.

Initiated by internal or external forces, digital presence can be intentional, unintentional, within control, and out of control. It facilitates almost every aspect of our personal lives, commerce, and government. A presence that is beyond the perimeter of a firewall unearthing a whole new realm of possible threats to individuals, businesses, organizations, governments, and nations.

OSINT (Open Source Intelligence) tools do an excellent job of aggregating and presenting available data to the public, but perform many separate functions. With so many tools, it can be a challenge to manage and maintain platforms. Some solutions effectively consolidate all of this disparate information, but they are not customizable and require resources to operate, manage, and support.

OSINT and other tools specializing in attack surface mapping fuel OPSEC (Operations Security) to help organizations defend themselves. It is a manual process to execute these tools, and sifting through their results takes a very long time. The time spent analyzing the results can be better devoted to investigating something else, like improving a process or tuning a rule. Due to the time it takes to effectively identify sources, risks, and threat vectors, comprehensive OSINT and OPSEC methods often get pushed to the backlog of a security team's workload.

The Dark Web has recently become a key source of intelligence, and several vendors now exist that specialize in Dark Web Analytics and Intelligence. The Dark Web is not as mysterious as some news outlets may portray though it is not without its risks. Accessing the Dark Web requires the use of tools like TOR or I2P. TOR works through bouncing your connection off of several random computers, called TOR Nodes. Anyone can set up a TOR Node. The threat here is that as your connection is bouncing off of these “TOR Nodes”, you are susceptible to being watched (traffic sniffing) or malware infection.

Other risks from the Dark Web are Distributed Denial of Service (DDoS) attacks from TOR relay traffic and bandwidth consumption, employees actively bypassing security policies, reputation damage, and blacklisting from operating a TOR Node. These risks often cause companies to forgo using the Dark Web as a resource. Despite these risks, the Dark Web contains information that does not exist anywhere else and should be used as a source to inform OPSEC. New variants of malware, tactics, and techniques, in addition to sensitive personal information about individuals and companies, can be found on the Dark Web. Having this knowledge is highly valuable in protecting any organization.

This whitepaper will dig into automating OSINT, OPSEC, and Dark Web searching; and how this 'Big Three' combination provides instant value to an organization's security posture.

# What is OSINT

OSINT stands for Open Source Intelligence which is the strategic use and analysis of publicly available resources. OSINT is not a new concept. It has been around for a very long time and has continued to evolve with the technological age. A modern-day example is posing a question to a community forum or looking up anything on the Internet. OSINT is not overly complex; it is used on a daily basis by many people without really knowing it.

OSINT can facilitate research, competitive intelligence, general information, or more specifically to this paper, IT Security. Security professionals are constantly using OSINT and OSINT tools to perform their day-to-day tasks and improve their companies' overall security posture. These tasks include validating company public-facing assets, certificates, unintended information disclosure, cloud security, and staying up to date with industry threats and trends.

As more data becomes increasingly available online, OSINT becomes invaluable to any organization for securing their assets and beyond. Public data, or open data, is data that is available to anyone. Private data requires payment and is often more verbose and easier to ingest into another platform, like a Security Information and Event Management (SIEM). This data is often in the form of IPs, Open Ports, Certificates, Domain Names, and Subdomains. Many public and private datasets exist. These data sets are generally used on an as-needed basis and often require manual effort to sift through to what matters to a company. Continuous analysis of this information is required but is often not the case due to lack of personnel or organizational knowledge to where the data exists.

# Why is OSINT Arduous?

As aforementioned, OSINT is a fundamental piece of every company. Whether they know it or not, OSINT methods/techniques are employed. OSINT tools are at times used in various ways addressing several different use cases. The use case highlighted and expanded upon in this paper will be for IT Security.

Security teams have an unbelievably important task assigned to them: to protect the organization's assets. A challenging task with multiple functional areas of Security ranging from Infrastructure, Applications, Operations, and beyond. It is not easy to staff and adequately budget for these positions usually resulting in tasks getting dropped or bumped down in priority.

