

Open and Dark Web Research: Tips and Techniques

How to access and analyze suspicious or malicious content without exposing your resources or your identity

Silo for Research

Silo for Research (Toolbox) is a secure and anonymous web browsing solution that enables users to conduct research, collect evidence and analyze data across the open, deep and dark web.

Silo for Research is built on Authentic8's patented, cloud-based Silo Web Isolation Platform, which executes all web code in a secure, isolated environment that is managed by policy, providing protection and oversight of all web-based activity.

Research teams can accomplish their goals without introducing risk to the organization or revealing intent. All web activity is logged and encrypted so compliance teams can be sure that the tools are being used appropriately.

Full Isolation:

All web code is executed on Silo servers, not end-user devices

Cloud-Based:

Turn-key, cloud-hosted solution that creates a clean instance every time

Managed Attribution:

Configure the browser fingerprint and egress location

Access Open, Deep or Dark Web:

One-click access to any destination without tainting your environment

Workflow Enhancements:

Integrated tools for content capture, analysis and storage

Complete Audit Oversight:

Encrypted audit logs of all web activity are captured in one place and easily exported

Table of Contents

Flash Reports

Investigating Site Ownership and History	2
Crypto Money Laundering on the Rise	7
Tracking Online Drug Dealers	10
Exif Data	16
Shodan	18

Product Brief

Silo for Research (Toolbox)	22
-----------------------------------	----

Infographic

Social Media Research Tools	24
-----------------------------------	----

Investigating Site Ownership and History

Analysts collecting publicly available information (PAI) encounter various sites and services with valuable information. While this information is of intelligence value, there are biases, agendas, and different reasons for the dissemination of such information.

To identify these reasons, analysts have to find information on the individuals/organizations behind the site/service which hosted, maintained, and funded them.

This information is commonly obfuscated, but accessible with proper research tools and tradecraft.

Resources Used for Site Ownership Research

Analysts can leverage the following sites and services:

- **WHOIS Records:** WHOIS records provide top level domain (e.g., russianmilitaryblog.com) information such as exact dates of registration, addresses, names, and phone numbers associated with the domain. In addition, it provides web host information.
 - URL Scan: <https://urlscan.io>
 - DomainIQ: <https://www.domainiq.com>
- **Advanced Search Engine Use:** Using advanced search engines and search engine parameters on uniquely identifying information found on the site or WHOIS records (i.e., emails, names, mail servers, other IP addresses, etc.) can provide additional information on the site or service administrator/s.
 - Carbon Date: <http://carbodate.cs.odu.edu>
 - Google Dorking: <https://www.google.com>

On the following pages we describe how to use these tools and give examples of information that can be gleaned from them.

For more information please contact osint@authentic8.com.

WHOIS Record Analysis: URLscan.io

URLscan.io conducts analysis of a domain, providing the end user with information on all HTTP connections made during the site’s retrieval, outbound links from the page, as well as detailed IP address information.

forums.airbase.ru
 148.251.51.134

URL: http://forums.airbase.ru/
 Submission: On May 24 via **2** manual (W**3** 24th 2019, 3:10:37 pm) from US

1 Summary

This website contacted **12** IPs in **4** countries across **11** domains to perform **40** HTTP transactions.
 The main IP is **148.251.51.134**, located in **Germany** and belongs to **HETZNER-AS, DE**. The main domain is **forums.airbase.ru**.

This is the first time this domain was scanned on urlscan.io!

2 structurally similar pages on different IPs, domains and ASNs found [Show Scans 2](#)

Live Information
 Certificates: 35 TLS certs observed from 2016-05-07 to 2019-05-24 [Show on crt.sh](#)
 Current Google Safe Browsing status: Clean

Domain & IP information

IP/ASNs	IP Detail	(Sub)Domains	Domain Tree	Links	Certificates
4	5	6	IP Address	AS Autonomous System	
14	148.251.51.134		24940 (HETZNER-AS)		
2	209.197.3.15		20446 (HIGHWINDS3 - Highwinds Network Group)		
1	205.185.208.52		20446 (HIGHWINDS3 - Highwinds Network Group)		
4	2a00:1450:4001:819::2002		15169 (GOOGLE - Google LLC)		
1 → 1	67.202.94.93		32748 (STEADFAST - Steadfast)		
1	185.225.208.133		13213 (UK2NET-AS)		

Screenshot [Live screenshot](#) [Full Image](#)

Detected technologies

- Twitter Bootstrap** (Web Frameworks) [Website](#)
- jQuery** (JavaScript Frameworks) [Website](#)
- Google AdSense** (Advertising Networks) [Website](#)
- Google Analytics** (Analytics) [Website](#)
- Nginx** (Web Servers) [Website](#)

Stats

40 Requests **21** Ad-blocked **0** Malicious **40%** HTTPS **46%** IPv6

Breakdown of URLscan.io result panels:

1. “Summary” provides a top level summary of what country the site is hosted in.
2. “HTTP” details how many HTTP connections are made during initial load.
3. “Links” details what other sites are linked to on the main page.
4. “IP/ASN” details the IPs of everything used upon initial load and the geographic location as well as ASN.
5. “IP Detail” contains the exact city/state/country an IP address is assigned to, and redirects.
6. “(Sub)domains” identifies how many subdomains a top level domain contains.

Example analysis of result panels:

Forums.airbase.ru, a russian military forum, uses hosting primarily in Germany, which is likely due to Germany’s strict data privacy laws. From the HTTP panel, the site uses Google Analytics for user tracking and also uses Yandex.ru for email. From the Links panel, a live “Telegram” chat is also available for users.

WHOIS Record Analysis: Hosting Research

Show hosting history for this domain:

forums.airbase.ru Check

1 Hosting Server History

What's this? This section provides a historical record of all servers this domain name was previously hosted on along with information about how many other domain names were hosted on that server at that time.

+ On 2018-12-19 the domain was hosted on [148.251.51.134](#). There are **4** other domains on this IP and **209** domains on this subnet.

+ On 2017-08-13 the domain was hosted on [148.251.51.134](#). There are **28** other domains on this IP and **615** domains on this subnet.

+ On 2016-08-20 the domain was hosted on [148.251.51.134](#). There are **28** other domains on this IP and **615** domains on this subnet.

+ On 2015-11-17 the domain was hosted on [148.251.51.134](#). There are **28** other domains on this IP and **615** domains on this subnet.

+ On 2015-10-24 the domain was hosted on [148.251.51.134](#). There are **17** other domains on this IP and **659** domains on this subnet.

+ On 2015-04-15 the domain was hosted on [148.251.51.134](#). There are **17** other domains on this IP and **659** domains on this subnet.


+ On 2015-01-16 the domain was hosted on [148.251.51.134](#). There are **17** other domains on this IP and **659** domains on this subnet.

+ On 2014-04-25 the domain was hosted on [148.251.51.134](#). There are **7** other domains on this IP and **267** domains on this subnet.

+ On 2013-08-31 the domain was hosted on [95.31.43.16](#). There are **5** other domains on this IP and **26** domains on this subnet.

2 Domains on this IP:

- [airbase.ru](#)
- [balancer.ru](#)
- [wrk.ru](#)
- [sologubov.ru](#)
- [psylab.info](#)
- [statexpert.org](#)



+ View All Domains
View IP Whois

Hosting Research provides the end user with historical information on the servers hosting the site. This can be useful as servers often host multiple sites from the same webmaster or have valuable information like the owner information available.

Breakdown of Hosting Research result panels:

1. Hosting Server History contains the historical IPs which hosted the site of interest, and details what other domains were on that server and the server's IP subnet.
2. "Domains on this IP" is opened when clicking on an IP. This details what other sites have WHOIS information that point to this IP.


Example analysis of the result panel:

Only one other IP aside from the current German IP has been used for hosting forums.airbase.ru.

