

A Look Under the Hood: Data Powering Attack Surface Intelligence



[SecurityScorecard.com](https://www.SecurityScorecard.com)

info@securityscorecard.com

©2022 SecurityScorecard Inc.

Tower 49
12 E 49th St
Suite 15-001
New York, NY 10017
1.800.682.1707



A Look Under the Hood: Data Powering Attack Surface Intelligence

Attack Surface Intelligence (ASI) provides the most contextualized global threat intelligence for you to drive actionable decisions to prevent attacks. SecurityScorecard's robust data lake is the power behind delivering enriched insights and ASI provides direct access to this underlying data across a number of our raw datasets, contextualized with deeper threat context than typical cyber risk intelligence tools.

In this white paper, understand how we collect the data that powers Attack Surface Intelligence and the tools we use.

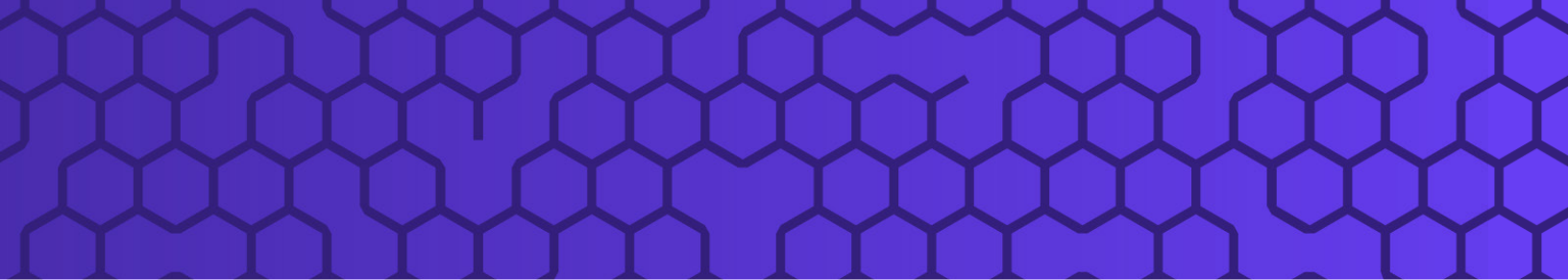
Data Sources and Destinations

Our data includes more than 3.9 billion routable IP addresses scanned every 10 days across more than 1,400 ports globally as the basis for the current IPv4 search index that you can query. We attribute these IPs to domains using the raw SecurityScorecard attribution data, which shows the world as the adversary sees it without filtering out cloud IPs, guest IPs, and manual company tags.

We additionally surface the industries and organizations for attributed IPs and correlate them with open ports, vulnerabilities, threat actors, ransomware group campaigns, and other data points.

Beyond attribution, there are millions of surfaced assets that we do not score due to the inability to attribute them to a specific company. For example, we may not identify a cloud IP address as being tied to an organization.

ASI data is collected in-house and also powers our Ratings platform as well as Automatic Vendor Detection (AVD) and Risk Quantification, except for the outbound links to industry research and reputable indicators of compromise (IOCs) we surface from across the Internet.



Through ASI, you search all these data points and track their connections to build an **understanding of the threat landscape from global to granular levels.**

Collection Tools and Methods

SECURITYSCORECARD'S DATA COLLECTION

DISTRIBUTED ACROSS SIX CONTINENTS, 70 COUNTRIES, AND DOZENS OF HOSTING PROVIDERS, INCLUDING OUR OWN AUTONOMOUS SYSTEM (AS).

SecurityScorecard owns all of the data available through ASI either through in-house purpose-built systems or through partnerships with internet sharing and analysis centers (ISACs), the intelligence and law enforcement community, alliances that we are members of, such as the [Cyber Threat Alliance](#), and Open-source intelligence ([OSINT](#)) collections.

Several commercial data sources serve to enrich the data, but we do not rely upon them to surface final results. We do not purchase data or aggregate any data from public sources.

To make this possible, we have built a comprehensive suite of data collection systems that are distributed across six continents, 70 countries, and dozens of hosting providers, including our own [Autonomous System \(AS\)](#).

Also leveraging the SecurityScorecard STRIKE team's in-house malware information sharing platform (MISP) and threat sharing feeds, we combine and enrich the data to present a unified view of the global attack surface and even for assets that are not tied to exposed IPs.



ARE AN ELITE SQUAD OF CYBERSECURITY EXPERTS, WITH DEEP EXPERIENCE IN CYBER INVESTIGATIONS AND ANALYSIS FROM PREVIOUS CAREERS IN INTELLIGENCE, MILITARY SPECIAL OPERATIONS, AND FORTUNE 500 THREAT INTEL TEAMS.

