Datasets for VIS cyber-insurance platform

SecurityScorecard, Inc.

New York, NY | April 6th 2017

Proposal Overview	4
About SecurityScorecard	4
Company History and Timeline	4
Ownership Structure	4
Staff	4
Core Data, Expertise, and Customers in Cybersecurity	4
Request and Delivery Mechanisms for Products and Services Describe Above	4
Data, Expertise, and Customers in Cyber Insurance	5
Product Vision	5
Roadmap	5
·	
Data Collection and Analytics	8
Proposal Scope SecurityScorecard datasets that correspond with the following categories listed in this RFP Internet Facing Infrastructure Security Health and Hygiene Any other IT best practices that are not followed Open Source Intelligence Internal Check Auto Discovery Company Domain: FoxConn.com	9 9 9 10 12 12 13
SecurityScorecard Proposal Data Content and Pricing	20
List of SecurityScorecard Core Data Modules	20
Banner Grabs Module Description Data Points Vulnerable Hosts (CVE Mapping) Description Data Points CMS Scans Description Data Points Document Metadata Description Data Points Booter Shells Module Description Data Points Defacements Module Description Data Points Defacements Module Description Data Points Defacements Module	21 21 21 22 22 22 22 22 22 23 24 24 24 24 24 24
DNS Analysis: DKIM Module Description	24 24

Data Points:	24
DNS Analysis: DMARC Module	26
Description	26
Data Points:	26
Github Dorking Module	26
Description	26
Data Points	26
End of Life Module	26
Data Points:	26
Leaked Passwords Module	27
Description	27
Data Points:	27
Hacker Chatter Module	27
Description	27
Data Points:	27
Sinkhole Feeds Module	27
Description	27
Data Points	28
SSL Certificates	28
Description	28
Data Points	28
TOR Nodes Module	28
Description	28
Data Points:	28
Sample data for Core Data Collection Modules described in section above	29
Banner grabs	29
SSL Certificates	29
Vulnerable Hosts and CVE Details	30
Leaked Passwords	30
Malware infections feed	31
Vendor Relationships	31
Attack Feeds	31
Hostname Subdomain Enumeration	31
TOR Node	32
Sample data schema for the banner grab module mentioned in prior section	32
Cartification by Authorized Official of SocurityScarcoard Inc	22

Proposal Overview

SecurityScorecard (SSC) proposes to provide data sets to Verisk Insurance Solutions (VIS) for the purpose of powering VIS's cybersecurity insurance platform. In addition to commentary on the specific data points requested in this document, we will first provide context on the scope, reach, and scale of SecurityScorecard's data collection and analytics platforms. SecurityScorecard currently offers, and VIS analysts may be familiar with, two distinct products in the marketplace – the SecurityScorecard platform and ThreatMarket. The combination of all collected data (ThreatMarket), and all contextualization/scoring/data analytics (SecurityScorecard) -- offered in this proposal -- will be made available as one SecurityScorecard data offering to VIS.

About SecurityScorecard

SecurityScorecard is the most comprehensive security rating platform that is able to determine the security risk posture of any organization. The company's proprietary SaaS offering helps enterprises gain operational command of their security postures and across that of all of their partners, and vendors. SecurityScorecard provides continuous, non-intrusive security piosture monitoring for every organization. The platform offers a breadth and depth of critical data points not available from any other single service provider, including a broad range of risk categories such as Application Security, Malware, Patching Cadence, Network Security, Hacker Chatter, Social Engineering and Leaked Information.

Company History and Timeline

SecurityScorecard was founded in 2013 by Dr. Aleksander Yampolskiy & Sam Kassoumeh. In 2015, the company raised \$12.5M in Series A funding, led by Sequoia Capital. An additional \$20M Series B, led by GV, was raised in 2016. There have been no prior acquisitions.

Ownership Structure

SecurityScorecard is a venture-funded privately held C Corp. Major investors include Sequoia Capital and Google Ventures.

Staff

- Full Time = 91
- Part Time = 0
- Contractor Staff = 8
- Intern = 2

Core Data, Expertise, and Customers in Cybersecurity

The core data utilized by SSC customers and partners in the cybersecurity world consist of OSINT described, collected and analyzed by the means enumerated in this proposal.

The SecurityScorecard team consists of leading industry experts in the fields of malware analysis, threat intelligence tradecrafts, cryptography, application vulnerability analysis, and network penetration testing. The comprehensive information security oriented nature of SecurityScorecard's collective background enables the organization to quickly identify and surface data points of interest to customers and partners.

Cybersecurity customers and partners include Symantec, Check Point, RSA, AlienVault, Carbon Black, Veracode, and others.

Request and Delivery Mechanisms for Products and Services Describe Above

• Initial data sets to be delivered through a file transfer

- Updates on a periodicity mutually agreeable with Verisk
- Updates may be provided through an API as an on-going service

Data, Expertise, and Customers in Cyber Insurance

SecurityScorecard has worked extensively with the largest cyber-underwriters, carriers, and re-insurers in the business, to determine which data analytics can be best used to speed up the underwriting process, and provide the basis for dynamic risk pricing by continuously monitoring a policyholder's cyber risk posture.

SecurityScorecard is working with two carriers to implement a system that incorporates the "Internal Check" described in **Section 3.**

Current insurance carrier customers include: Aetna, AIG, Allied World Assurance Company, Allstate, AON, Liberty Mutual and The Hartford.

Product Vision

SecurityScorecard built its data sensors, collection and analytics platforms for the purpose of detecting every service or system running on every IP address, and then assigning each IP address to the corporate entity that owns, controls, or operates it. Our discovery and attribution processes are **fully automated**, **and scale sufficiently to maintain an accurate mapping of the entire internet** and the entities operating therein.

Using the collected vulnerability data, SecurityScorecard defines 80 specific issue types that are correlated with known vulnerabilities. These issues include vulnerable services, weak ciphers, and malware infections, for example. (SecurityScorecard owns cvedetails.com that provides a widely-used front end for the NIST and Mitre databases). These issues are organized into the 10 factors that depict the most serious types of vulnerabilities that will be exploited by criminals.

In addition to detecting and rating the exposed vulnerabilities in every company's digital surface area, SecurityScorecard uses qualitative analysis to infer the quality of a company's operational security processes and practices – as there is a strong correlation between external observations and the internal actions that created these vulnerabilities. By observing the trend and timing of these results: up; down; faster; slower -- SecurityScorecard captures additional / enhanced predictive signals that improve risk modeling.

