487.5
The Dark Web, Breach Data, and International Issues

PLEASE READ THE TERMS AND CONDITIONS OF THIS COURSEWARE LICENSE AGREEMENT ("CLA") CAREFULLY BEFORE USING ANY OF THE COURSEWARE ASSOCIATED WITH THE SANS COURSE. THIS IS A LEGAL AND ENFORCEABLE CONTRACT BETWEEN YOU (THE “USER”) AND SANS INSTITUTE FOR THE COURSEWARE. YOU AGREE THAT THIS AGREEMENT IS ENFORCEABLE LIKE ANY WRITTEN NEGOTIATED AGREEMENT SIGNED BY YOU.

With the CLA, SANS Institute hereby grants User a personal, non-exclusive license to use the Courseware subject to the terms of this agreement. Courseware includes all printed materials, including course books and lab workbooks, as well as any digital or other media, virtual machines, and/or data sets distributed by SANS Institute to User for use in the SANS class associated with the Courseware. User agrees that the CLA is the complete and exclusive statement of agreement between SANS Institute and you and that this CLA supersedes any oral or written proposal, agreement or other communication relating to the subject matter of this CLA.

BY ACCEPTING THIS COURSEWARE, YOU AGREE TO BE BOUND BY THE TERMS OF THIS CLA. BY ACCEPTING THIS SOFTWARE, YOU AGREE THAT ANY BREACH OF THE TERMS OF THIS CLA MAY CAUSE IRREPARABLE HARM AND SIGNIFICANT INJURY TO SANS INSTITUTE, AND THAT SANS INSTITUTE MAY ENFORCE THESE PROVISIONS BY INJUNCTION (WITHOUT THE NECESSITY OF POSTING BOND) SPECIFIC PERFORMANCE, OR OTHER EQUITABLE RELIEF.

If you do not agree, you may return the Courseware to SANS Institute for a full refund, if applicable.

User may not copy, reproduce, re-publish, distribute, display, modify or create derivative works based upon all or any portion of the Courseware, in any medium whether printed, electronic or otherwise, for any purpose, without the express prior written consent of SANS Institute. Additionally, User may not sell, rent, lease, trade, or otherwise transfer the Courseware in any way, shape, or form without the express written consent of SANS Institute.

If any provision of this CLA is declared unenforceable in any jurisdiction, then such provision shall be deemed to be severable from this CLA and shall not affect the remainder thereof. An amendment or addendum to this CLA may accompany this Courseware.

SANS acknowledges that any and all software and/or tools, graphics, images, tables, charts or graphs presented in this Courseware are the sole property of their respective trademark/registered/copyright owners, including:

AirDrop, AirPort, AirPort Time Capsule, Apple, Apple Remote Desktop, Apple TV, App Nap, Back to My Mac, Boot Camp, Cocoa, FaceTime, FileVault, Finder, FireWire, FireWire logo, iCal, iChat, iLife, iMac, iMessage, iPad, iPad Air, iPad Mini, iPhone, iPhoto, iPod, iPod classic, iPod shuffle, iPod nano, iPod touch, iTunes, iTunes logo, iWork, Keychain, Keynote, Mac, Mac Logo, MacBook, MacBook Air, MacBook Pro, Macintosh, Mac OS, Mac Pro, Numbers, OS X, Pages, Passbook, Retina, Safari, Siri, Spaces, Spotlight, There’s an app for that, Time Capsule, Time Machine, Touch ID, Xcode, Xserve, App Store, and iCloud are registered trademarks of Apple Inc.

PMP and PMBOK are registered marks of PMI.

SOF-ELK® is a registered trademark of Lewes Technology Consulting, LLC. Used with permission.

SIFT® is a registered trademark of Harbingers, LLC. Used with permission.

Governing Law: This Agreement shall be governed by the laws of the State of Maryland, USA.

SEC487.5 | Open Source Intelligence (OSINT) Gathering and Analysis

The Dark Web, Breach Data, and International Issues

Welcome to SEC487 Open Source Intelligence (OSINT) Gathering and Analysis!

487.5 Table of Contents

This table is a reference for you to quickly move to certain topics in this 487.5 book.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. **The Surface, Deep, and Dark Webs**
2. The Dark Web
3. Freenet
4. I2P - Invisible Internet Project
5. Tor
6. Monitoring and Alerting
7. International Issues
8. Vehicle Searches
9. Putting It All Together

This page intentionally left blank.

Surface, Deep, and Dark Definitions

Three distinct layers of the internet:

- **Surface** - Easy to access
- **Deep** - More challenging
- **Dark** - Special techniques

Most images showing the distinction between them show levels of an ocean[2]

Surface, Deep, and Dark Definitions

Michael Bergman's 2001 white paper "The Deep Web: Surfacing Hidden Value"[1] may be the reason for the nautical references to the surface, deep, and dark webs. He wrote "*Searching on the Internet today can be compared to dragging a net across the surface of the ocean. While a great deal may be caught in the net, there is still a wealth of information that is deep, and therefore, missed.*"[1] His imagery sparked a plethora of sites making their own conceptual pictures of what Bergman stated (shown above). You can view these images in Google (https://sec487.info/73).

Bergman brings up a correct point: much of the internet's resources are outside of the normal world of ecommerce, emails, and social media that most people use on a daily basis. As one moves deeper into the areas of the internet that are less well-known, special knowledge is needed of where these sites are and, sometimes, special software is needed to access them.

Image from https://sec487.info/ou.

References:

[1] https://sec487.info/73, September 4, 2017.

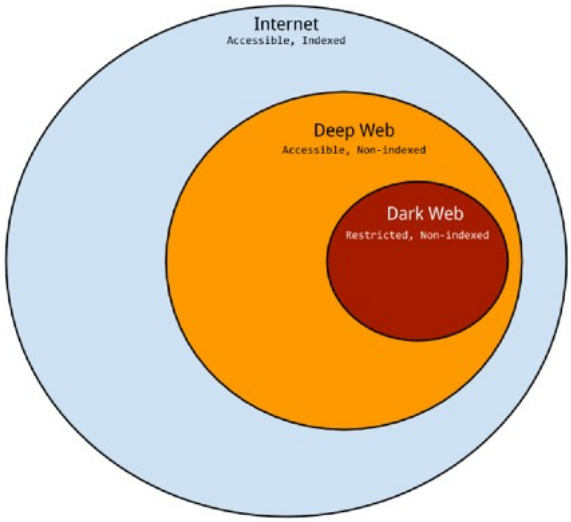
[2] https://sec487.info/74, September 4, 2017.

Levels of the Internet: Simple Version

Daniel Miessler has a blog post[1] that shows the relationships more simply

Dark web is inside the deep web, and deep web is inside the internet

It is all about accessibility: who can access what data from where



Levels of the Internet: Simple Version
Moving past the water references to these levels of the internet, Daniel Miessler (@DanielMiessler) created the above image. As Daniel points out in his blog,[1] these circles are not drawn to size based on usage, number of sites, or user base. They are representational and show that the internet is what all other networks use for the backbone of their communication. We should also mention that there are sites that may fit in multiple layers, but these are the exceptions to the rule.

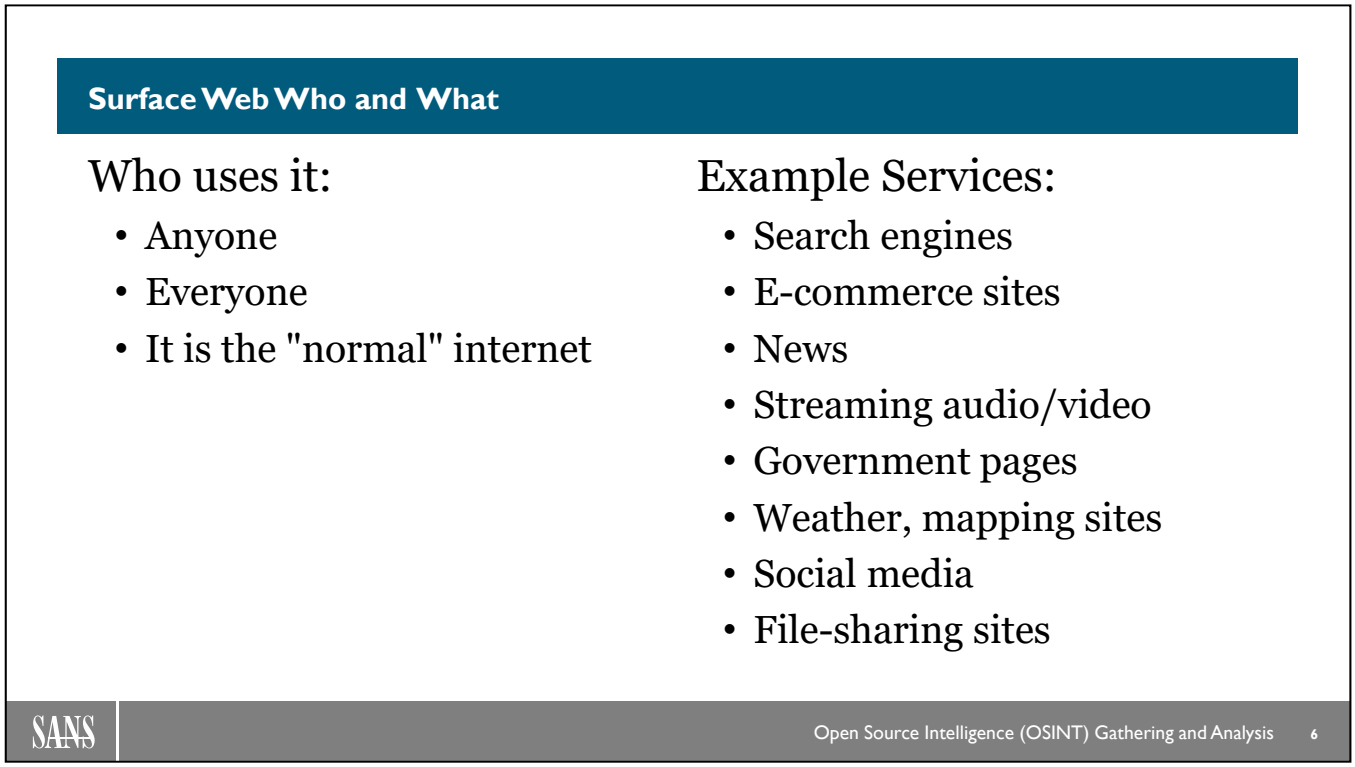
The "surface web" (or "internet" in the above image) is accessible by most users, except where personal, corporate, or national restrictions prohibit access. It is the place where DuckDuckGo, Yahoo, Bing, and Google have web crawlers that scan all the systems they can find, indexing their data for us to search. The surface web mostly shares data with whoever retrieves the web or system resources.

While those search engines scan a good deal of internet resources, they miss many more. This group of "missed services" forms the deep web. Why are these sites "missed" by the search engines? Wikipedia[2] notes many reasons why this data remains outside the reach of the crawlers, but some simple explanations are that these sites may require authentication, or they may have non-web content that web search engines don't collect.

Within the deep, unindexed web, we have the dark web. The resources in this part of the internet require additional knowledge and software to access. While the dark web is usually thought of as a place of evil and criminal activity, there are many other legitimate (and legal) reasons for using it.

Image from https://sec487.info/75, September 4, 2017.

References:
[1] https://sec487.info/75, September 4, 2017.
[2] https://sec487.info/76, September 4, 2017.



Surface Web Who and What

Most of the services that consumers use on the internet are found on the surface web. It contains all the sites and services you can find linked from search engines and social media and other sites. Examples of surface web sites are:

- Search engines - Google, Baidu, Yandex
- E-commerce sites - Alibaba, Amazon
- News - BBC, CNN
- Streaming audio/video - YouTube, Twitch.tv
- Government pages - nist.gov, gov.uk
- Weather, mapping sites - weather.com, Bing Maps
- Social media - Facebook, Twitter, Instagram
- File-sharing sites - Dropbox, Box

When people reference "the internet," they are usually talking about this layer of the surface/deep/dark web stack. It is simple to access and use.

Deep Web Who and What

Who uses it:

- Anyone
- Everyone
- Must know about or be able to find the resources without search engines

Gray area between surface and deep, with sites moving between them

Example Services:

- Business networks
- Unlinked web sites (using IP addresses instead of domain names)
- Banks, investments, financial
- Government information about you
- Social media (private)
- File-sharing sites

Deep Web Who and What

Deep web sites are accessible to anyone who knows where to look. These web sites and systems are generally not found in search engine results due to access restrictions. Sites in this area can be free or paid. They can require authentication or not. Mostly, they are services that are more challenging to find on the internet. There is a fuzzy area between the surface web and the deep web because, as search engines index more sites, some of the content in the deep web appears to shift to the surface web. Other sites move the opposite direction if they add authentication to protect search engines from accessing their data.

Examples of sites in the deep web are shown above.

Dark Web Who and What

Who uses it:[1]

- Journalists
- Oppressed people
- Criminals
- Governments/Law Enforcement
- Privacy-minded people
- Whistleblowers

Example Services:

- Anonymous browsing
- Anonymous communications
- Anonymous file sharing
- Unlinked web sites
- Forums about any topic
- Criminal marketplaces
- Specialized dark web search engines

Dark Web Who and What

The dark web requires people to use special software and have an understanding for how the particular dark web network operates. Most people think about criminals using the dark web and, while that is the case, they are not the only people to do so.

Overall, the dark web is a place for people to more anonymously communicate with others. If a person lives or works in a country where freedom of speech is not allowed, they can use the anonymity of the dark web to get their message out to others. There are marketplaces in the dark web where you can buy and sell all manner of legal and illegal items—from guns and drugs to services and counterfeit money.

Finding sites in some of the dark web networks can be challenging since many dark web services may be private or hidden. You may need to know someone to get an invite to certain resources. But other dark web networks allow users to move from the internet into the dark web, bounce from system to system, and then come back to the internet to surface content more anonymously.

References:

[1] https://sec487.info/7l, September 5, 2017.

OSINTing in the Dark Web
Some of the reasons why you, as an OSINT analyst, may need to visit the dark web include:

- **Curiosity** - What is it? Figuring out if it would/could help your clients for you to be in there.
- **Protecting Communications** - Sometimes OSINT analysts may need to reach resources in countries that have oppressive computer network policies. Using the dark web, they may be able to penetrate those networks and/or use the networks to retrieve useful information.
- **Anonymization of Traffic** - This is one of the main reasons many OSINT analysts use the dark web: it helps hide the source of their traffic. If you work at the "Acme Intelligence Agency" and try to visit web sites on the internet, they might block, monitor, and highlight your web activities. Using anonymizing services in the dark web, analysts can hide where their traffic is originating and access web resources without extra scrutiny.
- **Following Targets** - Depending upon your work, you may need to examine people and content that exists in one or more dark networks. Whether your target is a suspected terrorist, human trafficker, or counterfeiter, they may be operating within the dark web.
- **Financial Fraud** - With its marketplaces, the dark web is an ideal place for people who have stolen credit cards or bank accounts to sell their products.
- **Law Enforcement and Intelligence Agencies** - This follows from the above statement about your targets possibly operating in the dark web. Law enforcement and intelligence agencies both use the dark web to research crimes (past, current, and future).[1]

References:

[1] https://sec487.info/7m, September 9, 2017

[2] https://sec487.info/7n, September 9, 2017

Keeping Track of Sites: The Old Days

Back when the World Wide Web (WWW) was invented, we didn't have search engines. We did not need them, as there were not many web sites publishing content. As the WWW grew, people started creating lists of their favorite web sites and publishing them on pages such as the one you see above. This is the first web page on the internet,[1] and it is merely a list of organized resources that people can visit.

If you didn't know about these web pages with lists, you could have purchased the O'Reilly book *The whole Internet user's guide & catalog*,[2] where every year the "new sites" on the internet were cataloged.

Why the history lesson? Because the deep web is in a similar state.

Image from https://sec487.info/77, September 4, 2017.

References:

[1] https://sec487.info/77, September 4, 2017.

[2] https://sec487.info/78, September 4, 2017.

Same Problem, Different Layer

As the internet grew, innovators created search engines[1] to help us find content we wanted and that fixed the issue

We have the same problem with the deep web: where are all the resources?

There are web sites that list resources the author likes

So how do we find those lists of resources so that we can access?

Answers:

- Personal/Business Blogs
- "Gateway sites"
- Search engines can help locate references

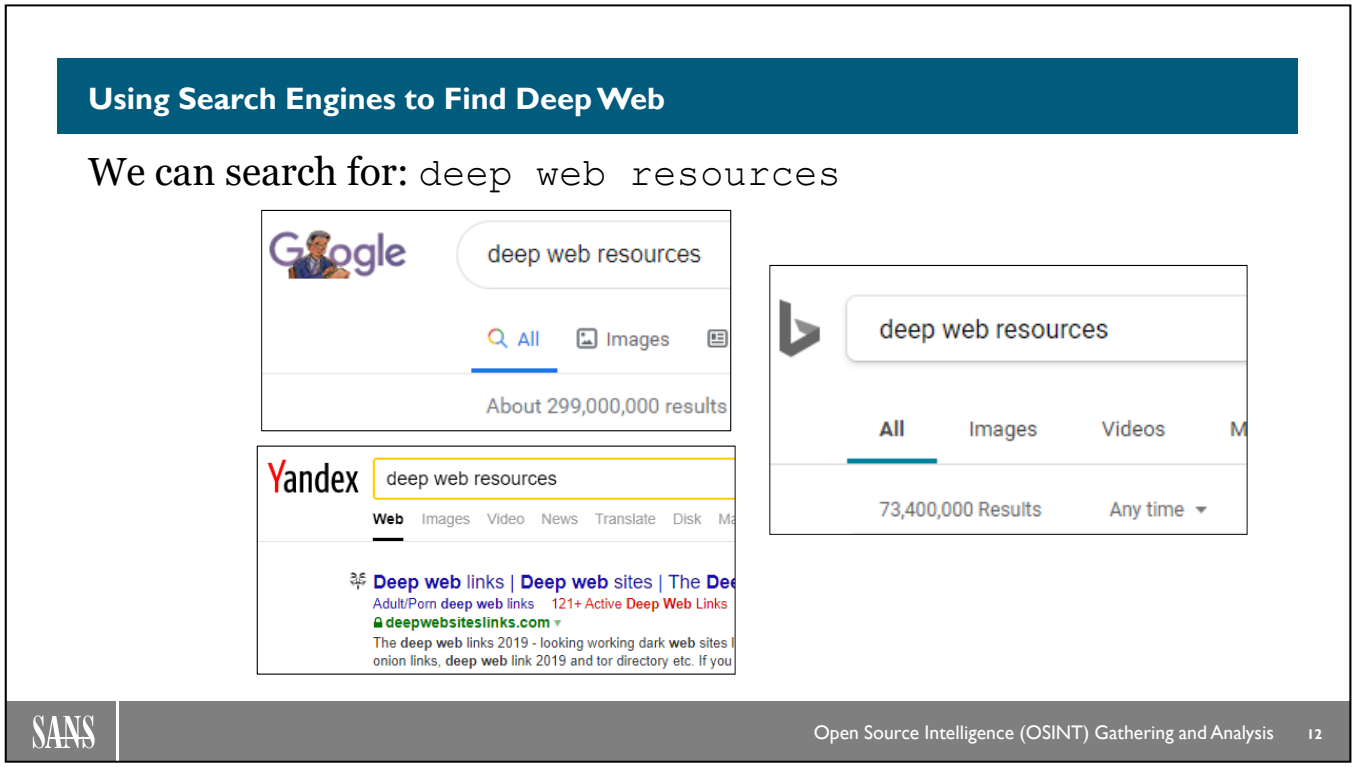
Same Problem, Different Layer

When the internet started, there were only a few sites. As its popularity and usefulness grew, people needed methods to catalog useful web resources. As mentioned, they created lists and hyperlinked to content. The deep web has this same issue. There are many resources on the internet that are not indexed by search engines that we need to keep track of. How do we do that? We do the exact same thing we did in the beginning: we make web pages that have hyperlinks to that deep web content.

Finding deep web resources can be accomplished by looking for these sites that list resources on the deep web. We find these references in blogs, web articles, and other content that IS indexed by major search engines. This is the blurry part between the surface web and deep web. We can use search engines to find blogs that have content that reference deep web resources, but we cannot locate those results directly in search engines.

Reference:

[1] https://sec487.info/ov, September 4, 2017.



Using Search Engines to Find Deep Web

Using Google,[1] Yandex,[2] and Bing[3] search engines, we searched for the string "deep web resources" and found tens of millions of possible pages that might list web pages for you to visit. Which would you start with? It really depends upon what your goal for going into the Deep Web is. Just looking around? Pick any of the top sites and surf. If you want something specific, you may have to dig deeper.

Some good starting points are:

- OEDB - https://sec487.info/7g
- MakeUseOf - https://sec487.info/7h

Images and References:

[1] https://sec487.info/7b, October 1, 2019.

[2] https://sec487.info/7c, October 1, 2019.

[3] https://sec487.info/7d, October 1, 2019.

The Resulting Sites and Good Links

There are an overwhelming number of sites you now need to check out

This is why deep web research needs to be directed or focused on a topic or target

OSINT deep web links:

- OSINTFramework.com[1]
- Awesome-OSINT[2]
- OSINT_Team_Links[3]
- IntelTechnique's Blog[4]

The Resulting Sites and Good Links

So, now all we need to do is visit the top web sites from these results, each of which will have hundreds of overlapping resources. Sounds like a waste of precious time, doesn't it? It is. Deep web research from an OSINT perspective is usually one of focus. We know what data we want, and we know that there must be a source out there that can give it to us.

A good example is the case of validating a healthcare worker's credentials. In one assessment, we were given a person's name and told she was a retired healthcare worker. We needed to validate this claim (and others) for the client. So, we visited the deep web and searched for United States medical license registries, entered the target's data, and eventually found her information.

The OSINT community is amazing at helping each other out. The slide notes four different sites that contain excellent, curated/organized pages of links to deep web resources that you can leverage in your OSINT work.

References:

[1] https://osintframework.com/

[2] https://sec487.info/7e

[3] https://sec487.info/3z

[4] https://sec487.info/if

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

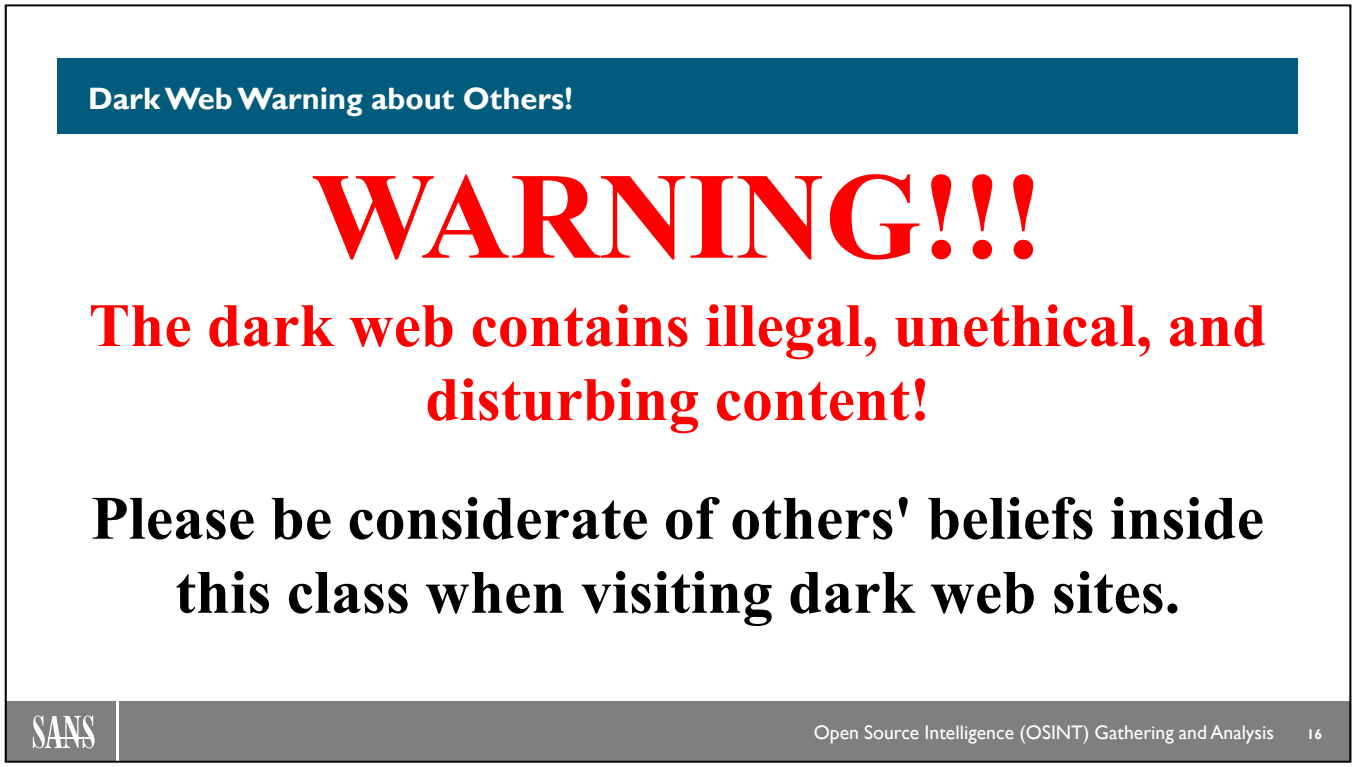
1. The Surface, Deep, and Dark Webs
2. **The Dark Web**
3. Freenet
4. I2P - Invisible Internet Project
5. Tor
6. Monitoring and Alerting
7. International Issues
8. Vehicle Searches
9. Putting It All Together

This page intentionally left blank.

Dark Web Warning for You!

Before we go any further, be warned that in the dark web, it is easy to find illegal, unethical, and disturbing content.

If your work takes you into the dark web, you need to ensure that you are protecting yourself, your organization, and your computer systems against threats. Knowing the risks and preparing for them will help when you come across something that you need to report or take action on.



Dark Web Warning about Others!

During this class on OSINT, we sometimes need to discuss content that may be disturbing or upsetting to some people. It may offend people's religious, moral, or other belief systems. We mention this content not to make people uncomfortable but to help them understand that, during OSINT work, we have no idea where the assessment will lead us. Some of those may be deep, dark, and dangerous places.

The upcoming section will contain information that may be upsetting to some students. There are no graphic pictures, but we will be mentioning some sensitive topics. Be warned and take whatever steps you need to be comfortable.

Dark Web Risks

Consider your work location before accessing the dark web

- Who else can view your screen?
- Is the sound on your computer on?
- What security precautions are on your computer operating system and in your browser?

Some risks to consider include:

- Viewing child sexual abuse material (CSAM)
- Malicious files
- Revealing information about you or your organization
- Attacks against your browser and computer system

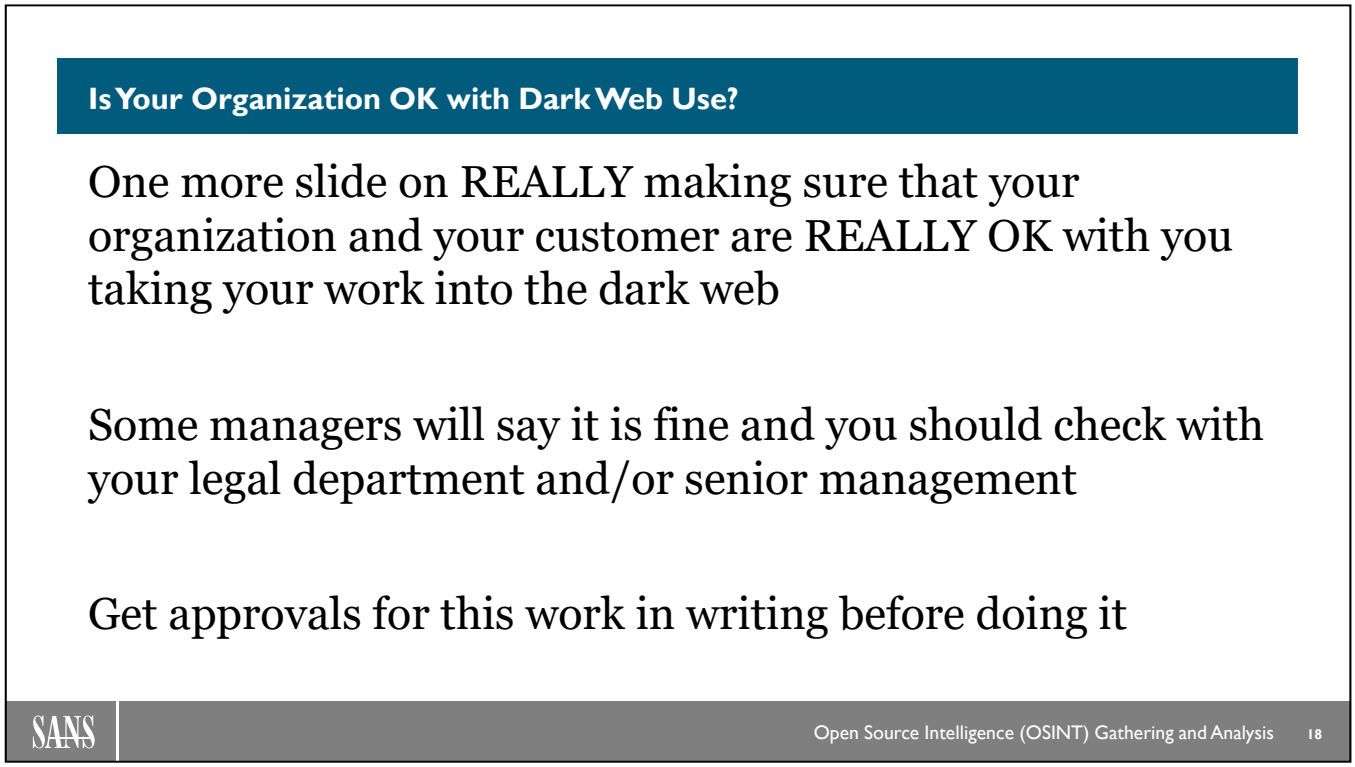
Dark Web Risks

Knowing that, at almost any point in your using the dark web, you could view and hear a video of something unpleasant or inappropriate for a work environment should raise concern. Plan ahead and ensure that you are in a safe place where you control who can see and hear what is on your computer screen. An investigator we knew worked at home and was performing work in the dark web. The content of the work was highly distasteful but important to her case. She took precautions to only view content from her home office and with her head phones on. While performing her work one day, her young son came into the room unexpectedly and watched a segment of the videos that she was viewing. Because of the headphones, she did not realize he was behind her for a couple minutes and, by then, he had seen a good amount of the content of her screen. Be aware of your surroundings.

We know that there are people with malicious intentions on the dark web. Before venturing into the dark web, ensure that your computer system and applications (such as web browsers) are hardened, patched, and running recent security software. Attackers on the dark web may embed malware into files that they offer for people to download, attack your computer browser and operating system, or try to identify you or your company. Be wary and protect yourself, your computer, and your organization.

If your assessments take you into those portions of the dark web where you may view child pornography, ensure that your team has a plan of what to do. In many countries, you may be required by law to report child sexual abuse material.

In the United States, you can report it to the National Centers for Missing and Exploited Children (https://report.cybertip.org/ and 800-THE-LOST (800-843-5678)) or contact your local Federal Bureau of Investigations (FBI) field office (https://sec487.info/fv).



Is Your Organization OK with Dark Web Use?

While many managers support their staff taking OSINT investigations wherever they need to go to get results, those managers may not have the authority to grant you that flexibility.

Before working in the dark web for your company, organization, and clients, ensure that you have written approval from someone in authority for you to perform this work, from what location(s), and using what system(s).

Getting this in advance can help prevent legal and ethical issues later.

Welcome to the Dark Web!

We finally made it. A couple things to note about the dark web. First, as previously mentioned, the dark web is an overlay network that sends and receives data on top of the normal internet. So, when we are using the dark web, we are also using the internet.

Second, while there are a variety of different dark networks,[1] we are going to focus on the most well-known three networks: Freenet, I2P (the Invisible Internet Project), and Tor.

Reference:

[1] https://sec487.info/7i, September 5, 2017.

Let's Compare the Three

Attribute	Freenet	I2P	Tor
Primary Goal	Anti-censorship, distributed file sharing	Anonymity and security for communications	Anonymity and bypassing restrictions
User Base	Small	Medium	Large
Access to Internet Sites?	No	Sometimes	Yes
Speed	More disk space	More bandwidth	Depends

Let's Compare the Three

Comparing the top three dark web networks is challenging since they are quite different. We will examine the networks from three perspectives: the primary goal of the network, how big its user base is, and whether it allows access to internet web resources.

Each of these overlay networks has a different primary goal. For the Freenet, it is a distributed file-sharing network that breaks up files into pieces and scatters those pieces in retrievable sections across its network. This makes it impossible to remove content that is uploaded to the Freenet.

I2P uses a series of temporal, one-way, encrypted tunnels to transfer data from one system to the next. Its goal is to protect the data traversing its networks and make it challenging to trace data back to a single source through network traffic analysis techniques.

Tor is used to access internet sites (and its own .onion services) in an anonymous fashion by proxying network traffic across many systems. User data hops from computer to computer and eventually reaches its destination. The destination site does not know where the original request came from and so the user's computer is anonymized.

When looking at the estimated numbers of users for each network, understand that Freenet has the smallest user base and Tor the largest, with I2P falling in the middle. These are rough estimates (as we shall see) because getting an accurate reading of the number of users traversing an anonymous network becomes quite a challenge.