This IP is 95.31.43.16, which also is used by a range of other domains — one of which, sologubov.ru has personal information on the individual behind forums.airbase.ru. This reveals the web host's full name, email, and ICQ number for further targeting.

the site of Alexander Sologubov

Cum his versare qui te meliorem facturi sunt



e-mail: mail + at + sologubov + dot + ru

ICQ: 274 - 647 - 579

Advanced Search Engine: Carbon Date

This advanced search engine automates advanced searches against web.archive.org, archive.md, Bing, bit.ly, Google, and Twitter to identify the earliest scrape/index or mention of a website on the web.

Breakdown of Hosting Research result panels:

1. “Estimated creation date” pulls the earliest date from the result set.
2. The result set shows the results from each source searched, and when available, a URL to the direct source itself.
3. The web.archive.org result is the earliest result set; with a URL you can follow to view the earliest iteration of the site.

Example analysis of the results panel:

The earliest mention of forums.airbase.ru was in October of 2003. To view the first ever scrape of this site by web.archive.org, use the URL in the “uri-m” field.

Advanced Search Engine: Google Dorking

Advanced Google search parameters and features are used in a technique called “Google Dorking”.

Users must combine various search parameters to effectively search and filter down results of interest to them.

The most commonly used Google Dorks are:

Intitle	This identifies any mention of search text in the web page title.
Allintitle	This will only identify pages with all of the search text in the web page title.
Inurl	This identifies any mention of search text in the web page URL.
Allinurl	This will only identify pages with all of the search text in the web page URL.
Intext	This will search for any mentions of search text.
Site	This will limit your results to only those within the site specified.
Filetype	This will limit your results to only the specified file type.
Cache	This will show the most recent cache of a site specified.
Around (X)	This will search for two different words within X words of one another.

Carbon Dating The Web

Predict the Birthday of a Webpage!

forums.airbase.ru

1 Estimated creation date: 2003-10-28T00:36:45

2

```

    {
      "self": "http://carbodate.cs.odu.edu/cd/forums.airbase.ru",
      "uri": "http://forums.airbase.ru",
      "estimated-creation-date": "2003-10-28T00:36:45",
      "earliest-sources": [
        "web.archive.org"
      ],
      "sources": {
        "web.archive.org": {
3
          "uri-m": "http://web.archive.org/web/20031028003645/http://",
          "memento-datetime": "2003-10-28T00:36:45",
          "memento-pubdate": "",
          "earliest": "2003-10-28T00:36:45"
        },
        "archive.md": {
          "uri-m": "http://archive.md/20140825015428/http://forums.ai",
          "memento-datetime": "2014-08-25T01:54:28",
          "memento-pubdate": "2014-08-25T01:54:28",
          "earliest": "2014-08-25T01:54:28"
        },
        "bing.com": {
          "earliest": ""
        },
        "bitly.com": {
          "earliest": "2014-03-21T04:07:24"
        },
        "google.com": {
          "earliest": ""
        },
        "last-modified": {
          "earliest": "2019-05-24T15:11:09"
        },
        "pubdate": {
          "earliest": ""
        },
        "twitter.com": {
          "earliest": ""
        }
      }
    }
    
```

The most commonly used Boolean logic search operators are:

AND	This will search for content mentioning two phrases anywhere.
OR	This is used in multi part search, and will search for content mentioning any combination of the first search term and two unique second search variables.
*	This will act as a wildcard and search for any word or phrase.
-	This will exclude any specific word or phrase (if using brackets or quotes). <i>Note: this is a dash sign.</i>
()	This will group specific terms or search operators together.

Example analysis using advanced Google Search parameters:

`site:sologubov.ru ICQ OR email`

This search will find mentions of ICQ or email on a site of interest, resulting in an ICQ number and email previously unknown to an analyst.

`site:forums.airbase.ru contact OR admin OR mod OR moderator OR donation`

This search will find uniquely identifying information that can be linked to a person, such as mentions of a moderator, a contact page, or a donation page (such as Paypal, Bitcoin, etc), resulting in multiple pages with mentions of the moderator and a donation page for their health bills.

`"95.31.43.16"`

This search will find exact mentions of forums.airbase.ru, resulting in mentions on another forum of Russian censorship of the servers IP address.

Conclusion

This workflow covers how to investigate the ownership and hosting information related to a site/service of interest. Results from the analysis include key identifiers such as server IPs, other related domains, and the webhost’s email address/name/ICQ number that can then be incorporated further into a finished intelligence product.

For more information please contact osint@authentic8.com.



CONNECT WITH US

+1 877-659-6535
www.Authentic8.com



PROTECT THE THINGS YOU CARE ABOUT FROM THE THINGS YOU CANNOT TRUST

Authentic8 enables anyone, anywhere, on any device to experience the web without risk. The Silo Web Isolation Platform by Authentic8 separates the things you care about like apps, data and devices, from the things you cannot trust like public websites, external users and unmanaged devices. Silo executes all web code in a secure, isolated environment that is managed by policy, to provide protection and oversight.

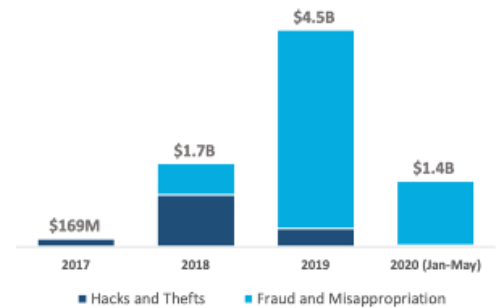
Today, the world’s most at-risk organizations rely on Silo to deliver trust where it otherwise cannot be guaranteed.

Crypto Money Laundering on the Rise

Since Bitcoin's historic rise, the [number of users participating in the cryptocurrency ecosystem](#) has reached nearly 69 million. This increase in active cryptocurrency users has also led to a surge of cryptocurrencies being used to launder money. According to the United Nations Office on Drugs and Crime, [money laundering costs the global economy](#) between \$800 billion and \$2 trillion per year — that's two to five percent of the world's gross domestic product.

Money laundering in the era of cryptocurrency is [more convenient than traditional laundering schemes](#), with multiple types of available crypto coins and a good amount of anonymity for traders to hide their true identities. Add the dark web, and you have a murky, constantly changing and decentralized environment that creates many additional challenges for investigators.

Blockchain Fraud Continues to Vastly Exceed Hacks and Thefts in 2020



Source: CipherTrace Cryptocurrency Intelligence

FinCEN Warns of Threats Posed by Virtual Currency Misuse

Criminals continue to exploit virtual currency to support illegal activity, money laundering and other behavior endangering U.S. national security. To help financial institutions, law enforcement and regulators who work with convertible virtual currencies (CVCs), the Financial Crimes Enforcement Network (FinCEN) [offers guidance](#) to assist organizations in identifying and reporting suspicious activity. The advisory highlights the risks associated with dark web marketplaces, peer-to-peer (P2P) exchangers, unregistered money services businesses and CVC kiosks. It also gives organizations a set of tools to help identify unregistered financial activity and suspicious virtual currency purchases, transfers and transactions.

The Need for Effective AML Programs

The FinCEN regulatory framework mandates that businesses develop, implement, and maintain an effective anti-money laundering program ("AML program") that is designed to prevent organizations from being used to facilitate money laundering and the financing of terrorist activities.

