Can we leverage data science to improve retail data security?

Presenters:

Ashwin Swamy, Omega ATC

Thomas Duncan, Omega ATC

CONEXXUS

solve forward

Agenda

- Housekeeping
- Presenters
- About Conexxus
- Presentation
- Q & A

Housekeeping

This webinar is being recorded and will be made available in approximately 30 days.

- YouTube (youtube.com/conexxusonline)
- Website Link (conexxus.org)

Slide Deck

- Survey Link – Presentation provided at end

Participants

- Ask questions via webinar interface
- Please, no vendor specific questions

Email: **info@conexxus.org**

Presenters

Conexxus Host
Allie Russell
Conexxus
arussell@conexxus.org

Moderator
Kara Gunderson
Chair, Data Security Committee
POS Manager, CITGO Petroleum
kgunder@citgo.com

Speakers

Ashwin Swamy
Director, Omega ATC
ashwin.swamy@omegaatc.com

Thomas Duncan
Security Strategist, Omega ATC
thomas.duncan@omegaatc.com

About Conexxus

- We are an independent, non-profit, member driven technology organization
- We set standards…
  - Data exchange
  - Security
  - Mobile commerce
- We provide vision
  - Identify emerging tech/trends
- We advocate for our industry
  - Technology is policy

CONEXXUS
solve forward

2018 Conexxus Webinar Schedule*

Month/Date	Webinar Title	Speaker	Company
May 24, 2018	QIR Program Update	Chris Bucolo Todd Rosen	ControlScan
June 21, 2018	Leveraging data science to improve retail data security	Thomas Duncan Ashwin Swamy	Omega ATC Omega ATC
July 26, 2018	Roadmap to a Vulnerability Disclosure Program	Joe Basirico	Security Innovation
August 23, 2018	Moving Toward Outdoor EMV	Linda Toth	Conexxus
September 2018	TBD	George Sconyers	Omega ATC
November 2018	Building a Scalable Security Engineering Team	Joe Basirico	Security Innovation

CONEXXUS
solve forward

NACS Show
October 7-10, 2018
Las Vegas, NV

TECHEDGE
In partnership with CONEXXUS

Conexxus thanks our 2018 Annual Diamond Sponsors!

mSHIFT.
Relevant. Mobile. Solutions.

Stuzo

DIEBOLD NIXDORF

GILBARCO VEEDER-ROOT

Cybera®
Simplify Security and Networks

CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

Aligning security with overall analytics strategy

Easing the burden of security management

Grounding data science in reality

CONEXXUS
solve forward

But first, let’s understand the problem.

This week at Tesla Motors – Insider Attack

From: Elon Musk
To: Everybody
Subject: Some concerning news
June 17, 2018
11:57 p.m.

I was dismayed to learn this weekend about a Tesla employee who had conducted quite extensive and damaging sabotage to our operations. This included making direct code changes to the Tesla Manufacturing Operating System under false usernames and exporting large amounts of highly sensitive Tesla data to unknown third parties.

The full extent of his actions are not yet clear, but what he has admitted to so far is pretty bad. His stated motivation is that he wanted a promotion that he did not receive. In light of these actions, not promoting him was definitely the right move.

However, there may be considerably more to this situation than meets the eye, so the investigation will continue in depth this week. We need to figure out if he was acting alone or with others at Tesla and if he was working with any outside organizations.

This week at Tesla Motors – Insider Attack

From: Elon Musk
To: Everybody
Subject: Some concerning news
June 17, 2018
11:57 p.m.

I was dismayed to learn this weekend about a Tesla employee who had conducted quite extensive and damaging sabotage to our operations. This included making direct code changes to the Tesla Manufacturing Operating System under false usernames and exporting large amounts of highly sensitive Tesla data to unknown third parties.

The full extent of his actions are not yet clear, but what he has admitted to so far is pretty bad. His stated motivation is that he wanted a promotion that he did not receive. In light of these actions, not promoting him was definitely the right move.

However, there may be considerably more to this situation than meets the eye, so the investigation will continue in depth this week. We need to figure out if he was acting alone or with others at Tesla and if he was working with any outside organizations.

An Analogous C-Store Attack

A bad actor (internal or external) attempts to log into the POS system, site controller, fuel controller, or Electronic Payment Controller.

Social engineering attacks are frequent at c-stores. Callers or visitors can frequently pose as vendors with minimum resistance.

If the bad actor succeeds in accessing a system, he or she may try to access particular files – either to exfiltrate sensitive data or simply to disrupt operations.

In the case of stealing sensitive data, the bad actor will attempt to export or send the data using a variety of methods, such as simply sending an email, covertly changing the web address where payments are sent, using a USB drive, etc.

CONEXXUS
solve forward

Human behavior leaves a trace

In both the Tesla and c-store example, multiple actions were taken by the bad actor – logging into one or more systems, accessing a file or program, and removing the data.

Hence, in these scenarios, any event logs related to user access, file behavior, and data transmission (USB drive, email, DNS settings) are related and collectively form a “fingerprint.”

The question is how to understand and detect these relationships.

So how does this relate to data science?

We must be able to filter, detect, and alert on specific events AND patterns of events.

