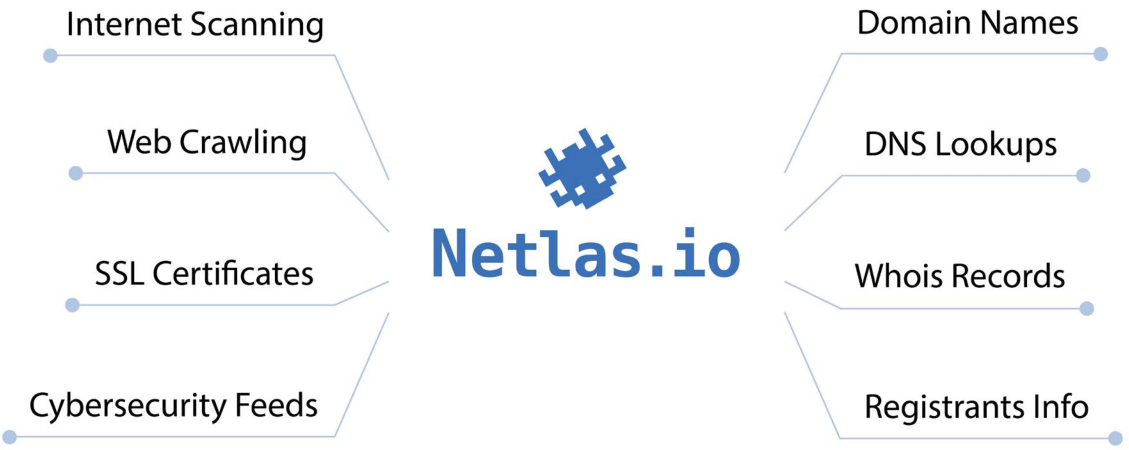
Security Research and Analysis with Netlas.io

Netlas.io

This document is designed to help you familiarize yourself with Netlas.io and its applications. It describes several of the most in-demand use cases for the service, falling into one of two categories: large-scale cybersecurity research or the security analysis of information systems.

Netlas.io is a technical atlas of the internet tailored for IT and cybersecurity professionals. Netlas contains up-to-date information from scanning the entire range of IPv4 addresses on the internet and hundreds of millions of domain names. It includes a database of internet network DNS records, WHOIS databases for both IP addresses and domain names, as well as an SSL certificate database. All this data is parsed and well-structured, enriched with tags, technology names, vulnerability data, geographical locations, and more.



Netlas.io is a search engine and a set of complementary tools. Netlas users access the data through the web interface, API, and command line. The web interface provides an effortless way to explore search results, and the intuitively understandable syntax of queries allows users to start working after just a few trial queries. Users leverage the API to integrate Netlas into their applications. The command-line interface enables the easy and swift development of scripts to automate work processes.

Table of Contents

Security Research and Threat Intelligence

Netlas is an invaluable source of data when it comes to global research in the field of cybersecurity, as well as identifying security threats on a scale encompassing the entire internet.

It would be wrong to say that Netlas (or any of our competitors) scans the entire Internet. It is not even possible to scan every IPv4 address on all 65535 ports in a reasonable amount of time. Thus, we scan the Internet against a specific list of commonly used ports.

> The number of ports to scan increases from time to time in proportion to our infrastructure. To determine which ports are being scanned, you can use the grouping feature in the search results. Grouping can be done by the «port» field (search link).

We are constantly improving the scanning technology and the scanners themselves. At present, Netlas identifies and parses over 25 application protocols, including:

- commonly used DNS, FTP, HTTP, SMB, NETBIOS, SNMP, SOCKS, NTP;
- email related IMAP, POP3, SMTP;
- DBMS protocols, including Elasticsearch, Memcached, Mongodb, MSSQL, MYSQL, Oracle, Redis, Postgres;
- IoT protocols, like AMQP, Telnet, Modbus, MQTT and S7;
- remote access protocols, including RDP, VNC, SSH, etc.