Combining this "Human factor" analysis with the "factorized issues" that are attributed to corporate entities -- SecurityScorecard can rate or score the cyber vulnerability risk associated with any given entity. This information can be used by third-party risk managers, cyber security practitioners, cyber insurance underwriters, and incident response teams, among others -- to gauge inherited cyber risk. This insight enables better business decisions -- before onboarding a vendor -- to selectively eliminate vulnerabilities in IT infrastructure, or to work collaboratively with policyholders or vendors to actually improve their security posture; in order to reduce the risk of the entire portfolio, or the corporate enterprise.

Roadmap

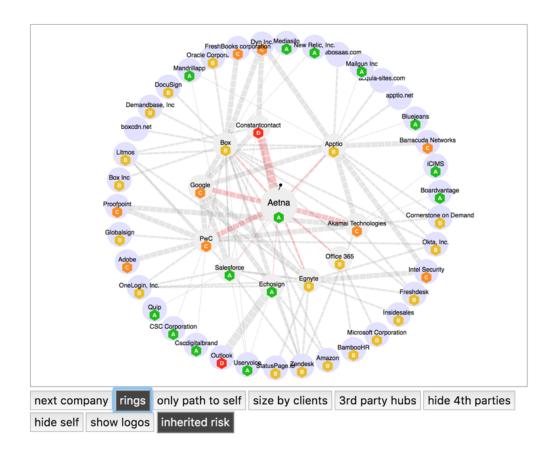
Key Product Roadmap Areas for Q3, Q4 2017

 Advanced Analytics. In a competitive marketplace, our key differentiator is the depth and breadth of our data collection. Our product roadmap focuses on leveraging that data to produce novel insights and on-demand analytics. For example, we will shortly launch our Breach Multiplier, which is a measure of breach correlated Issue Type diversity within a Portfolio of companies. This is a Confidential information shared under non-disclosure agreement

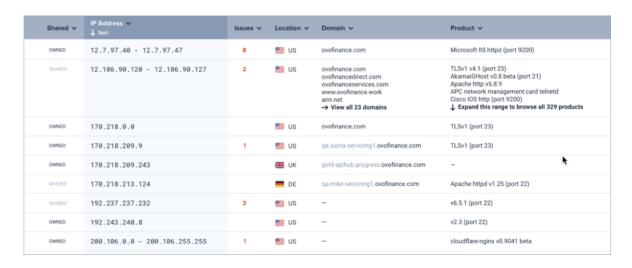
prescriptive measure within a Portfolio of companies that informs our clients on risk associated with specific issue findings.



2. Inherited Risk. As our network attribution model evolves, we are learning more about the relationships between companies that we monitor and their 3rd parties. As we identify these 4th parties of our clients, we are able to identify and monitor inherited risk. This risk comes in the form of security ratings, where a critical service provider in a client's supply chain is supported by a 4th party that has poor cyber security or is breached. In addition to service providers, the other critical dimension which we are tracking is infrastructure providers. Future analytics will provide insight into the overall security of 4th party service providers and the observable cloud or share infrastructure of any 3rd party.



3. Expanded Digital Footprint. Companies ratings on the SecurityScorecard platform are a function of all observable network infrastructure at any given time. The Digital Footprint within SecurityScorecard is the sum total of that observable infrastructure. We are working to expand that tool to provide richer insights into each company's digital footprint. For example, in the next iteration we will include observable protocols and products, along with version numbers and other details. There will also be classification of assets by shared and owned infrastructure. End points within the API and SecurityScorecard platform will also provide details on Issues, IPs, Domains and infrastructure by location.



- 4. Directed Remediation Scoring. SSC is introducing advanced scoring capabilities with the ability to direct security practitioners to specific vulnerabilities that correspond to the most immediate remedial actions for a company to take to raise their score. Key attributes and benefits of this scoring methodology include:
 - Statistical framework. Data-driven
 - Normed score more level playing field for small v. large companies
 - Global calibration (all domains, all industries) periodic (e.g. 1/Qtr)
 - User-defined cohorts, even across industries
 - Easy to calculate impact of issue remediation on score
 - Reduced volatility. Deterministic. Defendable. Reasonable. Explainable

Data Collection and Analytics

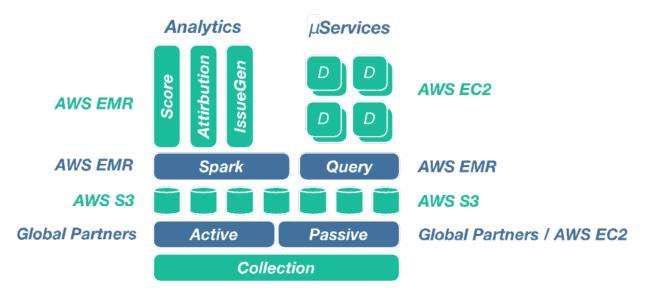
SecurityScorecard's data collection and analytics capabilities scale because they are completely automated, and built on cloud-based, horizontally scaling architecture. The SecurityScorecard platform can be quickly scaled to accommodate scoring of hundreds of thousands of additional businesses.

SecurityScorecard plans to score and map the digital footprint of 1 million businesses over the next year.

SSC collection methodologies combine cutting-edge gathering techniques for:

- internet-wide IPv4 mapping
- malware reversal and sinkhole efforts
- honeypot log collection and Internet-wide raw text scraping, normalization, and parsing,
- contextualized using techniques of raw banner analysis, CVE mapping & contextualization, and IP address ownership attribution.

These methods enable SecurityScorecard to provide a comprehensive portrait of external security risk in an automated, scalable fashion that leverages the technologies of the Amazon cloud.



Additional information:

How many organizations do you track on a daily basis?

110,000

What is the average history of these organizations that you have data on?

12 months

What granularity do you persist this historical data?

All SecurityScorecard data is persisted.

Do you calculate confidence intervals for the data points you collect?

Not at this time. Confidence intervals are defined as a 2017 roadmap item.

Proposal Scope

SecurityScorecard datasets that correspond with the following categories listed in this RFP

Short descriptions of the data returned in different categories are provided below. There are also additional examples of these data throughout the proposal.