One of the priorities that always takes precedence is perimeter security: Build up the walls so no one can get in! What happens when those walls have cracks? What happens when someone on the inside forgets to lock the gate? These are all issues that OSINT can help find!

Difficulties of OSINT:

1. Lack of staff
2. Lack of OSINT Tools knowledge
3. Lack of infrastructure to execute and calibrate OSINT tools to achieve results that matter

These are issues that every team (in Security or not) can relate to somehow. But just because they are complex does not mean they are impossible.

# What is OPSEC?

OPSEC stands for Operational Security and is the practice of taking an attacker or adversarial perspective on everything. This mindset helps address commonly overlooked issues and allows you to put on your own 'Black Hat'.

Operational Security is broken down into 5 Key Concepts:

1. Know the data (aka YOU or YOUR ORGANIZATION)
2. Know the threats (What/Who is out there? Why?)
3. Know the issues (What is wrong? What looks off)
4. Know the risk (What's your risk appetite?)
5. Know the strategy (If X happens, we have Y)

Knowing the data is knowing what is out there about you or your organization. This data can range from intellectual property to IP addresses to social media. Identifying this data and where it lives can help address the other Concepts of OPSEC. This information is often found using OSINT.

Knowing the threats can be broken down into understanding the kind of data that is publicly available and if it can be leveraged or exploited against your organization. The following is an example chain of events: (1) Identification of a Drupal server susceptible to a known vulnerability; (2) Exploitation of the known vulnerability; (3) Remote execution is achieved. If you do not know your data or what exists in the outside world, you will not know where to patch or what to patch.

Knowing the issues is an extension of understanding the data and threats. Your Security and IT team may understand how your infrastructure is internally configured, but if the outside world sees a different configuration, this is an issue that needs to be addressed. This issue can either be a misconfiguration, an insider threat, a shadow IT problem, or something malicious.

Knowing the risk is a result of obtaining all of the required information from the data, threats, and issues. Balancing an organization's risk appetite with the existing threats is key in the development and maintenance of security and privacy.

Knowing the strategy means that everyone on the security team is aware of the company mission and roles in achieving overall objectives all the while adhering to best practices in Confidentiality, Integrity, and Availability.

All of the above come together to form OPSEC. Organizations generally possess the concept but come up short in discipline or tools to achieve measurable and manageable visibility. OSINT can help address OPSEC challenges by presenting information that an attacker may possess and use against an organization.

## OSINT + OPSEC

The combination of OSINT and OPSEC facilitates a continuous view of an organization from a potential attacker's perspective and confirms the configuration of public-facing assets. Though this combination benefits of providing an outside-in view bring to light public-facing threats, getting this combination operational can be tricky.

As previously mentioned, OSINT tools are excellent sources of information, but there are so many of them. It is difficult to weed out what is relevant or applicable. Some tools can help by automating OSINT to fuel OSPEC. The Harvester, Recon-NG, and Spiderfoot are a selection of popular tools that have many available modules. However, they lack customization, require dedicated infrastructure, and require knowledge/time to execute/maintain. They also lack dedicated support for searching the Dark Web.

## The Dark Web

There are three 'versions' of the Internet that exist today; the clear or surface web, the deep web, and the dark web. The clear or surface web is the most popular among internet users as it is all the pages that are indexed and searchable. Popular services that offer this are Google, Bing, Wikipedia, or Yahoo. The deep web refers to the pages that are not indexed by search engines and often contain academic information, legal records, scientific reports, government databases, and libraries. The dark web refers to the pages that are not indexed by the search engines, are only accessible through specific browsers, and are often home to illicit activities.

The dark web is not as mysterious and dubious as the media portrays it, but it is not without risk. Connecting to the dark web involves using a TOR or I2P and then navigating on your own to the sites. The dark web is not indexed and ranked like the 'clear web.' For example, if you type in a website on the clear web using Google, Google aggregates results for you and displays them based on an algorithm. When you are on the dark web, you can still use Google, but your results may or may not include ".onion" URLs. Onion URLs (.onion URLS) are the URLs that identify dark websites. An example of this is the newly launched BBC mirror on the Dark Web, "www.bbcnewsv2vjtpsuy.onion." A person can only access these URLs via TOR or I2P. **Since the Dark Web is not indexed or easily searchable, you need to know where to browse to find what you are looking for; this lack of indexing helps mask illicit activities**.