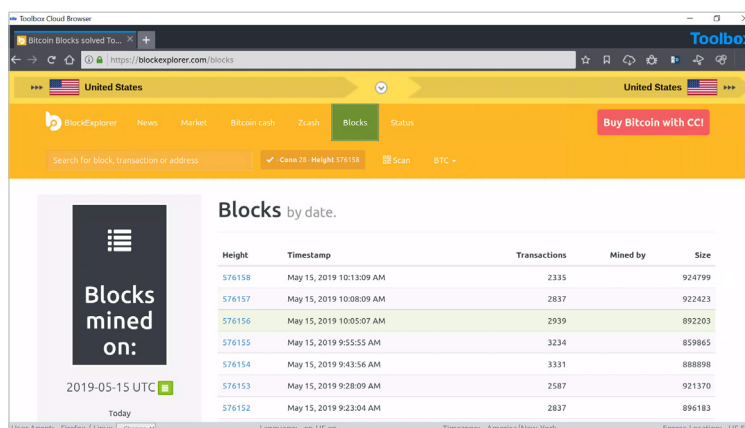
The minimum set of requirements for an AML program include:

- Establishment of policies, procedures, and internal controls designed to assure ongoing compliance (including verifying customer identification, filing reports, creating and retaining records and responding to law enforcement requests)
- Designation of individuals responsible to assure day-to-day compliance with the program
- Training for appropriate personnel, including training in the detection of suspicious transactions
- Ongoing independent reviews to monitor and maintain an adequate program

Without sufficient controls in place, financial institutions cannot reasonably assess and mitigate the potential risks posed by a customer’s source of funds, and criminals can exploit the U.S. financial system by engaging in illicit transactions. Individuals engaged in illicit activity will continue to exploit these vulnerabilities as long as the perceived risk of detection is less than that of using traditional financial institutions.

Tracking Cryptocurrencies

How can CVC transactions be tracked? One popular way is using blockchain technology. Blockchain is an open, decentralized ledger that records transactions between two parties in a permanent way without needing third-party authentication. For example, every transaction involving a Bitcoin address is stored forever in the blockchain; however, Bitcoin addresses are pseudonyms, meaning that the identity of the address owner (i.e., who receives bitcoin at that address) is generally unknown.



Height	Timestamp	Transactions	Mined by	Size
576158	May 15, 2019 10:13:09 AM	2335		924799
576157	May 15, 2019 10:08:09 AM	2837		922423
576156	May 15, 2019 10:05:07 AM	2939		892203
576155	May 15, 2019 9:55:55 AM	3234		859865
576154	May 15, 2019 9:43:56 AM	3331		888898
576153	May 15, 2019 9:28:09 AM	2587		921370
576152	May 15, 2019 9:23:04 AM	2837		896183

Bitcoin blockchains from blockexplorer.com

If a user’s address is ever linked to their identity, every transaction will be linked to that user. Below are examples of OSINT tools that allow investigators to search by block number, address, block hash, transaction hash or public key to find out more information on bitcoin transactions.


- <https://www.blockchain.com/explorer>
- <https://blockchain.info/>
- <https://www.chainalysis.com/>
- <https://bitcoinwhoswho.com/>

To prevent tracking on their transactions, money launderers have begun to use a system known as cryptocurrency tumblers. Cryptocurrency tumblers mix potentially identifiable currency with untraceable currency to make it harder to track.

Some addresses can be grouped by their ownership, using behavior patterns and publicly available information from off-chain sources. The challenge for forensic investigators, as usual, is to identify the persons behind the keyboard, which may be accomplished through a mixture of traditional investigative and digital forensic techniques.

Bitcoin Address Reports

Once a Bitcoin address is identified, it can be run through a blockchain tracking tool. Using bitcoinwhoswho.com, investigators can generate a report for bitcoin address 1Hz96kJKF2HLPGY15JWLB5m9qGNxvt8tHJ.

BTC Address	1Hz96kJKF2HLPGY15JWLB5m9qGNxvt8tHJ	# Website Appearances	9	
Wallet Name	000a027d20045b7d	Last Transaction IP	47.254.169.156, 52.60.49.56, 148.251.139.241	
Current Balance	112.11330270	Total Received	161472.17413649	
# Transactions	13121	# Output Transactions	Loading...	
First Transaction	13 Jun 16	Last Transaction	16 May 19	
Last Known Input	Loading...	Last Known Output	1BMjmWxy7h... 16 May 19	
Repeated Inputs From (50 most recent transactions)	...	Repeated Outputs To (50 most recent transactions)	1BMjmWxy7h... 13 1PqU8hFVgy... 9 1NSUvXctWw... 4	

Bitcoin address report from bitcoinwhoswho.com

Probable fields of interest:

- **Current Balance/Total Received:** This data point allows analysts to hypothesize which type of address they're dealing with. Due to the high volume of transactions, this wallet likely belongs to a bitcoin miner.
- **Last Transaction IP:** Analysts can view the last known IP to relay an output transaction involving a selected address. Repeated use of an IP can be used as a unique identifier.
- **Website Appearances:** Provides a view of any site where this exact bitcoin address appeared, which could be of value for identifying reputation/type of transactions.
- **Repeated Inputs From/Repeated Outputs To:** This data point allows analysts to view the 50 most recent Bitcoin addresses involved with incoming and outbound transactions associated with this address. By looking at the transaction history and frequently interacted-with wallets, investigators can engage in network and link analysis to identify patterns and possible relationships between the disparate Bitcoin addresses.



CONNECT WITH US

+1 877-659-6535
www.Authentic8.com



PROTECT THE THINGS YOU CARE ABOUT FROM THE THINGS YOU CANNOT TRUST

Authentic8 enables anyone, anywhere, on any device to experience the web without risk. The Silo Web Isolation Platform by Authentic8 separates the things you care about like apps, data and devices, from the things you cannot trust like public websites, external users and unmanaged devices. Silo executes all web code in a secure, isolated environment that is managed by policy, to provide protection and oversight.

Today, the world's most at-risk organizations rely on Silo to deliver trust where it otherwise cannot be guaranteed.

Tracking Online Drug Dealers

Drug Dealers Using Social Media to Sell Illegal Narcotics

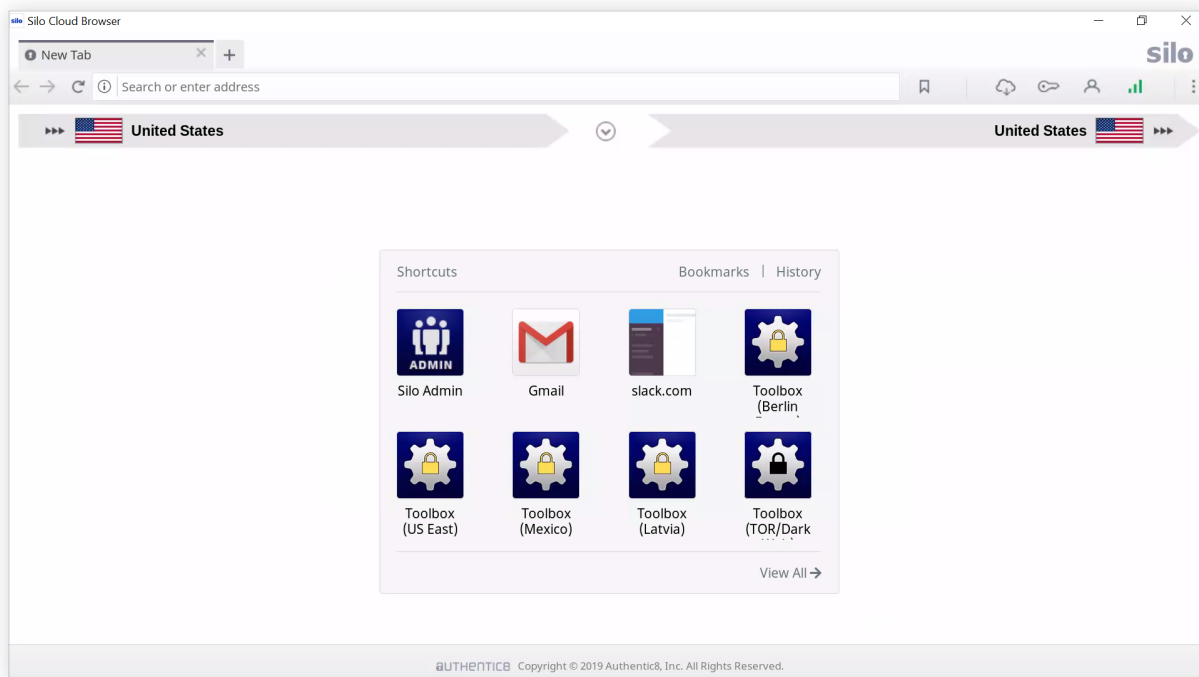
The rise in social media over the past ten years has led to drug dealers using various social media platforms to sell illegal narcotics on the open web. Investigators need a safe and anonymous browsing and research framework that allows them to investigate social media drug dealers without the risk of being identified or infecting their endpoint with malicious web code. This workflow will cover how the Silo Web Isolation Platform and managed attribution solution can be utilized to identify and investigate social media drug dealers anonymously.