Internet Facing Infrastructure

- o Firewalls, routers, servers, mobile devices etc.
 - SecurityScorecard collects and identifies all servers and services running across the IPv4/IPv6 space. Through proprietary means we identify services, then attribute them to specific organizations.
- O Software running on these external facing systems
 - By analyzing interne- facing banners, SecurityScorecard identifies software and version numbers.
- Any other devices and/or software that can be detected externally.
 - Along with looking at internet- facing infrastructure, SecurityScorecard detects endpoint devices including operating systems and browser versions within an organization through passive user-agent collection.

Security Health and Hygiene

- o Is internet facing infrastructure patched regularly?
 - When considering how an organization patches their infrastructure, we look for 1) currently open vulnerabilities 2) closed vulnerabilities which were open longer than best practices and 3) end of life/ end of service products.
- o Does any infrastructure have software/hardware with known vulns?
 - SecurityScorecard flags both open vulnerabilities and vulnerabilities that have taken longer than best practices to remediate. All SSC issues are correlated with CVEs. SSC owns cvedetails.com
- O Can you detect backup services for the above?
 - Anti-DDoS detection is an existing positive issue type that detects the provision of DDoS prevention services such as CloudFlare.

- O Does the organization have ongoing virus, bot, malware infections?
 - SecurityScorecard detects these infections and provides insight into the malware family, duration of infection, source IP, destination port, and destination IP.
- O How quick is the organization to clean up these infections?
 - SecurityScorecard provides insight into the length of time it takes for a malware infection to be cleaned up, among many other factors.
- o Is this organization vulnerable to typosquatting, email spoofing etc.?
 - SecurityScorecard detects typosquatting domains. For email spoofing, we provide insight
 into SPF records for example, whether or not they are registered to all owned domains,
 and correct SPF configuration.
- o Are there ports left open?
 - SecurityScorecard identifies open ports and specifically which services are running behind these.

Any other IT best practices that are not followed

Most of the listed issues identified by SecurityScorecard are the result of failure to follow best practices. SecurityScorecard uses inference to rate the organizational behaviors that affect a company's security posture, adding these analytics as additional inputs to SSC risk calculations.

List of issues rated:

- Network Security
 - SSL Certificate is Self-Signed
 - SSL Certificate is Expired
 - SSL Certificate(s) have been revoked
 - o Open Ports
 - Open DNS Resolver Detected
 - SSL Certificate Expiration is Longer Than Best Practices
 - SSH Software Supports Vulnerable Protocol
 - TLS Protocol Uses Weak Cipher
 - SSL Certification Uses Weak Signature
 - SSH Supports Weak Cipher
 - SSH Supports Weak MAC
- DNS Health
 - DKIM Misconfigured
 - DNSSEC Configuration is Invalid / Missing
 - Passive DNS
 - Passive DNS
 - Email SPF (Sender Policy Framework) Vulnerability
 - Malformed SPF Record
 - SPF Record Missing
 - SPF Record Contains a Softfail
 - SPF Record Contains Wildcard
- Patching Cadence
 - Exposed Vulnerabilities Found in the Last Observation
 Confidential information shared under non-disclosure agreement

- o End-of-Life Product
- o End-of-Service Product
- High Severity CVEs Patching Cadence
- Company Excels at Patching Vulnerabilities
- Low Severity CVEs Patching Cadence
- Medium Severity CVEs Patching Cadence

Endpoint Security

- Obsolete Operating System Endpoint Detected
- Obsolete Browsers Detected
- Vulnerable Browsers in Use by Enterprise
- Insecure Software
- Multiple Browsers Detected

• IP Reputation

- Attack Detected
- Malware Duration Longer than Average
- Malware Events, Last 30 Days
- Malware Events, Last Day
- Malware Events, Last Year
- P2P Activities
- SQL Injection
- Tor Exit Nodes

Cubit Score

- o Exposed Subdomain
- o IP Address Flagged in Malware Database
- Phishing Site Count
- Possible Typosquat Domains Detected
- Unsolicited Commercial Email
- "WHOIS" Directories

Application Security

- ClickJacking / Cross Frame Scripting (XFS) Vulnerability
- Blind SQL Injection
- Heartbleed Bug Vulnerability
- Content Management System vulnerabilities identified
- Session Cookie Missing 'HttpOnly' Flag
- o Cookie Missing 'Secure' Flag
- Cross Site Scripting (XSS)
- Session Cookie Missing 'HttpOnly' Flag
- o Directory Traversal
- Secure Coding Training Needed

Hacker Chatter

- Booter Shells Identified
- o Defacement
- Hacked News Mentions
- Hacker Chatter Mention

• Information Leak

- Insecure Secret Questions
- Sensitive Application Information Exposed (Github)
- Sensitive Application Information Exposed (Google)
- Credentials at Risk
- Social Engineering

- o Employee Satisfaction
- Corporate Email Used on Marketing Sites
- Corporate Email Used on Short-Term Lending Sites
- Leaked Company Emails Open to Spear-Phishing

Open Source Intelligence

Does this organization have leaked credentials on the dark web?

The SecurityScorecard platform continuously collects information from various deep web/darkweb/underground resources related to the circulation of leaked/compromised authentication credentials. Once credential sets are obtained, several validation steps take place to establish a confidence level on the discovery, and rate the findings accordingly.

Are blackhats looking to hack into this organization?

The SecurityScorecard platform continuously collects information from various deepweb/darkweb/underground resources related to the collection of relevant hacker chatter and discussion. SecurityScorecard maintains access to dozens of private and public hacking forums at any given time, with archives of discussion going back to 2007. The chatter resources are continuously crawled, cached, and aggregated. The chatter samples will surface into the SecurityScorecard platform whenever there is a mention of a corporate name, domain, or IP address.

Is there leaked data from this organization available?

The SecurityScorecard platform continuously collects information from various deepweb/darkweb/underground resources related to the circulation of leaked/exposed office documents (pdf, doc, ppt, etc), as well as the collection of leaking information from public OSINT resources. Discovered files are analyzed for metadata content, such as client-side software suites and embedded e-mail addresses, and are scored/rated accordingly.

In addition to its own collection program, SSC shares data with open source providers to enhance the volume of data for signal generation.

Internal Check

Do you have an appliance that can be plugged into a customer network to harness relevant data?

SecurityScorecard has begun to integrate its "outside view" of a company's digital assets with an "inside view" (network and user data) from 1 million customer gateways in partnership with several next-generation firewall and endpoint vendors. These datasets are combined at the report-output level today, with full data integration being specified. The combined output provides a holistic view of company security posture, including behavioural analytics on the performance and trend of operational security practices.