The Freenet and, for the most part, I2P are closed networks where you mainly access resources only on those networks and not the internet. Tor allows access to resources only found on its network and anonymous browsing of internet resources.

We will learn in the coming sections that these networks are much more complicated.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. **Freenet**
4. I2P - Invisible Internet Project
5. Tor
6. Monitoring and Alerting
7. International Issues
8. Vehicle Searches
9. Putting It All Together

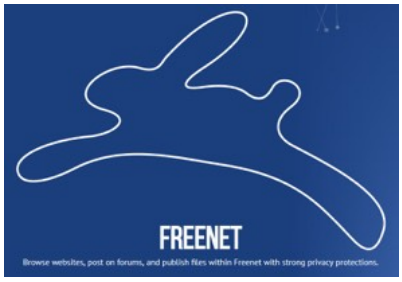
This page intentionally left blank.

First Network: Freenet

Distributed data store / local storage device[1]

When content is uploaded, it is broken into pieces and distributed across the nodes (other systems)

Upload a file and you are provided a key. Anyone who wants to retrieve that file must provide that key.



First Network: Freenet

Think of the Freenet project (https://freenetproject.org/) as a large storage device where content is spread across all the computers (nodes) of a network. Each person's computer has a local data store that contains encrypted fragments of popular files that it shares with the other nodes on the network. When you upload a file to Freenet, it is chopped up into pieces, encrypted, and then distributed to peer nodes (computers that are closest to your computer on the network). The uploader is provided a key to retrieve that file. They can provide that key to anyone else on Freenet, and that person can also retrieve that file.

Image from https://freenetproject.org/, September 6, 2017.

Reference:

[1] https://sec487.info/7j, September 6, 2017.

About Freenet's Storage

Decentralized network with each node only knowing about its direct peers

Censor resistant – You have no idea what is on your computer

Popular content can persist indefinitely on the Freenet

Because the data store is broken across all nodes, restrictive entities cannot shut down one computer and stop content from being shared

Since each node can have encrypted pieces of any file uploaded, that means if someone uploads child sexual abuse material or hateful content, all users could get pieces of the files

About Freenet's Storage

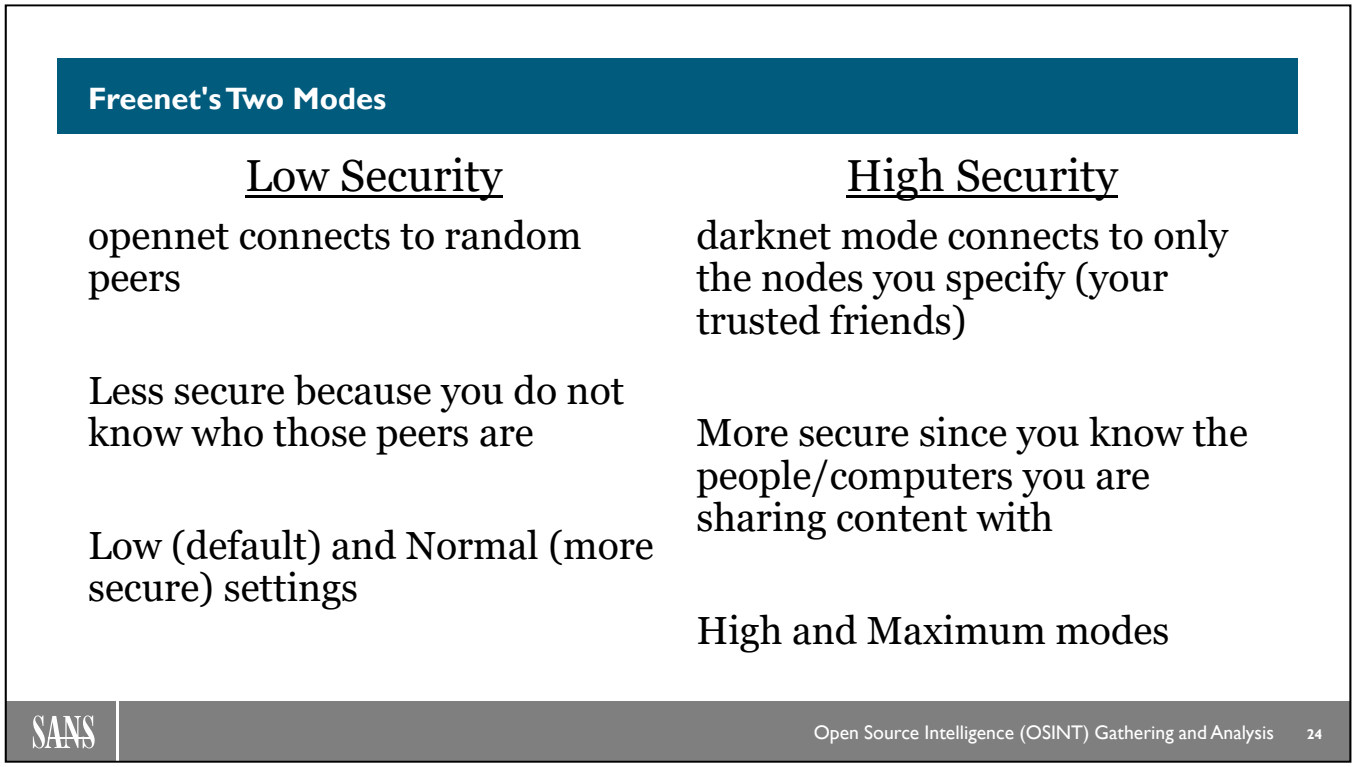
All nodes of the Freenet share pieces of the Freenet data. With each file being chunked into pieces and distributed to other nodes, each node could have content from any file. You have the option to set, increase, or decrease the encrypted storage space on your hard drive that you are willing to donate to the Freenet within the Freenet settings.

Be aware that if someone uploads classified or otherwise sensitive data to the Freenet, it will be encrypted and distributed to peer nodes.[1] Your computer may hold that content, but you may never know it. That is the power of the Freenet. No user knows what data their system contains; therefore, they cannot delete content from the Freenet and authorities cannot shut down the Freenet by removing some of the nodes.

If content is popular on the Freenet, it can be propagated between nodes forever, moving from node to node as people request the resources.

Reference:

[1] https://sec487.info/7o, September 9, 2017.



Freenet's Two Modes

When you start using Freenet, you will most likely wish to use it in the lower-security opennet mode, where Freenet will automatically connect to seed nodes (computers it knows about at install) and then start connecting to other random nodes as time goes on. This is less secure than darknet mode because you have no idea who owns/runs the other nodes that your system attaches to.[1] There are two sub-modes for opennet: low security (default) and normal. If you are using the opennet mode, consider moving to normal mode to increase your anonymity.

As you use the Freenet more, or as your friends begin to use it, you can change from opennet to darknet mode by specifying in your computer's Freenet settings to only connect to the nodes of your friends.[2] Since you presumably know these people, your level of trust increases, and it is much safer to use the Freenet. Darknet has two settings as well: high and maximum modes.

References:

[1] https://sec487.info/7q, September 9, 2017.

[2] https://sec487.info/7r, September 9, 2017.

Accessing the Freenet

Download the software from https://freenetproject.org/pages/download.html

Cross-platform (Windows, macOS, Linux installers available)

Excellent, wizard-driven installer to set your settings

Accessing the Freenet

Getting into the Freenet is accomplished by downloading the software from the freenetproject.org web site. The site has versions for macOS, Windows, and Linux systems. The installers are quite good and are wizard-driven, asking the user questions and providing them advice to set up the initial configuration.

Freenet Keys

Accessing data in Freenet is performed through a local Java-based applet called FProxy

Four types of keys:[1]

- **CHK** - Content Hash Keys (file)
- **SSK** - Signed Subspace Keys (web site)
- **USK** - Updatable Subspace Keys (versions of sites)
- **KSK** - Keyword Signed Keys

Accessible from a web browser at:
`http://127.0.0.1:8888/[Freenet Key]`

Example:

`http://127.0.0.1:8888/CHK@khy~PN5-P-Ze18SdPfb~eMVCcWv1RTDNaP3ketO3TwQ,-kcC5FBaNN7wZtcHgZGObkxk9221BL5r~wbbVTUhyHc,AAMC--8`

Freenet Keys

Content is given a key when it is configured/uploaded to the Freenet. When Freenet is installed on a computer, it installs the FProxy, which is then used to configure and access the Freenet via a web interface. In a web browser, you can visit http://127.0.0.1:8888/[key] as shown in the slide above.

There are four types of keys in Freenet:[1]

- **CHK** - "Content Hash Keys are for files with static content, like an .mp3 or a PDF-document. These keys are hashes of the content of the file."[2]
- **SSK** - "Signed Subspace Keys are usually for sites that are going to change over time. For example, a website that may need news to be updated or information to be corrected, added or deleted."[3]
- **USK** - "Updateable Subspace Keys are useful for linking to the latest version of a Signed Subspace Key (SSK) site. Note that USKs are really just a user-friendly wrapper around SSKs, which hide the process of searching for more recent versions of a site."[4]
- **KSK** - "Keyword-Signed Keys (KSKs) allow you to save named pages in Freenet."[5]

Reference:

[1][2][3][4][5] https://sec487.info/7k, September 6, 2017.

"Browsing" the Freenet

Browsing Freenet web resources involves requesting files from the data store and viewing them locally

There can be high latency with requests for data since pieces have to be reassembled from multiple sources

The longer your node is online and the more space you give to Freenet, the faster your browsing since your local data store will contain more content

Uses the FProxy on your system at 127.0.0.1:8888

"Browsing" the Freenet

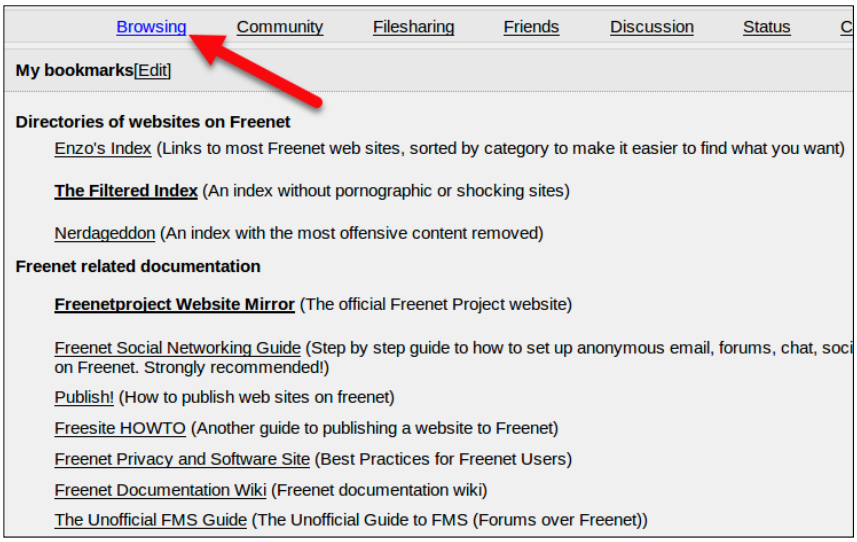
When you use the Freenet to access sites similarly to what you might do on the internet, it serves your browser the pages in a slightly different manner than a traditional web site may. You request a resource through the FProxy, and the proxy finds the pieces of that content on Freenet, requests and downloads them, then reassembles them so that they can be displayed in your browser. This takes time and makes Freenet a high-latency network.

Of course, if the content you want has already been downloaded to your computer, then viewing that content is fast since FProxy grabs it from the local cache. The longer your computer is on the Freenet and the more data you allow the Freenet to host on your system, the faster your browsing will be.

Browsing Default Bookmarks

Freenet comes with default bookmarks

They are excellent starting places and link to other content



Browsing Default Bookmarks

Where do you start browsing content in Freenet? Launch Freenet and then visit the FProxy in a browser at http://127.0.0.1:8888. From there, visit the "Browsing" item in the top menu bar (red arrow is pointing to it in the picture above). This will display the page in the picture and allow you to choose one of the bigger directory sites on Freenet. These sites contain links to other Freenet resources.

When you click on these links, Freenet requests the key for that site or resource. If it is not already on your system, it will bring your browser to a "Downloading" page where you will wait until FProxy grabs the resource.

"Surfing" the Freenet feels very similar to surfing the internet in the 1990s. It is slow. There are links to resources no longer available. In most cases, the web sites are text based.

Other Services in the Freenet

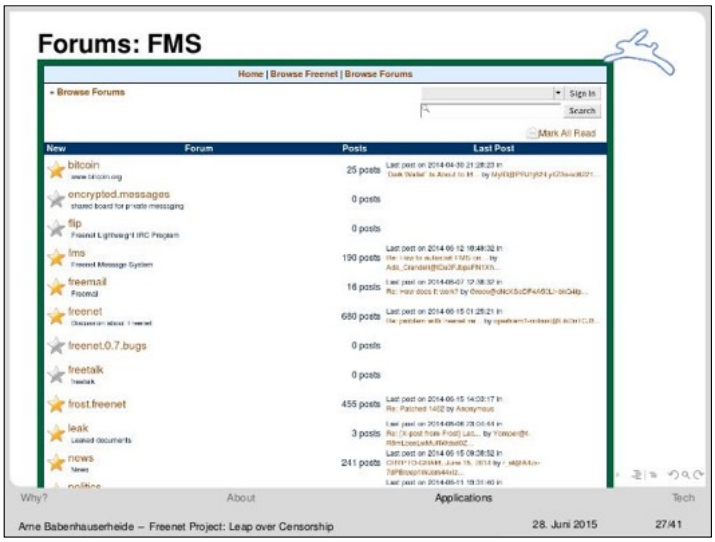
Email:

Create and use an email account

Forums:

Participate in discussions

Create your own web site



Other Services in the Freenet

Aside from file sharing and downloading, Freenet offers several other mechanisms to exchange information. It has email and forums, and you can event create your own web site. Directions for accomplishing these tasks can be found in tutorials, such as the one on https://darknetlive.com.

Image from https://sec487.info/7s, September 9, 2017. (URL no longer active)

Freenet Summary

1. Java based
2. Slow, retrieves file pieces for resources
3. Distributed across peers
4. Data may or may not be available
5. FProxy is your configuration and browsing tool
6. Two main modes: opennet and darknet
7. Web browse, file share, create content, email, forums

Freenet Summary

So, to recap this overview of Freenet:

1. Java-based application that is cross-platform compatible.
2. Slow. It retrieves file pieces for resources, which makes using it a little challenging unless you leave it up for many hours and allow it to store a lot of content on your computer.
3. Distributed across peers, making it hard to remove content and slower to access content.
4. Data may or may not be available depending upon how popular it is.
5. FProxy is your configuration and browsing tool that is installed when the Freenet program is installed.
6. Two main modes are opennet and darknet, with opennet being the default and darknet being the more secure.
7. Web browse, file share, create content, email, forums are many of the features of this dark web network.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

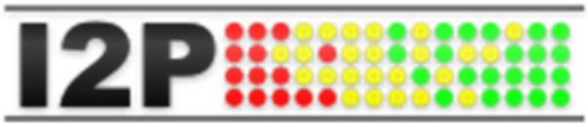
THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. Freenet
4. **I2P - Invisible Internet Project**
5. Tor
6. Monitoring and Alerting
7. International Issues
8. Vehicle Searches
9. Putting It All Together

This page intentionally left blank.

Second Network: I2P

The Invisible Internet Project (I2P) is an overlay network



Main purpose is to protect communications from eavesdropping, monitoring, and traffic analysis

https://geti2p.net/

Second Network: I2P

The Invisible Internet Project (I2P) is a complicated, highly encrypted network overlay for the regular internet. While there are instances where users can access hosts on the internet, most of the resources that I2P users access are only accessible from inside the network.

The goal of I2P is to allow people to visit network resources anonymously and prevent others (corporations, governments, oppressive parents, etc.) from eavesdropping on the traffic transmitted across I2P. It uses 3–4 layers of encryption, multiple hops, and one-way tunnels to achieve its goals.

Image from https://sec487.info/ow, September 9, 2017.

What Do People Do in I2P?

Torrents
- I2PSnark
- Vuze (used to be Azureus)

Forums

Email
- I2P-Bote
- Postman's service

Chat/IRC

Marketplaces
- Drugs
- Counterfeiting
- Weapons
- Pornography
- Coupon trading

What Do People Do in I2P?

I2P is much like the internet itself: people have a wide variety of activities that they can engage in and a good amount of content to download/view. Users can use the I2P-Bote, a "serverless peer-to-peer email application using a distributed hash table (DHT) for secure mail storage"[1] or use the I2PSnark client to access files on the I2P BitTorrent system. I2P websites are commonly referred to as "Eepsites," and there are many Eepsites to browse. From marketplaces that sell legal and illegal items and services to forums and chat systems to share information, the I2P network very much mimics the functionality found on the surface web.

Reference:

[1] https://sec487.info/7x, September 9, 2017.

I2P vs. Freenet

It differs from Freenet in that is more interactive

I2P uses a series of one-way tunnels to communicate with other systems

- Outbound = Send
- Inbound = Receive

"Unlike web sites hosted within content distribution networks like Freenet or GNUnet, the services hosted on I2P are fully interactive - there are traditional web-style search engines, bulletin boards, blogs you can comment on, database driven sites, and bridges to query static systems like Freenet without needing to install it locally."[1]

I2P vs. Freenet

In general, I2P differs from Freenet in that it works similarly to how most people use the internet. Instead of having to request and then download all the bits of encrypted content and then having your computer reassemble them into content that is displayed locally, if you want to browse to a web site in I2P, you enter in the address in your browser then hit send.

On the internet, when you're web browsing, your traffic will hop system to system until it reaches the resource you requested and then it'll hop back through a series of devices to get back to your system. (Yes, I know that is an over-simplification of how web browsing works. If you want to know what happens when you web browse, visit https://sec487.info/7u).

With I2P, your computer running the I2P software creates one-way tunnels to other nodes. It either sends or receives traffic through each tunnel. These encrypted tunnels allow data to be passed to and from the remote resources very fast and in a manner that is hard to trace.

The I2P main web site (quoted above) notes that the sites on I2P are similar to what you would find on the internet at large.

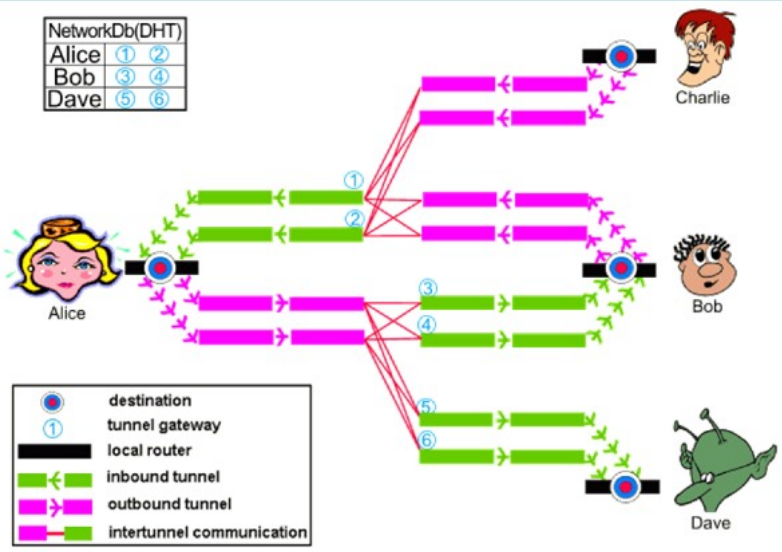
Reference:

[1] https://sec487.info/7t, September 9, 2017.

I2P Tunnels Illustrated

Traffic flows from outbound tunnels across multiple hops to inbound tunnels

It is then consumed and rerouted or used



I2P Tunnels Illustrated

I2P creates one-way tunnels on each node. A tunnel can either be outbound, sending traffic to other destinations, or it can be inbound and accept traffic. Each node has numerous inbound and outbound tunnels that are created and torn down every 10 minutes.[1] This prevents eavesdroppers from executing traffic analysis attacks against node traffic to determine which node certain content came from.

In the image above, traffic from Alice to Bob would flow from one of her outbound tunnels (pink) and hop to at least two other routers on the way to its destination at Bob's system. It would hit Bob's inbound tunnels (green) and then be decrypted and consumed by Bob's destination application (email, IRC, web browser, etc.). Return traffic would flow from Bob's outbound tunnels (pink) to Alice's inbound tunnels (green).

Note that Alice, Bob, Charlie, and Dave all have other tunnels that connect their systems to other I2P nodes. Those are just not shown in this diagram. Also, Charlie's inbound and Dave's outbound tunnels have been omitted for space.

Image from https://sec487.info/ox, September 9, 2017.

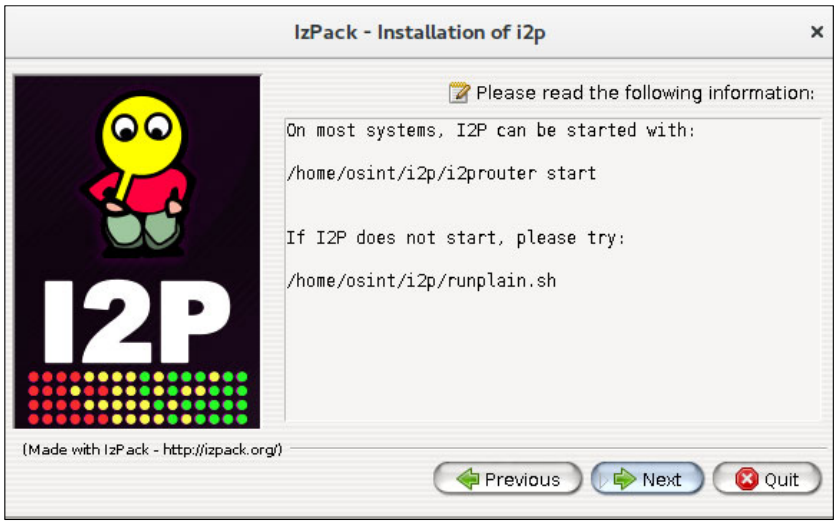
Reference:

[1] https://sec487.info/7y, September 9, 2017.

Installing I2P

Retrieved from the https://geti2p.net/en/download site

Java-based app that is cross-platform compatible (Windows, macOS and Linux)



Installing I2P

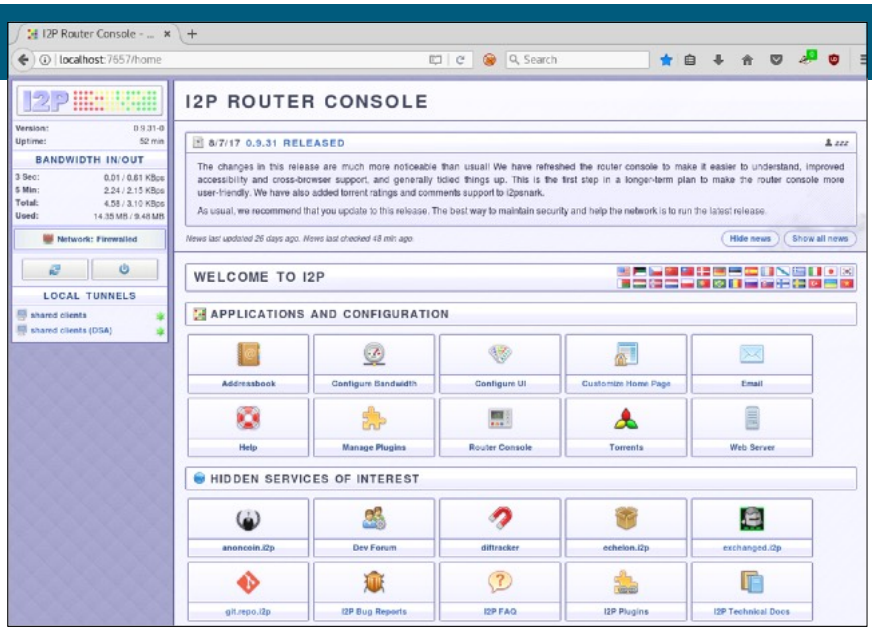
I2P has a Java-based installer that works on Windows, macOS, and Linux computers. After downloading it from the https://sec487.info/7v site, you should run it as a normal (non-administrative) user.

The installer is easy to use and will display the page above at the end of the installation. Take note of the location where the software was installed and the command to start the I2P router on your computer. In the picture above, we installed it into a Linux system into the home directory of a user named "osint." If we wanted to stop, start, or restart the service, we could run the `/home/osint/i2p/i2prouter [command]` (where we would type start or stop or restart instead of the [command]).

Running I2P

From a terminal or command prompt, "cd" into the directory of the installation and type:

```
i2prouter start
```



Running I2P

Once it is installed, you can launch the I2P service by executing it from a command line or terminal window, or you can click on the appropriate icons on your desktop (macOS and Windows). From the Linux command line, we would start the service using the "i2prouter start" command and parameter.

When the system begins, it should launch a web browser and point it to your local I2P proxy. I2P's web router console (shown above) is a great place to configure and begin using I2P.

You will want to visit the "Configure Bandwidth" feature to increase the bandwidth I2P can use and the amount you wish to share with other I2P users for routing their content. Remember that the more you share your bandwidth, the more traffic is sent to and from your system that comes from others. This increases your anonymity, as people analyzing traffic coming from your computer won't know if the traffic your system is sending originated from your computer or was forwarded from another.

It Will Take a While

When it starts up, I2P needs to initialize, establish tunnels, discover peers, check for firewalling...this takes time

You may need to open up firewalls (network and host) to allow UDP traffic to I2P through (/confignet page UDP Port section)

Configure apps to use I2P

Local proxies:

- Web browser (proxy through localhost:4444)
- IRC (proxy through localhost:6668)

It Will Take a While

There will be a delay between when you launch I2P on your system and when you can use it for your network communication. There are many underlying tasks that I2P must accomplish before you can begin to use it, and depending on your network and system setup, it could take a while.

Consider that you may also need to change network (router) and host-based firewall settings to allow I2P UDP traffic through. If the left side of the I2P Router Console shows an issue with the "Network:" setting, click on the Network text to view troubleshooting options. You may need to visit the /confignet page and scroll down to the UDP port section to find the port number for your I2P installation.

Finally, you will need to set up your applications to use I2P. Your web browser needs to have its local proxy setting altered to point to the I2P proxy at localhost:4444. To do this, open the preferences in your browser and navigate to the proxy settings. Ensure that traffic destined for "localhost" is not proxied through I2P, as you will not be able to access the I2P Console.

If you are going to be using IRC, you will need to set it up to use the local proxy at localhost:6668.[1]

Reference:

[1] https://sec487.info/7z, September 9, 2017.

Host Names in I2P

Systems may have one or more of the following host names or addresses:

Type	Example
516-byte	nightblade.i2p=nyErwSseXbsojcWtNkDyUul0YULtqr6qyWSzIp639Ygpe8juCdgPMLURVXcmlCvo~QPoHg6zt53KpgpGvB1-Wv2SGvc2MVs~o8USw3ius8fP1URphqcBbulK8Ci0bgknt0kD0AfxqfMz-p~xk1QEMxq2kZEoB3oyIIFnQlpb2ByS74Lx8iKzXTrwWk19I3Dvu4nIq8CBDDwu3lYoCD2kC-jT5pjgglverGPEGN4o55LYVTtfSg4gAJFZeaE4KjBR5P1z7cca6UDjGMWfR0iCa8P3qpkY2ODYpk~8w2xgBbgDq-8Hzik~uraHc598ccS8QpwB0f0Jw~2PZcTjOPdZ-239U6p3tESXa7FXzRBCujv4Bx6CVFRhCmBHpyFnCD-MugZ~vR6XFSS2XBsCT~duXKq94HH2n1iAWslG4Vu44ut1JVhDPFzp~Dk7wujB0tCo2HXH2icRQxOWe37foU4LZSJ4oMpFDACBzwSfcZdIPsVRxGttKQx4yzgffR1Q~Jl7AAAA
base32.b32.i2p	ukeu3k5oycgaauneqgtnvselmt4yemvoilkln7jpvamvfx7dnkdq.b32.i2p
vanity.i2p	stats.i2p

Host Names in I2P

The https://sec487.info/8h page discusses the I2P routing and host names in depth. For our purposes, recognize that you might see the I2P site names in one or more of the above three formats: 516-byte key, [52 characters].b32.i2p using Base32, or a vanity host name such as princessbride.i2p.

I2P Addressbooks and Name Resolution

The name -> address resolution occurs locally

It uses a hosts.txt or blockfile naming service for translation

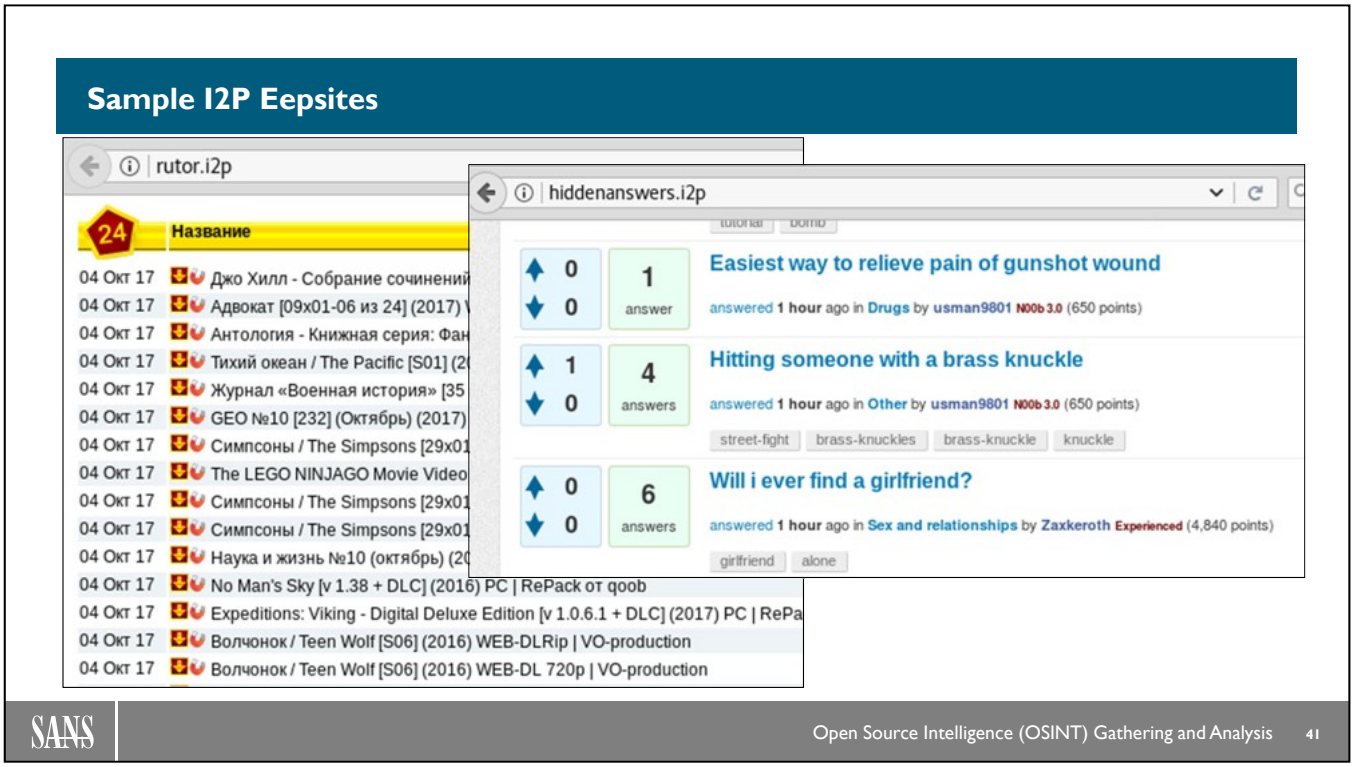
Other users' hosts.txt files are periodically merged with yours

Users configure subscriptions to others' hosts.txt files or addressbooks for regular merging

These actions increase the number of I2P sites and nodes your computer knows about

I2P Addressbooks and Name Resolution

Your I2P client has a single entry in its local hosts.txt file at the start. That is for the trusted http://i2p-projekt.i2p/hosts.txt file. Periodically, your local naming service will download the hosts.txt file from any site that your system is subscribed to. Upon the file downloading, it will merge with your existing hosts.txt file, increasing the number of sites your system knows about.



Sample I2P Eepsites

To give you a flavor for some of the content on I2P, we present the forum/message board hiddenanswers.i2p, where users post questions for others to respond to. And then there is rutor.i2p to show that not all I2P resources are in English. This is something to remember, depending upon who your targets are and the languages they know.

Images from http://rutor.i2p and http://hiddenanswers.i2p, October 3, 2017.

Additional Eepsites

Finding eepsites, blogs, and search engines within the I2P network can be challenging

We can use surface web sites to find them too!

Question is, do you trust these sites?

Search

eepsites
The premier search engine of i2p. Accurate results.

elgoog
Search i2p and onionland. Accurate results.

epsilon
Minimalist interface. Yet a fast search engine.

pointzero
.i2p search engine and links catalog. Slow.

i2pfind
Search for i2p and/or onion links. Accurate.

direct
Search engine and huge, link-O-rama archive.

Blogs

planet
Aggregator of trackers, blogs and other feeds.

killyourtv
Wiki-like blog. Howtos, repos, docs and services.

peek-a-boo
ReturningNovice's eepsite. Files about privacy.

str4d
A blog, guides, services, i2p links and more.

echelon
Blog and a software repo consisting of i2p-ware.

allofthis
A blog which reviews a variety of i2p sites.

bigbrother
The news source on i2p for world events, etc.

gateway
A blog available over onion, i2p and freenet.

sigterm
A blog about privacy, security and related stuff.

darkreboot
darkreboot's new i2p blog, mostly about love.