Silo Web Isolation Platform

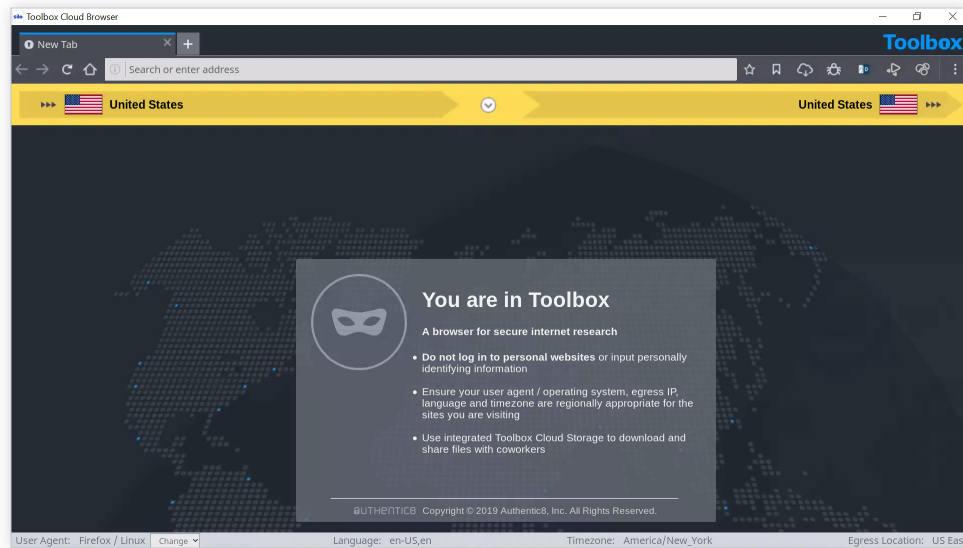
Silo for Research (Toolbox) is a secure and anonymous web browsing solution that enables users to conduct research, collect evidence and analyze data across the open, deep and dark web.

Silo for Research is built on Authentic8's patented, cloud-based Silo Web Isolation Platform, which executes all web code in a secure, isolated environment that is managed by policy, providing protection and oversight of all web-based activity.

Research teams can accomplish their goals without introducing risk to the organization or revealing intent. All web activity is logged and encrypted so compliance teams can be sure that the tools are being used appropriately.



Silo for Research allows users to spoof IPs to over 30 locales worldwide, manipulate their hardware and software fingerprints, and to collect, annotate, and store internet-based publicly available information (PAI). It includes tools for post-fetch language translation, web-code and traffic analysis, and linkage tracking.



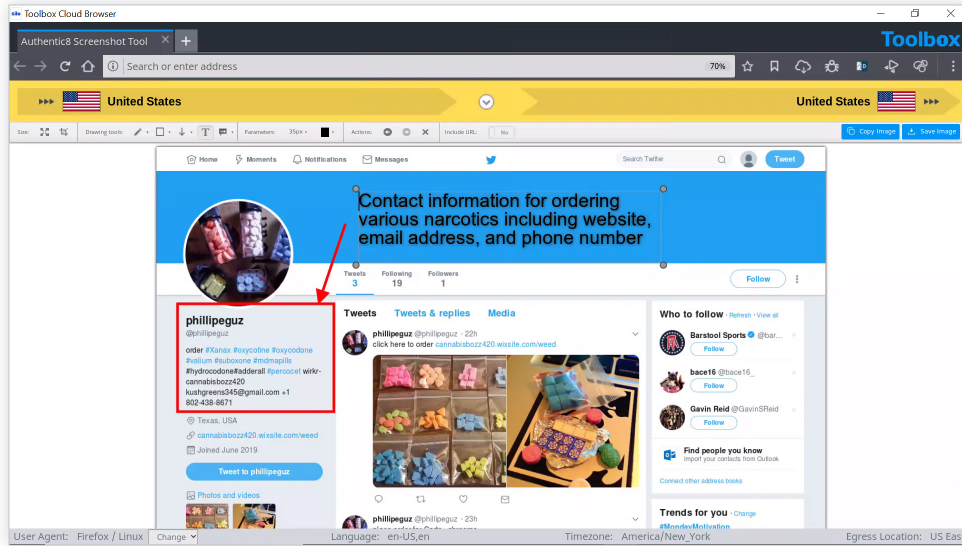
Identifying and Investigating Drug Dealers on Social Media with Silo for Research

The first step when conducting an investigation using Silo for Research is to select a regionally appropriate egress location and a user agent string that matches regional norms. This workflow will use the US and Google Chrome running on a Windows 10 machine as the user agent string. This process allows investigators to blend in as locals of that area.

When conducting research on social media, there are various data capture tools included with Silo for Research that can be used for gathering intelligence. This first is a video download tool; this tool allows investigators to simply download any video currently playing on their screen to save as evidence. The second is a screenshot tool that gives investigators the ability to take a screenshot of an entire page. The screenshot tool also gives investigators the ability to edit the screenshot by including boxes, arrows, and text to highlight important information, as well as the ability to include the URL of where the screenshot was taken. This allows investigators to easily return to that page to gather additional intelligence.

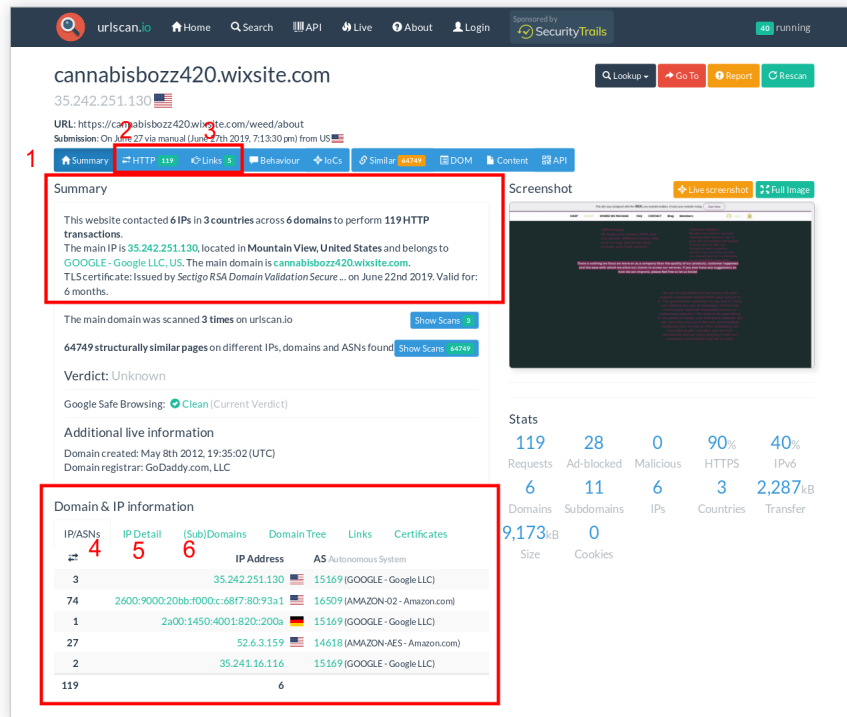
By conducting a search on Twitter for #xanax, the Twitter user @phillipeguz was identified as an account using Twitter to market and sell illegal narcotics. Shown on this profile is information on how to place an order; the information listed includes a website, email address, and phone number. This information can now be run through additional search engines to possibly identify the owner of the account.

FLASH REPORT: TRACKING ONLINE DRUG DEALERS



Resources for Site Ownership Research

WHOIS Records: WHOIS records provide top level domain information such as exact dates of registration, addresses, names, and phone numbers associated with the domain. Additionally, it provides web host information. @phillipeguz posted the website <https://cannabisbozz420.wixsite.com/weed/about> on their Twitter feed as a location to purchase the illegal narcotics. Using <https://urlscan.io>, a report was generated for this site.