The HTTP protocol is worth mentioning separately. To effectively gather data from web servers, we query them not only by IP address but also by domain names. Therefore, if multiple websites or web applications are hosted on the same web server, Netlas will query them one after another.

When Netlas queries a service, it always makes a standard request for that protocol, then parses and saves the received response. For example, for the HTTP protocol, a standard GET request for the root (index) page is executed. The data collected in this way is a valuable source of information for research. Especially with the ability to perform full-text search, search using regular expressions, and even search by fields unique to specific protocols. Analyzing the saved responses often allows for the identification of software, in many cases, even down to the specific version. Netlas identifies over a thousand applications in this way.

If the research target is an application not currently detected by Netlas, you can still perform detection yourself by making a search query in a specific manner. Alternatively, you can refer to one of our repositories on Github[1], where we periodically publish examples of search queries. Identification of an application or device can be achieved through a multitude of different indicators:

- port number and protocol combination;
- part of a text message in the banner;
- a specific value in any section of the response, for example, a specific header;
- unique field in the SSL certificate;
- favicon (for web applications);
- specific service publication URL;
- combination of various features.

In Netlas development, we have gone even further. In addition to the full user response, additional information is available, such as geodata, tags, WHOIS data, etc. These features significantly enhance the value of the data for research purposes.

Next, we will explore some scenarios of Netlas usage related to Security Research and Threat Intelligence.

[1] https://github.com/netlas-io/netlas-dorks

Security of IoT, Industrial Controllers and other types of devices

Today, a significant amount of research in the field of cybersecurity is dedicated to the security of the Internet of Things (IoT) and ensuring the security of critical infrastructure objects. This trend is due to many reasons. Partly because the protection of critical infrastructure objects falls under the close attention of state/national security authorities. Partly because there are some well-known cases of successful attacks where smart devices served as entry points in an organization's network. And so on.

By using Netlas, you can identify some of the critical infrastructure objects as well as many types of IoT devices. If these devices are accessible via the internet, there is a high likelihood that Netlas has already scanned them. This makes Netlas a valuable tool that can be successfully used in conducting research on the security of critical infrastructure objects and the Internet of Things.

Let's consider a general algorithm for investigating the security of critical infrastructure objects.

The first approach is to move from general to specific. The algorithm would be as follows:

1. Using publicly available information directories, the researcher compiles a list of enterprises of interest in the region.
2. Utilizing the Netlas attack surface discovery tool, the researcher identifies the network perimeters of the enterprises of interest. These are often named network ranges, occasionally autonomous systems, and less frequently individual IP addresses.
3. The researcher analyzes the results of the scanning conducted in the second step, focusing on the attack surfaces identified. At this stage, filtering based on ports, protocols, or specific protocol fields may be applied to highlight objects whose security will be further assessed.
4. The researcher conducts a more detailed examination of the objects identified in the previous step. By using the scanning results, it is often possible to determine the software version, search for known vulnerabilities, assess the

validity of authentication procedures, evaluate the cryptographic algorithms used, and assess the reliability of the management protocols. As a result, conclusions can be drawn about the risks of exposing these objects to the Internet.

5. Results are summarized and compiled into a report.

Here is an example query to Netlas that returns the scanning results of one of the critical infrastructure enterprises in Germany:

```
whois.net.name:"TPPA-NET"
```

At the time of writing, Netlas finds two networks with the same name. In the "description" field of both networks, the following information is provided:

```
This Space is statically assigned TPPA Turbine Power Plant Automation GmbH
Fritz-Karl-Henkel-Str. 7 67454 Haloch
```

An OpenVPN service and a couple of devices for remote access made by Synology were published on the TPPA-NET attack surface.

The second option is to move from the specific to the general. It involves searching for published SCADA and other industrial devices based on frequently used equipment. The algorithm in this case will be as follows:

1. The researcher compiles a list of devices whose security needs to be investigated.
2. Signatures are created for searching, and a search is conducted in the area of interest.
3. If possible, using the WHOIS data provided by Netlas, the researcher identifies the owners of the critical infrastructure objects found.
4. The researcher conducts a more detailed examination of the identified critical infrastructure objects, drawing conclusions about the risks of exposing these objects on the Internet.
5. Results are summarized and compiled into a report.