Additional Eepsites

We can use the surface internet to find eepsites in I2P that may be relevant to our investigations. The http://nekhbet.com and https://sec487.info/vo sites (and others) have lists of *.i2p domains with brief descriptions of what the site content or purpose might be.

Image from https://sec487.info/pv, January 6, 2019.

I2P Summary

1. Java based
2. Goal is to prevent eavesdropping and attribution
3. Series of temporal one-way tunnels
4. Data may or may not be available
5. I2P Console is your configuration tool
6. Web browse, file share, create content, email, forums

I2P Summary

So, to recap this overview of I2P:

1. Java-based application that is cross-platform compatible.
2. Goal is to prevent eavesdropping and attribution of network traffic to a host/node/destination by encrypting traffic and having it hop through other systems.
3. Series of temporal one-way tunnels that are either in- or outbound and only up for 10 minutes.
4. Data may or may not be available, depending upon if the resource hosting it is on or not.
5. I2P Console is your configuration tool and is installed when the I2P program is installed.
6. Web browse, file share, create content, email, forums are many of the features of this dark web network.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. Freenet
4. I2P - Invisible Internet Project
5. **Tor**
6. Monitoring and Alerting
7. International Issues
8. Vehicle Searches
9. Putting It All Together

This page intentionally left blank.

Third Network: Tor

The onion routing project, and written Tor, not TOR[1]

Main purpose is to protect how your data gets from your system to the destination and interfere with or defeat censorship, network blocking, and monitoring

https://www.torproject.org/

Third Network: Tor

By far the most used and most common dark web overlay network is Tor, with over 2 million users sending over 100Gb of traffic daily.[2] Originally named for The Onion Routing project and now referred to as just Tor (capitalized in that manner), it is mainly used to provide pseudo-anonymous network connections to allow people to access resources that may be censored, blocked, and/or monitored. The main internet web site for Tor is https://www.torproject.org.

Image from https://sec487.info/oy, September 9, 2017.

References:

[1] https://sec487.info/80, September 9, 2017

[2] https://sec487.info/oz, January 5, 2019

Who Uses Tor and Why?

Normal People/OSINT Analysts
- Safeguard privacy
- Avoiding censorship

Journalists and Activists
- Operating in oppressive locations
- Anonymous tips
- Counterculture info sharing

Law Enforcement/Intel Agency
- Sting operations
- Catch criminals
- Intelligence

Criminals/Terrorists
- Marketplaces selling illegal items
- Sharing stolen data
- Recruitment/ideology sharing

Who Uses Tor and Why?

As with many systems and tools, Tor can be used for good or for evil. Tor allows users to keep their communications anonymous and their sources hidden. As an OSINT analyst, using Tor may be a technique that you leverage to find what your target is doing in the dark web, or it might be a tool that you use to keep your work more private. It all depends on what you need and your assessment goals.

Tor has hidden services that are only accessible from within Tor. If desired, users can also browse the internet through Tor to hide their internet-facing IP address from the destination site.

Reference:

[1] https://sec487.info/81, September 9, 2017.

What Is Tor?

"Tor is a connection-oriented anonymizing communication service."[1]

"Users choose a source-routed path through a set of nodes and negotiate a 'virtual circuit' through the network, in which each node knows its predecessor and successor, but no others."[2]

"Traffic flowing down the circuit is unwrapped by a symmetric key at each node, which reveals the downstream node."[3]

What Is Tor?

The description of Tor written by the members of the Tor Project captures the essence of Tor in the above paragraphs. Let's break each down and discuss the important pieces.

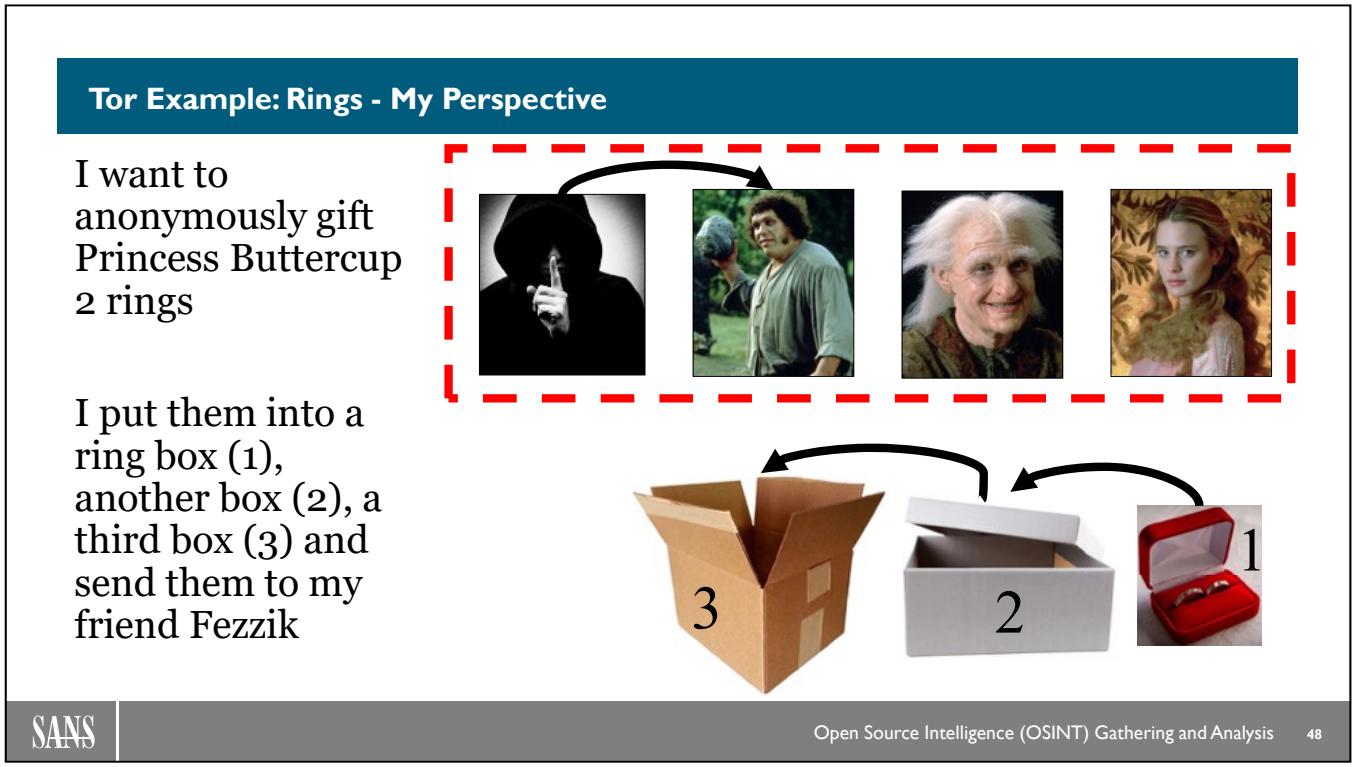
First, Tor is an "anonymizing communication service"[4] that runs on computer and mobile systems. It allows applications on those devices to tunnel across a network of nodes and access some service on the other side.

Tor creates virtual circuits for traffic to be routed through from the source to destination. These circuits are temporal and can be reconfigured by the user or by the Tor service automatically at a certain frequency. In this peer-to-peer network, each node only knows the one it received traffic from and the node it is going to give it to.

Data flowing through Tor is decrypted with a symmetric key along the virtual circuit. Each time this is performed, the content that is revealed is the next hop in that traffic's virtual circuit so that the current node can send the data along the path to the destination.

References:

[1][2][3][4] https://sec487.info/82, September 9, 2017.



Tor Example: Rings - My Perspective

For a simpler explanation of Tor, let's take a look at how I could anonymously deliver a pair of rings to my friend Princess Buttercup for her upcoming wedding. My goal is that she gets the rings but has no idea that they are from me. Additionally, I do not want the couriers who carry the rings to know that there is a precious cargo in the packages that they carry.

First thing I would do is to take the rings and place them into a red ring box (1). On the outside of this box I write "Deliver to Princess Buttercup." I then seal up the ring box and place it inside a larger box (2). I close that box and write on the outside "Deliver to Miracle Max." Lastly, I put that box inside a third box (3), close it up, and write "Deliver to Fezzik" on the outside.

I know every courier who will help these packages get to the destination (red dashed box above). Now I send the big box (with the other boxes inside) to Fezzik.

Tor Example: Rings - Fezzik

Fezzik gets the box (3) and opens it

He pulls box 2 out

It is addressed to Miracle Max

Fezzik gives the box to Max

Tor Example: Rings - Fezzik

Fezzik receives the box (3), opens it, and removes box 2, which is addressed to Miracle Max. Fezzik knows that I gave him the box and that he needs to deliver the box to Miracle Max (red dashed box above), but he does not know the ultimate destination of the packages. Fezzik also does not know the payload (the rings), only that he needs to give the box to Miracle Max, which he does.

Tor Example: Rings - Miracle Max

Miracle Max gets the box (2) and opens it

He pulls box 1 out

It is addressed to Princess Buttercup

He delivers it to her

Tor Example: Rings - Miracle Max

Miracle Max receives box 2 and opens it, revealing box 1 inside. Box 1 is addressed to Princess Buttercup. Max does not know what is in box 1 and only knows that Fezzik gave him a box and he needs to give the contents to Princess Buttercup (red dashed box above). Miracle Max does not know that I sent the box, and he delivers it to Princess Buttercup.

Tor Example: Rings - Princess Buttercup

The rings reach their destination and Princess Buttercup opens the box




She gets the payload/rings for her wedding

Tor Example: Rings - Princess Buttercup

Princess Buttercup receives and opens the red box and discovers the rings inside. She has the payload (the rings) but has no idea who sent it, just that Miracle Max delivered it (red dashed box above). My anonymous delivery worked!

Tor Example: Rings - Summary

How is it like Tor?

1. The source picks the path the data travels
2. Each relay decrypts the "outer layer" and forwards the content to the next recipient
3. The payload remains encrypted until it reaches the destination

The Ring Example

1. I picked the path the boxes travelled
2. Fezzik knew I gave him a box for Miracle Max. Max knew Fezzik gave him a box for Princess Buttercup. Princess Buttercup knew she got a box of rings from Miracle Max
3. Only Princess Buttercup knew what was inside the inner box

Tor Example: Rings - Summary

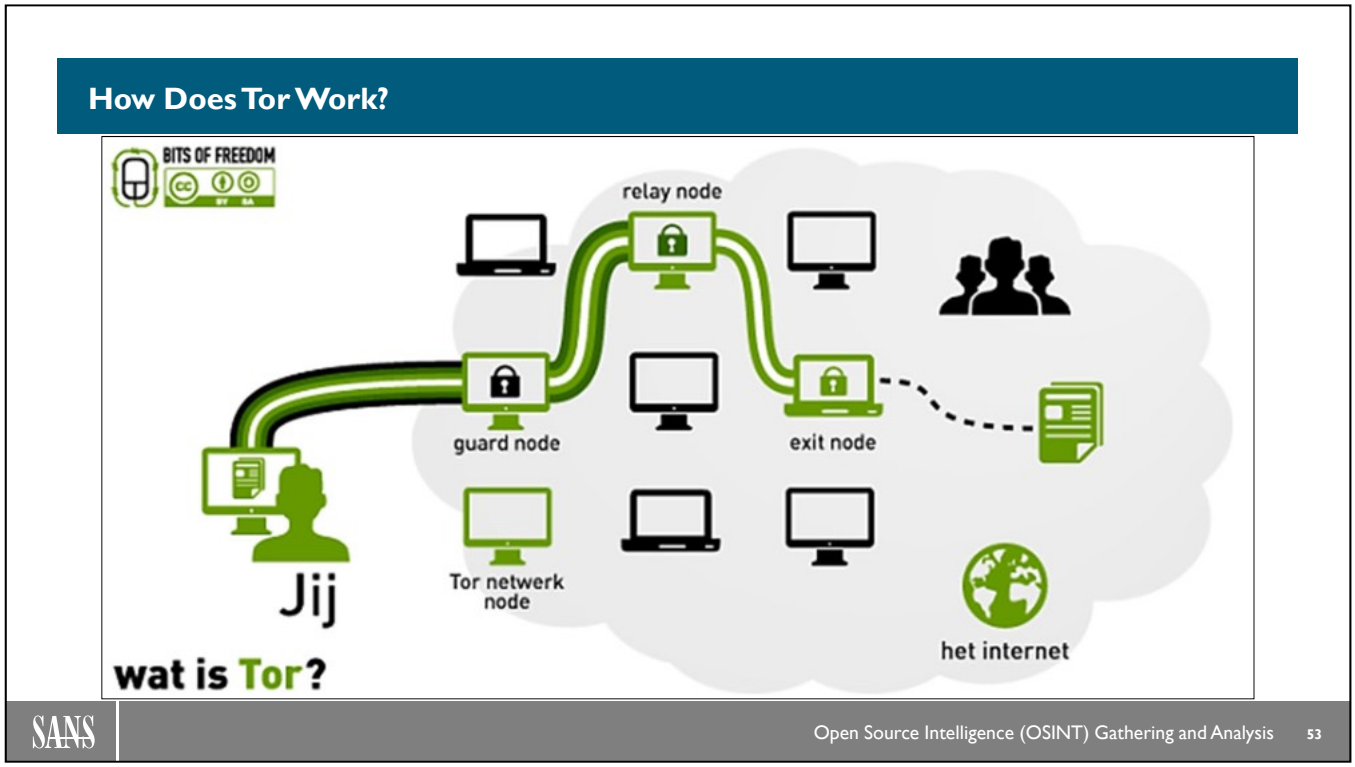
Comparing this story to how Tor works is quite simple.

First, as with my addressing each package to a certain person in the story, with Tor, the source determines the path to the destination.

Second, to maintain anonymity, only the source knows the entire path that the package (or network traffic) takes to the destination. Each step in the path unwraps the package and then sends it to the next location. In our story, each person unwrapped a package to determine the next address. With Tor, each node decrypts the traffic to reveal the next address in the path.

Finally, only the source and the destination know the content of the innermost packets since the layers of encryption protects it. In our story, only I, the source, and Princess Buttercup, the destination, knew that there were rings in the boxes.

The Electronic Freedom Foundation (EFF) has an interactive web site at https://sec487.info/87 that demonstrates how Tor works in a slightly different manner. It highlights how HTTPS should be used when sending traffic through Tor so that the final request from the exit (or last) node is encrypted when sent to the destination.



How Does Tor Work?

The image above[1] illustrates how data gets from your ("jij" is "you" in Dutch) computer on the left, through Tor, and arrives at the destination on the right. Notice that the data leaving your computer is wrapped in multiple layers of encryption (shown as thick lines). Your computer encrypts the traffic to each of the nodes before the data leaves the computer.

Your traffic flows along its source-routed path to an entry guard node.[2] These special Tor nodes help prevent data correlation attacks, whereby an attacker controlling both the entry and exit nodes of your traffic could correlate, intercept, and modify (if the data to the destination is not encrypted) your data.

The guard node decrypts the data and sends it to the next hop in the path you specified. The next (and each successive) hop is to a Tor relay node. Each node decrypts the data and passes along to the next node until it reaches an exit node. Data leaving Tor from an exit node is used to communicate with a service on the internet and may be encrypted (for example, HTTPS) or not (HTTP).

References and images:

[1] https://sec487.info/89, July 2, 2015.

[2] https://sec487.info/88, September 25, 2017.

Cautions When Using Tor

1. Careful using files downloaded from Tor; they may be malicious
2. Use HTTPS when possible
3. Stay in the dark web by using hidden services
4. Do not use personal information
5. Avoid using internet sites with logins from Tor

Cautions When Using Tor

Tor can be helpful in increasing user anonymity, but that does not mean it is a safer place than the internet. It suffers from some of the same problems we run into each day on the internet. Some of these are listed below, alongside several specific issues that privacy-minded users will want to keep in mind as they use Tor.[1]

1. Files hosted in Tor (such as PDFs, Word documents, JavaScript, or content downloaded from Torrent sites) might have malicious content in them. To protect your system from compromise or infection, ensure that you take appropriate antivirus, anti-malware, and anti-exploitation precautions before opening and using the files.
2. As mentioned previously, if you do not encrypt traffic to the destination site using an encrypted protocol such as HTTPS or SSH, your data may be interceptable and may be used against you.
3. One method to avoid data correlation attacks (where an attacker intercepts your data before it enters Tor and when it exits and then correlates user activities) is to never leave Tor. There are many internet-based services such as Facebook and DuckDuckGo that have native Tor services so that users need not exit Tor to use them. Keeping your traffic inside of the Tor network means that it is always encrypted, and attackers cannot correlate your traffic.
4. If you are trying to stay anonymous, do not use real personal information (or at least don't use YOUR personal information) when interacting with people and resources in Tor.
5. If you are using Tor to surf the internet more anonymously, avoid using internet sites that require authentication and user profiles, as they may be logged by the site and used to analyze your behaviors.

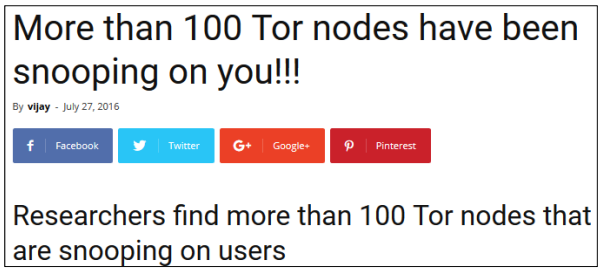
Reference:

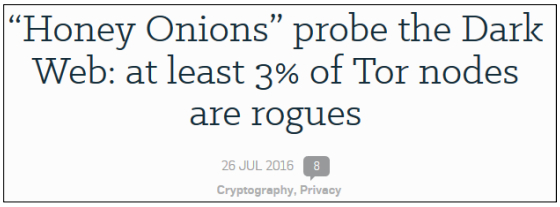
[1] https://sec487.info/89, July 2, 2015

Honey Onions (honions)

In 2016, Amirali Sanatinia and Guevara Noubir, from Northeastern University, performed some research on Tor nodes

They found 110 Tor nodes that were collecting information on the traffic passing through them for later analysis





Honey Onions (Honions)

While the Tor network can be helpful in increasing anonymity, since many criminals use it, law enforcement, intelligence agencies, and others may monitor traffic passing from one node to another to try to perform analytics on what systems are sending and receiving traffic from which other systems. In 2016, two researchers at Northeastern University found that 110 Tor nodes were collecting data that they should not be.

For our purposes, we need to be aware that no network is perfect and someone could always be watching and collecting what we do on the network.

Images from https://sec487.info/3b and https://sec487.info/3a, December 12, 2018.

Installing Tor

You can get Tor software from the Tor Project site[1] or from many macOS and Linux package managers

Standalone apps with Tor:

- Tor Browser
- Tails (Bootable USB)
- Orbot (Android)

Tor Browser is an easy, secure method of running Tor for web browsing[2]

Tails is a privacy-focused, live operating system that can be booted from or installed in a virtual machine[3]

Orbot is Tor for Android mobile devices[4]

Installing Tor

Aside from visiting the Tor Project page[1] to download the application, many package managers for Linux and macOS (brew and MacPorts)[5] also have methods of installing Tor onto a system. Tor has installers and some standalone applications for Windows, macOS, and Linux systems.

Three excellent projects that you may wish to use in your work (or personal lives) are the Tor Browser (for web browsing), Tails (for creating bootable USB devices and virtual machines), and Orbot (for use on Android mobile devices). Each of these projects makes Tor more accessible to users.

References:

[1] https://sec487.info/85

[2] https://sec487.info/83

[3] https://tails.boum.org

[4] https://sec487.info/84

[5] https://sec487.info/86

Running Tor

Tor is a proxy; a method to reach .onion and regular internet services

Many applications can use Tor as an upstream proxy

Launch Tor and send apps through or launch apps with Tor built in

If you are planning on accessing web services (hidden or not), it is highly recommended to use the Tor Browser[1]

Running Tor

When running from your computer, the Tor software acts as a local proxy, taking application information, and sending it into Tor. This process is sometimes referred to a "Tor-ifying an app" (configuring one normal application to tunnel traffic through the Tor network by using it as an upstream proxy). Tor can be used to access special "hidden" services (host names ending in the .onion top-level domain) or regular resources that can be accessed on the internet, such as https://www.facebook.com and https://www.interpol.int.

Some applications, such as the Tor Browser, come with versions of Tor built into them for ease of use and tight integration. The Tor Browser is a web browsing client built on the Mozilla Firefox ESR[2] code base and comes preinstalled with addons to help keep its users safer when browsing resources in Tor. The Tor project recommends using the Tor Browser if you are only surfing hidden services within Tor or using Tor to anonymously surf the internet.

References:

[1] https://sec487.info/8i

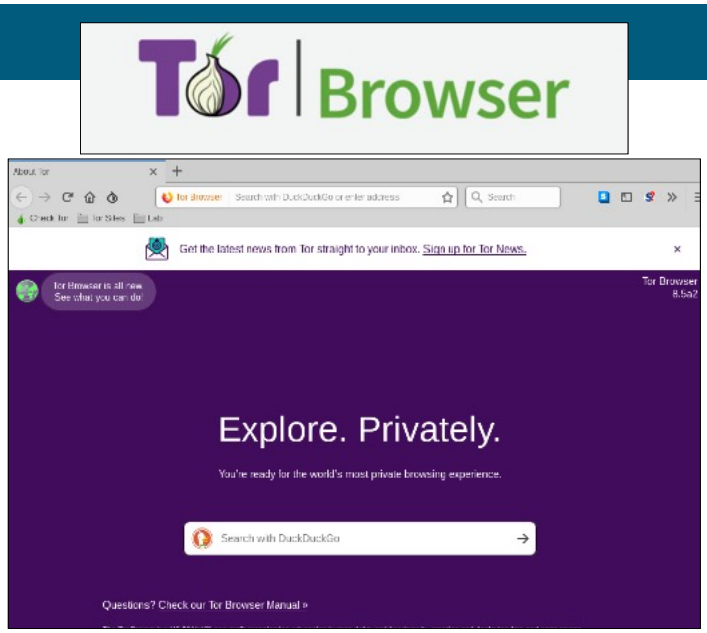
[2] https://sec487.info/83

Using Tor Browser

Uses Firefox and has installers for multiple operating systems

Ships with Tor software that launches automatically

Preconfigured with security and anonymity settings and addons



Using Tor Browser

The free Tor web browser is available at https://www.torproject.org/projects/torbrowser.html (https://sec487.info/8p) and will dramatically reduce the time and effort that you might need to put in to start browsing web pages through Tor or browsing Tor hidden services.

Detailed installation instructions for most major operating systems can be found on the project page, and once installed, the browser looks and feels exactly like Mozilla's Firefox browser. The benefits of using this self-contained browser for accessing Tor include:

1. It includes a copy of Tor that it will automatically start at launch.
2. By default, Tor Browser is configured in a more secure manner than normal Firefox.
3. Several helpful, preinstalled browser security addons (NoScript, HTTPS-Everywhere) keep you and your browsing safe and secure while you visit sites.

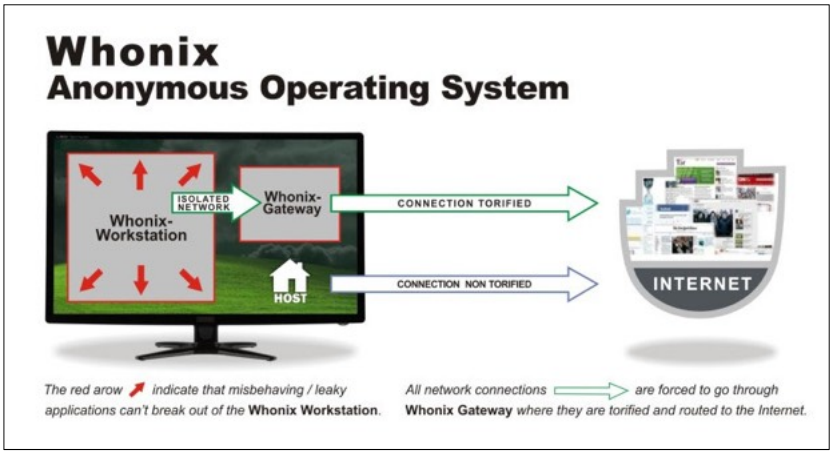
Depending upon how you installed the application, you can have Tor Browser automatically launch Tor locally on localhost:9150 and 9151. These listeners can be used by other applications to gain access to Tor.

Image from https://sec487.info/p3, October 1, 2018.

Whonix: Anonymous Operating System

For prolonged Tor use, consider the Whonix VM Gateway

Used in conjunction with a "Workstation" VM, it tunnels all the workstation's network traffic through Tor



Whonix: Anonymous Operating System

Tor Browser is a terrific choice for simple browsing in Tor or browsing internet sites through Tor. But what if you want to prevent your local Internet Service Provider (ISP) from viewing your web traffic, DNS requests, and other network services that may not be configured to send to Tor? You will probably want to use a gateway system such as Whonix.[1]

The Whonix system contains the Whonix Gateway Virtual Machine (VM) and one or more other workstation VMs where you will do your work. These workstation VMs are configured to be on a private virtual network and use the Whonix Gateway VM as their upstream network provider, just as your home computers may send their network traffic to your home router.

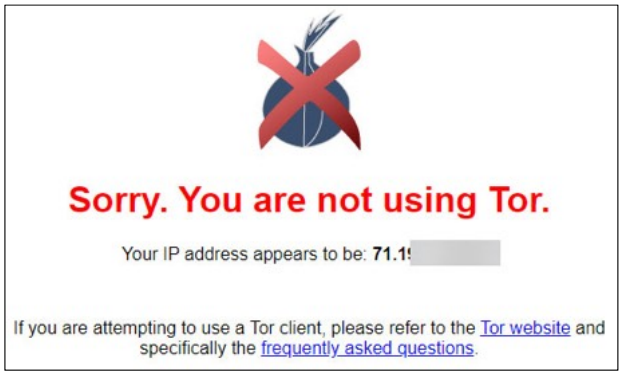
Using this VM, sending all data to another VM system prevents spurious network traffic from traversing the local network and the local ISP and helps to maintain your online anonymity.

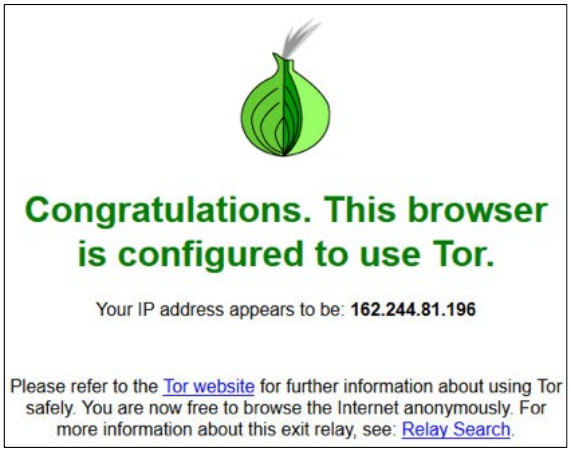
Image and Reference:

[1] https://sec487.info/8j

Checking If You Are Using Tor

https://check.torproject.org





Check If You Are Using Tor

Whether you are using the Tor Browser to surf Tor marketplaces or Google's Chrome browser sending traffic through Whonix to visit the social media page of your target, you are going to want to ensure that your traffic is actually being sent through the Tor network.

Imagine if you worked for law enforcement and had to browse to some web sites owned by your target. You would not want your law enforcement agency IP address to be logged as visiting their system. That would be poor OPSEC. So, you set up Tor and send your traffic through it. But how do you know it is actually going through Tor?

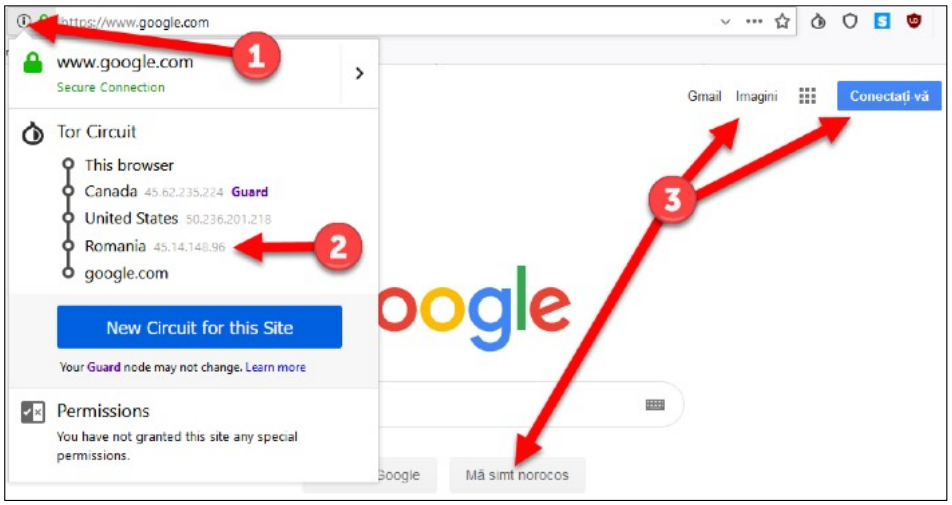
That is where the https://check.torproject.org web site comes into focus. Visit this site from a web browser, and it will show you if your IP address is a Tor exit node (indicating that you are browsing using Tor) or not (indicating you may have a misconfiguration). The above pictures are both taken from that web site when visiting without and then with Tor enabled, respectively.

Images from https://check.torproject.org, October 1, 2019.

Tor Surface Web Exit Nodes

When using Tor on the internet, sometimes the sites you visit customize your experience based on your location

Do you speak the language?



Tor Surface Web Exit Nodes

Using Tor to visit web sites on the surface web is as easy as typing in the URL bar. When you do this, Tor will figure out a route to the surface web and choose an exit node. These exit nodes (https://sec487.info/vk) are mostly located in the United States but can be in over 20 different countries. That presents us with an interesting problem. If I'm using Tor to visit surface web sites and increase my anonymity but exit Tor at an exit node in a country where I do not speak or understand the language, is this an issue?

It can be, depending on if the web site customizes my interface based upon the location of the IP address of the exit node. In the above image, we show that our Tor Browser traffic exited Tor from an exit node in Romania (arrow 2). Because Google.com, the web site we visited, customized our web page based upon the location of the IP address (Romania), we see Romanian words for the Google links and buttons (arrows at 3).

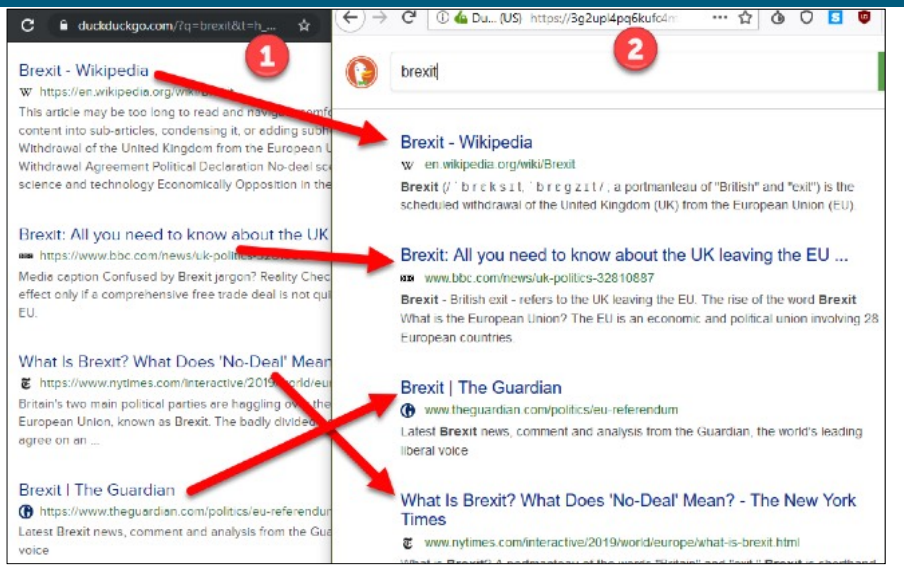
Of course, this may not be an issue for you as you might know the language, or be able to guess what the buttons and links do, or you can use a translator service.

Image from https://www.google.com, October 1, 2019.

Same Content, Different Layer

Some sites have Tor services so you can access their content from in Tor

DuckDuckGo is one of these and shows the same content from in or out of Tor



Same Content, Different Layer

One of the most risky things you can do from Tor is visit the surface web. That is because the exit node you use to reach the surface web can interfere with and monitor your traffic. To reduce the risk, many surface web sites have interfaces within Tor. When you use their Tor interface, you access the surface web content but stay secure because all your traffic is encrypted within Tor.

To prove that we are accessing the same data, in the image above, we performed a search for "Brexit" on the surface web (1) and from the Tor Browser within the DuckDuckGo.com onion service (https://3g2upl4pq6kufc4m.onion/, https://sec487.info/vm) in Tor on the right (2). You can see the content is extremely similar between these two searches. Other web sites also have Tor interfaces. Facebook.com's social media network can be reached and interacted with from their Tor site at https://www.facebookcorewwwi.onion/ (https://sec487.info/vn).