How high touch is this deployment?

Not high touch. For companies with a SecurityScorecard-partner gateway in place, there is NO deployment.

What kind of data can you collect internally?

Inbound / outbound traffic

- Threats in the network
- Malware
- Spam
- Phishing Attacks
- Confirm certain compensating controls
- User agent data

Auto Discovery

Can you identify subsidiaries of the named organization?

SecurityScorecard can identify associated corporate entities when there is any type of network topology connection between the organizations. We make use of financial filing intelligence and open source intelligence to map relations between organizations and their subsidiaries, such as holding companies and their assets.

Can you identify vendors of the named organization? Can you classify these vendors i.e, hosting provider, DNS providers, human resource vendors, ISP, email, DNS, data/content hosting?

Yes. The ThreatMarket IP attribution mapping system will propagate a display of all observed domain names to an IP address.

This can be done by searching on a domain name to identify all IP addresses associated with the domain, or searching by IP address to identify all known domains pointing to the domain.

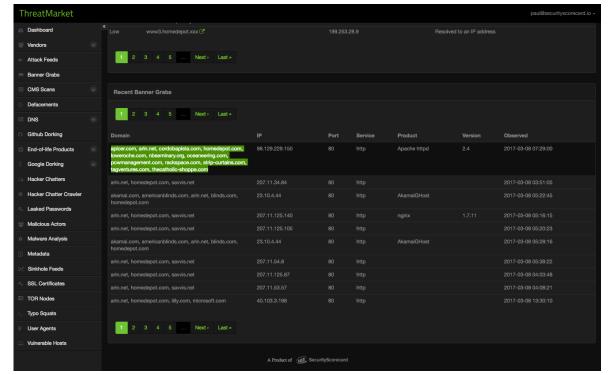
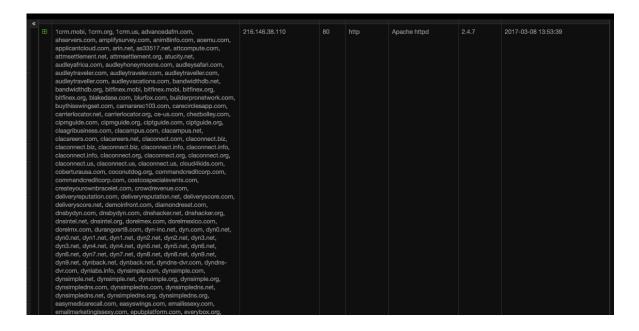


Fig 1- Shared Hosting / CDN IP Address

Figure 2 – DNS Nameserver for DYN that shows associated domains

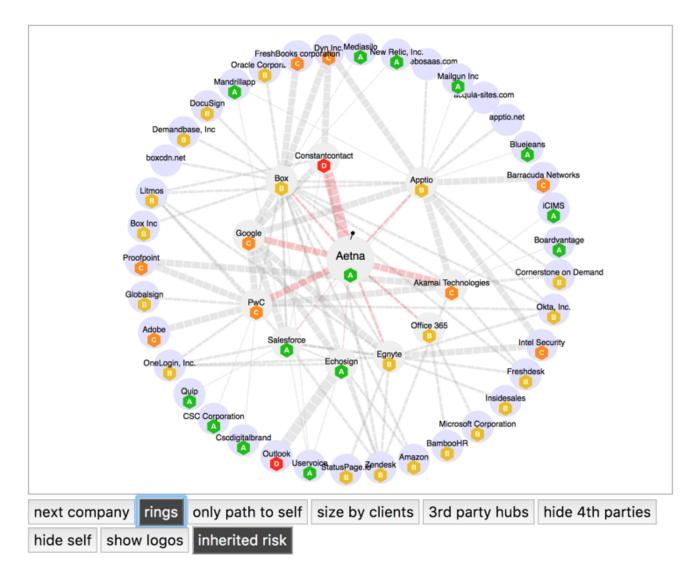


Can you identify multiple vendors for the same service? i.e. does the named organization host on AWS as well as Rackspace?

Yes. A domain name query will pick up all IP addresses attributed to the enterprise and identify the owner and user, as observed in Figures 1 and 2.

Can you identify vendors that your vendor is dependent on? i.e. how many degrees of separation can you map out to?

SecurityScorecard currently maps one degree of separation for 3rd parties, and will map one or two more degrees by Q3 this year.



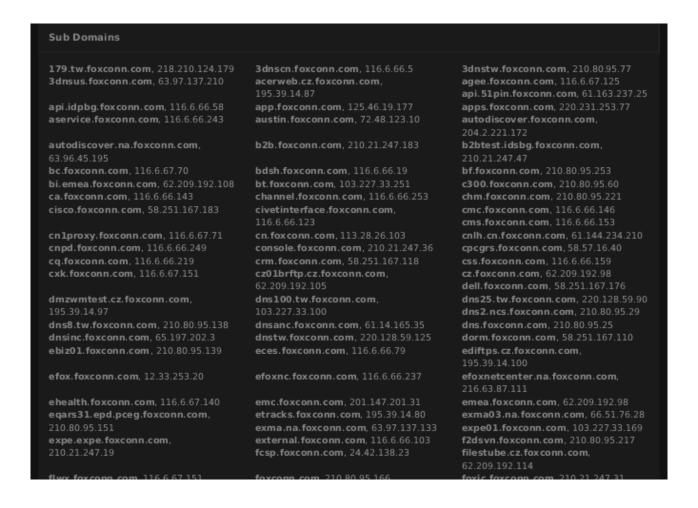
Can you map out the organizations IPv4 and IPv6 space?

SecurityScorecard has the ability to map out the complete digital footprint of the IPv4 internet space, and the ability to map out the complete IPv6 footprints of enterprises that announce their ranges and configurations.

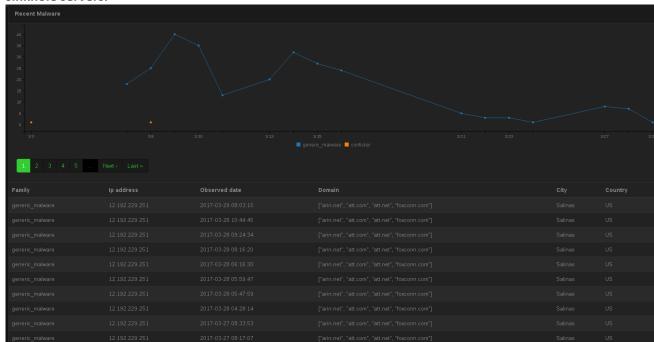
Company Domain: FoxConn.com

The screenshots below show a small sample of the data (first display screen only) that were pulled for foxconn.com (and all IP addresses attributed to them).