Global Internet Scanning Framework

What it provides

The framework provides the data that the threat actor sees:

- IP addresses
- Exposed port mappings
- Fingerprints of services, products, libraries, operating systems, devices, and other internet-exposed resources, including version numbers
- Common Platform Enumeration (CPE) IDs
- Common Vulnerability Enumeration (CVE) Version 2 IDs
- Nmap script output

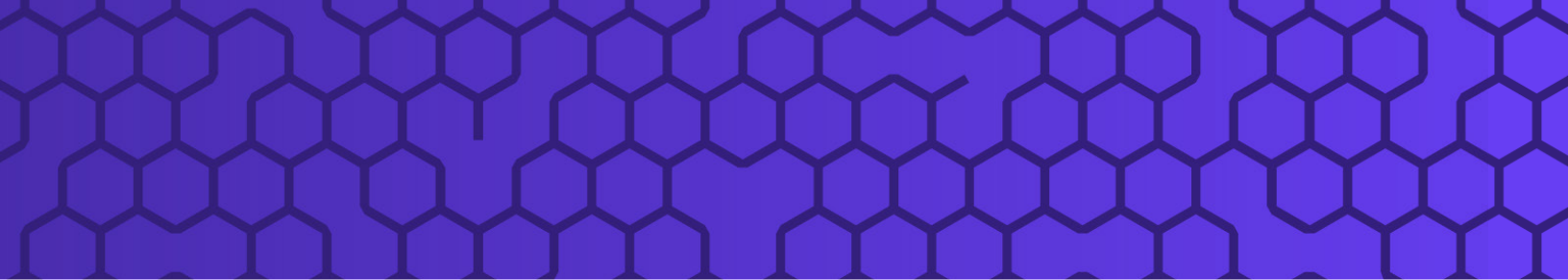
In addition to ASI, the scanning framework provides key data for issue types in the following security factors:

- Network Security
- DNS Health
- Patching Cadence
- Portions of Cubit Score
- Portions of Application Security

How it works

We perform a quick scan of the entire IPv4 address space to determine which TCP ports are open. A deep scan follows for detecting security-relevant information. The framework uses a global infrastructure of cloud VPNs to perform the scans, historically SOCKS5 proxies.

Support for UDP and IPv6 Scanning is in development.



SINKHOLES

SURFACES INFECTIONS FROM MORE THAN 150 MALWARE FAMILIES WITHOUT AN INSIDE SENSOR ON INFECTED SYSTEMS OR NETWORKS.

Malware DNS Sinkhole

What it provides

Our sinkhole provides Information about infections from more than 150 malware families without requiring an inside sensor on infected systems or networks.

How it works

Initially, an infected system makes a query to a DNS server to get the IP address of the domain.

When a domain is not sinkholed, the DNS server returns the IP for the domain. The infected system then directly contacts that IP on the assumption that it is a command-and-control (C2) server.

When a domain is sinkholed, the DNS server returns the IP of a sinkhole, and the infected system then sends requests to our sinkhole servers, not a C2 server.

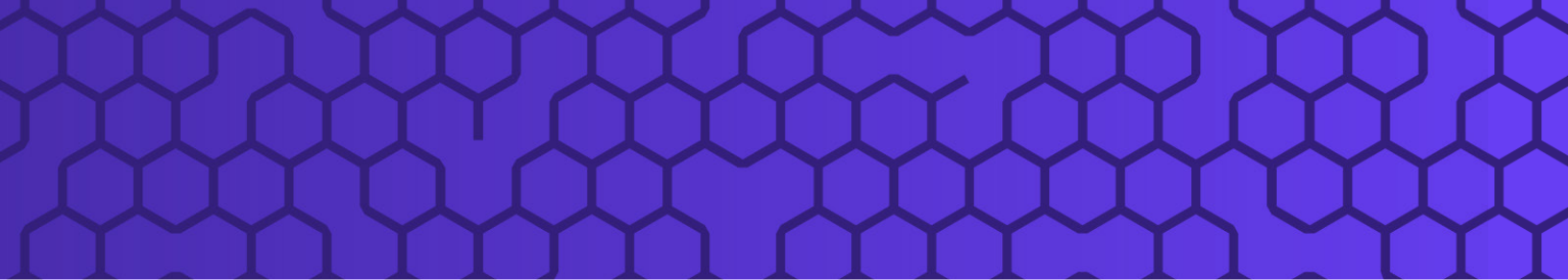
We buy a domain and set the IP address of the domain to our sinkhole infrastructure. Then we log the incoming requests from the infected systems. We receive the IP address that made the connection and also the entire HTTP request.

On average, we receive 2 billion requests from 14 million infected IPs per day. We sinkhole requests from 5,700 domains.

Malware Attribution System

What it provides

The system provides automated malware analysis,



classification, and tracking at a large scale. This includes rich information about malware families, threat groups, and campaigns, enabling us to generate issue types and track threat groups.

