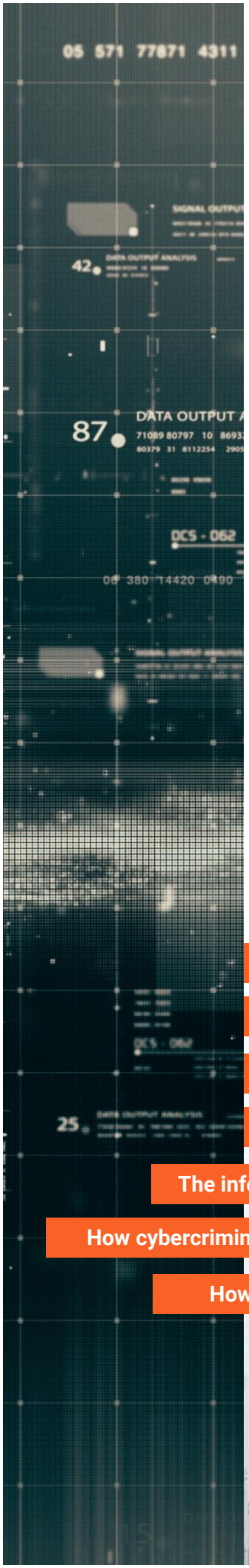
PICNIC™

OSINT E-BOOK

ESSENTIAL GUIDE TO OPEN-SOURCE INTELLIGENCE (OSINT)



…92% of all cyber attacks are specifically crafted from users' OSINT

Exploitation through publicly available information is the single largest threat to companies and their people today.

Known as Open-Source Intelligence, or OSINT, this public data reveals to hackers how they can compromise human targets via social engineering attacks and defeat the most powerful technical solutions.

The bad news for organizations is that the internet makes it easy for attackers to find information about them and their employees to craft convincing attacks.

The good news is that enterprise security teams can also use OSINT for defensive purposes in order to level the playing field and prevent attacks. With companies recognizing the important role this data plays, the global demand for OSINT tools is on the rise, with research predicting a market growth rate of 28.33% between 2022 and 2030. Fortunately, companies can now automatically harness OSINT like never before to protect their people and their assets.

We've created this e-book to explain OSINT, how it's used, and how security professionals can use Picnic's powerful new technology to take the advantage away from threat actors.

WHAT YOU'LL LEARN

1. What OSINT is
2. The history of OSINT
3. How people collect OSINT
4. The most-used OSINT tools
5. The information people can find with OSINT
6. How cybercriminals use OSINT for social engineering
7. How cybersecurity teams can use OSINT

1

WHAT IS OPEN-SOURCE INTELLIGENCE (OSINT)?

Open-Source Intelligence (OSINT) is information available through public data sources that someone can collect and analyze.

People can engage in OSINT gathering legally using tools that find data on:

- the "surface web," including search engines, blogs, and job postings
- social media
- databases containing public records

Additionally, malicious actors often use specialized intelligence tools and search engines for finding information on the dark web.

2

WHAT IS THE HISTORY OF OSINT?

Gathering OSINT is not a new phenomenon. However, the information available and the search processes have changed, especially as more people share data on the internet.

During World War II, the Office of Strategic Services established the first Research and Analysis Branch dedicated to collecting OSINT and using it for the war effort. Since then, global military and intelligence services have used publicly available data for their operations.

In the late 1980s, the US military first used the term OSINT, noting its tactical battlefield value. During the 1990s, OSINT became even more important to the US intelligence community, with the 1992 Intelligence Reorganization Act incorporating public information as valuable and the 1994 establishment of the Community Open-Source Program Office (COSPO) within the CIA.

As the internet became more accessible, so did OSINT. From websites with public government data and social media networks, almost anyone can search publicly available data legally and ethically.

Outside the confines of legality and ethics, threat actors use sophisticated tactics to gather data. For criminals, the definition of "public" also includes the dark web where malicious actors share stolen, otherwise-nonpublic personal information like credit card numbers, passwords, and social security numbers.

HOW DO PEOPLE COLLECT OSINT?

Since OSINT focuses on publicly available information, people can find it using paid and unpaid search methods. Further, their processes can be as simple as a Google search or as complex as creating a specialized tool.

Surface Web

The surface web is the internet that most people use. It's easy for the general public to search using standard search engines.