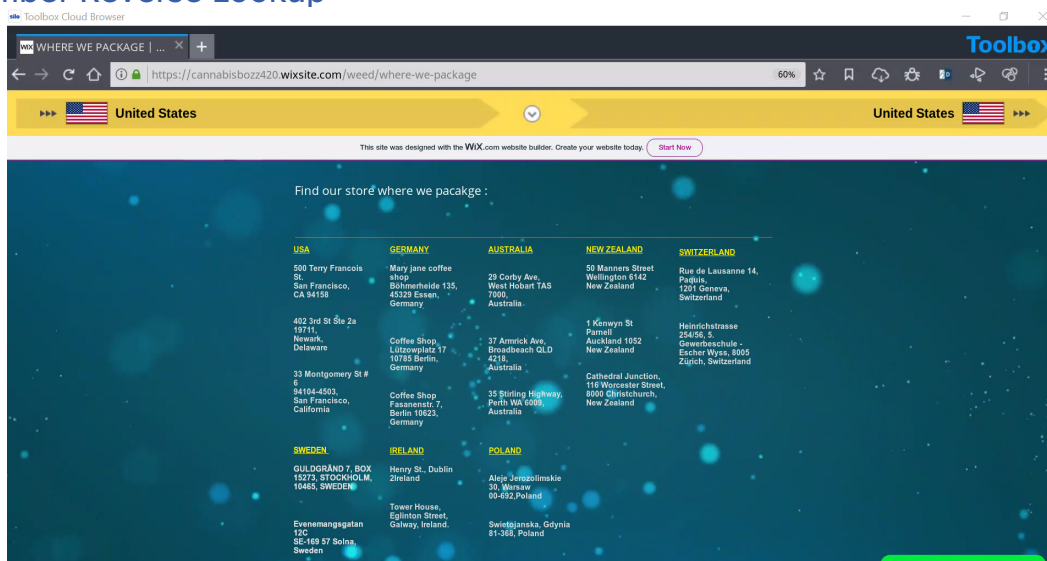
Breakdown of URLscan.io result panels:

1. “Summary” provides a top level summary of what country the site is hosted in.
2. “HTTP” details how many HTTP connections are made during initial load.
3. “Links” details what other sites are linked to on the main page.
4. “IP/ASN” details the IPs of everything used upon initial load and the geographic location as well as ASN.
5. “IP Detail” contains the exact city/state/country an IP address is assigned to, and redirects.
6. “(Sub)domains” identifies how many subdomains a top level domain contains.

Example analysis of result panels:

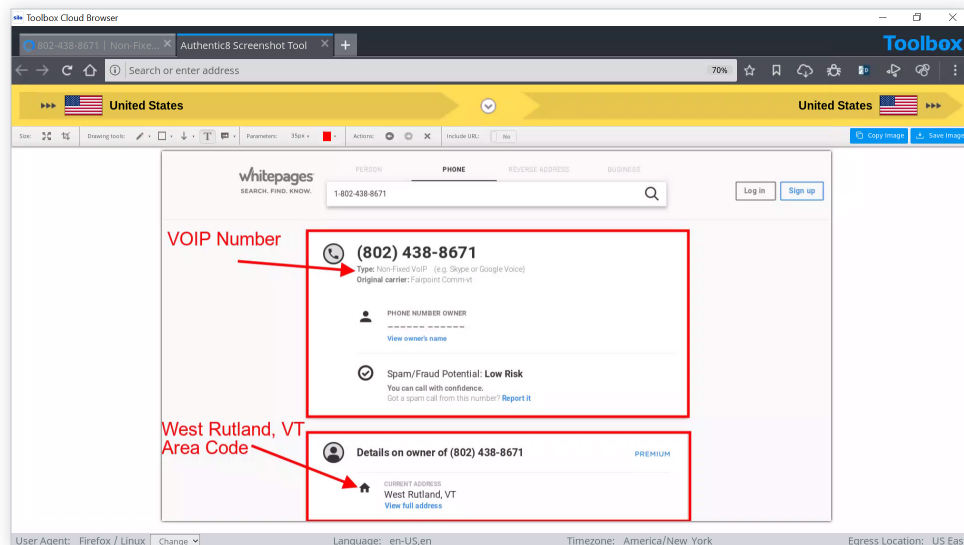
According to the generated report, <https://cannabisbozz420.wixsite.com/weed/about> uses hosting primarily in the United States but also has hosting in Germany. This means that the distribution could also include locations outside the United States. On the website, the site owners also listed packaging locations in the United States, Germany, Australia, New Zealand, Switzerland, Sweden, Ireland, and Poland. The following screenshot from their website depicts their packaging locations around the world. It appears that the domain was registered by godaddy.com. This information could be used to send out a subpoena or court order to godaddy.com to find out who registered the domain with them.

Phone Number Reverse Lookup



The phone number 1+802-438-8671 was also listed as contact information for ordering narcotics from this Twitter page. Having this number available is extremely valuable for the investigation. The number can be run through a reverse phone number search engine to identify the subscriber information. The following screenshot is from a report generated by <https://www.whitepages.com/phone/1-802-438-8671> for the listed phone number.

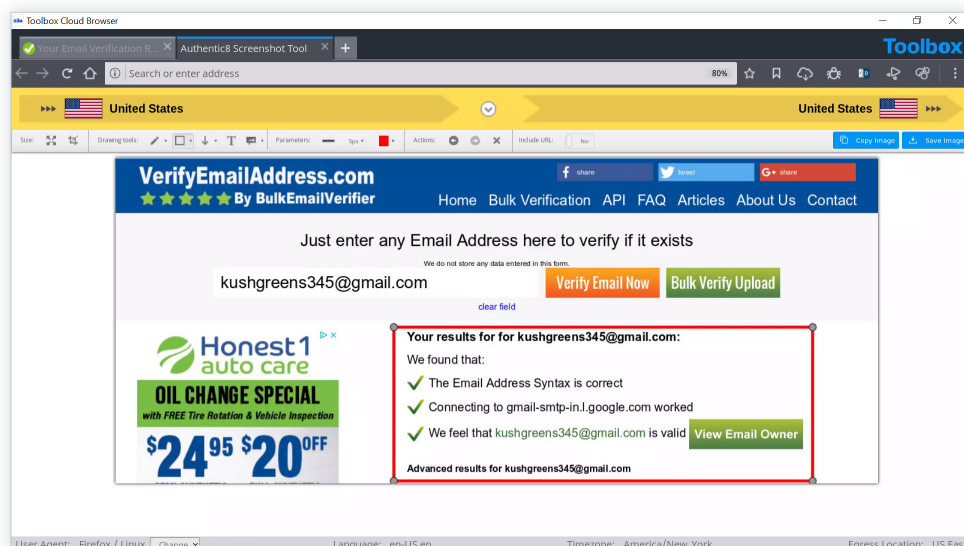
Example analysis of result panels:



Although there is no identity listed for the number and the number is associated with a voice over internet protocol (VoIP), there is some valuable information that can be pulled from the report. Seeing that the number has a Rutland, Vermont area code is telling: due to the website listing a packaging location on the East Coast, it is possible that the East Coast is their shipping HQ.

Searching for Additional Social Media Profiles by Email

The third piece of contact information listed on this Twitter page is the email address kushgreens345@gmail.com. Once a possible email address is identified for a target, it can be run through <https://verifyemailaddress.com> to verify that it is a legitimate email address. Once an email address is verified, a subpoena or court order can be sent to the email provider to identify who owns and operates that email address. The screenshot below depicts the results from <https://verifyemailaddress.com> for the email address kushgreens345@gmail.com, and it is in fact a legitimate email address.



Conclusion

With drug dealers increasingly utilizing social media to distribute illegal narcotics, investigators need a safe and anonymous method to investigate and capture social media data. This workflow covered how Silo for Research can be used by investigators to safely and anonymously investigate and capture data from social media drug dealers.