Tor Hidden Service Names - Version 2

Version 2 of the Tor software uses 16-character addresses ending in the .onion pseudo-top level-domain (TLD)

`3g2upl4pq6kufc4m.onion`

Hidden services announce themselves to Introduction Points[1]

"Tor generates an **RSA-1024 keypair**. The .onion name is computed as follows: first the **SHA1 hash** of the DER-encoded ASN.1 **public key** is calculated. Afterwards the **first half of the hash is encoded to Base32** and the suffix ".onion" is added. Therefore, .onion names can only contain the digits 2-7 and the letters a-z and are exactly 16 characters long."[2]

Tor Hidden Service Names - Version 2

Since Tor's .onion pseudo-top level domain (TLD) is not a valid surface web TLD, your computer cannot find Tor hidden services on the internet. They are, after all, "hidden" services. There are about 75,000 hidden services recorded in Tor each day.[3] Their .onion addresses are crafted using a public key that is SHA1 hashed and then Base32 encoded. The .onion TLD is added as a suffix, and you have the Tor hidden service name.

New hidden services make random connections to some relays and shares its public key. It then creates a hidden service descriptor by combining its public key along with summaries of the introduction points it shared its key with. This descriptor is then uploaded into a distributed hash table and, when the hidden service is requested using its .onion address, the hidden service descriptor describes how to reach the service.

References:

[1] https://sec487.info/8q

[2] https://sec487.info/8k, October 2, 2019

[3] https://sec487.info/8m, October 2, 2019

Tor Hidden Service Names - Version 3

Version 3 (v3) of the protocol works similarly to version 2 (v2)

From our perspective, the big changes are that v3 addresses are 56 bytes long instead of 16

v2 - `3g2upl4pq6kufc4m.onion`

v3 uses stronger, more recent crypto algorithms

v3 - `vps7nsnlz3n4ckiie5evi5oz2znes7p57gmrvundbmgat22luzd4z2id.onion`

Tor Hidden Service Names - Version 3

The Tor documentation for its hidden service names (https://sec487.info/8k) contains an overview of an update to the Tor protocol from version 2 to 3. Interested in the details of the version 3 protocol? Visit the https://sec487.info/wd page. It notes some of the main enhancements to the protocol:

"a) Better crypto (replaced SHA1/DH/RSA1024 with SHA3/ed25519/curve25519)

b) Improved directory protocol leaking less to directory servers.

c) Improved directory protocol with smaller surface for targeted attacks.

d) Better onion address security against impersonation.

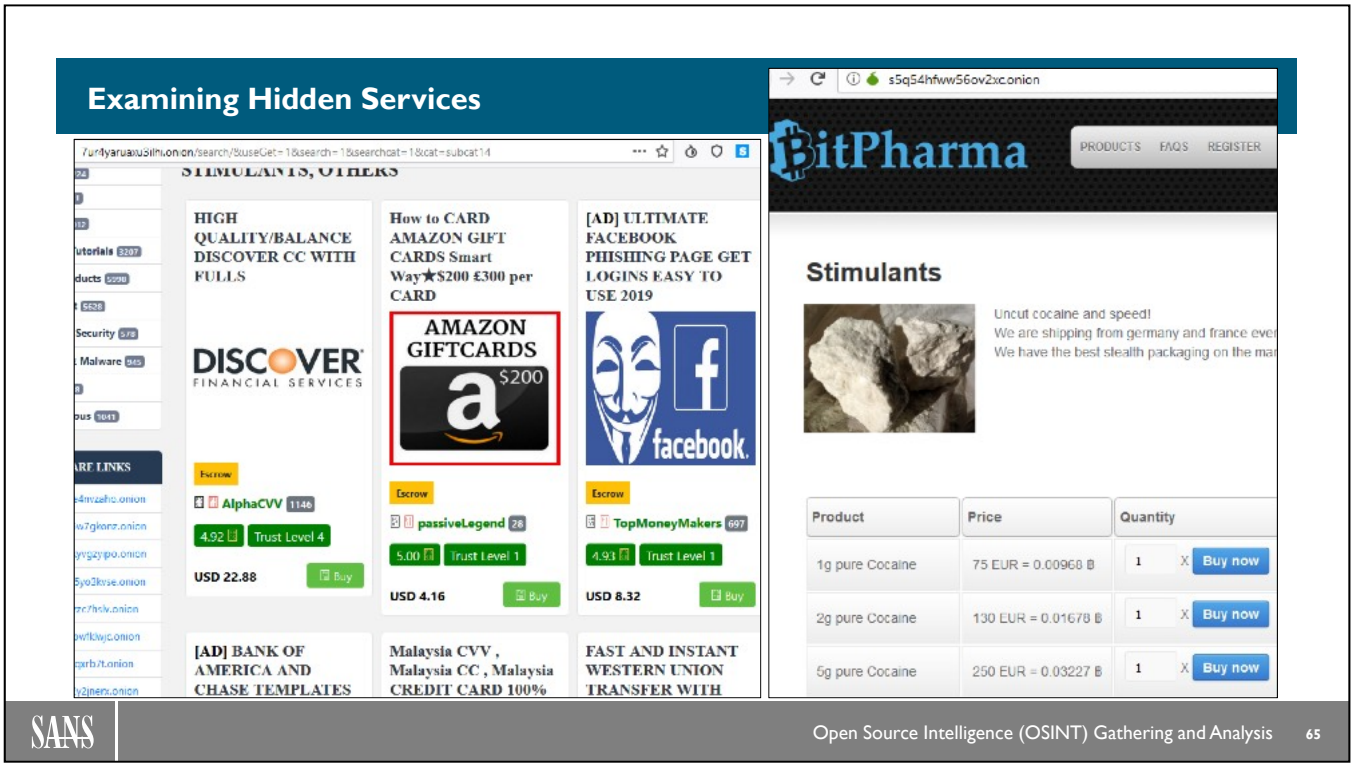
e) More extensible introduction/rendezvous protocol.

f) Offline keys for onion services

g) Advanced client authorization"[1]

From our perspective OSINT, version 3 hidden service names will be 56 bytes long instead of 16 bytes.

Reference: [1] https://sec487.info/wd, October 2, 2019.



Examining Hidden Services

While many Tor hidden services are innocuous and benign, the dark web contains marketplaces that feed on the illegal. You may have targets who are visiting these areas of Tor. As shown on the above slide, people sell nearly anything on the dark web, from credit cards and drugs to guns and "hitman" services. These marketplaces are prolific in Tor and can be located on the surface web using simple search engine queries like "darknet marketplace."

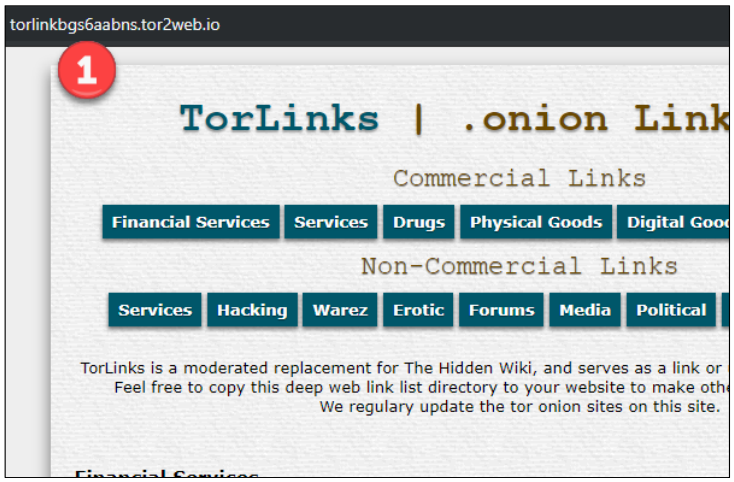
Images from http://s5q54hfww56ov2xc.onion/ and https://sec487.info/vl, October 1, 2019.

Accessing Hidden Services from Internet

Need to check a Tor site quickly in your non-Tor-ified browser? Use https://www.tor2web.org

OK with losing your anonymity?

Replace `.onion` with `.tor2web.io`



Accessing Hidden Services from Internet

While Tor provides anonymity for users visiting their hidden services (.onion sites), there is a small barrier to being able to access those sites. It might be that you are at a client site and they will not allow you to run a virtual machine or install the Tor Browser on their systems. Or perhaps you are at a place where Tor protocols are filtered at a firewall, so you are having challenges accessing Tor networks. Don't give up. There are several services on the internet that will allow you to browse hidden service sites without configuring Tor on your systems!

To use this service, replace the ".onion" in the Tor address with ".tor2web.io". In the example above, we took the http://torlinkbgs6aabns.onion/ site and visited it through the Tor2Web application by visiting https://torlinkbgs6aabns.tor2web.io in our surface web browser (1).

As you might imagine, our operational security decreases dramatically when we use these proxy services, as they can monitor where we are visiting and alter page content. In fact, we know they alter page content because the links on the returned pages have their links rewritten to use the hiddenservice.net, onion.guide, onion.city, onion.to, or onion.cab domain instead of the normal .onion or .i2p. So be careful when using them.

Image from https://sec487.info/w8, October 2, 2019.

Hunchly Daily Hidden Services Report

Justin Seitz runs crawler software in the Tor network and discovers new hidden services

Output can be emailed or downloaded

Sort and see patterns

	A	B	C	E
1	Last Contacted	Hidden Service	Language	Title
134	2019-10-01 09:01:47	ts4cwattzgsiitv7.onion	en	EasyCoin Bitcoin Wallet and free Bitcoin
135	2019-10-01 06:16:12	walletbjvdecnjgp.onion	en	EasyCoin Bitcoin Wallet and free Bitcoin
136	2019-10-01 21:48:55	easyko3zrfooqrpo.onion	en	EasyCoin Bitcoin Wallet and free Bitcoin
137	2019-10-01 23:10:47	easycoinsayj7p5l.onion	en	EasyCoin Bitcoin Wallet and free Bitcoin
138	2019-10-01 23:02:20	lhag24lbnfved7wu.onion	en	EasyCoin Bitcoin Wallet and free Bitcoin
139	2019-10-01 12:12:59	spacechadxxpkf6t.onion	en	Spacechadxxpkf6t[[.]]onion - The Bitcoi
140	2019-10-01 17:08:12	dwmjk4z7inup76jg.onion	uk	Реєстр актів харківської міської ради
141	2019-10-01 06:30:29	beast7ruvpc3qjhv.onion	en	!--!
142	2019-10-01 17:31:41	rbaco5flcou46wpd.onion	en	"Welcome To Dark Web Links & More!
143	2019-10-01 15:34:00	srx5xrpt5fuo33ot.onion	en	"Welcome To Dark Web Links & More!
144	2019-10-01 07:50:22	n2ha26oplph454e6.onion	en	"Welcome To Dark Web Links & More!
145	2019-10-01 20:35:23	jdpskjmgy6kk4urv.onion	en	"Welcome To Dark Web Links & More!
146	2019-10-01 15:48:36	youbroketheinternet.cheettyiapsyciev	en	#youbroketheinternet (YBTI)
147	2019-10-01 15:05:13	ybti.cheettyiapsyciew.onion	en	#youbroketheinternet (YBTI)
148	2019-10-01 05:30:44	*youbroketheinternet@loupsycedyglg	en	#youbroketheinternet (YBTI)
149	2019-10-01 21:44:45	cheettyiapsyciew.onion	en	#youbroketheinternet (YBTI)
150	2019-10-01 11:11:16	loupsycedyglgamf.onion	en	#youbroketheinternet (YBTI)
151	2019-10-01 17:10:01	zzq7gpluliw6iq7l.onion	en	$$ The Green Machine $$ - Index

Hunchly Daily Hidden Services Report

Justin Seitz runs several dark web (Tor) scanners that discover new hidden services each day. His tools note them and release daily Excel-formatted XLSX files via email (if you subscribe to https://sec487.info/w9) and via his Hunchly Twitter account (https://twitter.com/Hunchly). New hidden services can be important to monitor and examine for your work.

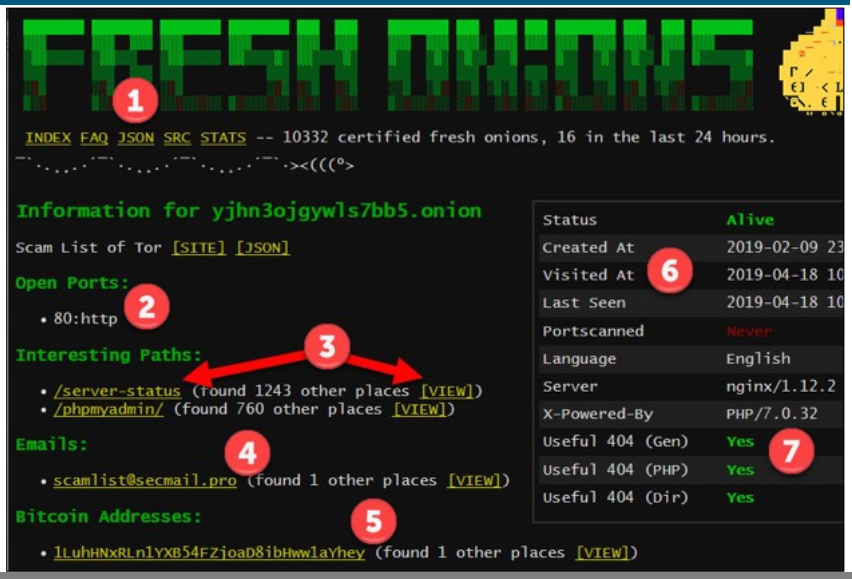
The files are also useful for finding similar sites. Above, we used Microsoft Excel to filter and sort the results of the "Up" tab (the one that notes what services were found to be running on a given day). By sorting the output by the title of the web site, we can see several groups (1, 2, and 3) where the site names are identical. This may be important for your investigations, as it could tie other onion addresses to a single owner.

Fresh Onions Hidden Service

Retrieves information about hidden services from a variety of sources

Gathers data from the onion system

Correlates SSH keys, web server banner, interesting locations



Fresh Onions Hidden Service

The Fresh Onions hidden service within Tor contains historical data about hidden services. It uses a variety of sites and techniques to discover new onion sites, scan them, and record bits of data that may help you in your work. We can use this site to see what hidden services may be connected either through using the same SSH key (indicating it may be a clone or created from an image), through web links in HTML pages, and through data such as emails (4) and bitcoin addresses (5). Each of these areas could be researched in depth to discover other servers that may be related to this system. For instance, if three different onion services all note that payments should be made to the same bitcoin wallet, that may inform you of a relationship between those sites.

This data that Fresh Onions collects is also available for downloading if you would like to retain a copy for your work or are interested in performing research on the hidden services it found. We placed a "1" on top of the JSON link in the above image. Click that, and the JSON record you are viewing will be downloaded.

Fresh Onions' Tor address changes but was http://vps7nsnlz3n4ckiie5evi5oz2znes7p57gmrvundbmgat22luzd4z2id.onion in October 2019.

Image from https://sec487.info/vp, October 1, 2019.

Fresh Onions Interesting Paths

Fresh Onions notes web places we may wish to visit as "interesting paths" (1)

In this case, the `/server-status`

There are 1,243 other systems with this same path (2)

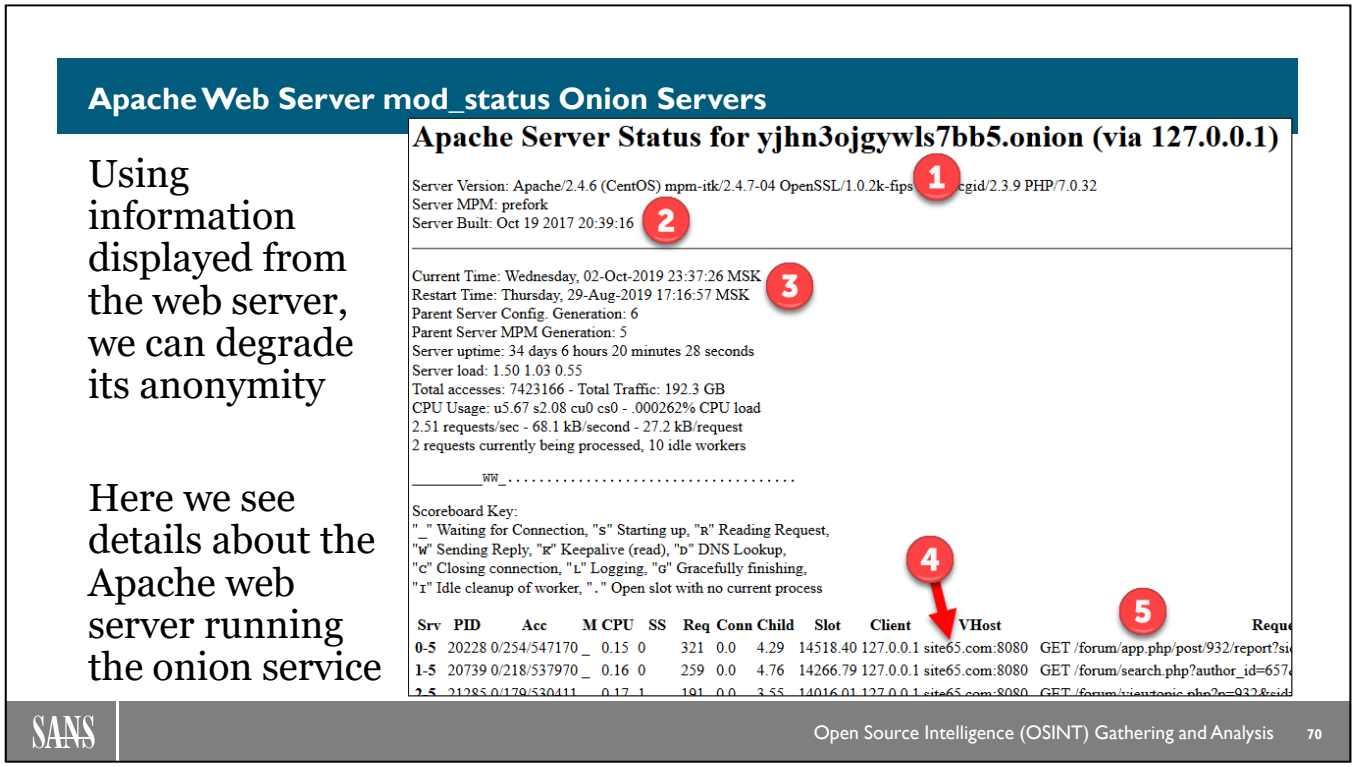
Fresh Onions Interesting Paths

Fresh Onions helps point out other places that we may wish to research, such as the Interesting Paths section of the web page. These are interesting directories that the Fresh Onions scanners found on a web server at the destination onion service. Let's visit the path and see what the web server shows.

There are other interesting web directories that will show up in this section. All of them should be researched and investigated, as they could tie servers together and be a source for leaked data. The directory /server-info may show parts of the web server's configuration file (something that should not be web accessible). Let's look deeper at this specific /server-status path in the upcoming slide.

For more information about these two specific web directories, check out Micah Hoffman's blog post on The OSINT Curious Project web site at https://sec487.info/wc.

Image from https://sec487.info/vp, October 1, 2019



Apache Web Server mod_status Onion Servers

When installing and configuring a web server like the free Apache Web Server (https://httpd.apache.org), it is common to harden the system's configuration (https://sec487.info/wa). This is because many of the debugging features that system administrators might enable when installing the server for the first time to check if certain features and functions work, if left enabled, would leak server and site data to people who should not see it. Disabling these web server features in places where anonymity should be high is important.

The previous Fresh Onions slide showed us the /server-status interesting path. Here, we visited that location. What we see destroys the anonymity of the onion server. We see:

1. The server is configured for the onion service.
2. The underlying operating system, Apache web server, and plugin versions, which may be vulnerable to exploitation.
3. The current time on the server, which could show us where the server is located in the world.
4. The Virtual Host (or VHost) is what web sites this server is hosting. Here, we see the "site65.com" site running on an onion service. So this server that is running the Tor site is also running the site65.com web site! If you wanted to learn more about this system, you could now pivot to the surface web and pivot to researching that domain.
5. We also view what pages and parameters people are submitting to the server. This discloses pages within the site and may reveal sensitive parameters like usernames, passwords, and session IDs that may be passed in the URL.

Image from https://sec487.info/wb, October 2, 2019.

Tor Summary

1. Proxy-based and cross-platform
2. Goal is to increase anonymity
3. Circuit created between relays and nodes
4. Traffic is source-routed and encrypted within Tor
5. Local files on your system and in apps are how you configure Tor and the Tor apps
6. Web browse, file share, create content, email, forums

Tor Summary

So, to recap this overview of Tor:

1. It is a proxy-based application that is cross-platform compatible.
2. The goal is to increase anonymity of the users and service owners, whether they are accessing services on the internet or within Tor.
3. Tor circuits are temporal and create hops through relays and nodes.
4. Traffic from client applications is source-routed and encrypted, as long as the information stays within the Tor network.
5. Depending on what software and platform you use to access Tor and its hidden services, you will have a variety of methods of configuring how circuits are created and destroyed, special hidden service authentication, and more.
6. Web browse, file share, create content, email, and forums are many of the features of this dark web network.

SEC487 Workbook

Please visit the course electronic workbook in the 487 virtual machine and begin the exercise named:

"Tor"

This page intentionally left blank.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. Freenet
4. I2P - Invisible Internet Project
5. Tor
6. **Monitoring and Alerting**
7. International Issues
8. Vehicle Searches
9. Putting It All Together

This page intentionally left blank.

Monitoring Data on the Internet

Blue team OSINT functions may include monitoring:

- Specific domains
- Keywords (sentiment, projects, etc.)
- People inside or outside their company

Who/what is your "target"?

- Your own organization?
- An adversary?
- Your organization's "secret" project?
- A honey token/file/folder?

Instead of doing manual queries, let's set alerts and have the data come to us

Monitoring Data on the Internet

There will come times when your organization or clients need you to monitor sentiment, a person, a group, or a term over time. This is where internet monitoring and alerting comes into focus for our OSINT work. Depending upon your client's/organization's goals, you might be tasked with:

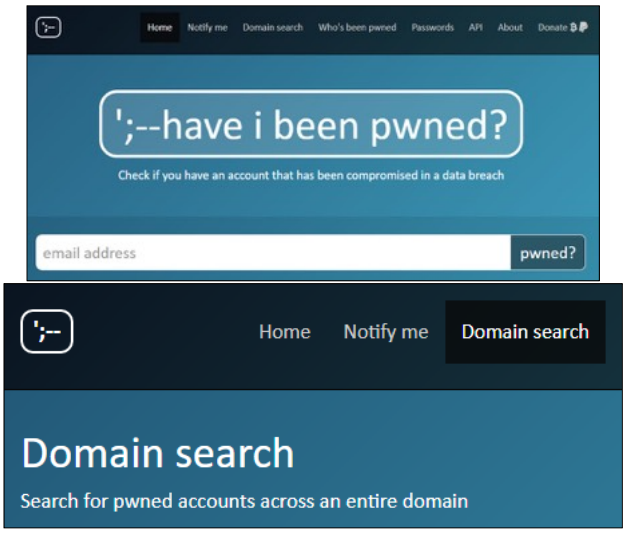
- Examining certain web sites for changes
- Finding out if your company's data has leaked to paste sites
- Monitoring activities of groups or people via their social media and news reports
- Seeing what your company's competitors are doing

Before setting up these alerts, you need to determine the terms, web sites, and groups that you will monitor and then choose platforms that can assist your efforts.

Compromised User Accounts - HIBP

The haveibeenpwned.com site, run by Troy Hunt (@troyhunt), collects dumped usernames and passwords

Can search emails, register to be notified, use API, or register @yourdomain and search for all your organization's user accounts



Compromised User Accounts - HIBP

Troy Hunt (@troyhunt) created and runs the HaveIBeenPwned.com web site. It is an amazing free resource. When web sites or applications get compromised and the usernames and passwords are stolen, Troy Hunt gets those dumped credentials and puts them into HIBP (Have I Been Pwned). While the application does not allow people see what the passwords for specific users are, you can search for email addresses to find out if that email was found in any dumps. Additionally, HIBP has an API that can be used to make scripted calls to the site in a rapid fashion.

If you have control over a domain (say, sec487.info), you can register to receive emails whenever new credentials are imported into HIBP that are from that domain. Work with an organization's security team to have them register all their email domains in HIBP with email alerts getting sent to the CIRT or SOC.

Image from https://haveibeenpwned.com, March 31, 2018.

Why Use HIBP?

OSINT

- Find valid emails
- Find emails that have been disclosed in breaches

Personal

- Find out when the credentials for people you care about may have been breached

For Blue Team/Privacy

- Know when you or your organization's data has been dumped
- Protect your users

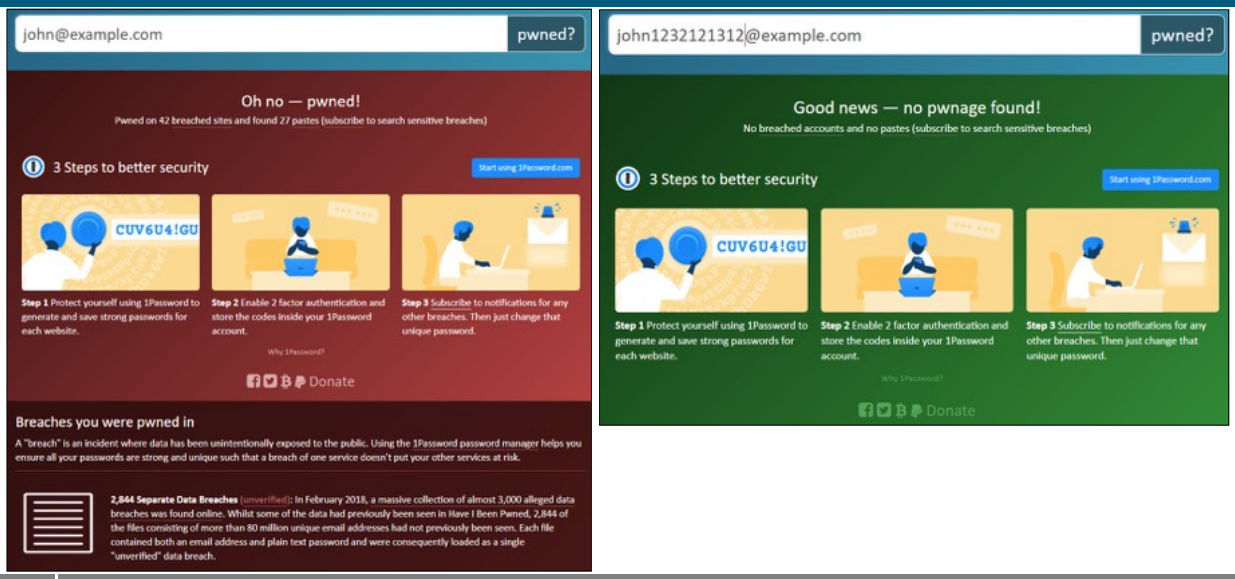
For Red Team/Offensive

- Locate possible accounts for compromise through password reuse
- Phishing email gathering

Why Use HIBP?

The HaveIBeenPwned breach search and notification system is a valuable asset to OSINT analysts, blue team, red team, and even to every one of us personally. The information we glean from this site and the rapidity that it sends out breach notification emails allows us to protect ourselves, our loved ones, and our users from attackers who may use their breached usernames, emails, and passwords against them.

HIBP Results



HIBP Results

In this slide we show a positive result, the email address WAS found in one or more breaches, on the left, and a negative result where the email account has not been found in breaches collected by HIBP, on the right. If the email address is discovered in one or more breaches, farther down the web page will be the names of the breaches where it was discovered and information about the scope and circumstances around each specific breach.

HIBP also has an API that can be used to rapidly retrieve data from the site. In 2019, to stem the abuse of the API, Troy Hunt changed the API from a free service to a paid one costing $3.50/month. More details about the API are found at https://sec487.info/w1.

Images from https://haveibeenpwned.com, January 6, 2019.

Password Lookup Sites and OSINT

When data is stolen from a system, it may be released to the internet

People collect these "dumps"

Each site dumped can contain:

- User names
- Emails
- Passwords

Could we look at these dumps in bulk and:

- If you know an email, find passwords
- If you know email, find other accounts for this user
- If you have password, find other accounts that use the same password

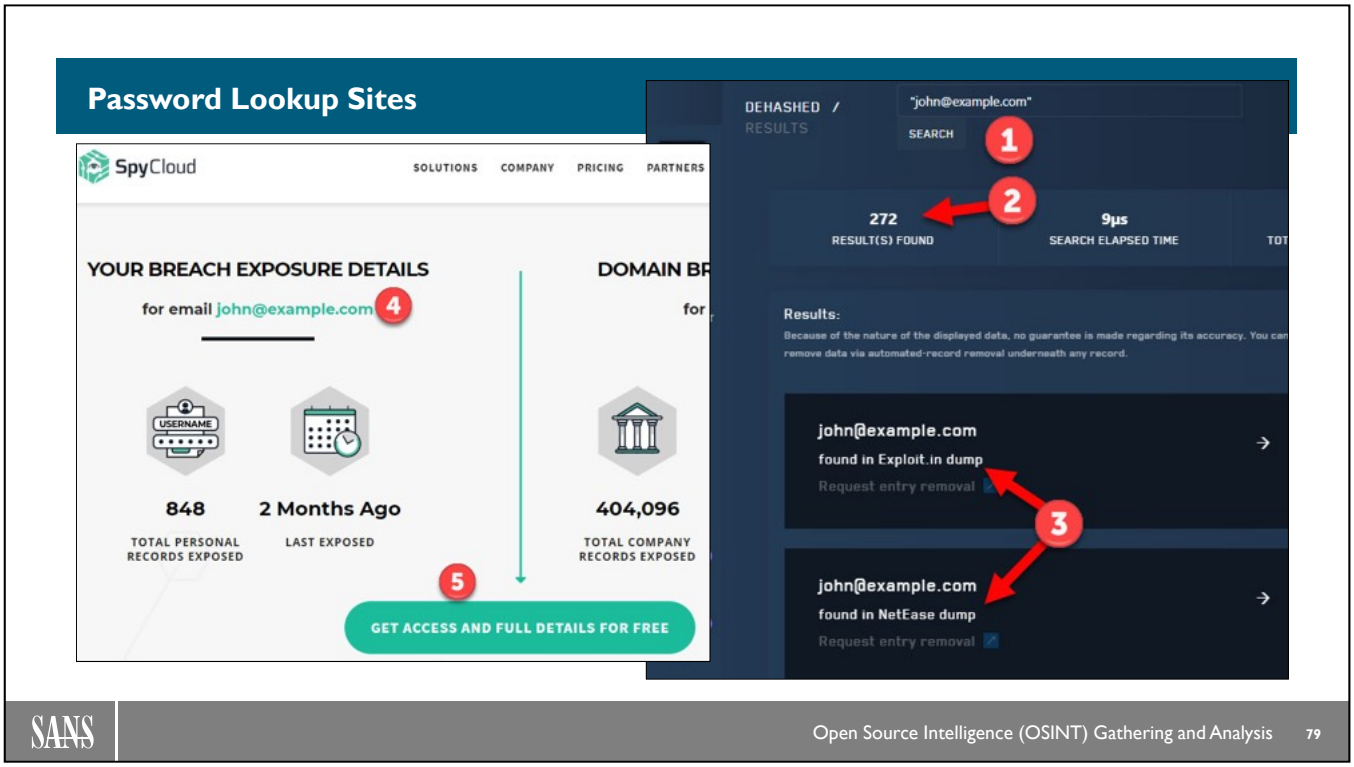
Is it ethical to use stolen data?

Password Lookup Sites and OSINT

When criminals break into a web application and steal user names, emails, and passwords, many times they will release that data on sharing websites. Others will retrieve the dumped data and store it. Later, when someone wants to compromise an account on a specific system or of a certain user, they will open those data dumps and search for an appropriate breached set of credentials. Reusing the user names and passwords on other web sites may yield valid logins if the user has used the same credentials across different sites.

We can use this information for OSINT too. Understanding how data is related can be helpful in our search for other accounts that a user has where they used a specific email. Or, if the user made a unique password that few others use, we might search in these dumps for other accounts (email addresses) where that specific password was used. This could give us unrelated email accounts where this target used the same password.

One last thing to think about is whether it is ethical to search this compromised data. Someone stole the data from a private system and has released it to the public. Are you and your organization OK with using stolen data in your OSINT?



Password Lookup Sites

We used the iconic "john@example.com" email address for searches on two breach collector sites (1) (4). The Spycloud.com and dehashed.com web sites allow users to perform free searches of their systems to find records from data dumps. If you want to see what records were found on the Spycloud web site, the "GET ACCESS AND FULL DETAILS FOR FREE" button (5) will generate an email to the email address you submitted with more details. This may not be desired if you are looking up other people's information on the site.

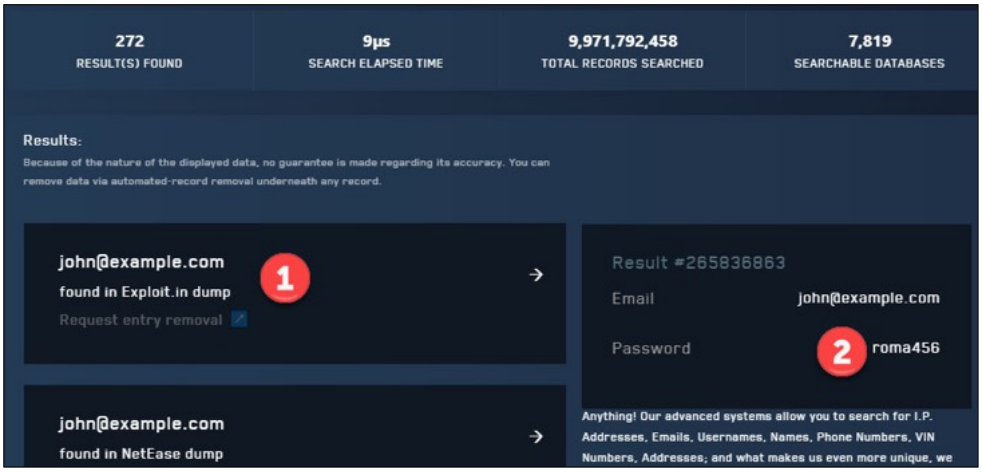
Dehashed.com shows 272 dumps the email was found in (2) (3) and will reveal the passwords if you sign up for a paid account.