Search Engines

When people want to find information, they usually start with generally available search engines. Most people are familiar with how these work. Google's search engine has become synonymous with looking up facts and data.

Some typical search engines include:

- Google
- Bing
- Yahoo!
- DuckDuckGo
- Startpage

Blogs

Blogs are regularly updated websites or web pages that people and organizations use to inform readers. An organization's blog might try to educate readers about topics related to its products or services. A personal blog often shares stories about someone's interests, like hobbies, books, music, television shows, or movies.

Job Postings

Most companies list job postings on their websites so that interested applicants can find them. Since companies use job postings to attract candidates, researchers can use them to:

- Locate corporate offices
- Find Human Resources contacts

Social Media

People and companies increasingly use social media. Many companies have social media marketing strategies that they use to make important announcements, like when they hire a new senior executive or acquire a new company. Similarly, people often share personal stories and information on social media sites.

For example, LinkedIn enables organizations to create digital business networks. However, since the company shares this information publicly, it becomes an OSINT source. As a career-focused social media site, people may be more "trusting" and open to connecting with others.

Some examples of OSINT gathering on LinkedIn include searching by company name for job roles like:

- Chief executive officer
- Chief financial officer
- Account executive

Someone could do a search for account executives at an organization, look at their connections, and then find a senior leadership team member's information.

Data Brokers/People Search Engines

Data brokers collect and sell personal or corporate data. While they often use public records to aggregate this information, they can also source it privately. As a paid service, they collect data from multiple locations that can include:

- Census records
- Electoral rolls
- Social media
- Court reports
- Purchasing history

Some examples of data brokers and people search engines include:

- PeopleFinderFree
- Truthfinder
- Spokeo
- US Search
- Whitepages

Custom Search Engines

More technical researchers can build custom search engines. With a custom search engine, a researcher can collect OSINT across multiple social media websites or filter searches by file type.

For example, the Google Programmable Search Engine is a platform enabling web developers to use Google search capabilities on their websites. However, researchers can use this functionality to search across specific websites and take multiple actions. When engaging in OSINT, researchers might create a custom search engine that enables a simultaneous search across various social networks that can isolate each network's results in their own tab. This streamlines their process, giving them a way to use the collected data more effectively and efficiently.

Specialized Search Engines

Specialized search engines enable researchers to expand their data collection. These provide search options and capabilities that typical search engines lack.

Some examples of specialized search engines include:

- Wayback Machine: cached website data providing historical information
- Searx.me: ability to export results and enabling researcher anonymity
- Exalead: unstructured data to find documents and audio files, including papers or webinars

Caller ID Databases

Caller ID databases enable people to do reverse lookups on phone numbers. While these traditionally only worked for landlines, more databases now provide services for cellular phones.

When researchers input a known telephone number, they can retrieve data like:

- Country
- Name
- Carrier name
- Carrier type

Third-Party Data Breaches

Whether researching legally or illegally, people can find public databases containing information about compromised email addresses and the passwords associated with them.

For example, cybercriminals often post this information on websites like Pastebin. Further, in response to increased data breaches, ethical services now exist, including:

- Have I Been Pwned
- Spycloud
- Scylla
- Leaked Source
- Ghost Project
- PSBDMP

While researchers need an email address to use these services, they provide valuable information by:

- Confirming that an email address is valid
- Providing insight into the breach that compromised the email

Since cybercriminals are not held to legal and ethical research requirements, they often download databases of publicly available and stolen databases, then run the data through analytics tools.

If they find a username and password for one service, like LinkedIn, they can try those credentials to gain access to a corporate environment.

Custom Tools

Gathering OSINT information from all these diverse locations manually isn't efficient. Often, researchers create or leverage custom tools. With these tools, they can more rapidly search across all potential locations and search engines.

Dark Web

What people call the dark web is really internet traffic directed through the Tor network that conceals users' location and network usage. This anonymity makes it more difficult to trace activity back to the user, including websites hosted on the network. Criminal activity thrives on the Tor network because the sites are not hosted on publicly viewable networks.

WHAT ARE THE MOST-USED OSINT TOOLS?

While threat actors may build their own tools, many ethical researchers leverage pre-existing research tools. Below are some of the OSINT tools often used to uncover publicly available data about people and technologies.

Maltego

Focused on discovering relationships, this gathers data like:

- Names
- Email addresses
- Aliases
- Companies
- Websites
- Document owners
- Affiliations

It uses several common public information sources, including:

- DNS records
- Whois records
- Search engines
- Social networks

Then, it provides charts and graphs that uncover the connections between the data points.