Similarly, you can search for IoT devices and other types of devices. Devices with web interfaces are particularly well-suited for this approach. The "http.title" field may contain information such as the device model, manufacturer’s name, software title, and more. However, many industrial devices can be discovered in scanning results using specialized protocols such as Modbus or S7.

Here are a few examples of search queries for finding industrial devices:

```
modbus.mei_response.objects.product_code:BMX
(Schneider Electric BMX series controllers)

http.title:"WinCC"
(Siemens WinCC Series WebUI)

\*.banner:Siemens
(Siemens Equipment by any protocol)
```

Indeed, one of the challenges in conducting such research is determining device ownership. Often, using WHOIS data, you can only identify the internet service provider rather than the ultimate owner of the device. However, in some cases, even this problem can be addressed. As an example, here is a snippet of response from one of the Siemens Advanced S7-300 series controllers collected using the S7 protocol somewhere in Romania:

```
"plant_id": "MHC Tomsani1",
"copyright": "Original Siemens Equipment",
"system": "Tomsani1_Grup2",
"module_id": "6ES7 315-2EH14-0AB0",
"module_type": "CPU 315-2 PN/DP",
```

If you google the value of the "plant_id" field, you will find a web page dedicated to a hydroelectric power station called "Tomsani1" on the website of the Institute of Hydroenergy Studies And Projects (ISPH S.A.).

Searching for vulnerable devices on the Internet

Penetration testers, bug bounty hunters, and other offensive cybersecurity professionals often use Netlas to identify vulnerable devices in the networks of interest. For this purpose, the Responses search tool includes a whole group of fields related to vulnerabilities in the mapping.

- cve.base_score
- cve.description
- cve.exploit_links
- cve.has_exploit
- cve.name
- cve.severity

You can use these fields to search for services based on a specific vulnerability, criticality, or even the presence of a published exploit. For example, the following search will return services that are likely susceptible to critical vulnerabilities in Google networks:

```
cve.severity:"critical" AND whois.net.organization:("Google"
OR "GOOGL")
```

However, this method has two significant limitations that should be understood:

- Netlas labels vulnerabilities only when the product and its version can be identified. Therefore, using filters from the "cve" group, you will be able to find vulnerabilities only for a limited number of products.
- Netlas labels vulnerabilities during the Internet scanning process. The scanning frequency for each IP address is determined. Therefore, using filters from the "cve" group for the most recent vulnerabilities, you will find only those devices that have been scanned since the vulnerability was published. However, a much larger number of devices may be vulnerable, as there has not been enough time since the vulnerability was disclosed for Netlas scanners to scan these devices.

Considering this, many of our users take a different approach to search for vulnerabilities using Netlas. Typically, the workflow looks as follows:

1. The specialist monitors various sources of vulnerability information. These sources may include developer bulletins, websites aggregating vulnerability data

(such as NIST NVD), social networks, and more. Selection criteria vary for each individual depending on their hacking specialization.

2. Upon discovering an interesting vulnerability, they create a search query for devices or software in Netlas, perform a search in the segment of the internet that interests them, and download the results along with contact information to be able to reach out to the system owner.
3. Depending on how the query is crafted, an additional step may be required. Often, it is necessary to identify specific software versions that are vulnerable. If the search query created on the previous step filters specific version, then the result is already achieved. However, sometimes it is impossible to determine the version from the information returned by the device. In such cases, the specialist needs to create an additional script to determine the software version. Netlas significantly helps narrow down the scope here. The specialist can run a verification script on the already-filtered list of IP addresses.

For example, let's consider CVE-2023-25135. The product affected by this vulnerability is vBulletin of certain versions. The Netlas search would look like this:

```
tag.vbulletin.version:<=5.6.9
```

This search takes into account the version, so there is no need for the development of additional tools.