## Traditional Dark Web Challenges

Using the Dark Web as a resource can be difficult as there are risks in connecting, navigating, and staying on top of what is available.

Connecting to the Dark Web is simple. Download TOR, open TOR, find an onion link (or several), browse, and repeat. It would be best to be on a VPN when connecting, though it is not required. In theory, these are all things that anyone can accomplish, but most organizations will shy away from this exercise as it can lead to exposure. The reason for this is that organizations do not own the routes/hops they are taking.

Navigating the Dark Web, in theory, is like browsing the Internet as you do today, but with additional layers of complexity. The URLs to browse often make no sense and contain numbers, which is by design for anonymity, plus there is no search engine to assist with your browsing. Some search engines such as Torch can help, but seeing as there are greater than 1 SEPTILLION V2 URLs and infinity more V3 URLs, using one of these services to find anything, in particular, is challenging. Resources like Reddit and the Hidden Wiki provide some links to .onion URLs, which are good places to start. However, once you find some browsable resources, you may be required to "Sign Up" to view the content of a forum or a site. You may also find things that are 'disturbing', and should report them or stay away if noted on a forum.

Staying on top of .onion URLs is also a considerable challenge. Sites often go down and reappear under a different moniker. Some tools can scrape links and Reddit for user feedback, but it could be a while before something you like resurfaces.

## OSINT + Dark Web + OPSEC = Improved Security Posture

Taking OPSEC principles into account, the Dark Web is a great place to think like an adversary as that is their preferred place of disseminating information, stolen goods, and tactics. Combined with general OSINT tools and techniques, you get a view of how an organization looks from the outside; this view can also apply to OPSEC. Together, the 'Big Three' can drastically shape an organization's security posture as it empowers them with ample information to do so, all the while providing a good place to start.

To further explain, if your organization is seeking to understand how they look from outside eyes, the first place to start would be OSINT tools (i.e., subdomains and social media). You should then search the Dark Web with that information through whatever means available to uncover mentions of any or all the following:

- Your Organization
- Vulnerable Component(s) Used within Your Organization
- An Executive or VIP
- Sensitive or Proprietary Information

Aggregating this information together with the mindset of 'What Do My Attacker's See' empowers your organization with insight into vulnerabilities or other possibilities of attack.

## Conclusion

In conclusion, the value gained from incorporating OSINT, OPSEC, and the Dark Web into your organization should significantly improve your security posture and visibility. OSINT, OPSEC, and the Dark Web separately provide great value, as every company has different use cases, budgets, visibility. But combined, they indeed are the 'Big Three.' ThreatNG was born out of this need to put OSINT, OPSEC, and the Dark Web data into everyone's hands providing an accurate outside-in view to manage risks beyond the perimeter.

ThreatNG uses the OSINT to analyze an organization's external-facing footprint, providing a starting point for OPSEC. An external footprint consisting of technical data to social media, employee and user sentiment, and more. This information can help Security and other teams within an organization such as Human Resources, Finance, IT, etc. Dark Web exploration and investigation is also provided within the platform with all the data and none of the risks. A complete "Big Three" solution that provides a holistic outside-in view, facilitates OPSEC configurations/management, improves security posture, and empowering everyone with all the information making "security everyone's responsibility".

---

ThreatNG is redefining digital risk protection and external attack surface management with a platform of unmatched breadth, depth, and capabilities in managing technical and business threats across the dark, deep, and open web. Living up to the company mantra ("Security Centric; Not Exclusive"), ThreatNG provides a configurable solution to target, discover, and assess digital assets across a definable ecosystem of organizations, subsidiaries, partners, third parties, supply chain, and customers. Bolstered and maintained by the open source intelligence (OSINT) experts at DarcSight Labs, ThreatNG empowers organizations of all types and sizes to uncover, understand, and manage their external digital threats.