Mitaka

Mitaka enables people to research using their web browsers. With the ability to search across more than seventy search engines, it returns information like:

- IP addresses
- Domains
- URLs
- Hashes
- ASNs
- Bitcoin wallet addresses
- Indicators of Compromise (IoCs)

Spiderfoot

A free tool, Spiderfoot is an application that red teams often use during their reconnaissance activities. Some information that it returns includes:

- IP addresses
- CIDR ranges
- Domains and subdomains
- ASNs
- Email addresses
- Phone numbers
- Names and usernames
- Bitcoin addresses

Spyse

Focused on detecting internet assets, Spyse collects and analyzes publicly available data about:

- Websites
- Website owners
- Servers associated with websites
- Internet of Things (IoT) devices

BuiltWith

BuiltWith provides information about a website's technology stack and platform. For example, it generates information that includes:

- Content management system (CMS), like WordPress, Joomla, or Drupal
- Javascript/CSS libraries, like jQuery or Bootstrap
- Plugin installed
- Frameworks
- Server information
- Analytics and tracking information

Intelligence X

As an archival service and search engine, Intelligence X enables researchers to obtain historical versions of webpages and leaked data sets, including controversial content.

Some examples of the data that Intelligence X retains include:

- Lists of compromised VPN passwords exposed on cybercriminal forums
- Indexed data collected from political figures' email servers
- Information from social media site data leaks

Ahmia

Ahmia enables dark web research by making Tor results visible without requiring users to install the browser. However, to open links and results, researchers still need to install the Tor browser to open links and results.

DarkSearch.io

As of January 2022, this service is available only to organizations who request private access. The platform allows researchers to run automated searches of the dark web without requiring them to use .onion versions or install the Tor browser.

Grep.app

Grep.app focuses on git repositories, providing a single search across:

- GitHub
- GitLab
- BitBucket

People use it when searching for code strings associated with:

- IoCs
- Vulnerable code
- Malware

Recon-NG

Recon-NG is a Python-based tool that enables researchers to automate redundant, manual tasks. It offers:

- Independent modules
- Database interaction
- Built-in functions for convenience
- Interactive help
- Command completion

Creepy

Another Python-based technology, Creepy is a geolocation OSINT tool that collects data from various online sources, including social media and image hosting sites. Users can

- Create maps
- Filter searches based on exact location and/or date
- Export data

theHarvester

With theHarvester, users can search for:

- Emails
- Subdomains
- IP addresses
- URLs

It offers both passive search and active DNS brute-forcing capabilities.

Shodan

Shodan is a search engine that both security teams and threat actors use to discover internet-connected devices and services.

The Shodan suite of products includes:

- Search engine
- Monitor to track devices
- Maps
- Collection of screenshots
- Collected historical data

TinEye

TinEye is a reverse image search tool that allows researchers to upload images or use URLs. With reverse image lookup, someone can find where a picture was taken so that they can find a physical location.

Metagoofil

With Metagoofil, researchers can scan a domain's documents and uncover the metadata. The tool provides information about files like:

- PDFs
- Word Documents
- Excel Spreadsheets
- PowerPoint Presentations

The metadata, or "data about data", can include information such as:

- User names
- Email addresses
- Printers
- Software

5

WHAT INFORMATION CAN PEOPLE FIND WITH OSINT?

While all OSINT information is publicly available, most people may not realize what is out there about them and how someone can find it. Even people who think they have a limited digital footprint would be surprised at what OSINT researchers can uncover.

Email Addresses

Today, most people have at least one personal and one professional email address. According to research, 90% of Americans have an email address, averaging 1.75 email addresses each. Typically, people use their email addresses to:

- Log into social media
- Access work resources
- Use ecommerce applications
- Register for media, like news, professional publication, and streaming services

Usernames

To maintain consistency, many people use the same username across different online services. For example, someone with an email jdoe@totallyfakemailservice.com might also use jdoe as a social media handle. Further, these are typically the same types of usernames that corporations use for generating user IDs.

With this information, cybercriminals can try to connect known usernames to compromised passwords as a part of credential-based attacks.

Addresses

Personal and professional addresses are easily discoverable. On its own, an address may not impact cybersecurity.