Images from https://sec487.info/ph and https://sec487.info/pi, October 1, 2019.

Dehashed Results

For paid accounts, Dehashed reveals passwords

We can then search Dehashed for those passwords



Dehashed Results

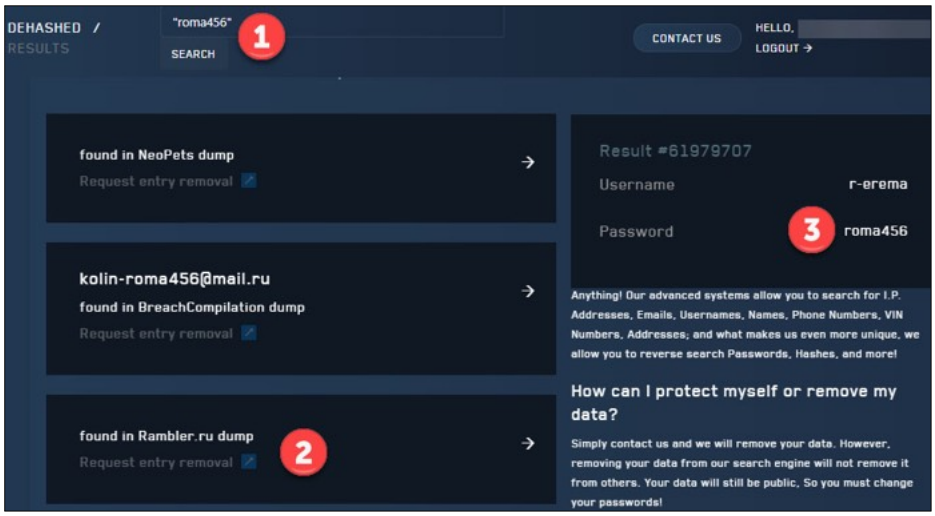
If you purchase a paid plan for Dehashed, you gain access to the clear-text or hashed (depending on the dump data) passwords for the user accounts. In this case, we clicked on the john@example.com account found in the Exploit.in dump (1) and revealed a password of "roma456" (2). Let's take that password and find other accounts that Dehashed knows about with the same credentials.

Image from https://sec487.info/pi, October 1, 2019

Dehashed Pivot on Password

We see that the"roma456" string was used in other emails and passwords

We would need to false-positive test these results to see if they are our target



Dehashed Pivot on Password

Taking the "roma456" password found in the john@example.com record from the previous search, we pivot and search on that string. Dehashed shows results where that string was found in email addresses and in other account passwords. This could indicate that these accounts were set up or maintained by the same person, or it may be a coincidence that another person chose that same password for their accounts. As OSINT analysts, we would then try to confirm or deny that those new accounts were our target's accounts.

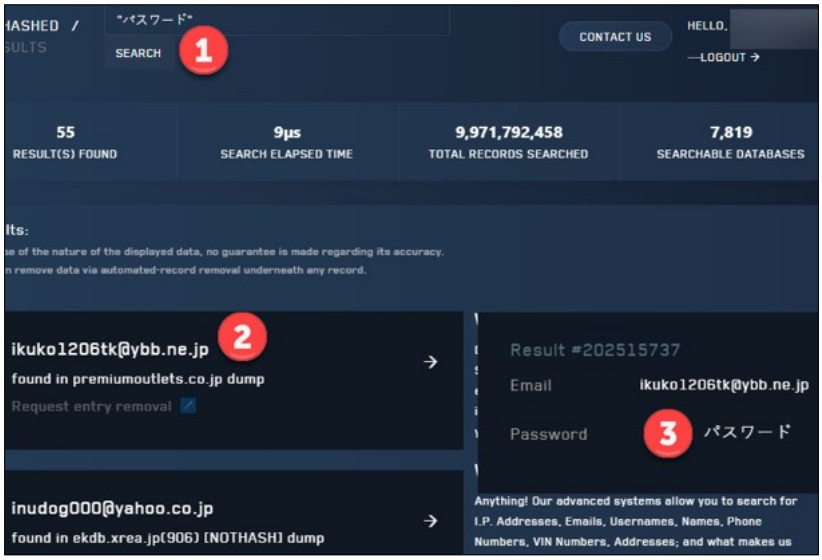
The image above shows us searching for the "roma456" string (1), with several results displayed below it. Clicking on the "Rambler.ru" dump (2), we reveal the password used by the account "r-erema" was our string, "roma456".

Image from https://sec487.info/w2, October 1, 2019.

Foreign Character Sets?

Breach data contains international character sets

Here we search Dehashed for the Japanese word for "password" (パスワード) (1) and receive Japanese TLD emails (2)



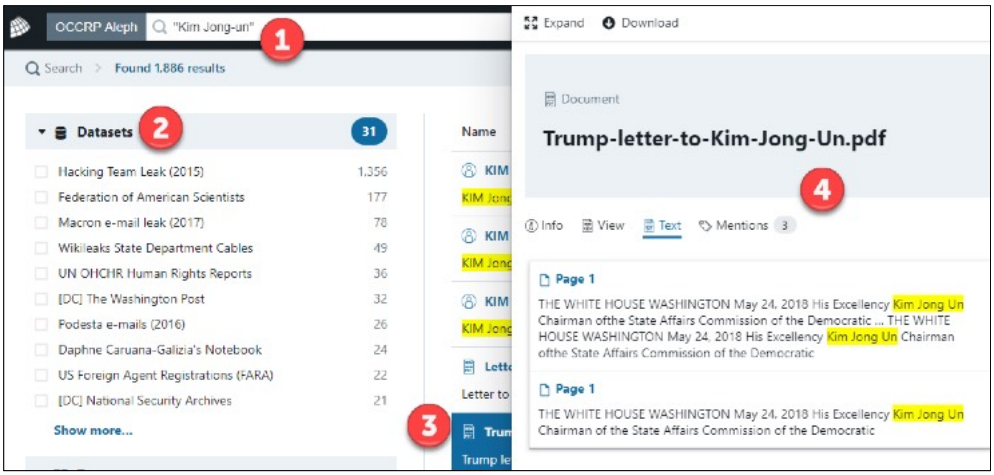
Foreign Character Sets

Many of you already interact with the world in non-English character sets, using either your native language or your target's language for your OSINT work. Dehashed and other breach sites contain these non-English characters, too. In the above slide, we translated the English word "password" into Japanese (https://sec487.info/w6). Then we used those characters to search for entries in Dehashed (1). We found 55 entries that had that string in their records. When we clicked on an entry (2), we found the password stored matched our translation for the word "password" (3). We also see the entries that were returned have the top-level domain (TLD) of ".jp," which is for Japan. It should come as no surprise to you that people in Japan use Japanese characters for their passwords.

Image from https://sec487.info/w5, October 2, 2019.

OCCRP Data

A journalist site with a database of sources, from FBI FOIA requests to Panama Papers to Russian court documents



OCCRP Data

Founded in 2006, the Organized Crime and Corruption Reporting Project (OCCRP) is designed to facilitate storage and sharing of sources for journalists investigating crime and corruption around the world. At its core, the online system contains thousands of breached emails, stolen documents, public data, and sources that can assist in investigative work. Search for people, events, locations, and other topics of interest and retrieve text matches, and then retrieve the documents and emails that the matches were discovered in.

In the above image, we searched the database (https://aleph.occrp.org) for the term "Kim Jong-un" (1). You can see the data sets that were returned (2) contain email leak data, breach data, and potentially stolen content too. Selecting a result (3), we can pull up the text of the document (4) and parse the data for information related to our OSINT work.

Signing up for a free account on the site allows greater access to the data, system alerts, and an API key for automated access to materials.

Image from https://sec487.info/w7, October 2, 2019.

What Are Paste Sites?

Paste sites are places for people to post text documents that they wish to share

Paste sites are a class of web application

Most of these sites allow anonymous posting, but some have accounts with advanced features

How it works:

- Someone posts content
- They get a URL
- They share the URL to the paste with others
- Others access and get content

Mostly public and relies on security through obscurity, as the URL to a given paste is usually hard to guess

What Are Paste Sites?

Plainly put, paste sites are places for people to post text content and share it with others in a largely anonymous fashion. While the https://pastebin.com web application is one of the more well-known paste sites, there are many other web applications that do this text-based content sharing.

Content posted to these web applications is made public, although some sites allow for hidden pastes. When someone pastes content into one of these sites, they get a unique URL to share with others. When those people use the URL, they see the original content. The URLs used usually are random characters like "Zv4baTPD" that, when placed into the correct URL, will retrieve the content requested.

On most of the paste sites, there is no requirement for authentication. Someone clicks a button to create new content, a page is generated, they paste the content, and then it gets posted so the public can read it. Some of these web applications allow users to purchase "pro" accounts with additional features. These use authentication.

What Do People Share?

- Programming code
- Lists of
  - Subnets/IP addresses
  - Hosts and domains
- Manifestos (Lulzsec)
- Stolen data[1]
  - Licenses, API keys, passwords
  - Internal documents
- IM/IRC conversations

What should they not share?

What Do People Share?

People can share nearly any text they wish on these paste sites, but in general we see more of the above types of content being shared on the sites. Senior SANS instructor Lenny Zeltser wrote a blog post about the content that he saw on the https://pastebin.com site and how he found people were using the site.[1]

Posting text content seems like it could not hurt anyone. After all, it is just text. There is no place for malware to hide and execute. But the content of some of the pastes can be quite damaging to people and organizations. The Pastebin.com web site displays the list of things that should not be posted to the site (shown on the right of the slide). If someone reports this content, Pastebin.com may take it offline and ban the user/IP address that posted it. But those consequences are very slight for people who violate these Terms Of Service. Most people post anonymously anyway, and IP addresses can be changed by using a VPN or Tor.

Image from https://pastebin.com/faq#1, November 29, 2017.

Reference: [1] https://sec487.info/ch

Pastes Are Not Permanent

Pastes are temporal

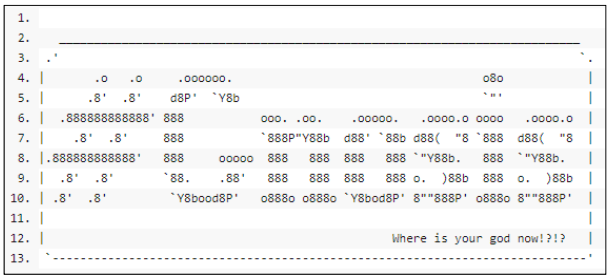
They can:

- Expire at the user's wish
- Be deleted (with paid accounts)
- Be removed for violating TOS

Public content sometimes gets backed up at archive.org

Paste that has been removed:
https://pastebin.com/Zv4baTPD

Archive.org version of it:
https://sec487.info/ci



Pastes Are Not Permanent

Pasted content can be posted with or without a self-destruction date/time when it will be automatically deleted from the site. Users with paid accounts might choose to delete their paste. Other times, the site maintainers may remove content that violates the site's Terms Of Service (TOS) or that someone else has request be taken down.

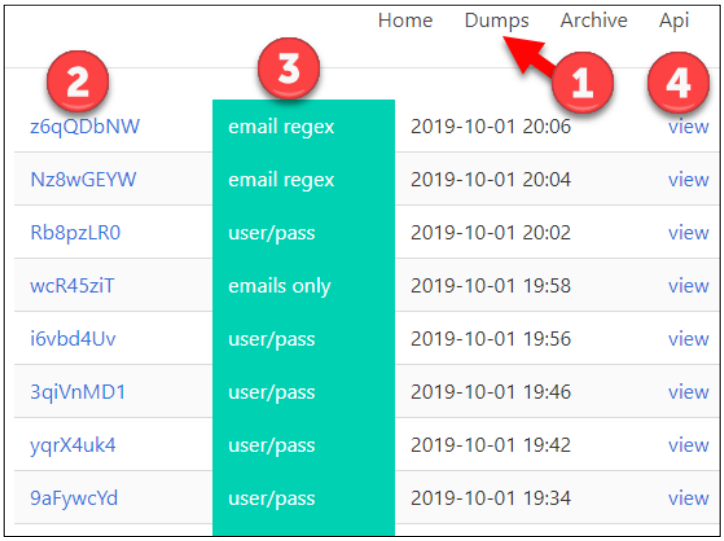
Remember that this content, most of the pastes anyway, are public and, as such, may be indexed by Google and other search engines. Google may contain a cached copy of the content that was removed. Alternatively, sites like archive.org also may have been asked to back up the content of the paste, as was the case in the 2010 Gnosis dump of the Gawker.com data. The original paste (http://pastebin.com/Zv4baTPD) is no longer hosted on the Pastebin.com site. However, looking that site up on the Archive.org one, we see that it was archived many times and contains the data that was removed (https://sec487.info/ci).

Pastebin Dump Collection Site

Since Pastebin pastes are public, other sites watch, harvest, and store them

Third-party sites like **psbdmp.ws** index and archive pastes for us to retrieve and analyze

Free, unauthenticated API



Pastebin Dump Collection Site

The https://psbdmp.ws site scans, monitors, and retrieves data from Pastebin.com. It works because most of Pastebin's pastes are public and can be retrieved. This site indexes and stores copies of the pastes. We can visit the "Dumps" page (arrow 1 above) and see a list of the recent pastes that the site has found (2).

Clicking on the hyperlinks in the "2" column on the left will take the user right to the Pastebin site. If the paste is still viewable, it should load. Column 3 above shows the types of data that psbdmp.ws has detected in the paste. This is a rough guess based upon the matching of terms and patterns. Finally, clicking the "view" links in column 4 above would take the user to the cached copy of the paste that is archived on the psbdmp.ws site.

As shown above, the site also allows users to retrieve data via their public API (https://psbdmp.ws/api). The API is URL-based and requires no authentication. It returns JSON, which makes it an excellent resource for tools. We not only can retrieve whole dumps via the API but can also search their dumps. The https://sec487.info/w0 link will search for the "barack@whitehouse.gov" email in the archives and return JSON results with pastes that can then be retrieved.

Image from https://psbdmp.ws/dumps, October 1, 2019.

Searching, Scraping, and Monitoring

`https://pastebin.com/search?q=webbreacher`

In 2011, the Corelan team created the pastenum[1] pastebin enumerator tool

The pastehunter[2] project uses the Elastic Search product and a Kibana dashboard to store and display retrieved data

Justin Seitz (@jms_dot_py) created a great blog post[3] called "Building a Keyword Monitoring Pipeline with Python, Pastebin and Searx," which is worth a read if you are technical and want a robust monitoring system

Pastebin.com has a pro account where you can set up to 15 alerts

Searching, Scraping, and Monitoring

When we think about accessing the data on Pastebin.com, our techniques fall into roughly three categories: we can search for content, we can scrape content from the pages, and we can persistently monitor the site to see if a term shows up in new pastes.

Searching the paste sites like pastebin.com and others, can be as simple as using the site's own search fields, as is shown above. Sometimes the sites' pastes are indexed by search engines (which we will see in a coming slide), and we can use DuckDuckGo or Google to search them. Alternatively, we can use tools such as the Corelan team's pastenum or the pastehunter tool to search for and download content of our choosing.

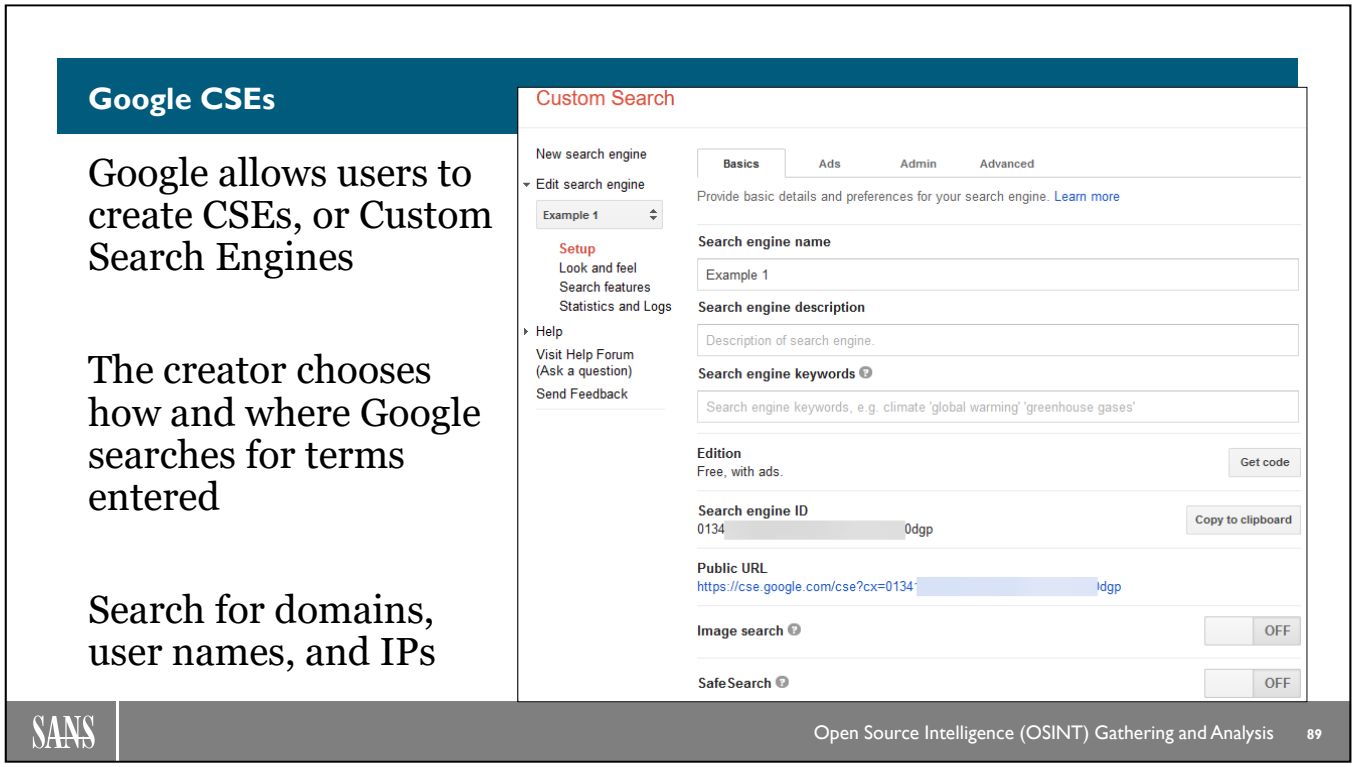
Finally, if you are looking to monitor these sites long term to see if certain terms are submitted in pastes, then you have a couple choices. Some of the sites, such as Pastebin.com, offer paid accounts that allow users to enter keywords that they will be alerted on if the site sees those terms in new pastes. The other, more technical solution, is using Justin Seitz's Python and Searx system to constantly monitor a site.

References:

[1] https://sec487.info/cj

[2] https://sec487.info/ck

[3] https://sec487.info/cl



Google CSEs

Google allows its users to create Custom Search Engines (CSEs) to perform specific searches of certain sites (https://sec487.info/vx). The creator of the CSE sets a variety of parameters when configuring the CSE. Anyone with a valid Google account can create their own CSE and tailor it to return results that are important to them. As shown in the slide above, many of the features that can be enabled and customized in the CSEs are the same parameters that you might customize in a Google Dork.

Stephanie Proto (@sprp77) has a HUGE love of all things CSE. She has created hundreds of CSEs and tweets new ones often at https://sec487.info/py.

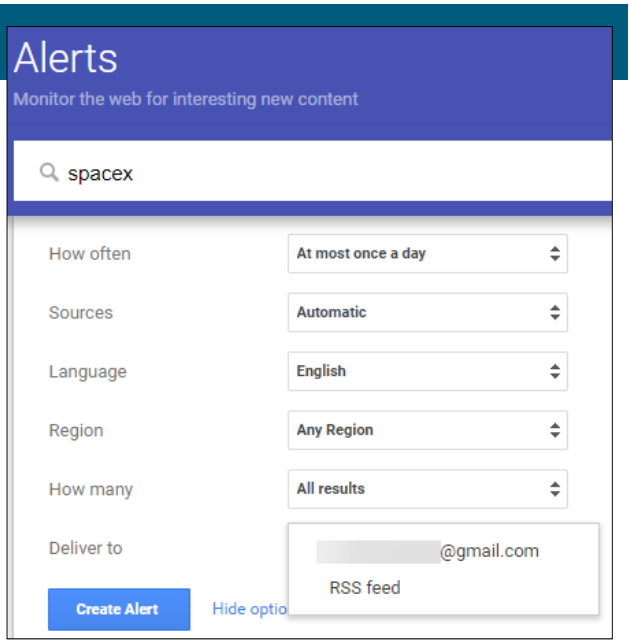
The main OSINT issue we have with these Google Custom Search Engines is that each one is a "black box" that you do not know how it is configured and what sites it searches. If you think CSEs might be useful for your team and clients, consider making your own.

Google Alerts

Sign into a Google account

Set up alerts so, when Google indexes your keywords, you get emails

Can choose what language, region, and sources the data may be harvested from



Google Alerts

Create a Google user account (email address), and you can start using its alerts to get notified when Google indexes content that you are interested in. Enter search terms, such as "sec487" or "Inigo Montoya" and, when Google finds that content online, it will alert your Gmail account to the new hit. You can customize the language, region, and source where the data must come from for the alert to fire. This can help narrow the searches so that they give you and your team more-relevant data.

In fact, you can enter any valid Google search in this field. So if you wanted to be alerted when a new PDF file named "secretproject" was indexed, a Google dork of `filetype:pdf intitle:secretproject` could be used.

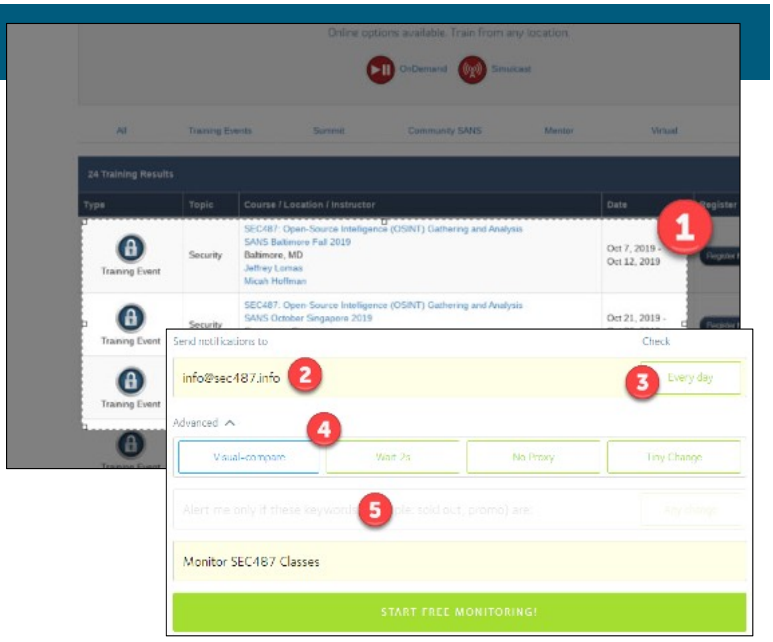
Image from https://sec487.info/p4, October 1, 2019.

VisualPing.io

https://visualping.io

Can visually examine a section of web page or examine text for changes

Free account includes 65 free checks per month (2 tasks, once a day)



VisualPing.io

Similar to Google Alerts, VisualPing.io sends an email when content on the internet changes. VisualPing can watch a web page and look for changes either in the code or via a visual inspection. As shown in the slide above, you can specify that VisualPing observe a certain region of a web site for changes. When the data is altered, you get an email.

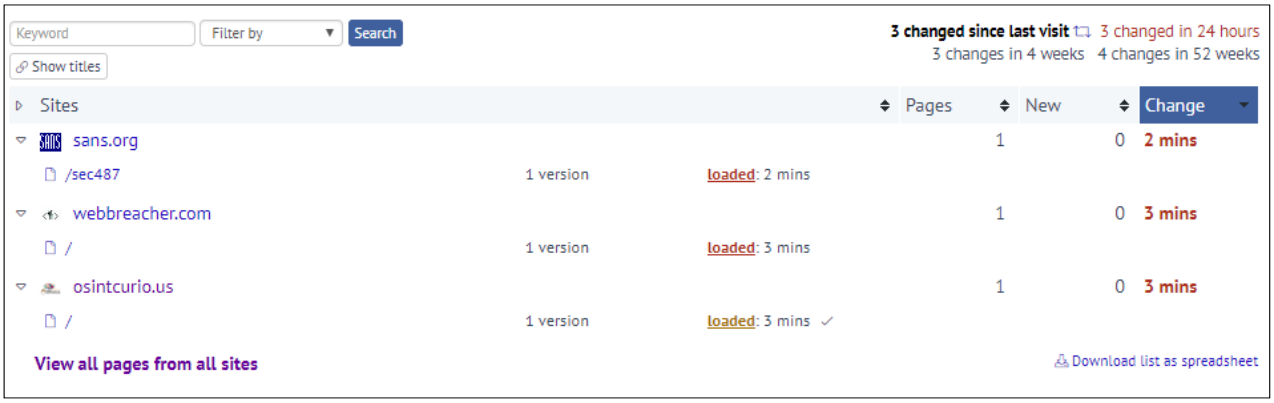
The free level of alerting limits accounts to 65 checks per month. This works out to roughly two checks per day for a month, or you can monitor something more often and use up those checks faster. There are paid versions of this site that give you more alerts.

In the above image, with an authenticated account, we started an alert to monitor the https://sans.org/sec487 class page (1). We chose the email that we wanted to send alerts to (2) and the frequency (3). Then we set how much change needed to occur for the alert to fire (4) and could optionally add in additional strings/words to look for in the page (5).

Images from https://visualping.io, October 2, 2019.

Versionista.com

Similar to VisualPing, and still free, Versionista.com monitors web pages



Versionista.com

Another web page monitoring service is the free https://versionista.com. It works similarly to VisualPing and periodically scans a web site you designate for changes.

Image from https://versionista.com/home, October 2, 2019.

IFTTT.com

IF This Then That can monitor web sites, user accounts, and much more

Free accounts on web or mobile

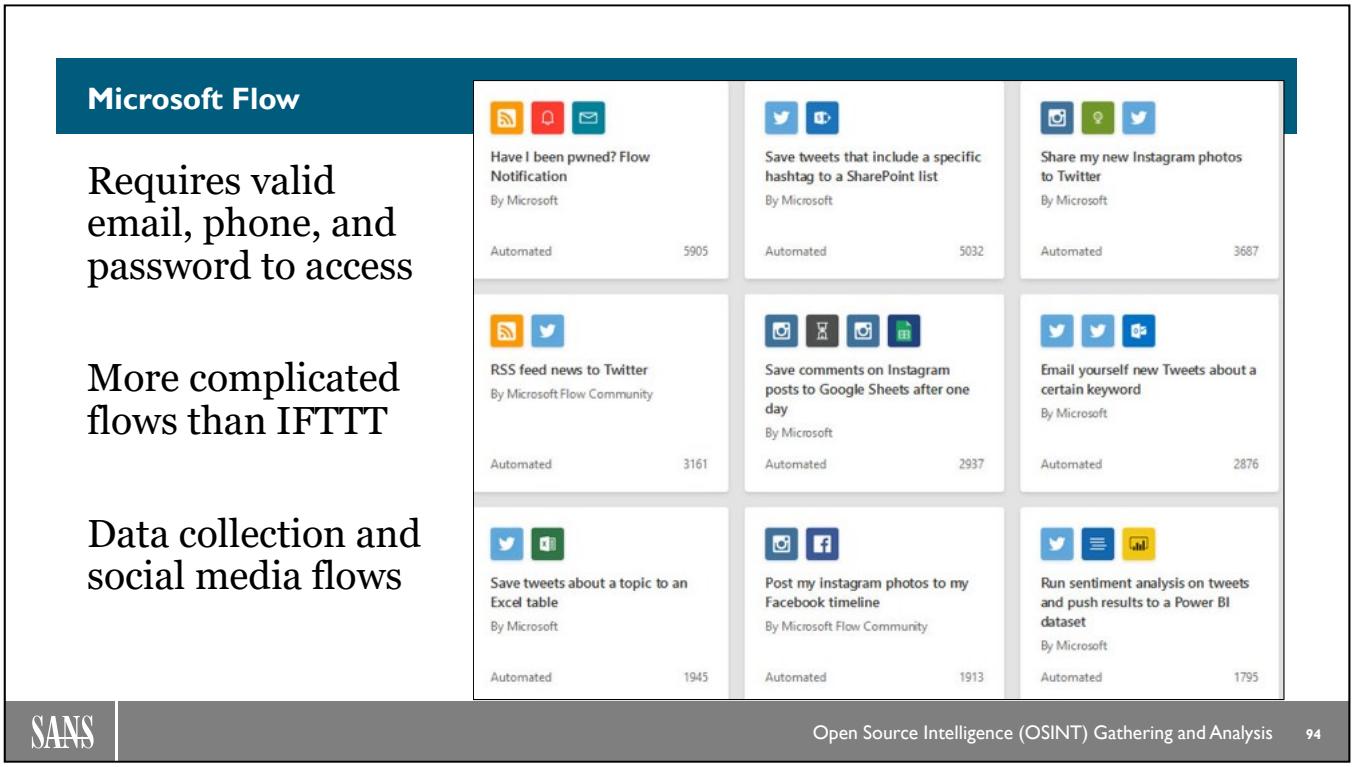
Can choose from existing applets or make your own

IFTTT Twitter Applets

IFTTT.com

The web site If This Then That (https://ifttt.com) can assist OSINT analysts in monitoring and alerting of content on the web through its applets. Applets are small pieces of code that do something when some action occurs. IFTTT is accessible through the web interface and through mobile applications on Android and iOS devices. This greatly increases the number of services that can be used. For instance, you could get an Android alert on a mobile device when the temperature in Mexico City climbs to a certain level. Or post to a certain Slack group when a user posts an Instagram picture. IFTTT integrates with IOT devices and many other web services. There are applets that are created (such as those shown in the slide) that you can choose to start with or create your own.

Image from https://ifttt.com/, March 31, 2018.



Microsoft Flow

Similar to IFTTT, Microsoft's Flow application (https://sec487.info/w3) connects your accounts and automates events based upon a trigger. These "flows," as Microsoft named them, can do much more than what IFTTT's system can do, as they can leverage Microsoft's cloud applications in Office 365 and cloud storage like OneDrive. What this means for you is that you can interact with events, store data, and alert on activity, all within a single tool.

While free to use, a Flow user account requires a valid email and phone number when you sign up. Once this account has been created, you then can choose from hundreds of preconfigured templates (shown in image above) or create your own. Flows can be automated (where they complete based on preconfigured settings), instant (where you manually trigger an event), or scheduled (to fire at a certain date and time interval).

One of the biggest differences in Flow and IFTTT is the complexity of the triggers and actions that can be performed. Instead of taking action on a single trigger, Flow allows the creation of triggers with conditional statements in them. An example might be, if a new tweet with the #sec487 hashtag is tweeted and it is tweeted from a target account, then record the tweet details in an Office 365 Excel spreadsheet.

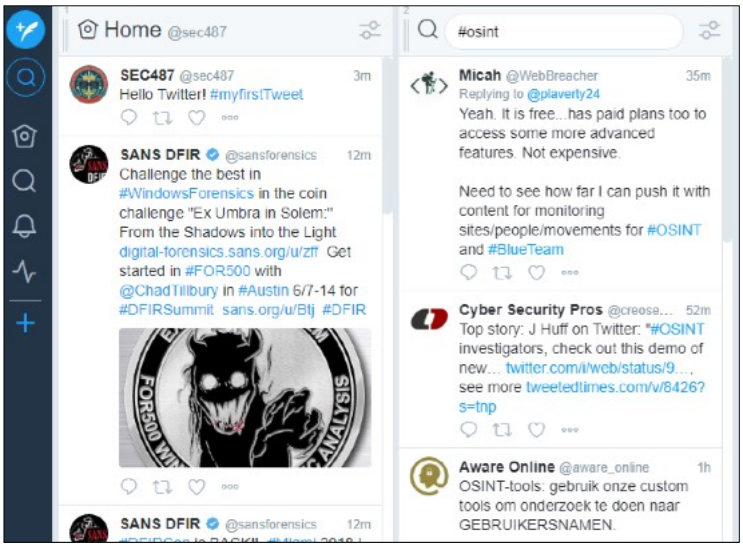
Image from https://sec487.info/w4, October 2, 2019.

Tweetdeck Monitors Twitter

Keeping tabs on topics, people, locations and hashtags in Twitter?

Use the streaming Tweetdeck application to display content in columns

Requires valid Twitter account(s)



Tweetdeck Monitors Twitter

The https://tweetdeck.twitter.com site amplifies the data you can see and interact with on Twitter. With a valid account, you can create columns to monitor hashtags of interest, locations, people, and much more. Content can be set to stream so, as new tweets are posted, they show up in your feeds. With both a light and dark theme for displaying the data, this is not only helpful but pretty too!