* Subdomain Collection - Full Analysis of Subdomains



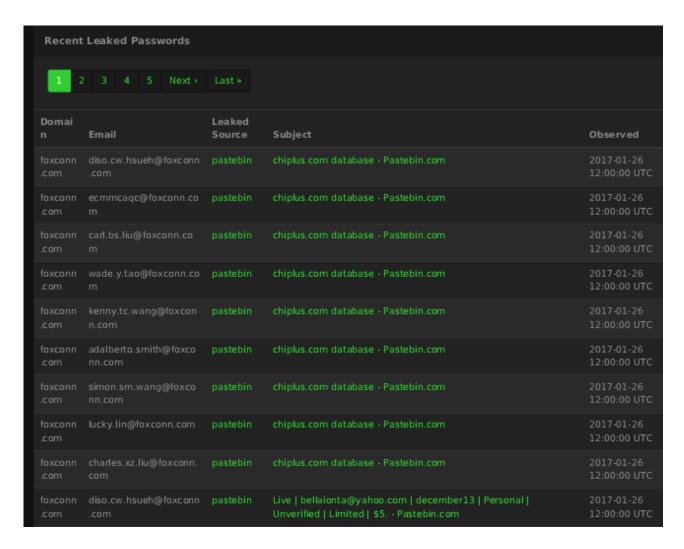
* Malware Detected in Sinkhole - Infected machines beacons are intercepted by SecurityScorecard sinkhole servers.



* Detected CVEs - Matched from IPv4 banner text + SSL data

⊞		CVE-2016-0800	DROWN	03-16-2017
⊕		CVE-2016-0800	DROWN	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2016-0800	DROWN	03-16-2017
⊕		CVE-2014-3566	POODLE	03-16-2017
⊕		CVE-2014-3566	POODLE	03-16-2017
⊕		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2016-0800	DROWN	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊕		CVE-2016-0800	DROWN	03-16-2017
⊕		CVE-2014-3566	POODLE	03-16-2017
⊕		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2016-0800	DROWN	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2015-0204		03-16-2017
⊞		CVE-2016-0800	DROWN	03-16-2017
⊞		CVE-2016-0800	DROWN	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2014-3566	POODLE	03-16-2017
⊞		CVE-2015-0204		03-16-2017
	<u> </u>	·		

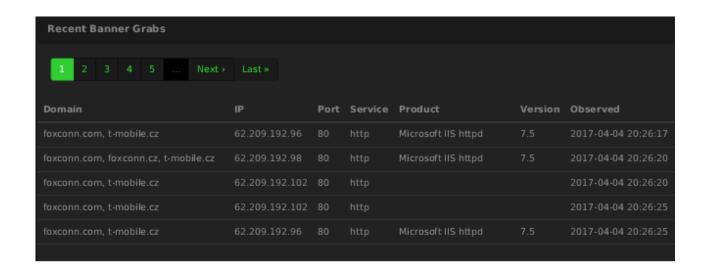
^{*} Leaked Passwords - Obtained from Pastebin & Hacker Underground



* Detected Typosquats - Resolving typos and nonstandard tld extensions

	7		
Lo	w zm.foxconn.international ♂	46.252.18.143	Resolved to an IP address
Lo	w zm.foxconn.work 🗷	69.172.201.218	Resolved to an IP address
Lo	w webmail.foxconn.nl 🗷	185.15.248.85	Resolved to an IP address
Lo	w webmail foxconn company 🕜	216 239 133 253	Resolved to an IP address

^{*} All Open Banners on IPv4



Domain	IP	Port	Service	Product	Version	Observed
foxconn.com, t-mobile.cz	62.209.192.102		http			2017-04-04 20:26:26
foxconn.com, foxconn.cz, t-mobile.cz	62.209.192.98		http	Microsoft IIS httpd		2017-04-04 20:26:27
foxconn.com, t-mobile.cz	62.209.192.96		http	Microsoft IIS httpd		2017-04-04 20:27:03
foxconn.com, t-mobile.cz	62.209.192.102		http			2017-04-04 20:27:20
foxconn.com, foxconn.cz, t-mobile.cz	62.209.192.98		http	Microsoft IIS httpd		2017-04-04 20:25:26
foxconn.com, foxconn.cz, t-mobile.cz	62.209.192.98		http	Microsoft IIS httpd		2017-04-04 20:26:05
foxconn.com, foxconn.cz, t-mobile.cz	62.209.192.98		http	Microsoft IIS httpd		2017-04-04 20:26:07
foxconn.com, foxconn.cz, t-mobile.cz	62.209.192.98		http	Microsoft IIS httpd		2017-04-04 20:26:12
foxconn.com, t-mobile.cz	62.209.192.102		http			2017-04-04 20:26:08
foxconn.com, t-mobile.cz	62.209.192.102		http			2017-04-04 20:26:09
foxconn.com, t-mobile.cz	62.209.192.102		http			2017-04-04 20:26:11
foxconn.com, t-mobile.cz	62.209.192.114	443	https			2017-04-04 20:33:36
foxconn.com, t-mobile.cz	62.209.192.107	443	https			2017-04-04 20:33:30
foxconn.com, t-mobile.cz	62.209.192.108	443	http			2017-04-04 20:33:33
foxconn.com, t-mobile.cz	62.209.192.114	443	https			2017-04-04 20:33:33
foxconn.com, foxconn.cz, t-mobile.cz	62.209.192.123	80	http	Apache httpd	2.2.22	2017-04-04 20:33:35

* Open Ports of Interest (Telnet, FTP, RDP, VNC, Databases, etc)

Issue Details

DOMAIN

IP ADDRESS A

DETECTED SERVICE

PRODUCT

VERSION

LAST OBSERVED
DATE

foxconn.com

218.4.91.63

MS-WBT-SERVER

Microsoft Terminal

- 3/11/17

* Russian Hacker Chatter Mention of FoxConn RDP login + IP address

TIO TUROC ADIT :

707394583 (04:52:54 15/02/2017)

Nu vivod ot 17\$ vozmojen chto za h?