For more information, please contact osint@authentic8.com.



CONNECT WITH US

+1 877-659-6535
www.Authentic8.com



PROTECT THE THINGS YOU CARE ABOUT FROM THE THINGS YOU CANNOT TRUST

Authentic8 enables anyone, anywhere, on any device to experience the web without risk. The Silo Web Isolation Platform by Authentic8 separates the things you care about like apps, data and devices, from the things you cannot trust like public websites, external users and unmanaged devices. Silo executes all web code in a secure, isolated environment that is managed by policy, to provide protection and oversight.

Today, the world's most at-risk organizations rely on Silo to deliver trust where it otherwise cannot be guaranteed.

What is Exif Data?

When a digital image is captured, metadata specific to that image is stored. This information is called Exchangeable Image File Format data, or Exif data. Some examples of Exif data are date, time, and file size. This information can be extremely useful when conducting image analysis. Analysts can exploit Exif data to find the location of the image, camera make and model, and other information that is valuable to the intelligence production cycle.

Incorporating Exif Data

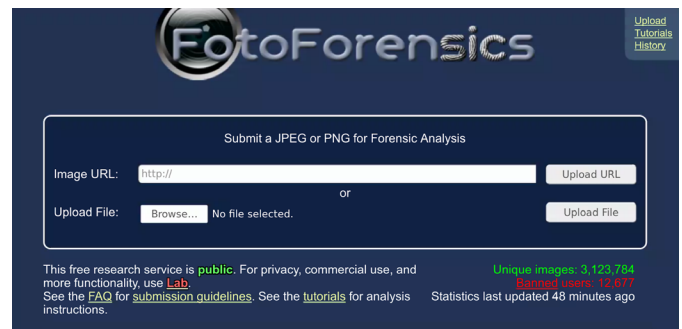
To find Exif data, an analyst can use a number of different tools. [FotoForensics](#) is the service used for the workflow described here. An analyst can take the image of the cargo ship and upload it into FotoForensics to analyze the available Exif data. This image was found in a [ship-spotting forum](#).

User-uploaded images to forums will likely have their Exif data. However, if the analyst tries to pull Exif data from an image on social media, there will likely be little to no data present. Social media platforms have begun to strip Exif data off of user images to protect user privacy.

Once on FotoForensics, the analyst will have two options for image analysis. The analyst can paste an image URL or upload a file for analysis.

For this workflow, the analyst can save the above image of the cargo ship, and then upload the .jpg file into FotoForensics.

When the upload is complete the analyst should select the metadata field (#1). The analyst can then scroll down and begin to review information pertinent to the investigation.



FotoForensics user interface: <http://fotoforensics.com>



FotoForensics, post image upload with metadata selected.

After reviewing the Exif data collected by FotoForensics, a few pieces of information stand out. In the above images, the analyst can glean what type of device was used to capture this image (#2). This information can be useful when investigating a party of interest that may have a standard issue camera for reconnaissance.

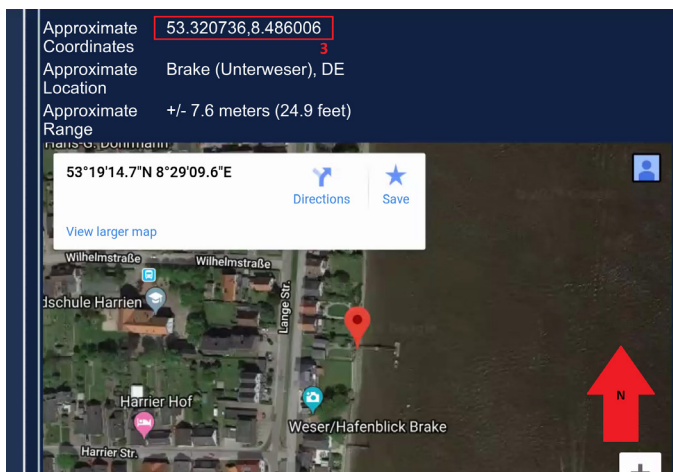
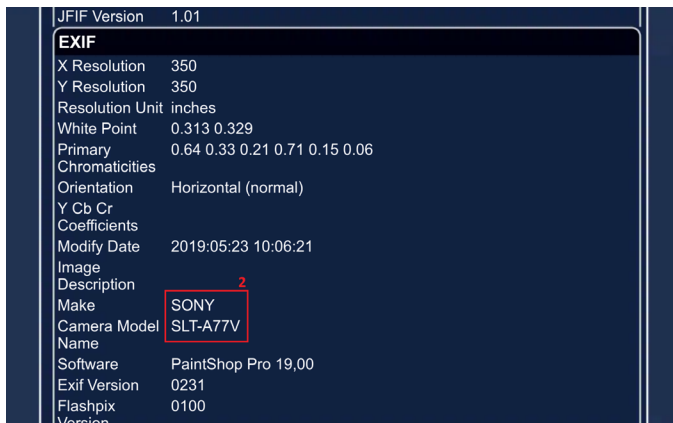
FotoForensics also provides the analyst with an approximate latitude and longitude coordinate (#3). This coordinate can be further incorporated into a targeting packet or reconnaissance mission.

Overall, the information captured from Exif data can greatly enhance a unit's analytic ability. The exploitation of images, whether of an adversarial object or person or of a location, can help the analyst to further understand their battlespace or objective.

Conclusion

This workflow covers how to extract and incorporate Exif data into the intelligence product. The analyst found a .jpg file of a cargo ship and leveraged FotoForensics to conduct Exif data analysis. Results from the analysis included key identifiers such as equipment used and location data that can then be incorporated further into a finished intelligence product.

For more information please contact osint@authentic8.com.



What is Shodan?

Shodan is “the world’s first search engine for internet-connected devices.”¹ But what exactly does this mean?

Most search engines are text indexes, meaning they allow search for content based on keywords. However, the task of scanning, indexing the ports and services running, and then searching for internet-connected devices at the scope and scale of the internet has been largely impossible to do.

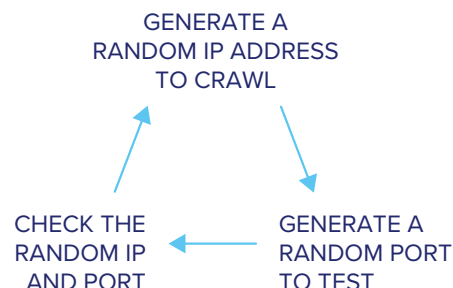
With Shodan, it is now possible to identify nearly any internet-connected device, such as industrial control systems running specific software, internet-of-things devices like smart TVs, FTP servers with sensitive information, and even Very Small Aperture Terminals (VSATs) on naval vessels.

How Shodan Works

Shodan maintains servers across the globe that scan the internet-connected devices and harvest the banner of whatever is running on the server.

The diagram at right shows how these servers crawl.

These internet-connected devices return different banners depending on the different service running on it.





Example Search Returns

Two examples are below, one for an IP camera and one for an FTP server (FTP runs on port 21):

Basic Shodan Searches/Filters

Document Error: Unauthorized

62.112.117.205
OA0 MGTS
 Added on 2019-05-07 10:56:51 GMT
 Russian Federation, Odintsovo
 Technologies: IIS\;confidence:50 

```

HTTP/1.1 401 Unauthorized
Server: Cam-Webs
Date: Tue May 7 13:20:55 2019
WWW-Authenticate: Basic realm="Megapixel_IP_Camera"
Pragma: no-cache
Cache-Control: no-cache
Content-Type: text/html
  
```

188.225.26.71

vds-olgafirsova.timeweb.ru
hosting & vds
 Added on 2019-05-28 17:02:17 GMT
 Russian Federation

```

220 (vsFTPD 3.0.2)
230 Login successful.
214-The following commands are recognized.
ABOR ACCT ALLO APPE CDUP CWD  DELE EPRT EPSV FEAT HELP LIST MDTM MKD
MODE NLST NOOP OPTS PASS PASV PORT PWD  QUIT REIN REST RETR RMD  RNFR
RNT0 SITE SIZE SMNT STAT STOR STOJ STRU SYST TYPE USER XCUP XCWD XMKD...
  