The Bellingcat organization created an in-depth guide to using Tweetdeck for OSINT (https://sec487.info/vq). It discusses how to use filters and set up columns, use Twitter lists, and pull regional content into this application.

SEC487 Workbook

Please visit the course electronic workbook in the 487 virtual machine and begin the exercise named:

"HaveIBeenPwned"

This page intentionally left blank.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. Freenet
4. I2P - Invisible Internet Project
5. Tor
6. Monitoring and Alerting
7. **International Issues**
8. Vehicle Searches
9. Putting It All Together

This page intentionally left blank.

Going Global

Your OSINT work may require you to use resources outside of the United States. Some of you students taking this course may be learning it in a non-US country or may be returning to one after training. In either case, we need to think about how to perform OSINT collections outside of the US.

There are some challenges that you might run into, the first of which is that people in some countries do not publish much data on the internet. Their governments do not put documents online or make them easily searchable via web pages. This lack of data can cause assessments to stall.

We also need to find out what social media applications and web sites people in the country you are researching use to share data and connect with friends. Many times, this data is out there for us to find if we know where to look for it. Within the United States, many people focus on certain social media applications and figure the entire world must use those, too. This is absolutely not reality. While people around the world may choose to use one of the big social media sites, there are many regional and country-specific social media outlets that may be of interest to our investigation.

Finally, we need to consider potential language barriers when researching content in other countries. Perhaps you can read the language that your target country uses and maybe you cannot. There are some resources we can use to help decode pages written in languages that we are not familiar with. However, we may have issues interacting with web applications in other countries if we cannot send the appropriate characters to the sites.

International start.me Page

Bruno Mortier's start.me page shows international resources

Each section of the site is organized by country

Great starting point

http://osintframework.de

International start.me Page

We mentioned Technisette's start.me page (https://sec487.info/startme) earlier in the course as well as Emmanuelle Welch's (https://sec487.info/ca). They are excellent, well-categorized resources for your OSINT work. More to add to this list are the international OSINT resources on Bruce Mortier's (@digintelosint) start.me page (https://sec487.info/ia).

Bruno has international resources scattered throughout this well-categorized web of pages, making this a valuable resource of resources for your international work.

Image from https://sec487.info/ia, October 1, 2019.

Language Tools

The best translator is a linguist

They understand context, idioms, and nuances automated tools do not (yet!)

There are web sites and apps that can help translate text, images, and spoken language

Examples of translator programs:

- Mobile apps
  - Microsoft Translator
  - Google Translate
- Web sites
  - https://translate.google.com
  - https://www.bing.com/translator
  - https://www.deepl.com
  - https://translate.yandex.com

Language Tools

One of the first barriers to performing effective international OSINT can be language. While we will talk about electronic methods of performing adequate translation of images, sounds, and text, the reality is that having a person who understands the target language to view, listen, or read the content you are trying to understand is the best method for translation. Humans understand context. They can interpret idioms and slang. A trained linguist or native speaker is the best translator.

Many companies cannot afford to have a linguist on staff or retainer and must use tools available to get their translation jobs completed. In those cases, there are mobile applications and web applications that can assist. Microsoft and Google make excellent apps and web sites for translation. There are, of course, other sites and apps that you can use that may be better at converting a specific language or medium, but for our general purposes, these Google and Microsoft products will do fine.

Images from:

https://sec487.info/p6, January 5, 2019

https://sec487.info/p7, January 5, 2019

Google Translate

Google's web translator can extract and translate text from images (1) and text input in the form. Select the languages (2) you want, insert the text, use the microphone on your computer (3), or use alternative means to input content (4)

Google Translate

Online translators are simple to use. Copy text or characters into their form fields, and they do all the hard work of translating the content into another language. They vary in how many and which languages they can translate from and to, and you may need to try several to find one that works for your needs.

In the above examples, we used the phrase "Life is pain. Anyone who says differently is trying to sell you something." (from *The Princess Bride*) in the Google Translate application. We translated English to Arabic (2).

OPSEC Alert! Pasting your OSINT case content into an online tool can violate customer confidences, give sensitive data to a third party, and breach terms of contracts. Ensure that you understand the guidelines of your assessment before using online tools.

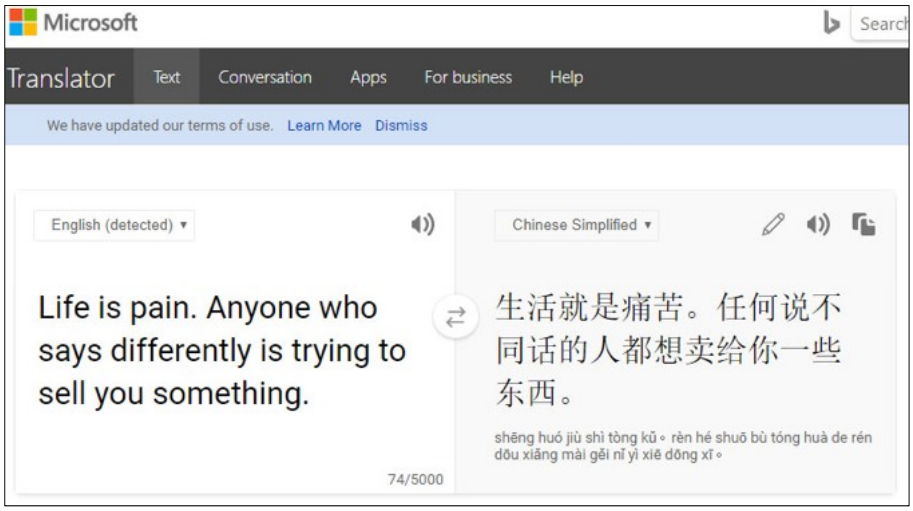
Images from https://sec487.info/vz, October 5, 2019.

Microsoft Translator

Microsoft has its own Bing Translator

Fewer languages translated than Google

Fewer input modalities too



Microsoft Translator

Microsoft's Bing Translator also translates a large number of languages. It is not as extensive as Google in both the number of languages supported and the number of methods of input it allows. However, it uses a different translation engine than Google and can provide a different perspective on text you find in your OSINT work.

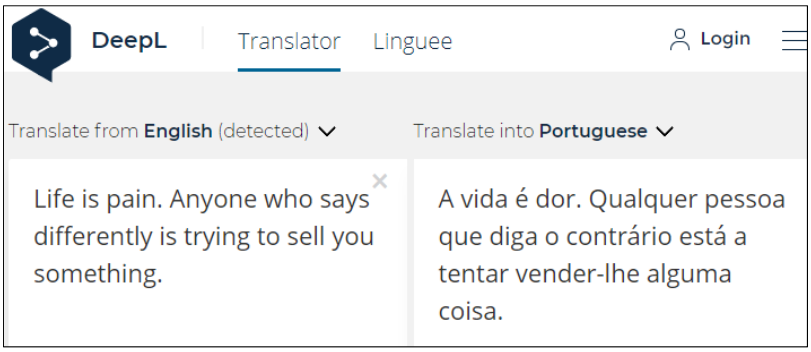
Image from https://sec487.info/vy, October 5, 2019.

DeepL Translator

Free and paid web-based translator

Reported exceptional translation (by them and SEC 487 students)

Less than 10 languages supported



DeepL Translator

There are options for online language translation aside from Google and Microsoft. For the small number of languages it translates, the DeepL Translator (deepl.com) reports that it works more accurately than its competitors.[1] Several SEC487 students confirm this as well. With both paid and free accounts, you may wish to try it.

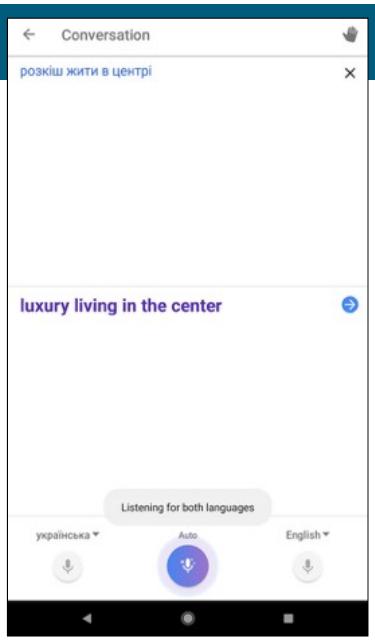
Image from https://sec487.info/yv, October 5, 2019.

Reference: [1] https://sec487.info/pn

Audio Translation

For translation of audio content:

1. Launch translator mobile app (or web site with microphone enabled)
2. Enable app's "listen" for spoken content translation feature
3. Play the content you want translated out loud



Audio Translation

For short audio sound clips, the mobile translation applications from Microsoft and Google (shown above) can be useful. The simple process for using them is outlined above, although there are obvious variations of this process (such as using an online web site for the "listening" and playing the audio from your phone into the computer's microphone).

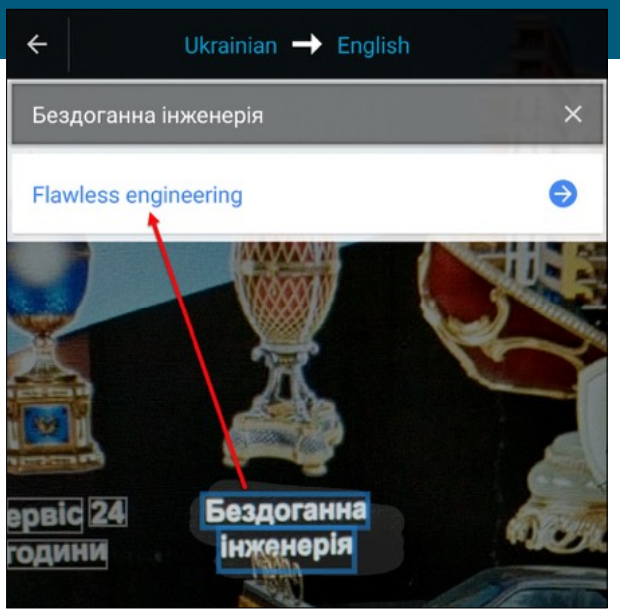
Picture/Photo Translation

Mobile translator applications can take photos of text and translate it

Some provide live translation:

1. Launch app
2. Point at text for translation
3. Text is translated

Some web translators do this too



Picture/Photo Translation

Using the Google mobile app (shown above) or the Microsoft Translator app, we can point our devices at something we would like to translate and, either take a photo of it or just keep the content in frame, and the application will convert the text for us.

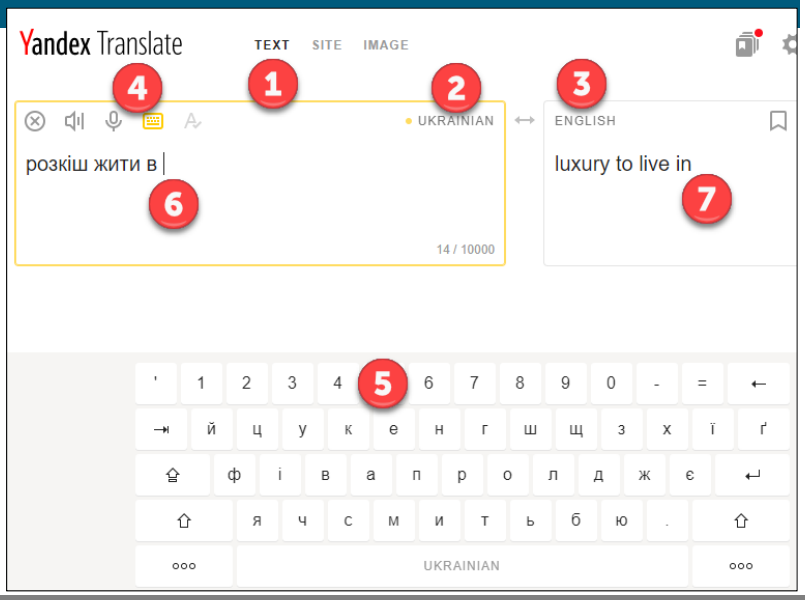
Google and Yandex (shown next) both perform this "image to text" extraction and translation from their respective web applications.

Yandex Virtual Keyboard

Yandex has a virtual, on-screen keyboard to "type" the characters you need in other character sets

Google also has this feature[1]

The notes have an image with text to translate



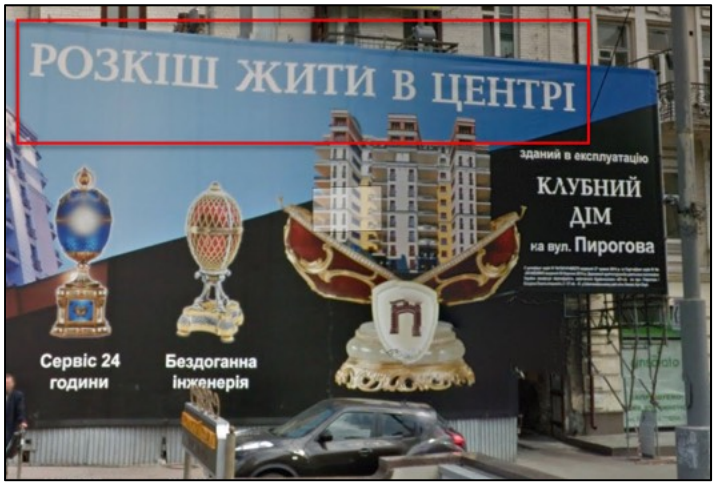
Yandex Virtual Keyboard

Below we have a billboard on a street in Ukraine (https://sec487.info/dk) and can see the characters we need to translate but don't have any idea how to bring up a keyboard using the specific character set on our mobile device or computer. Yandex.com's Translate page (https://translate.yandex.com) has a virtual keyboard feature that may come in handy. Google has one, too.[1]

Above, we visited the site and selected to translate text (1). Then we selected which languages to convert from (2) and to (3). Clicking on the virtual keyboard (4) makes the keyboard at (5) appear. Yandex only has this feature for some languages. From here, we click on the letters that make the word we see in an image, and it is displayed at (6) with the translation on the right side at (7).

Image from https://sec487.info/wv, October 3, 2019.

Reference: [1] https://sec487.info/wu

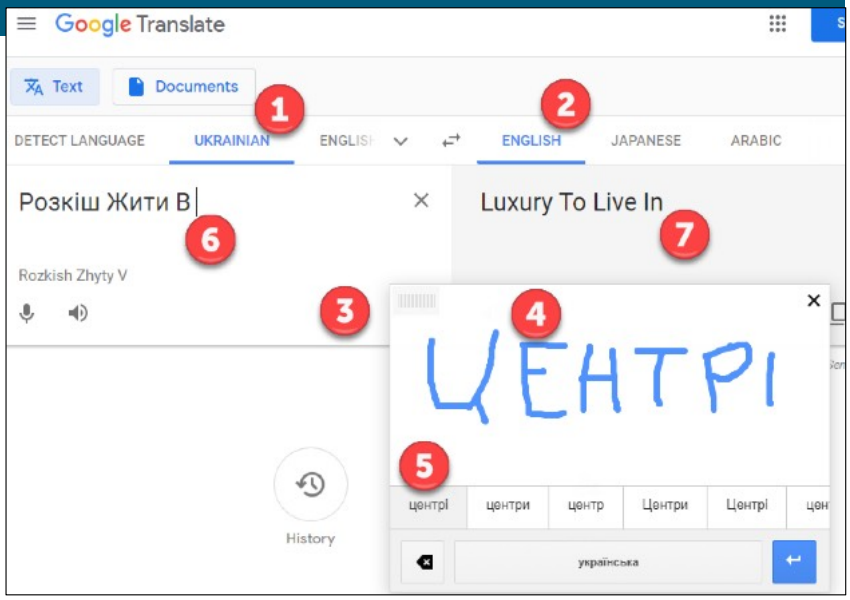


Character Translation

Don't know how to type in the foreign language?

Just draw characters

Google has a handwriting feature that allows you to draw text to be translated

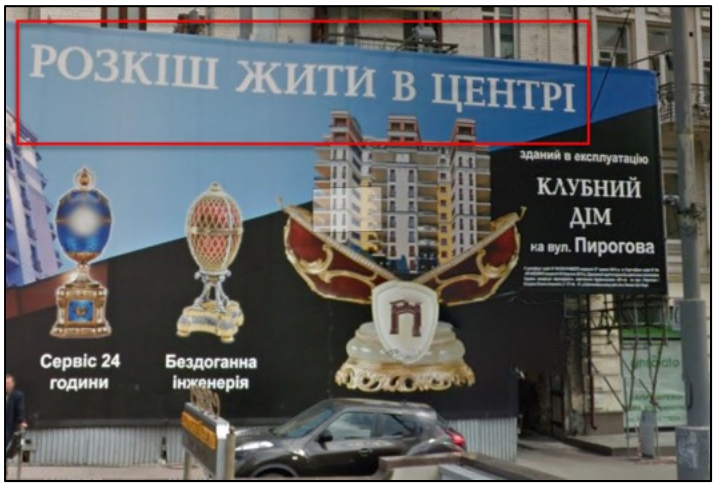


Character Translation

Sometimes the translations we seek are for content discovered in images or other places we cannot just copy the text from and past into the translators. In these cases, if you are unfamiliar with typing in the target language, you can use character-drawing tools within the web applications of the translators to create the letters to translate.

Below we have a billboard on a street in Ukraine (https://sec487.info/dk). Above, we used the character-drawing functions of Google Translate to select the languages (1 and 2), choose to draw characters (3), draw the characters (4), and then select the correct letters (5) to be placed in the upper section to be translated (6). You can finish the drawing yourself by visiting https://sec487.info/wu and drawing the above content and seeing what the text reveals.

Slide image from https://sec487.info/wt, October 3, 2019.

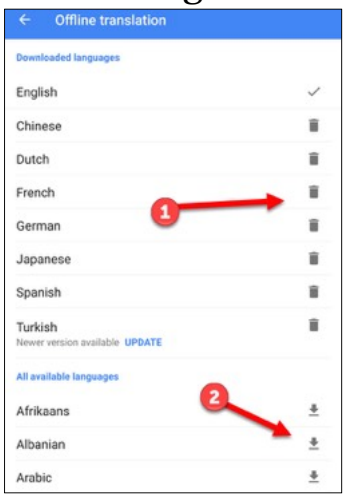


Offline Translation

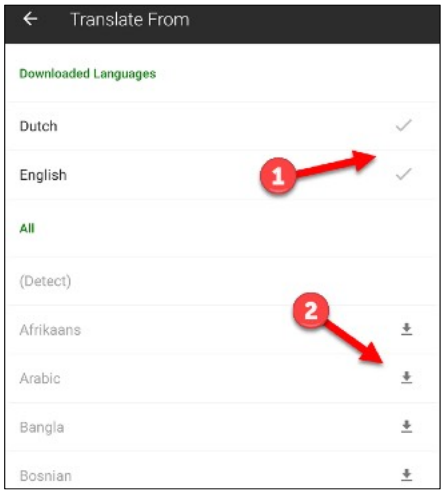
Both Microsoft's and Google's mobile applications can download language packs and be used offline (no network)

Download the language and switch device to airplane mode

Google



Microsoft



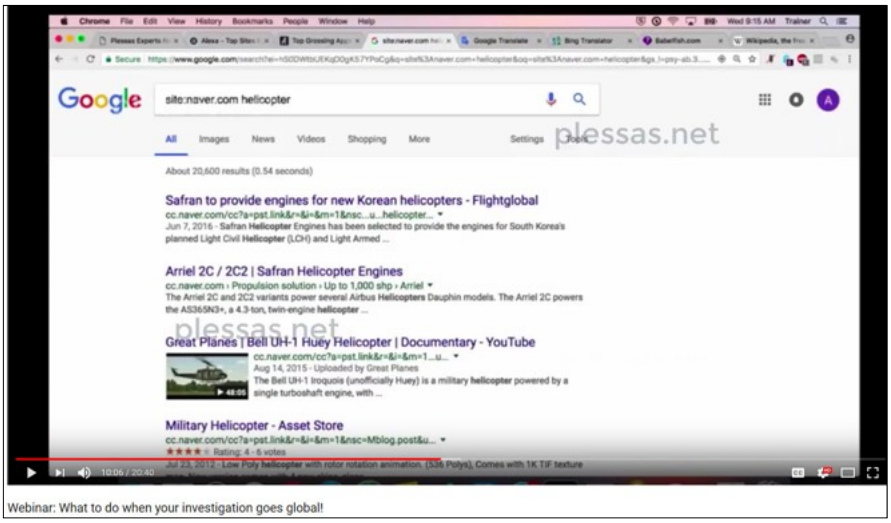
Offline Translation

There are many reasons for not sending your interesting OSINT content to an online site to be translated. In these cases, we can use the same mobile device applications in an "offline" mode to perform the translations using previously downloaded language packs. In the slide above, we see Google and Microsoft's language pack options. The "1" arrows point to those packs that have been downloaded in the device, and the "2" arrows point to those additional language packs that can be downloaded.

International OSINT YouTube Video

Kirby Plessas (@kirbstr) recorded a webex on international OSINT

It has excellent tips that we will explore



International OSINT YouTube Video

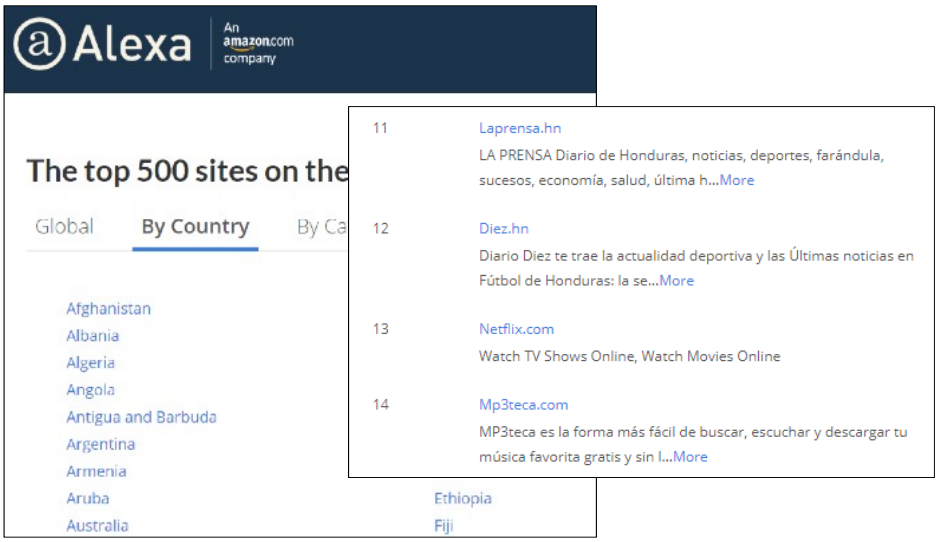
Kirby Plessas (@kirbstr, http://www.plessas.net/site/) recorded a webex on "What to do when your investigation goes global" (https://sec487.info/dm) in November 2017. It provides excellent suggestions and processes for working with foreign content in your OSINT assessments.

Let's take a look at some of the major issues she raises in the video.

Popular Sites

Alexa ranks popular web sites around the world

It also organizes them by country



Popular Web Sites

People within different countries may visit different web sites. If you are researching content in another country, you should examine what sites the people in that region use to get their news, share their personal lives, and buy things. Your OSINT may be richer and more complete.

To find the popular web sites that may be used in a given country, visit the Alexa.com analytics web site (https://www.alexa.com/topsites, https://sec487.info/p8). Choose a country, and your browser is redirected to a page that hosts the top 500 web pages used by people in that country. In the image above, we selected entries 11–14 from Honduras' most popular sites (https://sec487.info/dn).

Images from https://sec487.info/dn, October 1, 2019.

APPLyzer Mobile App World Charts

Understanding what mobile apps are popular in a given country can help you find your human targets faster

APPLyzer has several app stores that it pulls data from

APPLyzer Mobile App World Charts

Looking for a human target's social media traffic on sites that are not used in their country will waste your valuable OSINT time. The APPLyzer.com web site taps into several of the most popular application stores and marketplaces that are used on mobile devices to bring you what is popular in a given region.

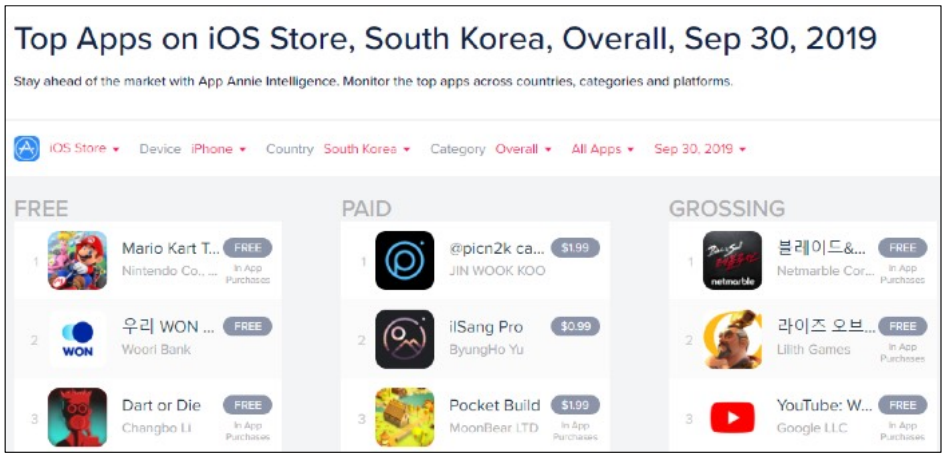
In the above image, we chose to research apps on the Google Play store (1) in Sri Lanka (2). Selecting the "Top Overall" dropdown from the category field then showed us the top apps for that day in that marketplace. Now we know that many users in Sri Lanka use WhatsApp and Messenger for their messaging services (4). We could shift our OSINT data gathering to those platforms in search of our targets.

Image from https://sec487.info/vw, October 1, 2019.

Popular Mobile Apps

App Annie provides "app market data and insights" [1]

Choose the country, and it will reveal popular mobile applications used there



Popular Mobile Apps

The App Annie web site (https://www.appannie.com/) is an analytics site that tracks mobile application popularity. It allows people to view the top applications by platform (Apple iOS, Google Android, Windows Phone Store, Amazon Appstore, etc.) and by country. Select the country you wish to view, and you can see the top mobile applications used there. Above, we show the iOS apps popular in South Korea on September 30, 2019.

App Annie has free and paid accounts, along with anonymous access.

Image from https://sec487.info/v-, October 1, 2019.

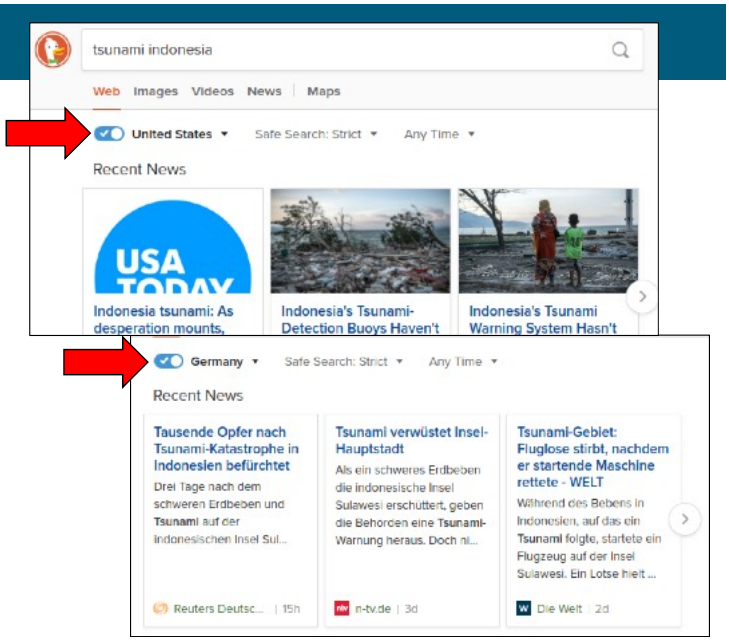
Reference:

[1] https://sec487.info/fu

Searching by Region

DuckDuckGo, Bing, and Google usually show search results from your geographic region

Each has settings to alter this and help get regional content that you may not see in your search results



Searching by Region

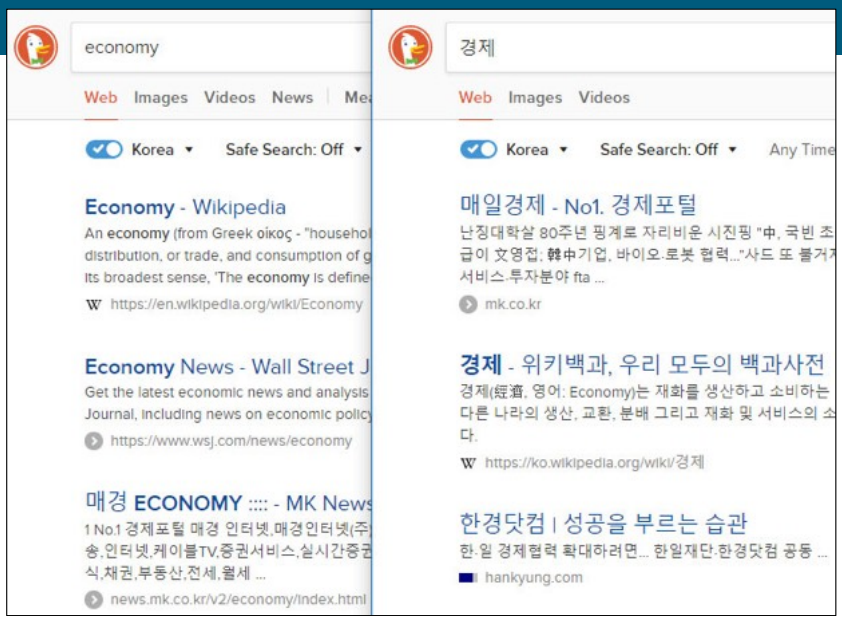
Each of the major search engines allows users to choose a different region to search for relevant information. Why might we want this localized data? Consider if there is a protest in a foreign country. Choosing their region will give you results from within their country instead of how the rest of the world saw that event. Of course, perhaps you WANT to get the rest of the world's perspective on an issue or event. In that case, you can choose several different regions and examine the articles, reports, and web sites that are returned.

The places in each application to shift to a new region can be found in the search settings options for Google and Bing. DuckDuckGo displays the region option in the search results (see above).

Search Local Language

Mashing local language and regional results

We can use languages local to the area we are targeting to retrieve search results that may have stayed hidden if we use English



Search Local Language

A final tip mentioned in Kirby Plessas' YouTube video reveals that you will get different results when you search for a term in English in a foreign region versus in the language that is native to that area. Shown above are the results of DuckDuckGo searches for the word "economy" in English and Korean. While we only show the first three results; rest assured that there was a large difference in the results from both searches.

International Risk

Understanding international travel risk can help us inform our clients

Helpful sites include:

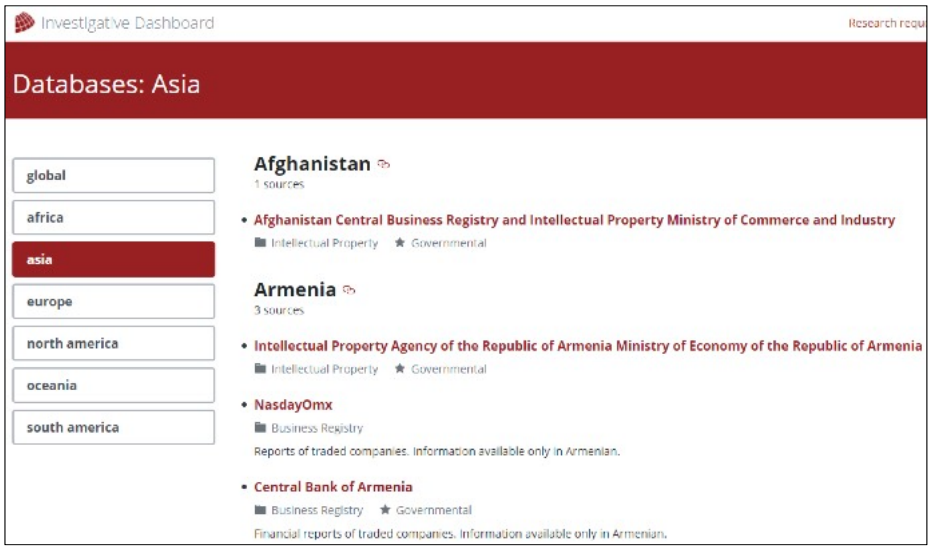
- travel.state.gov
- safearound.com
- internationalsos.com

International Risk

When our clients need to understand what the risks of doing business in or travelling to a distance country are, we can turn to some internet resources to help us understand the risks. These web applications many times have cultural information to help people understand the customs and practices of the target location. They describe crime levels and types, scams, risks to certain populations such as women or specific religions, laws that may be important (like do not bring gum to Singapore), and more.