707394583 (04:53:16 15/02/2017)

Prochekanie

707394583 (04:53:21 15/02/2017)

Znachit

968559 (04:53:35 15/02/2017)

128.30.30.172:3389@foxconn; foxconn ADM не заходит

707394583 (04:53:56 15/02/2017)

Adm ne vnisui

* DNS Settings

SPF Record Contains a Softfail: Softfail attributes in SPF makes spoofing and phishing email possible. Show Details Share

SPF Record Missing: A missing SPF record has been detected for a domain.

Show Details Share

All above processes are currently automated. The collection process is kicked off whenever a customer adds new domain that SSC is not already monitoring. The infrastructure is set for scaling throughout the year of 2017 as new domains are continuously entered into the platform, with the expectation of hundreds of thousands of new entries in the near future.

SecurityScorecard Proposal Data Content and Pricing

SecurityScorecard data is licensed on an annual basis. The monthly recurring costs stated are derived by dividing the annual cost by 12.

SecurityScorecard proposes to charge Verisk \$33,333 per month for the modules listed in the section below (\$400,000 per year).

List of SecurityScorecard Core Data Modules

Below please find the list of the datasets available as of December 2016. We are constantly improving and evolving the dataset, and will provide updates as improvements are introduced.

- IP Mapping for a domain
- Banner grabs
- CVE Vul Hosts
- SubDomains
- CMS Platforms
- Malware sinkholes/Honeypot /Tor Node feeds
- Attack Feed
- DNS Analysis DKIM and DMARC Modules
- End of life Product modules
- SSL Certs
- Github
- Leaked Password Modules
- Hacker chatter modules

Monthly dumps of datasets keyed by IP Address (entire IPv4 space):

- Banner grabs
- CVE Vul Hosts
- SubDomains
- CMS Platforms
- Malware sinkholes/Honeypot /Tor Node feeds
- DNS Analysis DKIM and DMARC Modules
- End of life Product modules
- SSL Certs

Keyed by domain for the list of 50K domains monthly dumps of:

- Github
- Leaked Password Modules
- Hacker chatter modules

(A sample data file accompanies this proposal as an attachment).

Descriptions for Core Data Collection Modules

Banner Grabs Module

Description

This module is a regularly updated collection of the active global IPv4 address space, with contextualized IP address mapping for attribution of assets beyond assigned ASN ranges. The Banner Grabs Module is searchable on any data point, and allows for instant identification of public facing services.

Data Points

The Banner Grab Module contains the following data points:

timestamp, IP address, domain names (hostnames + attributed domains), open ports, running services, identifiable products, identifiable version numbers, and searchable raw banner text that beacons from every identified open port.

Vulnerable Hosts (CVE Mapping)

Description

This module contextualizes the data points collected from the Banner Grabs Module and maps the attributed domain names, IP addresses, and identified services to the Common Vulnerabilities and Exposures (CVE) database. The Vulnerable Hosts Module is searchable by domain name and timestamps.

Data Points

The Vulnerable Hosts module contains the following data points:

timestamp, IP address, domain names (hostname + attributed domains), idenfified CVE number, identified vulnerable service

Note: Please see schema for banner grabs data in the following section xx below.

CMS Scans

Description

The CMS Scans Module conducts analysis on all Wordpress, Joomla, and Drupal web applications that are identified during IPv4 collection and domain name attribution. The CMS Scans Module analyzes the web applications to determine the versions, themes, plugins, and security configurations and also maps vulnerabilities to the Vulnerable Hosts Module.

Data Points

The CMS Scans Module contains the following data points:

timestamp, domain name, web application names and versions, and identified CVE numbers

Document Metadata

Description

The Document Metadata Module conducts analysis on all document files that are identified as publicly accessible within analyzed domain names. The Document Metadata module searches for PDFs, .doc, and other related office files and identifies metadata encoded within the documents that reveal information about the internal security posture of an enterprise.

Data Points

The Document Metadata contains the following data points:

timestamp, company name, URL of document, document type, usernames identified within document, email addresses identified within document, client side software details of the user that created the document.

Booter Shells Module

Description

The Booter Shell Module archives the leaked SQL databases of DDoS botnet services. A booter shell/webshell script is a malicious script that is placed on compromised web servers to create a server based botnet. The Booter Shells Module is searchable on domain name and IP address, and surfaces URLs of historically compromised servers that have been used in DDoS/Booter Service botnets.

Data Points

The Booter Shells Module contains the following data points:

Source (botnet name), URL of booter webshell

Defacements Module

Description

The Defacement Module is a historical archive of compromised web applications collected from various OSINT resources. The presence of an archived defacement indicates that the web server has been successfully targeted for exploitation in the past. The archived incident may have led to loss of customer and user data, administrative data, and a cause of reputational damage. The Defacements Module is searchable by domain name or IP address.

Data Points:

The Defacements Module contains the following data points:

Timestamp, defaced URL, source of report

DNS Analysis: DKIM Module

Description

The DKIM Module uses DNS records to map the DKIM configurations and service providers of email protocols within an enterprise. The DKIM Module is a collection 'DKIM selectors' that have been mapped to identifiable domain names. Visibility into the public DKIM signatures provides insight into the security and compliance configurations of enterprises, and allows for the identification of third party Email Service Providers (ESPs). The DKIM Module is searchable by domain name.

Data Points:

The DKIM Module contains the following data points:

Domain Name, Selector, Key Type, Key Length, Key Strength, Key Status, Key Flags, Timestamps



DNS Analysis: DMARC Module

Description

The DMARC Module uses DNS records to map DMARC configurations of domain names in order to determine the presence of basic e-mail policy and compliance controls.

Data Points:

The DMARC Module contains the following data points:

Timestamp, domain name, DMARC status

Github Dorking Module

Description

The GitHub Dorking Module crawls public code repositories for the presence of leaking credentials, source code, API keys, or other data of value. The GitHub Dorking Module identifies code of interest and attempts to the map discovered code to the developing enterprise. The GitHub Dorking Module is searchable by domain name.

Data Points

The GitHub Dorking Module contains the following data points:

Timestamp, Domain, Repository URL, Owner, File URL, Caption

End of Life Module

Description

The End of Life Module maps identified products and services within an enterprise to the 'End of Life' product support dates released by manufacturers. When a product has reached the 'End of Life' cycle, vendors no longer offer updates or support. The presence of legacy and End of Life systems within an enterprise network allows for vulnerabilities to accumulate over time, creating a larger attack surface for potential exploitation. The End of Life Module is searchable by domain name.