However, when aggregated with a name or IP address, ethical and criminal actors can use the information to build a relationship with a target.

Phone numbers

When researchers collect and aggregate OSINT, phone numbers become even more valuable. When connecting a person's name and phone number, someone can spoof, or create a fake version of, that phone number as part of an attack. For example, when a smishing attack sends a text message that appears to come from a trusted contact, the target is more likely to take the action that the attacker requests.

IP Addresses

When someone obtains an IP address, it gives them the ability to do a reverse lookup that gives them a lot of information about the server hosting a domain, including:

- City
- State
- Zip code
- Open ports

HOW DO CYBERCRIMINALS USE OSINT FOR SOCIAL ENGINEERING?

The first step to a successful social engineering attack is to gain a target's trust or buy-in. People may be skeptical enough to ignore an email from a Nigerian prince, but they're far less likely to ignore an email from their boss or human resources department.

Cybercriminals leverage OSINT so that they can build their attacks around information that will prompt someone to take an action that's against their best interests. Further, cybercriminals collect and correlate various data types so that they can build out robust attacks. They rarely just use one type of data, like an email address.

Email Attacks

Phishing, spear phishing, and whaling are all typically email-based social engineering attacks. However, they use OSINT in subtly different ways.

Phishing

With a phishing attack, cybercriminals send out high volumes of fake emails, pretending to come from a legitimate entity. In this case, they really only need the email domain of the entity they want to impersonate.

For example, in a sophisticated attack targeting Office 365 credentials, cybercriminals imitated the domain for the US Department of Labor. They created domains like dol-gov.com, using a legitimate dol.gov domain for replies.

The emails sent fake bidding instructions with a PDF that redirected the target to a phishing site where the criminals collected credentials.

Spear Phishing

With a spear phishing attack, cybercriminals might start by doing a LinkedIn search to find someone new to an organization in a high-visibility position, like a Chief Executive Officer (CEO). Once the cybercriminals have this information, they can search LinkedIn for people who will work directly with the new CEO.

They find the organization's domain and make a fake, or spoofed, version of it. For example, fakcompany.com would be fakecompany.io. With this fake domain, they create a form that hides the ".io" so that it looks like it's from the organization's legitimate domain.

Building on this, they can then find examples of past statements that the new hire made for the email's text. They email the form to the targets that they found on LinkedIn, requiring them to supply login credentials when they complete it.

Between 2013 and 2015, cybercriminals used a spear phishing attack to steal $100 million from Google and Facebook. In this case, they created a fake computer manufacturing company, then sent invoices to targeted employees under the guise of being the legitimate services provider. Instead of paying the real provider, the companies directed the deposits to the cybercriminals' bank accounts.

Vishing

Also called "phone phishing" or "voice phishing" attacks, cybercriminals call their targets to deploy the attack. During a vishing attack, cybercriminals will often incorporate pretexting, creating a situation that lures the target into taking action.

Many cybersecurity awareness training modules include pretexting scenarios where someone calls a new employee, pretending to be from human resources. For this attack to work, cybercriminals need to do their OSINT research.

For large organizations that might have upwards of 100 global new hires per week, this scenario provides cyber attackers a significant return on investment. To be successful, attackers need a few different types of OSINT data. First, they need to find people on LinkedIn who recently announced that they joined an organization. Next, they need to find the VOIP data for the organization's phone system so that they can spoof it. Then, they create a fake HR portal that sends data directly to them. They call the new employees, telling them that to get paid they need to confirm payment data by clicking on a link that they're sending while on the phone. When the targets enter their credentials, the cybercriminals collect it.

In 2020, attackers compromised 130 Twitter accounts with a vishing attack. Twitter classified this as a phone spear phishing attack, saying that cybercriminals called employees and tricked them into revealing account credentials.

7

HOW OSINT ENABLES CYBERSECURITY TEAMS

The good news for organizations is that their security teams can also use OSINT. The information itself is benign. The danger or benefit comes from how someone uses it.

When organizations use OSINT to protect themselves, they can follow the same processes as threat actors. When security teams have access to the same publicly available information that malicious actors have, they can mitigate risk by reducing their digital footprint or implementing additional security controls.

Discover Public-facing Assets

Most security teams leverage OSINT to detect assets connected to the public internet. For example, many security teams use Shodan to detect IoT devices so that they can implement controls or protections.