```

¹ <https://github.com/polarityio/shodan>

FLASH REPORT: SHODAN

Shodan allows for advanced search using filters. Filters are entered in a simple format: a filter, a colon, and the search value, with no spaces between these three components.

Filter format	<code>filtername:value</code>
Filter example	<code>City:Moscow</code>

If searching a value that includes a space, double quotes must be used.

Filter example	<code>City:"Saint Petersburg"</code>
----------------	--------------------------------------

Examples of Shodan's most useful geographic filters:

Country using 2 letter geocode	<code>country:XX</code>
City using city name	<code>city:cityname</code>
Geographic coordinates in a bounding box	<code>geo:top-left-lat,top-left-long, top-right-lat,top-right-long</code>
Region	<code>region:region-name-or-state</code>

These filters are useful when attempting to identify something of interest in a specific AOR.

For example, a search for webcam `City:Incirlik` would find webcams, with some hopefully located near Incirlik Air Base.

Examples of software-focused filters:

Firewall port	<code>port:XX</code>
Product name	<code>product:XX</code>
Product version	<code>version:XX</code>
Product vulnerability CVE	<code>vuln:XX</code>

These filters are useful when searching for a particular technology, like a database, a file server, or vulnerable software.

For example, a search for `port:21 country:"RU" "login successful"` would find file transfer protocol (FTP) servers in Russia that do not require logins. This could yield valuable unsecured information if found in a location of interest, or can be used as a non-attributable temporary data transfer point.

Examples of organization-focused filters:

Device hostname	<code>hostnames:XX</code>
Organization assignment	<code>org:XX</code>
Network CIDR range	<code>net:XX</code>

Examples of Shodan's temporal filters:

Results before a given date	<code>before:00/00/0000</code>
Results after a given date	<code>after:00/00/0000</code>

Finding Open Databases

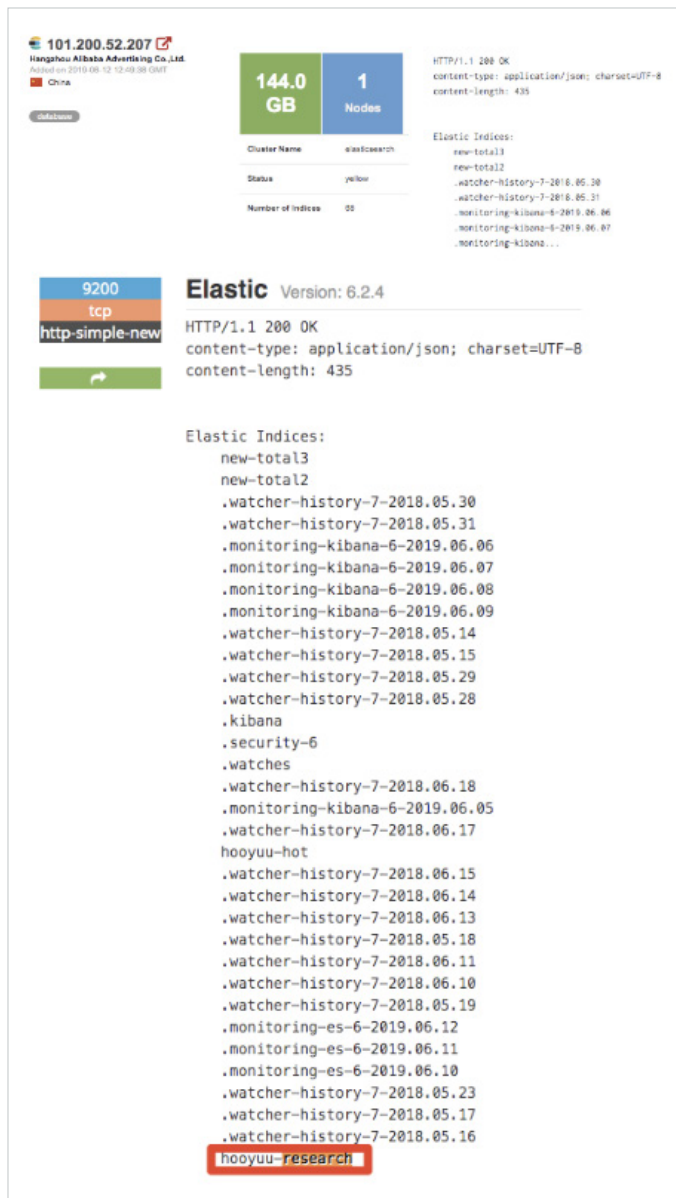
A few databases openly list their indices: MongoDB, ElasticSearch, and CouchDB.

Below are the baseline searches that allow you to quickly identify open databases with potentially valuable information sources.

Example database searches:

Elasticsearch databases	<code>product:elastic port:9200</code>
MongoDB databases	<code>product:MongoDB</code>
CouchDB databases	<code>product:couchdb</code>
Kibana visualization of Elasticsearch	<code>kibana content-length: 217</code>
Gitlab software repos	<code>http.favicon.hash:1278323681</code>
Rsync utilities	<code>product:rsyncd</code>
Jenkins software automation	<code>jenkins 200 ok</code>

Combining these search filters and other key phrases allows analysts to identify high value and unsecured information.



Example search for Elasticsearch databases in China mentioning “research”:

```
product:elastic port:9200 country:cn
research
```

This results in identifying an IP address hosting an open elasticsearch index with mentions of research. In this case, the research is about “Hooyuu,” a Chinese social media site.

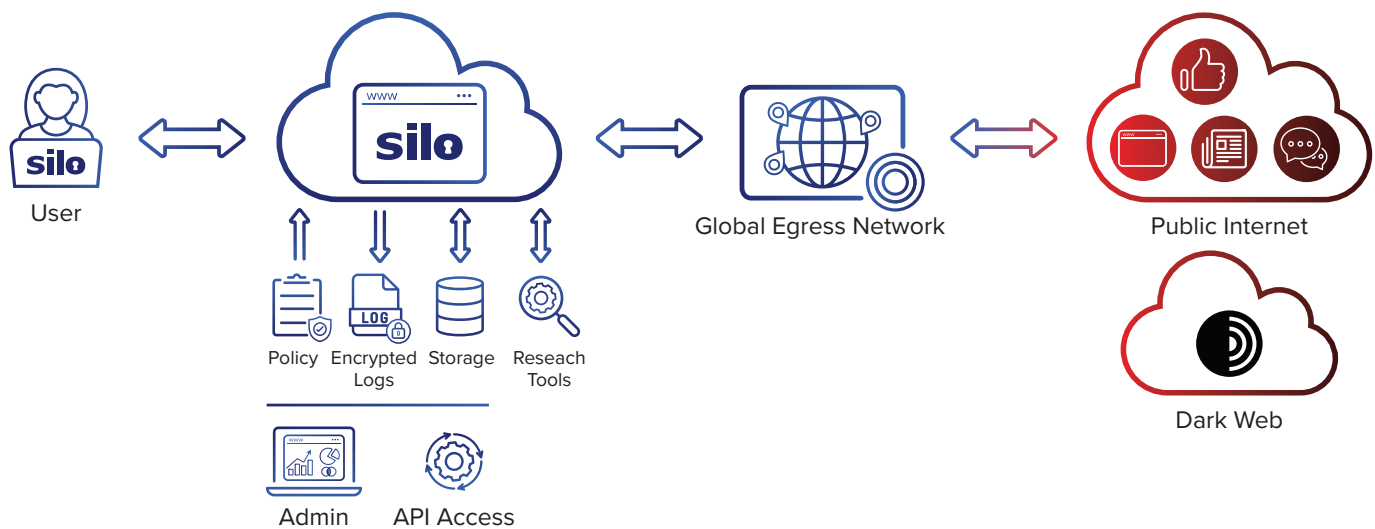
The others range from what looks like security research, notifications, and some form of alerting.