Output includes:

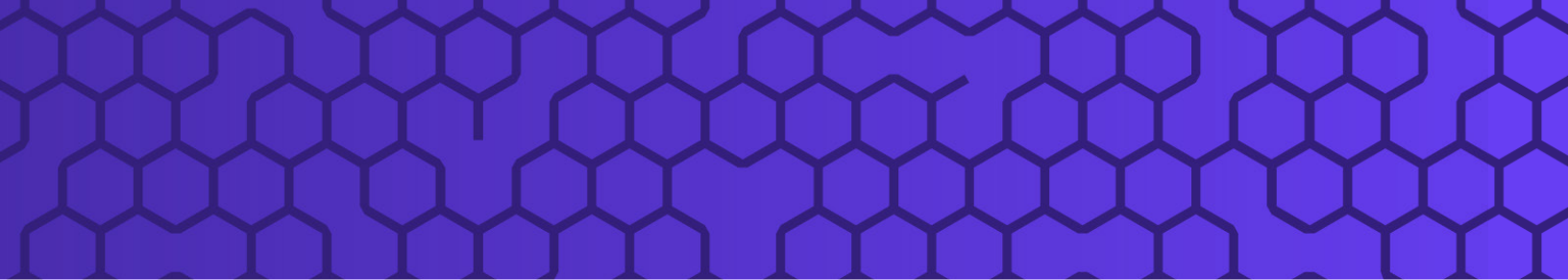
- More than 120 tracked threat groups
- More than 350,000 classified malware samples
- More than 2,000 malware samples with weaponized exploitation
- More than 1,000,000 classified command-and-control (C2) IP addresses

In addition to ASI, the system provides key data for the IP Reputation factor.

How it works

The system digests and analyzes thousands of malware samples and indicators of compromise (IOCs) each day. Human experts and automated malware analysis tools contribute knowledge for attributing a high number of malware samples.

This tool links new and existing unlabeled malware samples to adversary infrastructure, infected customer domains and IPs, and emerging threats, fed by a mix of in-house threat research, malware analysis, partner data, and the open-source intelligence (OSINT) framework.



Leaked Breach Records Service

What it provides

The service provides 68 terabytes of processed data, including user names, passwords, social media URLs, Social Security numbers, credit card numbers, addresses, IPs, and more, all leaked in breaches of a variety of files and sources. It provides more than 600 normalized .jsonl data files and 10 billion records. This results in over 20 different signals provided to the platform.

The SecurityScorecard platform uses this data for findings specific to breaches and data leaks.

Analysts can use this data for penetration testing and determining how a hacker might have infiltrated a target company.

How it works

The service crawls hacker forums for posted leaked data. The corresponding leaked data is then downloaded and uncompressed as necessary. The files are then parsed for extraction of the different fields, such as username, email, and password. We then normalize this data and store it for use in our platform.

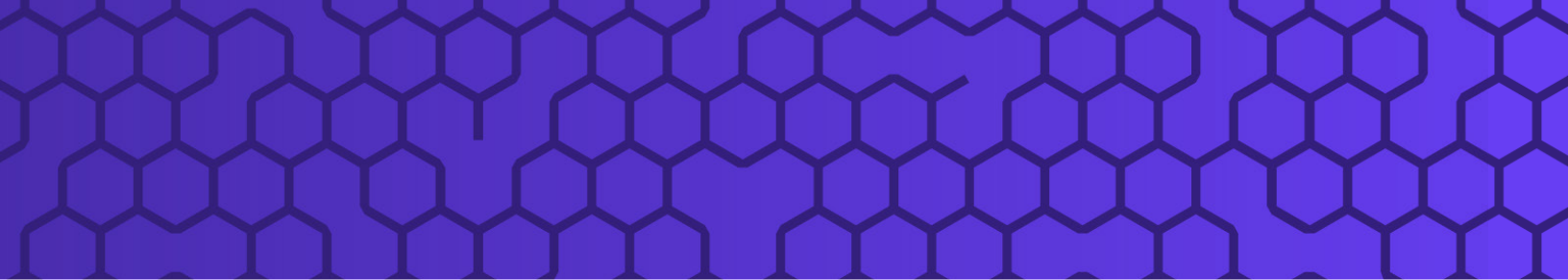
Ransomware Leak Site Crawler

What it provides

The crawler provides findings specific to ransomware attacks, breaches, and credential leaks. It also provides insight into how the ransomware groups operate, what industries they target, who the victims are, and what data is stolen.

LEAKED BREACH RECORDS

**68 TERABYTES OF PROCESSED DATA
USED FOR PENETRATION TESTING
AND DETERMINING HOW A HACKER
MIGHT HAVE INFILTRATED A TARGET
COMPANY**



Information on more than 40 active ransomware groups is actively tracked and updated daily.