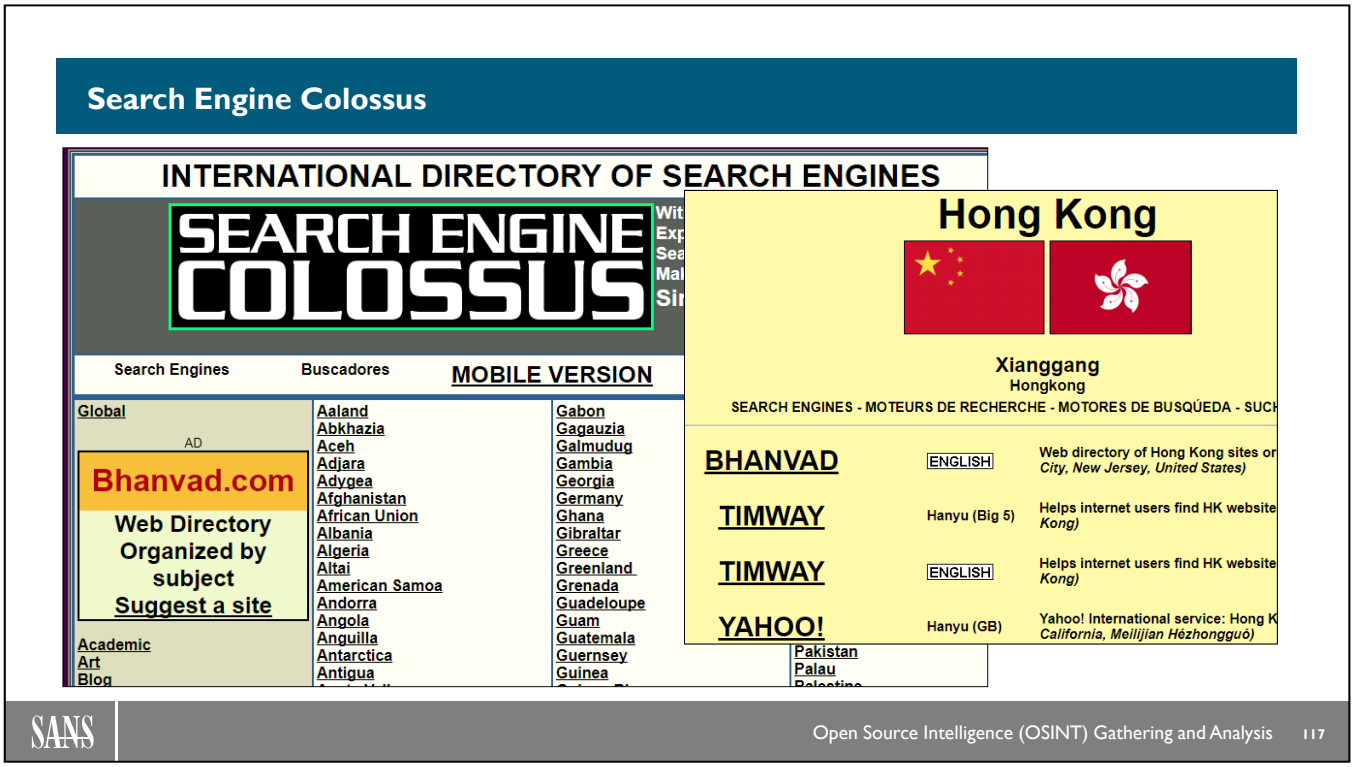
Image from https://sec487.info/pm, January 6, 2019.

International Investigative Databases



International Investigative Databases

The web site for the Investigative Dashboard (https://sec487.info/an) has a large number of databases that can be used to enhance an international OSINT effort. Links take users to databases from around the world that host data on businesses, people, governments, and more.



Search Engine Colossus

Created in 1998, the Search Engine Colossus site (https://sec487.info/dp) contains organized links to search engines that may be used within different countries. Navigate to the country to wish to search and view the search engines that are in use in that region.

The "about" page on the site[1] describes the site's purpose succinctly:

> "Search Engine Colossus: International Directory of Search Engines can be looked upon as an effort to give the internet "structure". This WWW roadmap allows surfers to efficiently gain access to the far reaches of the net!
>
> "It is hoped that visitors will have an informative, fascinating internet surfing experience as they gain better understanding of their fellow world citizens.
>
> "Search Engine Colossus first went online in April 1998 and it now has about 2500 listings from 317 countries and territories around the world!"

Reference:

[1] https://sec487.info/p9

SEC487 Workbook

Please visit the course electronic workbook in the 487 virtual machine and begin the exercise named:

"International Issues"

This page intentionally left blank.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. Freenet
4. I2P - Invisible Internet Project
5. Tor
6. Monitoring and Alerting
7. International Issues
8. **Vehicle Searches**
9. Putting It All Together

This page intentionally left blank.

Vehicles, Owners, and Activities

Vehicles and sometimes their cargos have serial numbers identifying owners

These allow us to track where the vehicles have been and, possibly, where they are going

Researching these vehicle IDs can yield interesting pivot points in investigations

Vehicle, Owners, and Activities

In most parts of the world, vehicles such as cars, motorcycles, planes, and ships need to be registered with one of more government offices. This can be done for purposes of tracking ownership in case they are stolen, for tax purposes, or for regulatory reasons. For consumer vehicles such as cars and trucks, registrations contain information about who owns the vehicle, what country or region it is registered in, may contain the Vehicle Identification Number (VIN), and usually will contain the make, model, and year of the vehicle. With ships and planes, there may be additional data collected.

If you come across these identifiers in your OSINT assessments, performing searches on the vehicle registration and identification data can yield excellent data that you can pivot on.

Image from https://sec487.info/pa, October 12, 2017.

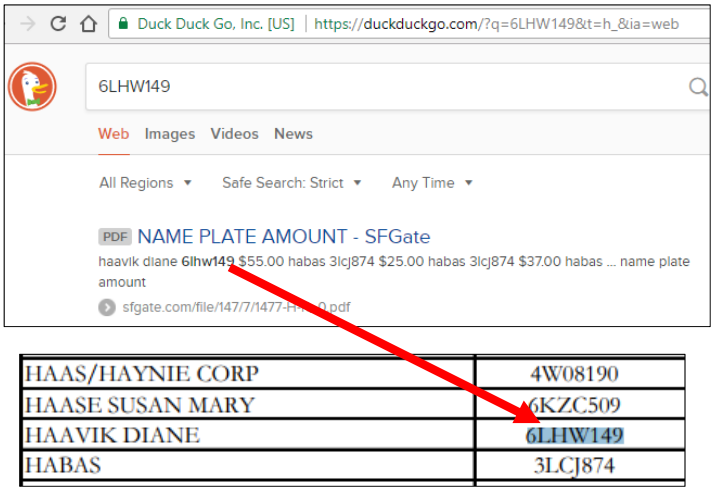
Vehicle License Plates

DuckDuckGo it! (or Google)

Docs with the plate will show in results

Some places allow you to buy the info

Some commercial people search engines show owners



HAAS/HAYNIE CORP	4W08190
HAASE SUSAN MARY	6KZC509
HAAVIK DIANE	6LHW149
HABAS	3LCJ874

Vehicle License Plates

Let's start with consumer-level vehicles such as cars, trucks, motorcycles, and sometimes bicycles. Most of these vehicles, once registered with the local or regional government, get a unique license plate that needs to be displayed. We can collect this data and, depending upon where the vehicle is registered, perform a reverse lookup on the plate to find the owner and additional data.

By far, the easiest place to look up license plates is on a search engine. Looking up a plate such as "6LHW149" in DuckDuckGo shows an excellent link as the top result. Without entering a country or state, DuckDuckGo found a match. Turns out that this California license plate was found on the San Francisco Municipal Transportation Agency's document as a vehicle who's owner over-paid citations (https://sec487.info/93).

Using search engines can yield valid results, but the best places to find valid data is, in the United States, the state level. In other countries, check who the registering authority for vehicles is and see if they allow searching of records. In the United States, some states allow for license plate lookups, but it may cost a small fee and may require you to be a private investigator or law enforcement agent. Additionally, some of the commercial-tiered people search engines may show this data, but again there is a cost and restricted access.

Private investigator (PI) access to license plate data varies from state to state within the United States. Some states are more permissive with the data, trusting the PI has a good reason to request the data. Other states may require a compelling reason before they release the information.

FAXVIN.com License Plate Lookup

For certain states in the USA, you can use the FAXVIN site to map a license plate to a VIN

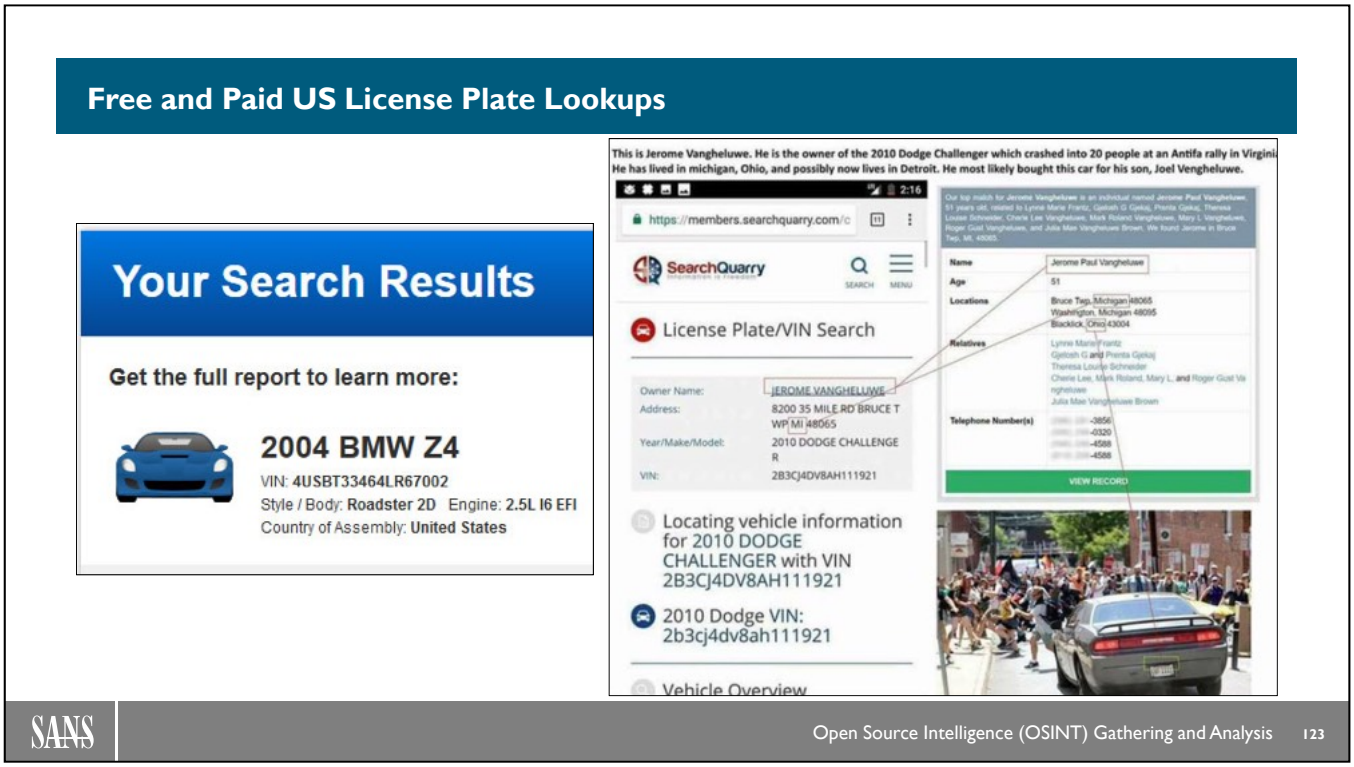
Enter the plate and get the VIN!

Plate searched: **GWF8627**, New York State

FAXVIN.com License Plate Lookup

The faxvin.com web site will resolve a United States license plate to a Vehicle Identification Number (VIN). In the example above, we searched for the New York license plate GWF8627 and discovered the VIN associated to it (1). The site will then allow you to perform additional paid searches based on the VIN.

Image from https://sec487.info/wq, October 3, 2019.



Free and Paid US License Plate Lookups

One fairly good and free license plate lookup tool for United States vehicles is at Autocheck.com.[1] Visiting this page brings up search fields where you can enter in a license plate, such as 5HUH323 in California, and the search results should show, in this case, that the plate is registered to a 2004 BMW Z4 car. Note the results also display the Vehicle Identification Number (VIN) of the vehicle, which can be used to perform additional searches.

The SearchQuarry.com site also allows free license plate searches. It has a paid feature where you can obtain additional data about the licenses plate or VIN. This was used in the aftermath of an attack on a crowd of protesters in 2017 in Charlottesville, Virginia. A person drove his car into a crowd of protesters. Someone took his license plate number and used their subscription to SearchQuarry.com to retrieve additional details about the person (image on right in slide).[2]

Data aggregators such as IRBSearch and TLO (mentioned earlier in the course) also have license plate search capabilities but are more restrictive in who can access the data and have a fee associated with it.

References and images:

[1] https://sec487.info/94, October 12, 2017.

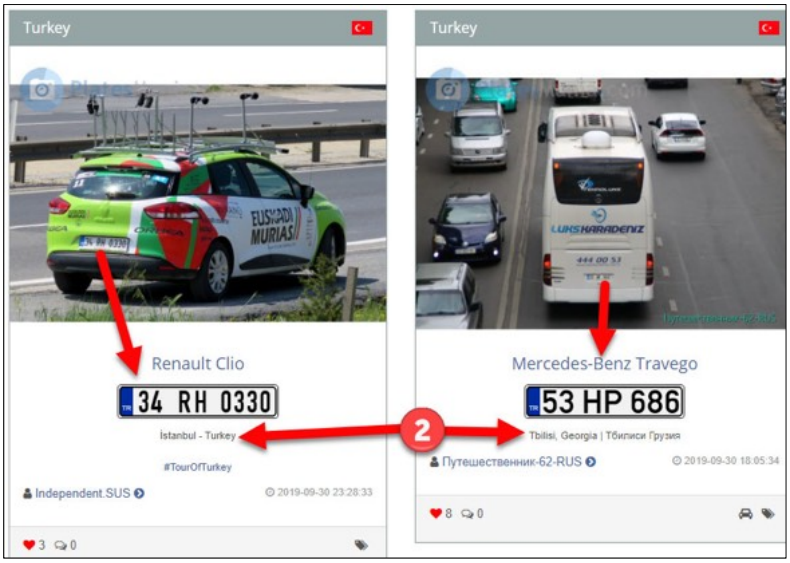
[2] https://sec487.info/95, October 15, 2017.

PlatesMania.com

International catalog of license plates and makes/models of vehicles

Search by model, plate, or just browse

Dates, times, locations of vehicles



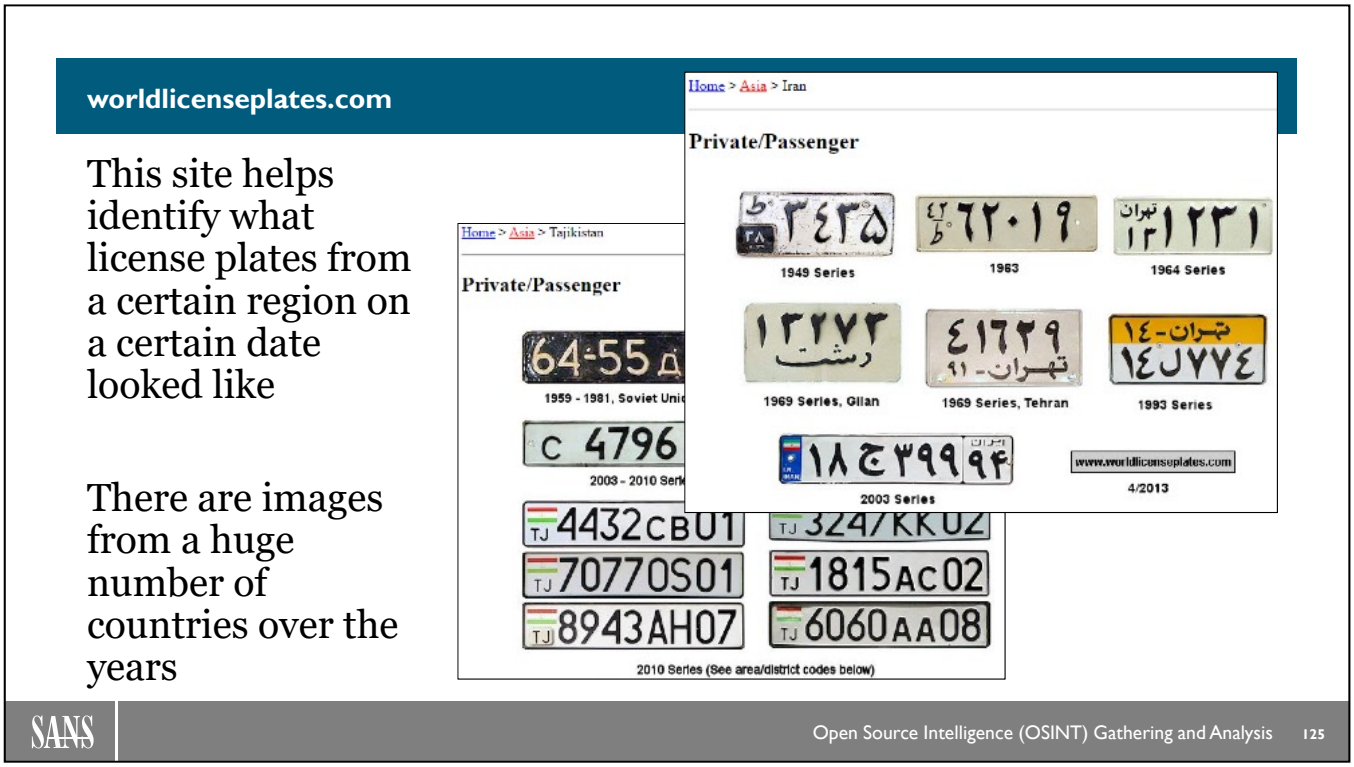
PlatesMania.com

The PlatesMania.com website is an amazing resource for crowd-sourced license plate tracking. Its international user base takes pictures of vehicles, from buses and tractor trailer trucks to motorcycles and cars, and then posts them to the web site. We can search the site data based upon make and model of the vehicle we are looking for, license plate, and location. So if you had a target that drive a recreational vehicle (RV) with a Norway license plate of ZH 50694, you could search the site (https://sec487.info/vt) and discover the VIN, where it was seen, and when it was registered.

The data recorded about a vehicle varies. As in the example above, that RV had a VIN in the system but the Mercedes-Benz Axor truck registered in 2005 in Tbilisi, Georgia, does not. Instead, it shows what the mileage was on a certain date and has another picture of where that vehicle was seen.

The site allows us to search by make and model of a vehicle and see where it has been seen. Alternatively, we can look at a country and see what vehicles have been located there—possibly telling us what vehicles are popular in that country (https://sec487.info/vu).

Image from https://sec487.info/vv, October 1, 2019.



worldlicenseplates.com

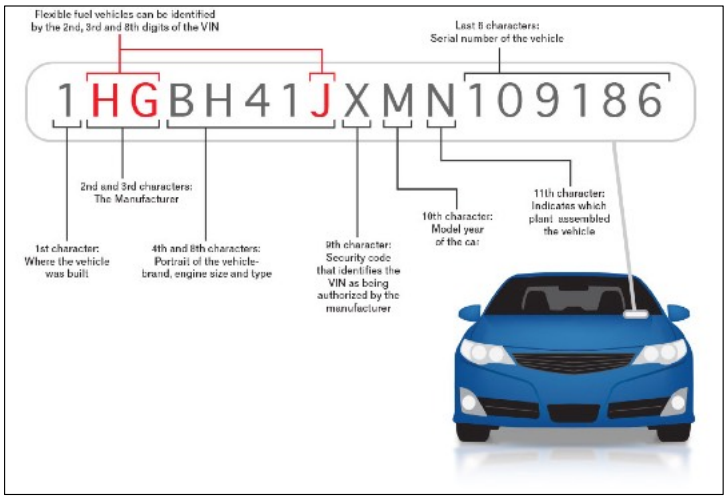
Performing OSINT on a license plate and need to understand if it is a valid format, what country it might be from, or what time frame it may have been issued from? The http://worldlicenseplates.com/ site will be useful. It has a historical view into international license plates.

Image from http://worldlicenseplates.com/, October 1, 2019.

VINs (Vehicle Identification Numbers)

Unique ID numbers for motor and towed vehicles engraved/etched/bolted to the vehicles in multiple places

It is a series of codes that provide details about where the vehicle was made, attributes of it, and a serial number



VINs (Vehicle Identification Numbers)

A VIN is a unique, coded string of characters that is engraved, etched, and otherwise embedded into one or more surfaces of a (usually motorized) vehicle. Each position in the string represents some facet about the vehicle,[1] where it was manufactured, features of the vehicle, and more. Since each VIN is unique, that allows us an opportunity to investigate or track a specific vehicle through time and across owners.

Image from https://sec487.info/71, October 12, 2017.

Reference:

[1] https://sec487.info/fw

VIN Lookup Sites

If you obtain a VIN, you can search for vehicle details and perhaps the owner on several sites

- https://www.decodethis.com
- https://www.autodna.com
- https://www.vindecoderz.com
- https://www.searchquarry.com

Some sites will decode the VIN and offer to pull additional records about accidents and ownership for additional fees

Well-known paid VIN search:

- https://www.carfax.com/

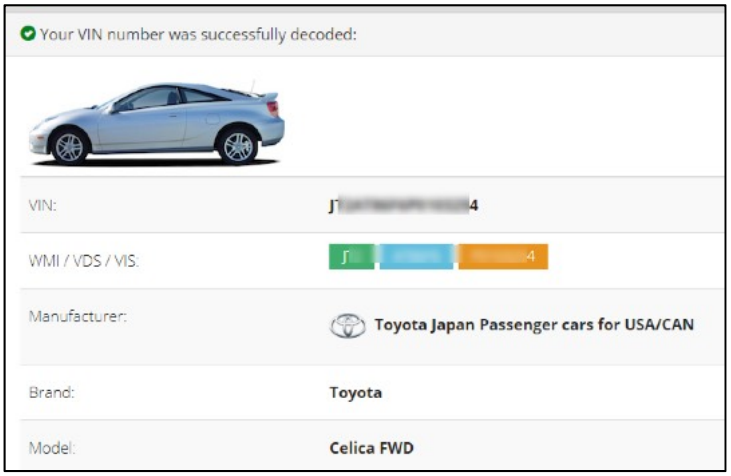
Consider "googling" a VIN to find additional data

VIN Lookup Sites

Since decoding the VIN is a simple process, there are web sites that can help perform this function. The slide shows several web sites that will decode the VIN and show the vehicle's data (make, model, year of manufacture, etc.) and sometimes stock pictures of the vehicle. While most of the VIN decoder sites are free to use and provide this basic data, some will try to entice users into getting a more complete report that may tap into other databases holding owner information and accident/repair records. These are paid services.

Additionally, always use a search engine (Google, DuckDuckGo) to retrieve data on a VIN as you don't know what will show up in the search results. An example would be searching for 2B3CJ4DV8AH111921 in a general search engine. It pulls up some interesting content that may be relevant to your assessment.

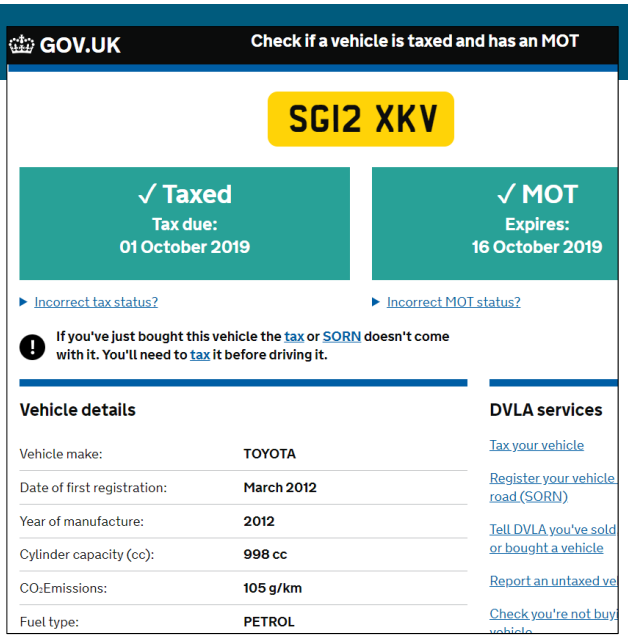
Image from https://sec487.info/pb, October 15, 2017.



United Kingdom Car Tax Check

The UK government has a web site that allows for searching of vehicle license plates to discover tax information and vehicle details

It does not reveal owner data



United Kingdom Car Tax Check

The United Kingdom government has a web site for retrieving information about whether vehicles have been registered and their owners have paid their taxes. Enter a license plate number of a UK-registered vehicle, and the site will show the latest tax information along with vehicle details.

Image from https://sec487.info/pq, January 6, 2019.

United Kingdom Car Tax Check

Free site in the UK allows search of UK license plates to discover vehicle details, taxes, maintenance, and mileage

No owner data revealed

United Kingdom Car Tax Check

The cartaxcheck.co.uk site in the United Kingdom allows people to search for information on UK-licensed vehicles. Enter in a license plate, and the results may contain maintenance history, mileage, vehicle specifications, and tax data. This site does not display owner information.

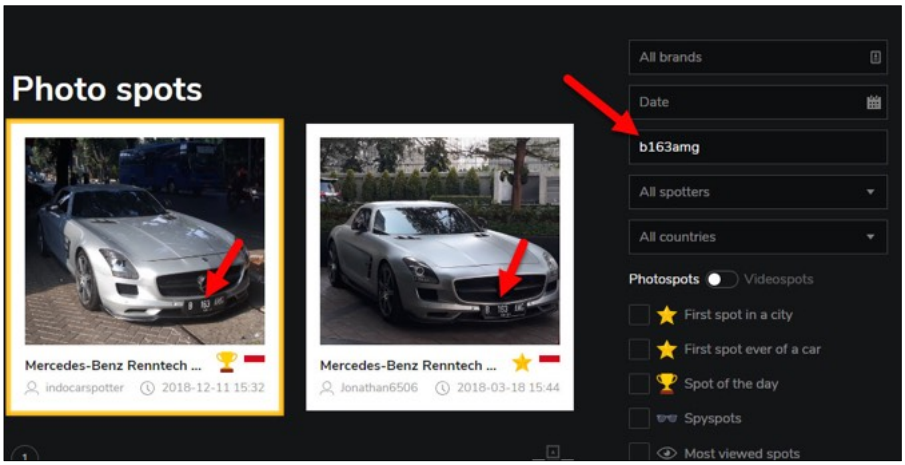
Image from https://sec487.info/pp, January 6, 2019.

High-End Vehicle Search

Searching for the license plate of a high-end car?

Visit autogespot.com

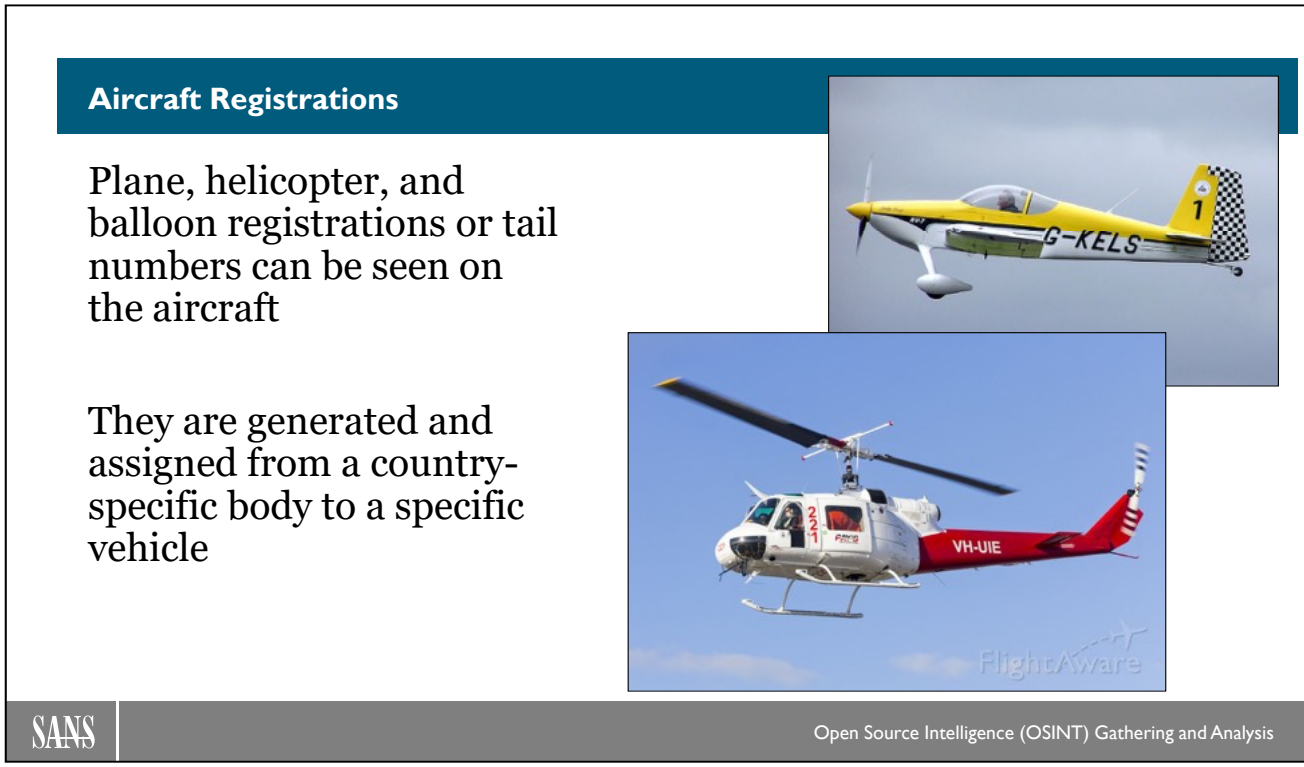
International scope



High-End Vehicle Search

If you have the license plate of a high-end (expensive and high-performance) car, you can use the filters on the Auto Gespot site (autogespot.com) to find if people have located it and submitted it to the site. The images of vehicles are user-submitted and mostly not filtered or redacted.

Image from https://sec487.info/pt, January 6, 2019.



Aircraft Registrations

Each aircraft, whether balloon, helicopter, or airplane, should[1] have a registration number clearly marked on its body. While it can appear in almost any place on the vehicle, on airplanes, it used to appear on the tail and, hence, is referred to as a tail number. These strings are not usually numbers but a combination of both English letters and numbers, forming a unique string.

The owners of the aircrafts, when registering their vehicles with their home governments, will receive the vehicle's registration number. Every country controls its own registrations, usually carried out by a government agency.

Images:

https://sec487.info/pd, October 15, 2017.

https://sec487.info/pc, October 15, 2017.

Footnote:

[1] It is true that all aircraft should have a registered number displayed on the craft. However, certain covert government and military groups, criminals, and others may not place the correct tail numbers on their crafts, may remove the numbers entirely, or may use another plane's information in order to carry out stealthy or nefarious business. We are not covering those cases.

Tail Numbers Are an International Standard

The first character of a tail number is the country of registration

Some countries have more than one allocation table entry

The tail number G-KELS is registered in the United Kingdom (shown on the right)

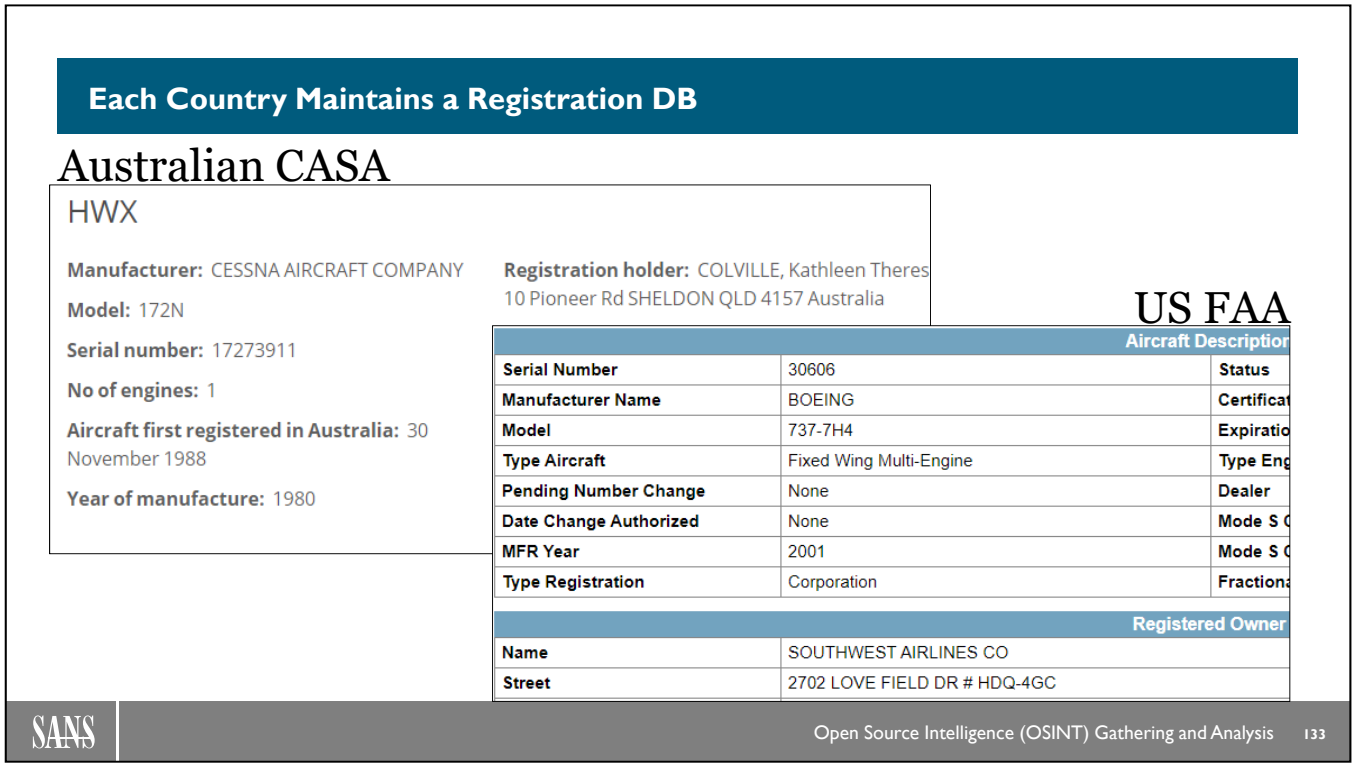
	F
F	France *(and its Overseas departme*
	G
G	United Kingdom *(and its overseas t*
	H
HA	Hungary
HB	Switzerland
HB (HB0, HB3Y, HBL)	Liechtenstein (uses prefixes allocat
HC–HD	Ecuador
HE	Switzerland
HF	Poland
HG	Hungary

Tail Numbers Are an International Standard

The International Telecommunication Union (ITU) is the governing body for airplane tail numbers.[1] Each country has one or more entries in the allocation table shown in the slide. This is the first character of an aircraft's registration number, which must be displayed on the vehicle.