Data Points:

The GitHub Dorking Module contains the following data points:

Timestamp, Manufacturer, Product Name, IP Address, EOL Information URL, End of Service Date, End of Life Date, End of Sales Date

Leaked Passwords Module

Description

The Leaked Passwords Module identifies email addresses that have been identified circulating within the public realm, and/or within the hacker underground. The data sources are continuously updated and consist of forums, Pastebin-style sites, and archives of leaked/breached public databases. The Leaked Passwords Module is searchable by e-mail address or domain name.

Data Points:

The Leaked Passwords Module contains the following data points:

Timestamp. Email, Domain, Source, Leak Name

Hacker Chatter Module

Description

The Hacker Chatter Module is an archive of forums posts dedicated to the discussion of hacking, fraud, and other underground economic activities. The Hacker Chatter Module is continously updated, constantly crawling, indexing, and archiving posts made to dozens of private and public underground resources. The Hacker Chatter Module also archives and indexes the leaked/breached SQL databases of many hacking forums, providing insight into administrative activities, private message discussions, and Malicious Actor Profile information. The Hacker Chatter module is searchable by domain name (API), or wildcard keyword searches in the UI.

Data Points:

The Hacker Chatter Module contains the following data points:

Timestamp, Message Body, Message Subject, Mentioned Company, Mentioned SQLi attacks, Mentioned XSS Attack, Mentioned Credit Card Number(s), Mentioned Emails, Mentioned Paswords, Mentioned IPs

Sinkhole Feeds Module

Description

The Sinkhole Feeds Module consists of malware infected IP addresses that have been collected by the ThreatMarket Malware Sinkhole infrastructure. Malware communications are analyzed in order to identify the Command and Control (C2) infrastructure, and a DNS sinkhole methodology is employed to take over control the C2 domain names. When the C2 has been taken over, incoming beacons from infected machines that arrive into the ThreatMarket Sinkhole and are mapped back to the affected enterprise. The Sinkhole Feeds Module is searchable by domain name and timestamp.

Data Points

The Sinkhole Feeds Module contains the following data points:

Timestamp, IP Address, Domain Name, Hostname, Malware Family, Sinkhole Domain, Requested URL, UserAgent, Geolocation Information, Operating System Metadata

SSL Certificates

Description

The SSL Certificates module is a collection of all SSL/TLS data that was observed during the mapping of the IPv4 network space. The SSL certificates are mapped to vendor domain names, and analyzed from the standpoints of compliance, security, and best practices. The SSL Certificates Module is searchable by domain name.

Data Points

The SSL Certificates Module contains the following data points:

Timestamp, Domain, IPAddress, CommonName, StateOrProvince, Locality, Country, OrganizationName, KeyType, KeyLength, NotValidBefore, NotValidAfter, Md5, Sha1, Domain, IssuerOrganizationName, IssuerOrganizationUnit, CipherTlsFallbackScsv, CipherForwardSecrecy, CipherRc4Allowed, CertSignatureMd2, CertSignatureMd5, CertSignatureSha1, CertHasCrl, CertHasOcsp, CertRevoked, CertWeakDebianKey,

CertChainValid, SecureRenegotiation, CertSelfSigned, CertServerKeyLess_1024, CertExpired, KeyExNoAuthen, KeyExNoEncrypt, KeyExExportAllowed, ProtoSslv2, ProtoSslv3, ProtoTls1_0, ProtoTls1_1, ProtoTls1_2, Cipher_40Or_56, Cipher_128Or_168, Cipher_256

TOR Nodes Module

Description

The TOR Nodes Module is a historical collection data that archives all Exit Nodes and Relay Nodes identified as participating within the TOR network. The TOR Nodes Module is searchable by domain name and IP address.

Data Points:

The SSL Certificates Module contains the following data points:

Timestamp, IP Address, Nickname, Domain Name, Onion Relay Port, Directory Port, Identify, Digest, Publication, Statuses, Bandwidth, Policy, Exit Node Status

Sample data for Core Data Collection Modules described in section above

SSC's observational data can be made available in a variety of formats including batch dumps in XML or JSON as well as through a REST API. We provide here a few samples, used purely for illustrative purposes.

Banner grabs

Banner grab data includes detected product and version information associated with services detected running on a port at a given IP Address. Additionally, SSC attributes this IP Address to a domain or organization.

```
<BannerGrab>
  <Port>389</Port>
  <Version>2.2.X - 2.3.X</Version>
  <Extra></Extra>
  <Product>OpenLDAP</Product>
  <Name>ldap</Name>
  <Domains>
    <Domain>fuse.net</Domain>
  </Domains>
  <RawBanner>&#xA; Context:
dn: cn=autoserversetup,dc=ipadserver,dc=local

                                                 cn: autoserversetup

                                                                            objectClass:
container

Result limited to 20 objects (see Idap.maxobjects)</RawBanner>
  <IPAddress>66.161.136.19</IPAddress>
  <DateObserved>2015-11-09T18:02:13Z</DateObserved>
</BannerGrab>
```