Below is an example of a search that does not consider the version:

```
tag.name:"microsoft_exchange"
```

As we can see, Netlas tags MS Exchange servers but does not determine the software version. In this case, if a specialist needs to identify the version of MS Exchange, they must do it using third party tools. For MS Exchange, there are several scripts available on GitHub from different developers that address this issue in various ways. The specialist can download the list of addresses from Netlas where Exchange is identified and run one of the version-detection scripts against this list.

> To assist our users, the Netlas team periodically publishes ready-made search queries for the most critical vulnerabilities on social networks. By joining us on one of your chosen social networks, you can significantly save time on developing a search query:

- https://twitter.com/Netlas_io
- https://t.me/netlas

Search for cameras and other telemetry devices

Let's say you need to get information about events happening in a certain region without physically going there. For example, it could be an area affected by a disaster or located in a conflict zone. Live cameras are one of the best ways to conduct such reconnaissance of an area.

Netlas stores data on more than one and a half million web cameras distributed worldwide. To access them, you can use the tag category "Web cameras":

```
tag.category:"Web cameras"
```

To find cameras in a specific zone, you should use geo-tagging. Referring to the mapping of the responses search tool, you will see that Netlas supports the following geolocation filters:

- geo.continent
- geo.country
- geo.city
- geo.subdivisions, etc.

Combining these with other filters allows you to find relevant objects in the desired area. Here are a few examples of search queries:

```
tag.category:"Web cameras" AND geo.country:US
tag.category:"Web cameras" AND geo.city:Paris
```

If you are interested in not all cameras but devices from specific manufacturers, some of them have separate tags. You can also create a custom query to discover devices not tagged by Netlas. Here are a few examples of queries to find cameras not tagged by Netlas:

```
http.title:"Avigilon"
http.headers.server:"GeoHttpServer"
http.headers.server:"hipcam"
```

> The topic of searching for cameras is further detailed in an article[2], specifically dedicated to this subject.

For searching, you can also work with various IoT protocols, such as AMQP, SNMP, MQTT, etc. Telemetry devices that support the HTTP protocol can be found using HTTP header fields, for example, the server header or additional fields like x_powered_by. Here is an example query to find Arduino boards:

```
"http.headers.x_powered_by:"Arduino"
```

Here are examples of searching for message brokers, which are typically used to collect messages from telemetry devices. In this example, the search is conducted on the University of Toronto network:

```
whois.net.organization:"University of Toronto"
(mqtt:* OR amqp:*)
```

[2] https://publication.osintambition.org/how-to-find-online-cameras-with-netlas-io-c68cdf5f327f

Reputation systems, Threat Intelligence Feeds

Using Netlas, you can implement various reputation scoring algorithms. Netlas data can significantly enhance threat intelligence feed providers and cybersecurity analysts' algorithms for detecting suspicious hosts.

1. You can rely on Internet scanning data when deciding whether a host is suspicious or not. Hosts containing numerous vulnerabilities are more likely to be compromised and may host agents or proxy services (e.g., web shells) through which attackers can carry out attacks. During Internet scanning, Netlas tags vulnerabilities that the scanned applications may be susceptible to if the software version is determined. You can also assess the reputation of a host, for example, based on how frequently the software is updated. To do this, you can compare Netlas scanning results taken at different time intervals.
2. There are numerous reputation scoring algorithms based on the comparison of devices or services. Using available Threat Intelligence data, you can query Netlas for the scanning results of malicious devices, identify distinctive features (create a signature for malicious nodes), and search for similar nodes in the Netlas database. This way, other instances of malicious services or previously unidentified parts of the attackers' infrastructure can be discovered.
3. There are also reputation scoring algorithms based on the idea that the reputation of a whole should be defined by the reputations of its parts. In other words, the more malicious hosts there are in a certain network segment, the worse the reputation of that segment should be. Data on malicious nodes should be obtained from third-party sources, such as Threat Intelligence Feeds. WHOIS and DNS data libraries, which Netlas constantly collects and publishes for its users, become useful in this case. Having data on the relationship between domains and IP addresses, as well as data about the relationship of IP addresses to domain zones, countries, providers, networks, autonomous systems, etc., allows you to score the reputation of these entities. For example, you can assign reputation levels to countries depending on the concentration of malicious hosts or assign ratings to providers based on the number of malicious hosts they serve.