For more information please contact osint@authentic8.com.

Silo for Research

Secure and Anonymous Online Investigations

Silo for Research (Toolbox) is a secure and anonymous web browsing solution that enables users to conduct research, collect evidence and analyze data across the open, deep and dark web.

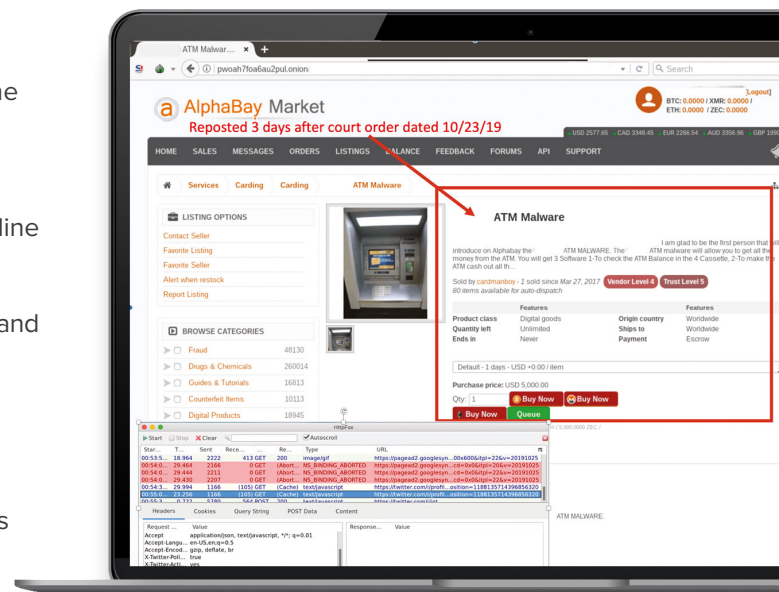


Silo for Research is built on Authentic8’s patented, cloud-based Silo Web Isolation Platform, which executes all web code in a secure, isolated environment that is managed by policy, providing protection and oversight of all web-based activity.

Research teams can accomplish their goals without introducing risk to the organization or revealing intent. All web activity is logged and encrypted so compliance teams can be sure that the tools are used appropriately.

The world’s most at-risk enterprises and government agencies rely on Silo for Research to conduct secure and anonymous online investigations:

- **Criminal investigations:** Comply with chain-of-custody policy and securely collect evidence on the open, deep, or dark web
- **Cyber threat intelligence:** Access and analyze suspicious or malicious content with 100% isolation from corporate infrastructure
- **Financial investigations:** Keep your online fraud investigations anonymous and secure, even on the dark web
- **Open-source intelligence (OSINT):** Disguise your identity with a managed attribution platform and global egress network

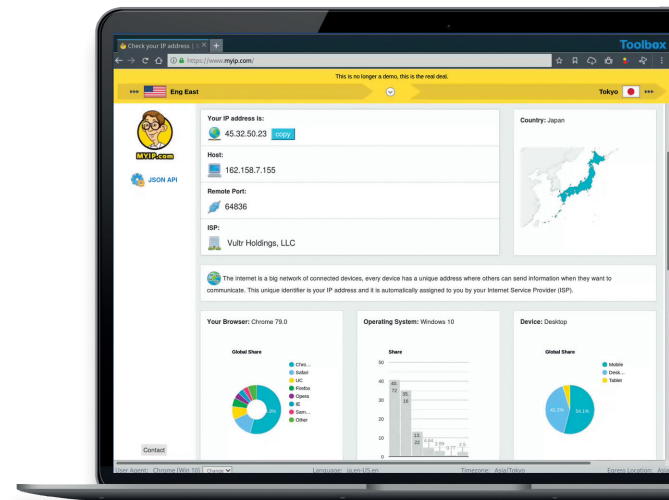


Features and Benefits

Features	Benefits
Full Isolation: All web code is executed on Silo servers, not end-user devices	Potentially unsafe content never touches your organizations' assets
Cloud-Based: Turn-key, cloud-hosted solution that creates a clean instance every time	Seamless, immediate deployment with on-demand access from anywhere
Managed Attribution: Configure the browser fingerprint and egress location	Blend in with the crowd to not trip off your intent to others
Access Open, Deep or Dark Web: One-click access to any destination without tainting your environment	Maintain policy, while providing a secure way for users to interact with any destination
Workflow Enhancements: Integrated tools for content capture, analysis and storage	Improve time to insight for analysts with integrated tools
Complete Audit Oversight: Encrypted audit logs of all web activity are captured in one place and easily exported	Simplify analyst compliance and audit, and improve case documentation

Use the Silo Web Isolation Platform to maintain data security, and respect the privacy of your users. Our compliance:

- FedRamp In Progress
- HIPAA
- PII
- PCI
- GDPR
- CCPA



PROTECT THE THINGS YOU CARE ABOUT FROM THE THINGS YOU CANNOT TRUST

Authentic8 enables anyone, anywhere, on any device to experience the web without risk. The Silo Web Isolation Platform by Authentic8 separates the things you care about like apps, data and devices, from the things you cannot trust like public websites, external users and unmanaged devices. Silo executes all web code in a secure, isolated environment that is managed by policy, to provide protection and oversight.

Today, the world's most at-risk organizations rely on Silo to deliver trust where it otherwise cannot be guaranteed.

Top 5 Social Media Research Tools for Online Investigations

Silo for Research (Toolbox)

Secure and Anonymous Online Investigations

Silo for Safe Access (Cloud Browser)

Put IT Back in Control of the Web

silo
By Authentic8

www.authentic8.com



INSTAGRAM SEARCH APPS

1 BILLION
users on Instagram

- Google Reverse Image Search www.google.com/imghp?hl=en&tab=ri&ogbl

Paste or upload images to find similar images as well as other websites where that exact image has appeared. Great for finding additional social media pages for a subject.



TWITTER SEARCH APPS

330 MILLION
users on Twitter

- Tweet Beaver tweetbeaver.com

Receive a complete analysis on an individual's Twitter profile, including tweets, replies, retweets, hashtags used, sources of tweets (Android, iPhone, online), and geotagged tweets.

- SocialBearing socialbearing.com

Comprehensive suite of searching options. Make links across variables.



REDDIT SEARCH APPS

542 MILLION
users on Reddit

- Track Reddit www.trackreddit.com

Receive notifications on keyword searches to identify who is discussing certain topics.

- Reddit Insight www.redditinsight.com

Search by Reddit username to receive a complete overview of that username, including when the account was created, email address, posts, and most common subreddits where the account has posted.

Contributors

Nick Finnberg started his intelligence career as an Army National Guard Intelligence Analyst, dedicating four years to the counterdrug task force and specializing in large scale money laundering investigations and open source intelligence gathering. At Authentic8, Nick provides consultation and training to analysts as they enhance their anonymous online investigation capabilities.

Nick Espinoza is a career technologist working at the intersection of the Silicon Valley and the US Defense and Intelligence community. His work has been focused on the collection and structuring of information for companies such as Palantir, Recorded Future, and Authentic8 on a range of missions such as insider threat, network defense, targeting and more.

Alec Feltri is a former US Army Intelligence professional with experience in tactical, strategic, and targeting intelligence processes. His current position with Authentic8 allows him to best combine his experience with the cutting edge of technology.



CONNECT WITH US

+1 877-659-6535

www.Authentic8.com



PROTECT THE THINGS YOU CARE ABOUT FROM THE THINGS YOU CANNOT TRUST

Authentic8 enables anyone, anywhere, on any device to experience the web without risk. The Silo Web Isolation Platform by Authentic8 separates the things you care about like apps, data and devices, from the things you cannot trust like public websites, external users and unmanaged devices. Silo executes all web code in a secure, isolated environment that is managed by policy, to provide protection and oversight.

Today, the world's most at-risk organizations rely on Silo to deliver trust where it otherwise cannot be guaranteed.