Locate Information Outside Organization Boundaries

Sometimes, employees share information on social media without realizing that a little personal information can lead to an attack that leads to a breach.

For example, an employee might list their telephone number on LinkedIn. With this information, skilled attackers can implement a successful vishing or smishing attack that could compromise both the personal and corporate accounts of the employee.

When security teams have visibility into this risk, they can implement preventative measures that reduce risk, in this case working with the employee to remove the phone number before it can be leveraged in a social engineering attack.

Identify External Threats

When security teams have OSINT tools, they can monitor dark web forums for stolen credentials that compromise the organization's security.

According to research, 70% of users tied to breach exposures from 2021 or earlier were still reusing the exposed credentials. Further, more than two out of three people use the same passwords across multiple accounts, meaning a compromised personal password could impact someone's professional login credentials.

Security teams that can find and link employee personal and professional leaked credentials can use this information to make sure these credentials are no longer being used.

Enhance Penetration Tests

Penetration tests look for weaknesses in an organization's security program. As part of this process, penetration testers start with the reconnaissance phase to map out the attack surface of the target.

This involves running OSINT, looking for accidental sensitive information leaks across social media, data brokers, and other publicly available data locations. Then they leverage this information to aid their ethical social engineering attacks.

With regular OSINT monitoring, security teams can reduce the number of findings by proactively identifying and mitigating these risks.

Design Adversary Emulations

When security teams engage in adversary emulations, they follow threat actor tactics, techniques, and procedures (TTPs) to test their defensive controls.

For example, when security teams want to emulate a remote desktop protocol attack they need to follow the same steps that attackers do. Many security teams focus on the steps that attackers take once they gain access to systems because they lack the OSINT visibility to emulate attackers' social engineering and credential theft capabilities.

When security teams can effectively obtain publicly available data, like information employees post on social media, they can create more realistic emulations. By identifying employees that attackers might target, they can implement controls that proactively address these risks.

For large organizations that might have upwards of 100 global new hires per week, this scenario provides cyber attackers a significant return on investment. To be successful, attackers need a few different types of OSINT data. First, they need to find people on LinkedIn who recently announced that they joined an organization. Next, they need to find the VOIP data for the organization's phone system so that they can spoof it. Then, they create a fake HR portal that sends data directly to them. They call the new employees, telling them that to get paid they need to confirm payment data by clicking on a link that they're sending while on the phone. When the targets enter their credentials, the cybercriminals collect it.

In 2020, attackers compromised 130 Twitter accounts with a vishing attack. Twitter classified this as a phone spear phishing attack, saying that cybercriminals called employees and tricked them into revealing account credentials.

7

Picnic: Automated OSINT Monitoring and Remediation for Enhanced Cybersecurity

Picnic is the first technology platform that allows organizations to fully and automatically harness OSINT for defensive purposes.

The platform provides enterprise security teams with the capability to instantly emulate attacker reconnaissance on the entire OSINT footprint of their organization and its people across the surface web, social media, data brokers, breach repositories, and the deep and dark web. At the same time, Picnic's technology continuously hunts and flags any exposed data and PII that would be of value to threat actors, identifies likely human targets and pathways to compromise, streamlines external data footprint cleansing, and enhances existing security controls to prevent attacks.

Since attackers have OSINT exposure too, Picnic also monitors for suspicious domains and other attacker infrastructure before these can be leveraged against an organization's people.

With these preemptive and continuous capabilities, organizations gain an unprecedented level of visibility and control over their OSINT footprint and can substantially reduce a threat actor's ability to use OSINT successfully against them.

For large organizations that might have upwards Picnic's technology marks a decisive moment in the history of OSINT, as it takes away the asymmetrical advantage threat actors have had until now.

Attackers need OSINT to craft their attacks. The public data vulnerabilities revealed during a cybercriminal's reconnaissance are ultimately what lead to phishing, credential compromise, ransomware, malware, and the like.

Picnic's platform addresses this problem head-on by providing enterprises and their people with the power to automatically know the full extent of their OSINT exposure, proactively remediate their human risk, and preemptively neutralize the pathways to compromise that their public footprint reveals. In this way, they can detect and prevent attacks before they happen on a scale not previously possible.

PICNIC™


SIGNAL OUTPUT ANALYSIS_
SIGNAL OUTPUT ANALYSIS
SIGNAL OUTPUT ANALYSIS_
DCS - 062
DCS - 062
DCS - 062