Examining specific reputation-scoring algorithms goes beyond the scope of this document. Moreover, we have listed only three possible directions for the development of algorithms, but, of course, there are more.

Threat hunting

The volume of information that Netlas stores for each discovered node allows for effective use of the search engine within threat hunting, primarily for identifying infrastructure owned by malicious actors. Utilizing Netlas, a specialist can quickly identify the signature of a specific malicious software and promptly stop interaction with malicious infrastructure.

One of the most complex examples of such usage is the search for Command and Control (C&C) servers. These devices, employed by malicious actors to control infected nodes, typically hide or masquerade as something legitimate. Researchers need to consider various factors and their combinations. Possible indicators include:

- fields in the service response headers;
- title and content of the returned web page (for web);
- favicon (for web);
- used ports and protocols;
- SSL certificate and secure connection settings;
- specific error codes or error messages, etc.

Search by SSL certificates is often the most effective. Since certificates are not issued for each instance separately, their match indicates the use of the same software. Verification can be done based on JARM fingerprint for TLS and hashes for SSL. Using this approach, you can find, for example, more than a thousand Metasploit servers. In addition to hashes and fingerprints, the search can be carried out using fields such as subject, issuer, and so on.

All the features mentioned above are available for search on Netlas. By identifying several such features and assessing the significance of each, the researcher can create a search query that will identify C&C servers using Netlas.

Similarly, services used for penetration testing, frameworks for emulation or execution of attacks, phishing attack frameworks, sites infected with crypto-miners, and much more can be found.

> Some examples of search queries that demonstrate Netlas' capabilities in searching for malicious infrastructure[3] are provided in the Netlas Cookbook repository published on GitHub. The Netlas Cookbook repository is a collection of recipes for automating work with Netlas. It contains sample search queries and code that can be used to automate many common tasks.

[3] https://github.com/netlas-io/netlas-cookbook?tab=readme-ov-file#search-for-servers-with-malicious-software

Security Analysis

The previous pages describe cybersecurity research cases that can be conducted on a large scale across the entire Internet. However, there are additional Netlas use cases dedicated to the security analysis of information systems and organizations.
These use cases are of greater interest to developers of security analysis systems and providers of security analysis services. All these usage scenarios aim to build the attack surface of an organization or information system, track its changes over time, and assess the security of the attack surface.

Netlas tools and data serve as excellent complements to traditional assessment tools commonly employed by security service providers for evaluating the security of information systems worldwide.

Since Netlas contains comprehensive data about services published on the Internet, it effectively addresses tasks such as identifying the attack surface, tracking changes in the attack surface, detecting Shadow IT, and finding phishing resources. The advantages of Netlas scanning technology should also be considered. The superfast and non-intrusive Netlas scanners can be safely used to scan objects of any scale and criticality. The substantial scanning speed allows for updating data frequently enough, even for enterprise companies.

Attack Surface Identification

The attack surface is the network resources of an organization or information system that may be accessible from the Internet. To build it using Netlas, several methods can be employed. Here are just a few of them:

- IP and Domain WHOIS lookups: Our search engine supports searches through WHOIS records for both domains and IP addresses. The company name can be used as an initial piece of information. By utilizing the appropriate field in the query, it is possible to discover network resources owned by that specific company.
- Forward and reverse DNS lookups: The next type of searches is based on the DNS protocol. By examining DNS records for IP addresses or domains, a specialist can discover numerous connections. For example, this could involve the coincidence of MX records between two domains or the mention of a specific IP address in a TXT record.
- Links and tags in web content: Another important connection between entities on the Internet is cross-references. They can be found using a search within the body of responses. Similarly, tracking codes such as Google Tags or Facebook Pixel can be utilized.