Reference and image:

[1] https://sec487.info/96, October 15, 2017.



Each Country Maintains a Registration DB

Keeping track of all the aircraft registered in a country is usually a governmental role, and many of these organizations allow people to look up tail numbers of that country's planes through a web site. Links to many of the country-specific aviation web sites can be found at https://sec487.info/fx. The slide above shows both the Australian government's Civil Aviation Safety Authority (CASA)[1] and the United States' Federal Aviation Administration (FAA)[2] lookup pages as examples.

References and images:

[1] https://sec487.info/wi, October 2, 2019

[2] https://sec487.info/98, October 2, 2019

Aircraft Tracking Sites

There are several aircraft-tracking sites that are amazing at historical and near-real-time data display

Our most-useful sites:

- https://www.adsbexchange.com
- https://opensky-network.org
- https://flightaware.com
- https://www.flightradar24.com
- https://planefinder.net

Aircraft Tracking Sites

There are several free, excellent aircraft-tracking web sites on the internet. The sites listed above allow users to search for specific planes using their tail numbers or airline flight numbers and will display both historical and current data about them. The historical data regarding where a plane has flown and when is an amazing resource to access. These sites show additional content, such as plane owner, images of the plane and its internal configuration, and more.

Each also shows planes in flight in near-real-time, so you can see where a certain plane is in the air and approximately when it will arrive at a specific airport.

Image from https://sec487.info/fy, October 15, 2017.

ADS-B

Automatic Dependent Surveillance - Broadcast

System of GPS satellites, ground stations, and radios in aircraft

Aircraft broadcast their altitude, heading, location, and other data to ground stations and other ADS-B receivers

ADS-B

The Automatic Dependent Surveillance - Broadcast (ADS-B) system is newer technology aimed at increasing awareness of where aircraft are and getting more real-time data into the cockpits (such as weather). Because these aircraft radios broadcast ADS-B signals, ground stations and personal radios can receive them. Several web sites aggregate this data and plot the aircraft on a map and show historical information about where a certain aircraft travelled.

You can even sign up to get a free receiver and send the received data to a web site collector,[2] or you can create your own device using a Raspberry Pi computer.[3]

Image from https://sec487.info/we, October 2, 2019.

References:

[1] https://sec487.info/wf

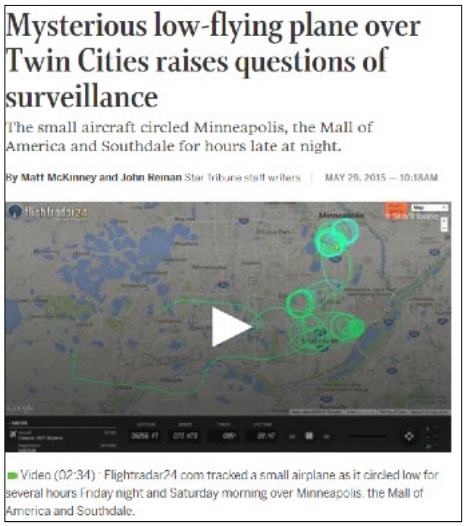
[2] https://sec487.info/wg

[3] https://sec487.info/wh

Law Enforcement Planes Tracked

There are reports of people using the flight-tracking services to discover law enforcement vehicles overhead

This can mess up a surveillance operation



Mysterious low-flying plane over Twin Cities raises questions of surveillance

The small aircraft circled Minneapolis, the Mall of America and Southdale for hours late at night.

By Matt McKinney and John Reinan Star Tribune staff writers | MAY 29, 2015 — 10:18AM

Video (02:34) : Flightradar24.com tracked a small airplane as it circled low for several hours Friday night and Saturday morning over Minneapolis. the Mall of America and Southdale.

Mystery surveillance plane that circled Twin Cities was part of secret FBI fleet

Aircraft that circled metro area and Mall of America had high-tech surveillance equipment.

By John Reinan and Matt McKinney Star Tribune | JUNE 3, 2015 — 9:38AM

This FBI surveillance plane flew over the Twin Cities several times recently. Surveillance equipment and its power cord can be seen on the fuselage.

Law Enforcement Planes Tracked

The ease with which people can use the flight-tracking web sites and discover government/law enforcement vehicles is empowering and frightening. When some aircraft is circling overhead, anyone can turn to these web sites and potentially retrieve tail numbers, vehicle flight history, and owner information. This has happened in the United States and was publicized in several articles in 2015 when an FBI plane circled Minneapolis while performing law enforcement recon activities (https://sec487.info/pr and https://sec487.info/ps).

Images from https://sec487.info/pr and https://sec487.info/ps, January 6, 2019.

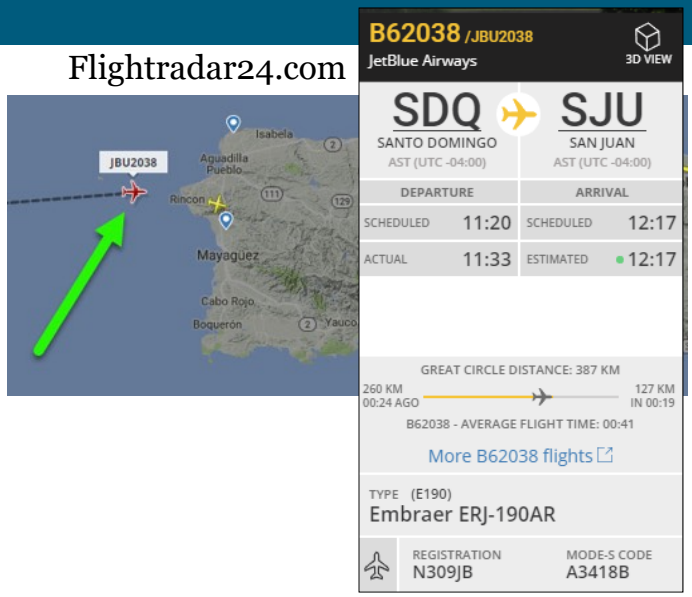
Where Is an Aircraft Now?

As mentioned, these flight-tracking sites allow us to see where a specific plane is in (near) real time

An example is JetBlue's JBU2038 flight headed into San Juan, Puerto Rico in the image to the right

Flight stats shown

Flightradar24.com



Where Is an Aircraft Now?

Your customers may need intelligence on where a certain plane is flying. The flight-tracking web sites present this data in beautiful visual interfaces. An example from flightradar24.com is shown above from October 15, 2017 at 12:00pm Eastern US time. The JetBlue airline flight JBU2038 flying from Santo Domingo to San Juan, Puerto Rico is just making landfall.[1] The red plane (arrow pointing to it) in the image[1] is animated on the web site and is moving from west (left) to east (right) in the image. We can see its scheduled and estimated flight times, the tail number of the plane, and the type of plane.

Reference and images:

[1] https://sec487.info/99, October 15, 2017.

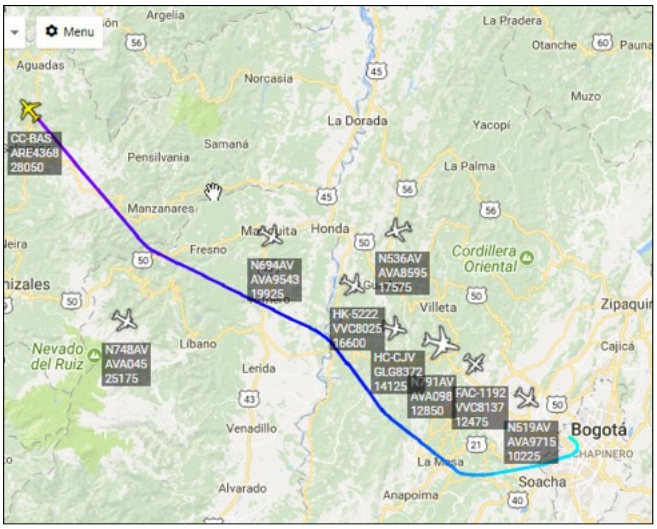
ADS-B Exchange (global.adsbexchange.com)

Unfiltered access to Automatic Dependent Surveillance — Broadcast (ADS–B) data

Shows aircraft that are:

- Government/Military
- "Interesting" (experimental, unique, and test aircraft)

Free, community-driven



ADS-B Exchange (global.adsbexchange.com)

The ADS-B Exchange web site shows much of the same type of data that the other sites mentioned do. However, it differs from those web applications in that "all data sent in from the community is, in turn, made available back to the community through various archives and APIs [https://sec487.info/pe]. Best of all, for non-commercial use, all of this data is freely accessible to anyone!"[1]

It has customizable filters allowing the viewing of the specific Automatic Dependent Surveillance - Broadcast (ADS–B) data you wish to see. You can filter out certain types of planes or carriers, look for categories of aircraft, or watch everything. Data is sent to the site by a network of community and commercial ADS-B receivers around the world. Some of our favorite filters to use are the "Military" and "Interesting" ones that display military aircraft and other aircraft such as helicopters and fire-fighting planes.

Reference:

[1] https://sec487.info/i3, June 16, 2018.

Historical and Future Aircraft Information

Flight tracking history for N119NN

Route	Flight No	Date	Playback
Logan International Airport (BOS) - Los Angeles Intl (LAX)	AA2542	2019-10-03 00:45:00	
Los Angeles Intl (LAX) - Logan International Airport (BOS)	AA1423	2019-10-02 21:31:00	
Daniel K. Inouye International (HNL) - Los Angeles Intl (LAX)	AA798	2019-10-02 14:00:00	
Los Angeles Intl (LAX) - Daniel K. Inouye International (HNL)	AA143	2019-10-02 02:55:00	

Activity Log

UPCOMING FLIGHTS

Date	Departure	Arrival	Aircraft	Duration
Friday 04-Oct-2019	*07:30AM CEST Frankfurt Int'l - FRA*	*08:25AM CEST Hamburg - HAM*	A321	*0h 55m*
Thursday 03-Oct-2019	*07:30AM CEST Frankfurt Int'l - FRA*	*08:35AM CEST Hamburg - HAM*	A321	*1h 05m*

PAST FLIGHTS

Date	Departure	Arrival	Aircraft	Duration
Wednesday 02-Oct-2019	07:24AM CEST Frankfurt Int'l - FRA	08:29AM CEST Hamburg - HAM	A321	1h 05m
Tuesday 01-Oct-2019	08:10AM CEST Frankfurt Int'l - FRA	09:19AM CEST Hamburg - HAM	A319	1h 09m
Monday 30-Sep-2019	07:36AM CEST Frankfurt Int'l - FRA	08:52AM CEST Hamburg - HAM	A321	1h 16m
Friday 27-Sep-2019	07:30AM CEST Frankfurt Int'l - FRA	08:40AM CEST Hamburg - HAM	A321	1h 10m
Thursday 26-Sep-2019	07:33AM CEST Frankfurt Int'l - FRA	08:35AM CEST Hamburg - HAM	A320	1h 02m
Wednesday	07:33AM CEST	08:40AM CEST		

SANS Open Source Intelligence (OSINT) Gathering and Analysis 139

Historical and Future Aircraft Information

The FlightAware[1] site has an interactive slide display that allows a user to retrace where a flight was, at what velocity it was travelling, and the altitude of the plane over the course of its flight (shown above).

Additionally, FlightAware (and the other sites)[2] has data on upcoming and previously flown flights for many planes. This might be useful if you find that a target is posting about a flight number. Visit these sites and find out where the plane will be going or went on a certain date and time. Did the plane malfunction and have to make an emergency landing somewhere other than the expected destination? Was there a storm the day of travel and the plane was delayed? These types of incidents and activities will be shown on this historical page.

References and images:

[1] https://sec487.info/wi, October 2, 2019.

[2] https://sec487.info/wj, October 2, 2019.

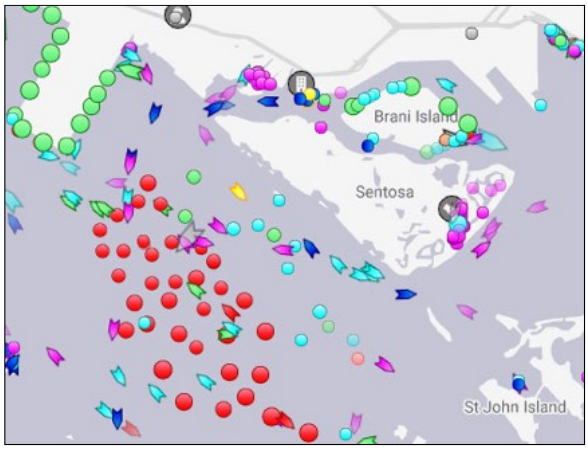
Ships and Watercraft

Vessels moving on the seas can be tracked in a similar fashion to aircraft using AIS

Free sites that do this include:

- http://shipfinder.co
- http://www.myshiptracking.com
- https://www.marinetraffic.com
- https://www.vesselfinder.com

https://www.marinetraffic.com



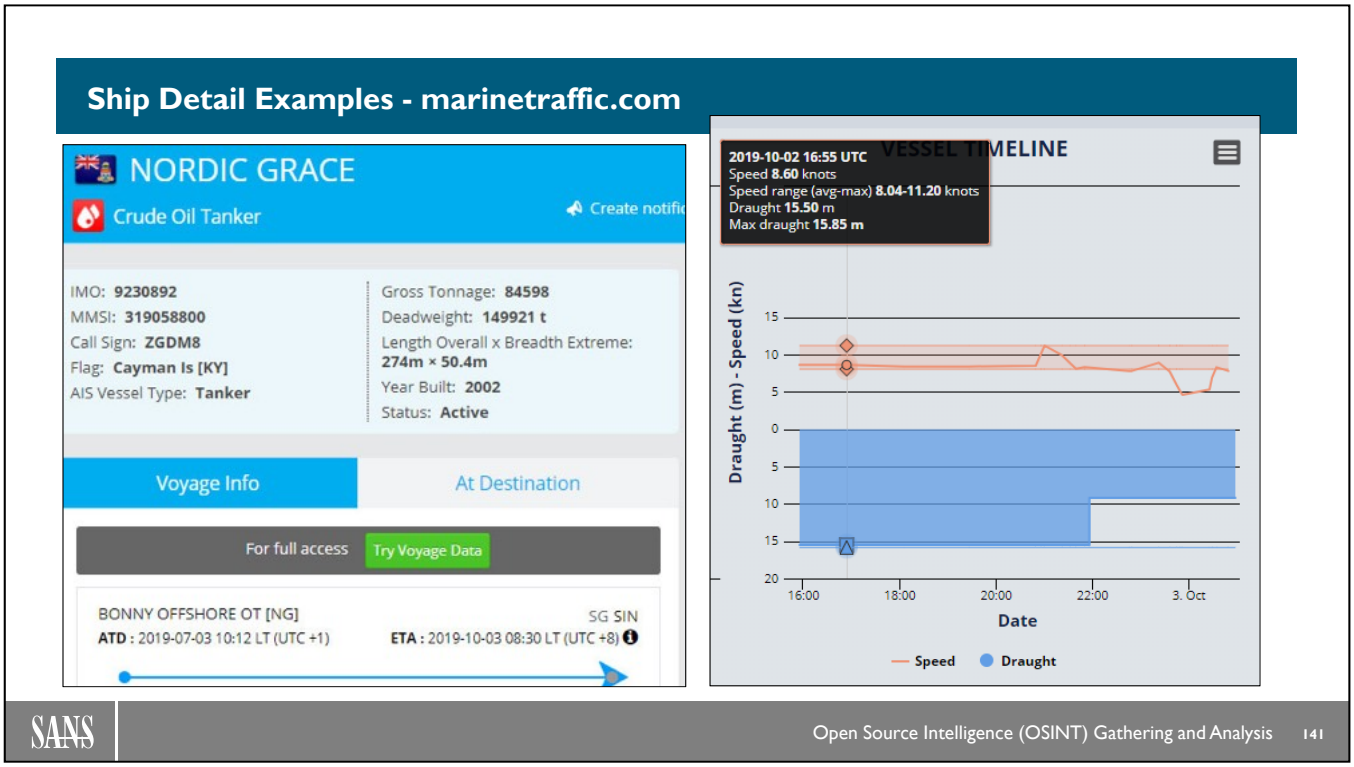
Ships and Watercraft

As with aircraft, we have several free sites that track ships and watercraft across oceans and around shorelines. They mostly rely on Automatic Identification System (AIS)[1] devices sending VHC radio signals that report the ship's location, heading, and ship name/call sign. This information can be received by other ships with these AIS devices, land-based antennas, and satellites.

The web sites located in the slide all have dynamic, beautiful maps that can be customized to show certain types of vessels (pleasure craft versus military versus cargo). OSINT analysts interested in a region can navigate on the map to that place and observe what ships are reported. Alternatively, for some vessels, we can search for a specific ship and view details about it.

Image from https://sec487.info/wk, October 2, 2019.

Reference: [1] https://sec487.info/9d



Ship Detail Examples - marinetraffic.com

The images above illustrate some of the data recorded about ships within these tracker web sites. In the left image, the display shows detailed data about the NORDIC GRACE vessel that was underway near Sentosa Island off Singapore's coast. The right image displays historical movement: speed and draught. There are many other data points on these pages, including previous course and past ports of call.

Depending upon your OSINT goals, this real-time or historical information may provide the necessary answers you are looking for. While each of the web sites mentioned on the previous slide show similar content, there are variations.

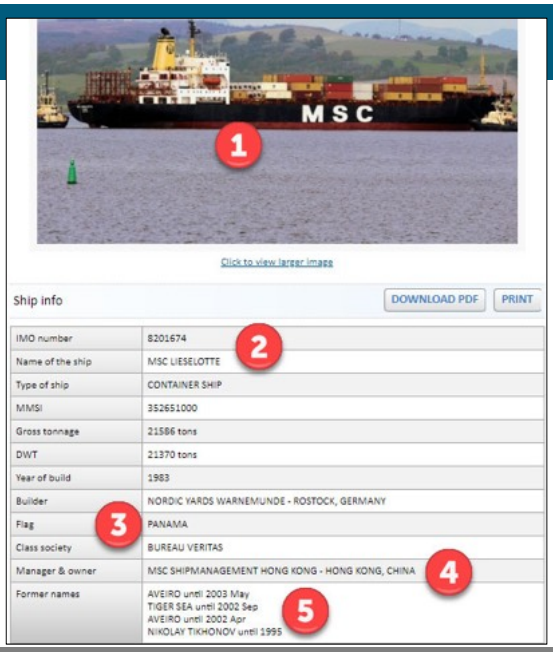
Images from https://sec487.info/wl, October 2, 2019.

Maritime Jobs (1)

Rae Baker (@wondersmith_rae) wrote a blog about investigating maritime activities[1]

She mentioned the maritime-connector.com site for looking at jobs posted, seafarer "resumes," and ship information

As anonymous user, you can search for ships and parse seafarer details



Maritime Jobs (1)

Rae Baker (@wondersmith_rae) published a blog post about investigating maritime activities and the people who work on and run ships and boats. In addition to mentioning the marine-traffic.com site (previously shown), Rae mentions the maritime-connector.com site,[1] which is meant to help seafarers who work on ships find their next job.

In this slide, we show an image of the MSC LIESELOTTE ship (1), including its IMO number (the International Maritime Organization number is a unique number given to ships)[2] and name (2), where it is registered (3), who owns and manages it (4), and what other names it was known by (5). More details on the ship can be found on the site.[3]

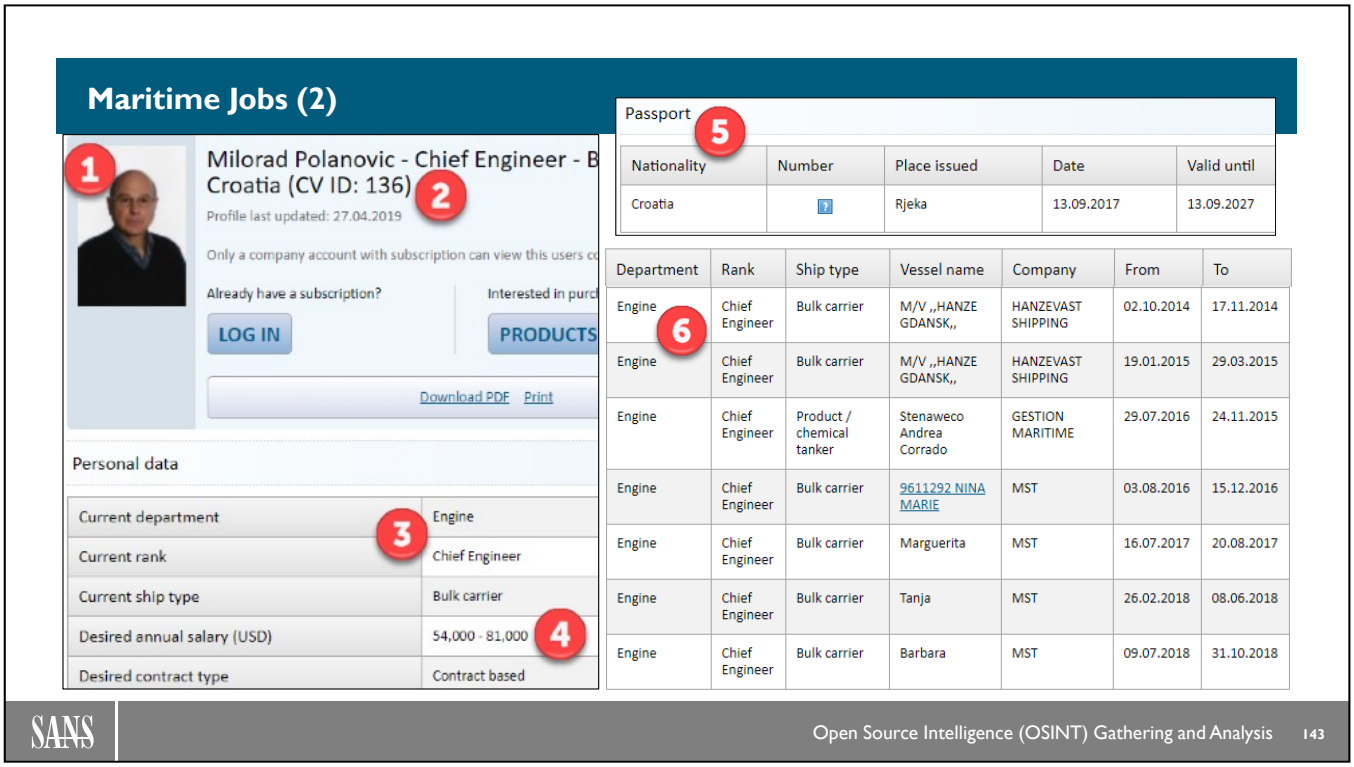
Image from https://sec487.info/wn, October 3, 2019.

References:

[1] https://sec487.info/wm

[2] https://sec487.info/wo

[3] https://sec487.info/wn



Maritime Jobs (2)

Pivoting from ships to the people who work on them, the maritime-connector.com site has a plethora of data on this topic. As an example, let us examine a random sailor, Milorad Polanovic (https://sec487.info/wp), shown above. We see an image of him (1), what type of role he is in and where (2), what he does on ships (3), and what salary in USD he would like (4).

As we scroll down the page, we see certificates, passport details (5), education and vaccination information, and his service record, which details what he did on what ship owned by which company during which dates for decades (6).

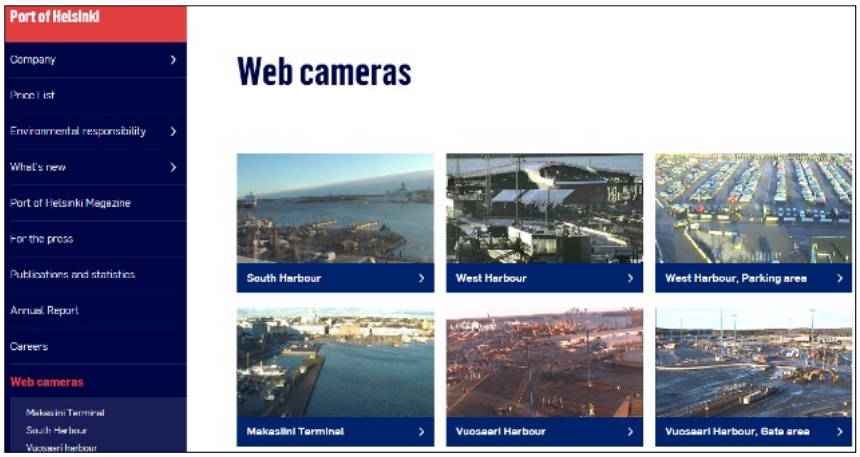
Images from https://sec487.info/wp, October 3, 2019.

Harbor Web Cameras

Marinas/ports many times have public web cameras we can tap into

Watch vessels in harbor and activity

Live, streaming images



Google: `harbor OR marina webcam OR "web cam"`

Harbor Web Cameras

A SEC487 student shared that many docks, harbors, ports, and marinas have webcams that live-stream activity. A Google/DuckDuckGo/Bing search for `harbor OR marina webcam OR "web cam"` and then your city or country of choice should help to locate the cameras for your target area. We performed this search without a location (https://sec487.info/vr) and received hundreds of millions of results (many of which are mostly likely false positives).

Above is an image showing six cameras that the port of Helsinki in Finland has online. Also keep in mind that other cameras may be pointed at these water areas. Using already discussed sites like insecam.com and worldcam.eu could locate additional cameras of interest.

Image from https://sec487.info/vs, October 1, 2019.

Course Roadmap

- Day 1: Foundations of OSINT
- Day 2: Gathering, Searching and Analyzing OSINT
- Day 3: Social Media, Geolocation, and Imagery
- Day 4: Networks, Government, and Business
- **Day 5: The Dark Web, Breach Data, and International Issues**
- Day 6: Capstone: Capture (and Present) the Flags

THE DARK WEB, BREACH DATA, AND INTERNATIONAL ISSUES

1. The Surface, Deep, and Dark Webs
2. The Dark Web
3. Freenet
4. I2P - Invisible Internet Project
5. Tor
6. Monitoring and Alerting
7. International Issues
8. Vehicle Searches
9. **Putting It All Together**

This page intentionally left blank.

Where Are We?

For the past week you have read, listened to, and used a variety of tools and techniques

You might have learned some new tricks or maybe evolved your existing OSINT process

Now let us put all your learning to the test!

Where Are We?

Where are we? Well, for some of you who have never performed an OSINT assessment before, you are on the verge of doing just that. The Solo CTF is about taking all of the processes, tools, and techniques you learned and implementing them on live targets. For those who have existing OSINT experience, challenge yourselves to try some different tools or experiment with a new note-taking technique. This Solo CTF is about your personal growth.

The Rules Are Simple

Treat this as a real assessment

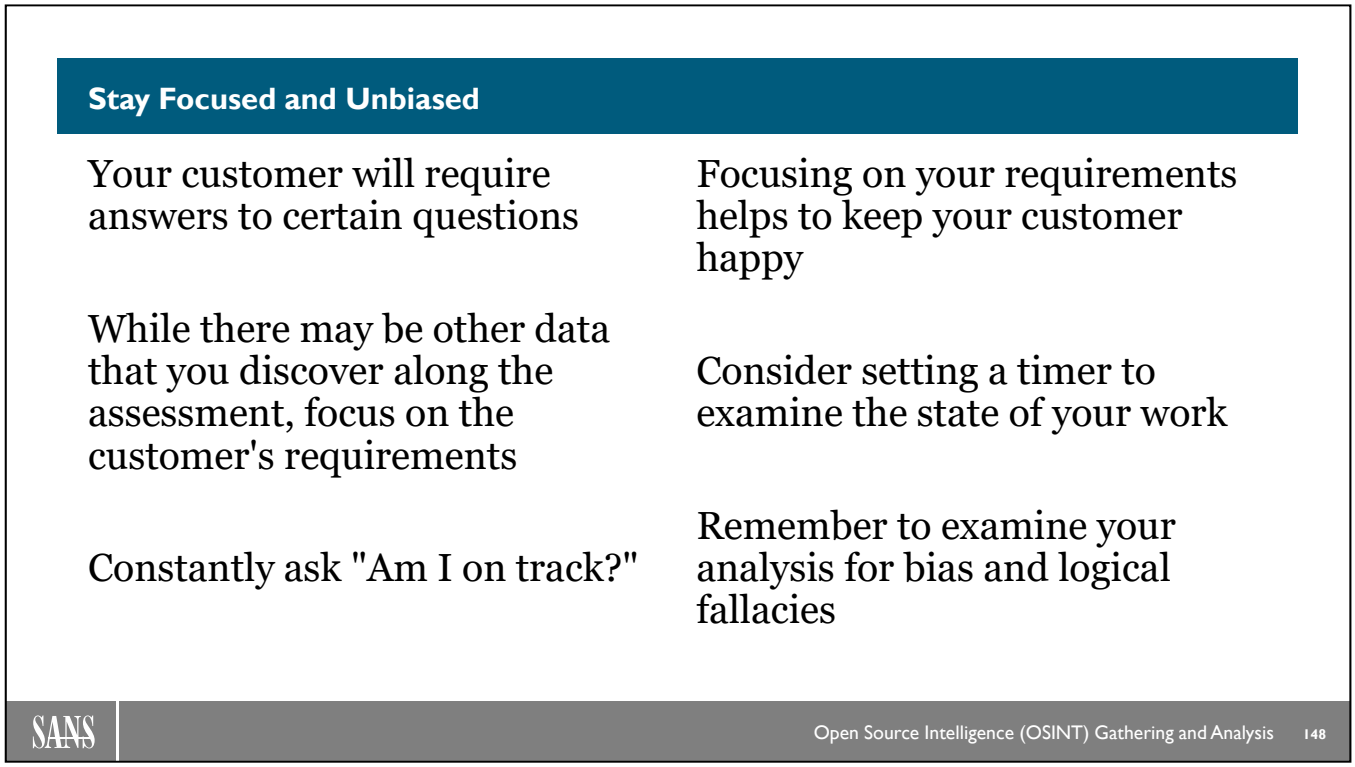
Before we head over to the lab section to get the details, let us re-examine several of the concepts that will make your investigation successful

Concepts to work through:

- Sanitize your system
- Determine requirements
- Examine techniques
- Use tools
- Analyze your data
- Repeat as needed
- Create your report/output
- Use your time wisely

The Rules Are Simple

The Solo CTF is the final exam before the final, final exam (our Day 6 work). Treat this final lab as you would a real assessment. Start at the beginning of the course and work through all we have done. The slide presents some reminders of what we covered in this course. Take a moment to refresh your mind concerning some of these concepts before moving to the exercise.



Stay Focused and Unbiased

It is easy to get overwhelmed in data and analysis and wander off-track during OSINT assessments. What starts as a single phone number turns into a user name that has 10 social media accounts. You dive into those accounts trying to harvest the data and realize that there is other interesting data to pursue. Time passes rapidly during assessments, especially if you are engaged in the work. Stay focused on the customer requirements and goals. Find the data and analysis to help your customer and, as hard as it can be, try to minimize side investigations into interesting but unrelated data.

Experienced analysts use their time wisely and ask themselves, "Am I on track with the data and analyses I have?" If you find that your answers are anything other than "Yes," consider pausing your current data gathering and analysis and take up a different one. Until your time management skills are where you need them to be to be successful, consider setting a timer for yourself. It can be every 30 minutes, each hour, or longer. When that timer goes off, step back from your current line of research and examine where the entire assessment is. Then you can adjust to meet your goals.

One little note to remember to examine your data and analyses for bias and fallacious logic/interpretation.

Before We Go

Tomorrow the instructor will be an observer

Ask them about ANYTHING (course-related, please ;))

When you have completed the lab, you are free to leave until tomorrow's Day 6 event!

Remember to join the OSINT community. Share your knowledge and expand all our skills

Before We Go

We have reached the end of the course. The Solo CTF and the Day 6 events are the only things left. We hope that you have learned much this week. If you have been saving your questions, now is a great time to ask them. Tomorrow, the instructor will be in "observer-only" mode and will not answer questions about tools and techniques.

After questions are answered and the Solo CTF is completed, you are free to leave. Make sure to join the OSINT community, contribute to making the tools better, give talks at conferences, and share your knowledge so that we all can grow more effective at our craft.

SEC487 Workbook

Please visit the course electronic workbook in the 487 virtual machine and begin the exercise named:

"Solo CTF"

This page intentionally left blank.