In addition to ASI, the crawler provides key data for the following security factors:

- Information Leak
- Hacker Chatter

How it works

We crawl various ransomware leak sites on the dark web and clear web for ransomware leaked data. The crawler parses each leak site to extract the compromised victim names and or victim sites.

Honeypot System

What it provides

The honeypot system provides information about the structure and intentions of advanced persistent threats (APTs) and ransomware groups that actively exploit a particular vulnerability.

In addition to ASI, the system provides key data for the IP Reputation factor.

How it works

We deploy an intrusion detection system (IDS) and varieties of honeypots (industrial control system, traditional, and more) in cybercrime hotspots and cloud virtual machines to attract hackers. We log IP addresses of the parties who access them.

We also log each request that the system receives and match it with the signature from the IDS. We update and maintain 35,000+ signature patterns. This

helps us to understand the nature of the request and the intention of it.

We deploy vulnerable applications in several systems to actively monitor the APTs that scan them and try to intrude using those vulnerabilities.

Chrome-based Web Crawler

What it provides

This crawler gathers information on all major sites' infrastructure, vulnerabilities, vendor dependencies, screenshots, and more.

It detects supply chain dependencies in software and provides data for more than **50 issue types in the the Application Security factor**, such as:

- Expired copyright
- Communication with site with expired certificate
- Server certificate issued by country on denylist
- Certificate key is smaller than recommended size
- Insecure JavaScript library
- Website defacement
- Site requests data over insecure channel
- Links to insecure websites (HTTP)/Redirects to insecure website
- Fail to load page components
- Server error
- Non-standard links detected
- Browser logs contain debug message
- WebSocket requests contain sensitive fields or PII

How it works

We crawl 13 million domains each week with full Chrome browsers. The crawler can detect sites, such as Facebook and Google, that require JavaScript to work. It runs passive scanning and dynamic application security testing (DAST) on every site .

The crawler detects 14 million unique cookies and more than 20 million unique browser console logs each month.

MISP Threat Sharing

What it provides

MISP (malware information sharing platform) Threat Sharing is an open-source resource that provides key information in ASI related to:

- Threat actors
- Malicious IP reputations

The platform serves the government agencies and the threat intelligence community in a number of ways:

- Enablement of additional malicious reputation to aid in prioritization and detection of active or imminent cyber-threats that impact all [critical infrastructure sectors identified by Homeland Security](#).
- Management and collector for all open source intelligence (OSINT) sources.
- Advanced Persistent Threat (APT) campaign tracking, including correlation between seemingly unrelated cross sector campaigns. Support for Common names and vendor contextual synonyms. (APT)
- Support, sharing, and evangelization of threat intelligence standards such as STIX, CAEP, CYBOX, MITRE ATT&CK.
- Support for intelligence sharing controls such as Homeland Security's [Traffic Light Protocol \(TLP\)](#).

- Enablement for SecurityScorecard integrations into products such as FortiNet.
- A platform for Threat Intelligence analysis, researchers, and incident responders
- Support for external collaboration with advanced research teams.

How it works

1. Threat Data is ingested into MISP, manually and automatically, multiple times each day.
2. Our researchers review, notate and enrich data.
3. MISP creates correlations between data attributes--indicators of compromise (IoCs)--to add context by connecting IoC's to their source events, including threat actors.
4. ASI ingests this data in real time to enrich malicious reputation and provides actionable insight.

Take Attack Surface Intelligence for spin and experience the data for yourself.

Free Account

[SIGN-UP](#)

Platform Log in

[SEARCH NOW](#)

ABOUT SECURITYSCORECARD

Funded by world-class investors including Evolution Equity Partners, Silver Lake Waterman, Sequoia Capital, GV, Riverwood Capital, and others, SecurityScorecard is the global leader in cybersecurity ratings with more than 12 million companies continuously rated. Founded in 2013 by security and risk experts Dr. Aleksandr Yampolskiy and Sam Kassoumeh, SecurityScorecard's patented rating technology is used by over 30,000 organizations for enterprise risk management, third-party risk management, board reporting, due diligence, cyber insurance underwriting, and regulatory oversight. SecurityScorecard is the first cybersecurity ratings company to offer digital forensics and incident response services, providing a 360-degree approach to security prevention and response for its worldwide customer and partner base. SecurityScorecard continues to make the world a safer place by transforming the way companies understand, improve, and communicate cybersecurity risk to their boards, employees and vendors. Every organization has the universal right to their trusted and transparent **Instant SecurityScorecard** rating.



FOR MORE INFORMATION, VISIT [SECURITYSCORECARD.COM](https://www.securityscorecard.com)
OR CONNECT WITH US ON [LINKEDIN](#).

SecurityScorecard.com

info@securityscorecard.com
©2022 SecurityScorecard Inc.

Tower 49
12 E 49th St
New York, NY 10017
1.800.682.1707