Since identifying the attack surface is a standard task, usually performed in the initial stages of penetration testing projects, we have developed the Attack Surface Discovery Tool to address this challenge. This tool is highly intuitive and allows users to visualize attack surfaces in the form of a graph, which can include domains, IP addresses, networks, autonomous systems, WHOIS objects, and more.

Working with this tool is extremely straightforward; users do not need to input commands or apply filters; Netlas automatically performs searches for each graph object. Users only need to choose one of the possible ways to discover relationships, and the results will be instantly displayed as new nodes in the graph.

> More information about the Attack Surface Discovery Tool can be found in our blog article[4] or in this short video[5].

[4] https://osintteam.blog/netlas-io-attack-surface-discovery-tool-6fbd6b3e9706

[5] https://www.youtube.com/watch?v=98s-Iu5MyRw

Shadow IT and phishing domains

The term "Shadow IT" generally refers to those parts of the attack surface that are not taken into account by security subdivision. These can include external services used without coordination with the cybersecurity department, resources from subsidiary companies, cloud storage, and more. Such elements can be risky, as attacks on them may go unnoticed initially.

With Netlas, you can conduct a search for Shadow IT. Here are several heuristic examples suitable for such searches:

- Search for domains in neighboring zones. Similar domains, such as netlas.io and netlas.com, may arise in different situations, such as the opening of a regional branch of a company or some experiments. Using the search syntax features of the Netlas search engine, specialists can find such domains and determine whether they belong to a known part of the infrastructure or fall under Shadow IT.
- Search for third-level domains can help identify services used by company employees. Cloud services such as CRM, collaborative organization services, and others often create a workspace for their clients using third-level domains. For example, netlas.cloud-service.com. Using Netlas DNS search tool is an effective way to identify such cases.
- Content analysis. The content of a website sometimes refers to the branding content of another resource rather than being stored directly. For example, this could be the company's logo. Netlas responses search tool allows for the detection of such references.

These same heuristics can be utilized in the search for phishing sites. The challenge lies in the fact that it is quite difficult to distinguish resources falling into the category of Shadow IT from phishing resources without human intervention. Technically, the data of a phishing website often duplicates the data of one of the legitimate sites of a company.

There are numerous other methods for finding phishing sites and Shadow IT. A specialist can employ fuzzy searching, identifying domain names similar to a specified one, or verify certificates if they suspect a site is mimicking an original resource. Sometimes, positive results are achieved through searching for favicons.

More information about searching for Shadow IT and phishing sites can be found in our article dedicated to this topic[6].

[6] https://medium.com/osint-ambition/how-to-detect-scam-and-shadow-it-domains-with-netlas-io-f72085e6f18b

Non-intrusive scanning with Netlas.io

Netlas is a non-intrusive scanner. This means that during the scanning process, Netlas does not attempt to perform actions beyond what the system is designed for. In other words, there are no attempts at authorizations, password guessing, or any non-standard requests. If Netlas encounters an obstacle (such as a login form), the scanning stops at that point.

The main advantage of this approach is complete safety and the ability to scan any resources on the Internet. Everything discovered by Netlas is publicly available. Therefore, by using Netlas, you can perform security analysis on any objects. The beauty of it is that you do not even need permission since you are not interacting directly with the object of analysis; you are only utilizing the already-gathered data.

For example, you can conduct additional security checks on potential business partners before sharing confidential company data with them. To do this, you can identify the attack surface of the business partner and perform a security analysis: ensuring the absence of critical vulnerabilities, verifying that the software is updated regularly, and confirming adherence to the principle of minimizing the attack surface–meaning critical services and unnecessary information are not exposed publicly.

Conclusion

This white paper refers to security research and analysis use cases only. These are the most common use cases. However, there are many other areas of application for Netlas, including:

- performing marketing research;
- collection of contact information for sales;
- identification of the site visitor's company by IP address;
- OSINT investigations;
- restricting access from VPN/Proxy/Tor nodes (Netlas provides this data);
- development of automation for SOAR and other security systems;
- data enrichment in SIEM systems, etc.

Please note that data and tools availability are subject to subscription.

To receive more detailed advice on using the service, leave a request on the website at https://netlas.io/sales