SSL Certificates

SSC's performs comprehensive analysis on SSL server configuration as well as site certificates.

```
<ScanSSLResult>
  <IP>
    <ID>49915484</ID>
    <ip_address>104.99.19.218</ip_address>
    <reverse_dns>a104-99-19-218.deploy.static.akamaitechnologies.com</reverse_dns>
    <country>US</country>
    <city></city>
    <latitude>38.0</latitude>
    <longitude>-97.0</longitude>
  </IP>
  <Port>443</Port>
  <ObservedDate>2016-06-07T23:59:37Z</ObservedDate>
  <CAListed>false</CAListed>
    <GUID>0c8e61ea-931e-5108-a118-a6db3a661563</GUID>
  </CertChain>
  <ChainValid>false</ChainValid>
  <Cipher256>true</Cipher256>
  <Cipher128or168>true</Cipher128or168>
  <Cipher40or56>true</Cipher40or56>
  <CipherSuites>
    <SSLv2>
      <CipherSuite>SSL2_DES_192_EDE3_CBC_WITH_MD5</CipherSuite>
      <CipherSuite>SSL2_RC4_128_WITH_MD5</CipherSuite>
      <CipherSuite>SSL2_DES_64_CBC_WITH_MD5</CipherSuite>
      <CipherSuite>SSL2_RC4_128_EXPORT40_WITH_MD5</CipherSuite>
    </SSLv2>
    <SSLv3>
```

```
<CipherSuite>TLS_RSA_WITH_3DES_EDE_CBC_SHA</CipherSuite>
    <CipherSuite>TLS_RSA_WITH_AES_256_CBC_SHA</CipherSuite>
    <CipherSuite>TLS_RSA_WITH_DES_CBC_SHA</CipherSuite>
    <CipherSuite>TLS_RSA_WITH_RC4_128_MD5</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_RC4_128_SHA</CipherSuite>
    <CipherSuite>TLS_RSA_WITH_AES_128_CBC_SHA</CipherSuite>
  </SSLv3>
  <TLSv1_0>
    <CipherSuite>TLS_RSA_WITH_3DES_EDE_CBC_SHA</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_DES_CBC_SHA</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_AES_256_CBC_SHA</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_AES_128_CBC_SHA</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_RC4_128_MD5</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_RC4_128_SHA</CipherSuite>
  </TLSv1_0>
  <TLSv1 1>
   <CipherSuite>TLS_RSA_WITH_AES_256_CBC_SHA</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_3DES_EDE_CBC_SHA</CipherSuite>
    <CipherSuite>TLS_RSA_WITH_AES_128_CBC_SHA</CipherSuite>
  </TLSv1_1>
  <TLSv1_2>
   <CipherSuite>TLS_RSA_WITH_AES_128_CBC_SHA</CipherSuite>
   <CipherSuite>TLS_RSA_WITH_3DES_EDE_CBC_SHA</CipherSuite>
    <CipherSuite>TLS_RSA_WITH_AES_256_CBC_SHA</CipherSuite>
  </TLSv1_2>
</CipherSuites>
```

The above excerpt shows a portion of the data SSC collects, including a breakdown of ciphersuites supported by a target server per protocol. In the above example, the certificate chain presented by the server was analyzed and was unable to create a valid chain back to a trusted root CA.

Vulnerable Hosts and CVE Details

SSC tracks a large set of current vulnerabilities and reports vulnerable IPs, which are also attributed to a domain or organization. The following sample data has been anonymized.

```
<VulnerableHosts>
  <IpAddress>[redacted]</lpAddress>
  <Domain>[redacted]</Hostname>
  <Latitude>49.75</Latitude>
  <Longitude>6.1667</Longitude>
  <Country>LU</Country>
  <City>Arlington Heights</City>
  <DateObserved>2016-06-05T23:59:54Z</DateObserved>
  <VulnName>HEARTBLEED</VulnName>
  <Cve>CVe-2015-0160</Cve>
  <Port>443</Port>
  <Url></Url></VulnerableHosts>
```

Leaked Passwords

Leaked account names and passwords are collected from a variety of sources across the Internet and attributed to a domain or organization.

```
<LeakedPasswords>
<Email>josh.wasserman@goldmansachs.com</Email>
<Date>2016-06-03T12:00:00Z</Date>
<Tld>goldmansachs.com</Tld>
<LeakSource>pastebin</LeakSource>
```

Malware infections feed

SSC researchers are constantly reverse-engineering current malware in order to discover malware-associated domains which are promptly registered by SSC as malware sinkholes. This enables SSC to discover IPs that have been infected by specific families of malware and associate them with a domain or organization.

```
<MalwareInfections>
  <asn>8048</asn>
  <timestamp>2015-10-24 15:05:04</timestamp>
  <data_source>asdoijasdoij.ru sinkhole</data_source>
  <notes>source_port: 4414, dest_addr: 38.229.12.17</notes
  <category>bot</category>
  <address>186.94.151.48</address>
  <family>conficker</family>
  <company>goldmansachs</company>
  </MalwareInfections>
```

Vendor Relationships

SSC discovers automatically shadow IT and cloud-services used by other companies. The data will include the relationships of vendors used by a specific company through an API call, and an explanation of how that relationship was obtained.

Attack Feeds

SSC monitors for common attacks such as SSH brute-forcing attempts and exposes that data.

Hostname Subdomain Enumeration

SSC employs a number of methods to discover all subdomains associated with a Top-Level Domain.

```
</VendorSubdomains>
<VendorSubdomains>
 <Subdomain>portal.stratafs.com</Subdomain>
 <UpdatedAt>2016-06-09T00:31:34Z</UpdatedAt>
 <Domain>stratafs.com</Domain>
 <IpAddress>54.67.0.105</IpAddress>
</VendorSubdomains>
<VendorSubdomains>
 <Subdomain>remote.stratafs.com</Subdomain>
 <UpdatedAt>2016-06-09T00:31:34Z</UpdatedAt>
 <Domain>stratafs.com</Domain>
 <lp>Address>162.219.21.82</lp>
IpAddress>
</VendorSubdomains>
<VendorSubdomains>
 <Subdomain>work.stratafs.com</Subdomain>
 <UpdatedAt>2016-06-09T00:31:34Z</UpdatedAt>
 <Domain>stratafs.com</Domain>
 <IpAddress>162.219.21.85</IpAddress>
</VendorSubdomains>
```

TOR Node

SSC tracks all TOR nodes and, where possible, attributes nodes running inside of corporate environments.

Sample data schema for the banner grab module mentioned in prior section

```
"dataset": "banner_grab",

"doc": "A method of software and service identification performed by scanning computer systems on a network.",

"fields": [
{
    "name": "id",
    "type": "integer",
    "doc": "ID of the record."
},
{
    "name": "ip_address",
    "type": "string",
    "doc": "IP address of the host where banner was discovered."
},
{
    "name": "port",
```

```
"type": "integer",
  "doc": "Port number on which banner was discovered."
  "name": "name",
  "type": "string",
  "doc": "Name of service disclosing banner."
  "name": "version",
  "type": "string",
  "doc": "Version associated with banner contents."
  "name": "extra",
  "type": "string",
  "doc": "Additional information included with banner."
  "name": "domain",
  "type": "string",
  "doc": "Domain associated with banner."
  "name": "product",
  "type": "string",
  "doc": "Software that is serving the banner."
  "name": "raw_bannner",
  "type": "string",
  "doc": "Banner at the time it was acquired."
  "name": "date_observed",
  "type": "datetime",
  "doc": "Timestamp of observation."
]
```

Certification by Authorized Official of SecurityScorecard, Inc.

I have reviewed the content of this proposal and authorize it's submittal to Verisk.

Sam Kassoumeh, Co-founder and COO