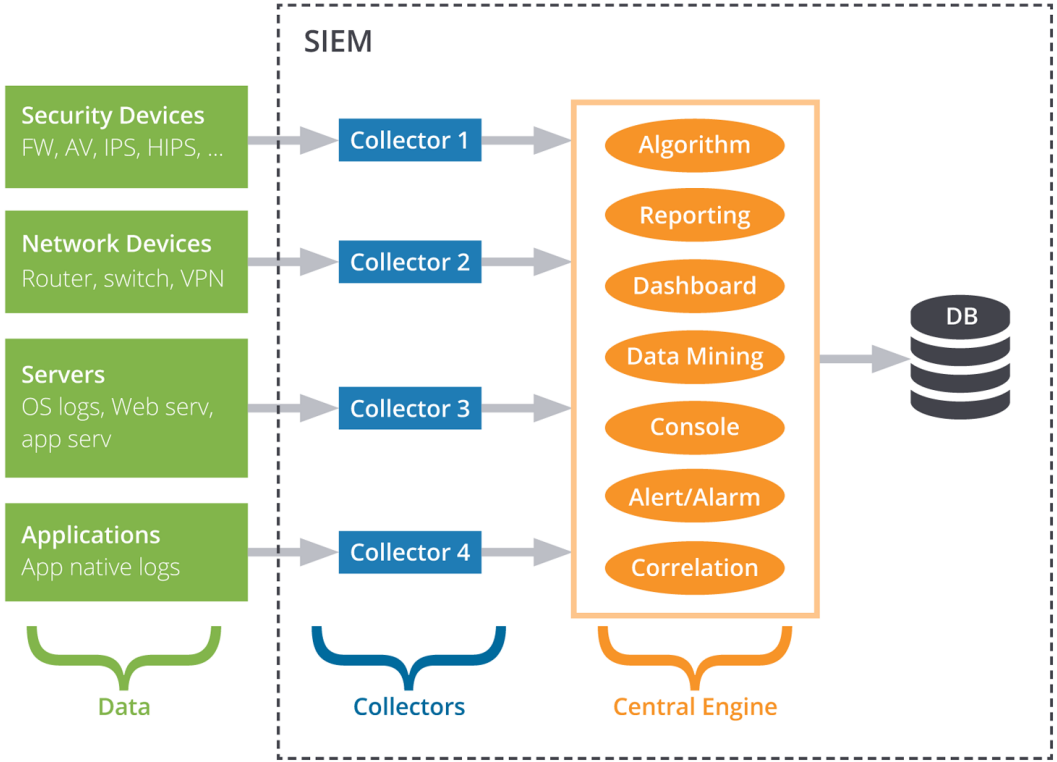
It’s crucial to see log data from multiple perspectives…

CONEXXUS
solve forward

Common problems managing a SIEM program

Endpoints and “vectors” yield a discrete set of alerts:

- No understanding of relationships between machine activity by time.
- No understanding of relationships between different processes.
- Events are not understood in historical context.

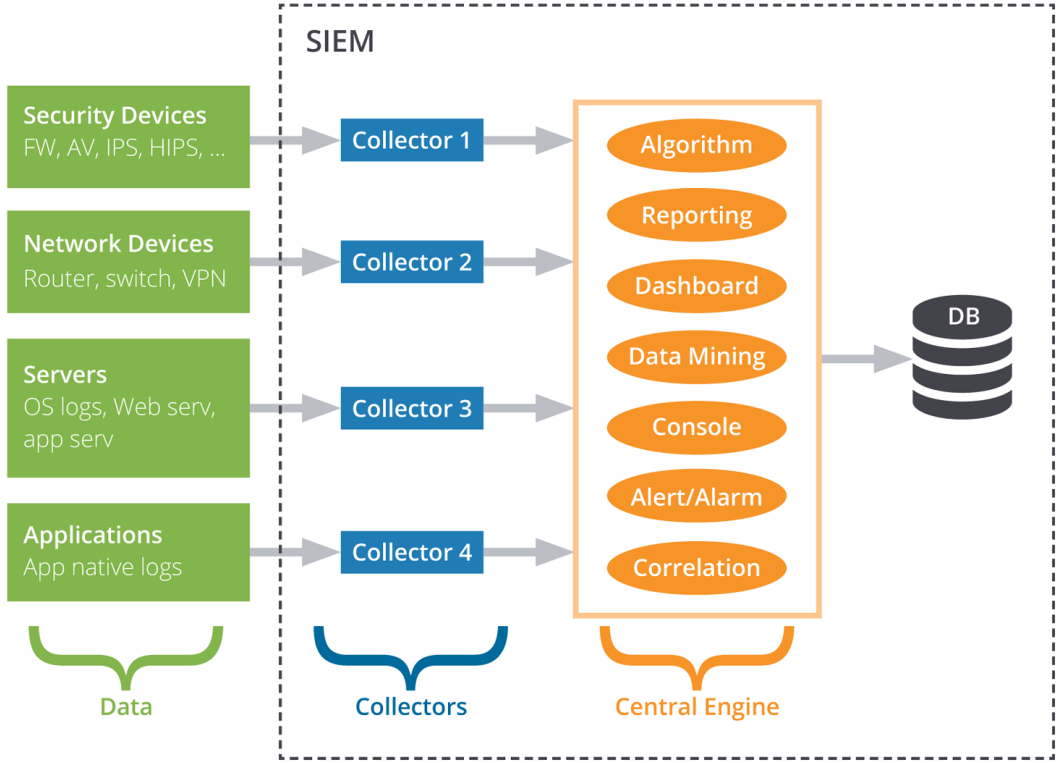


CONEXXUS
solve forward

Data Science and SIEM – how they work together

Centrally collecting logs is not enough:

- SIEM tools must be able to correlate events in a time series format.
- SIEM tools must work alongside other analysis tools that provide methods for regression and clustering.
- SIEM tools must allow the security teams the ability to make models smarter.



CONEXXUS
solve forward

Common problems managing a SIEM program – events are not easily relatable by time

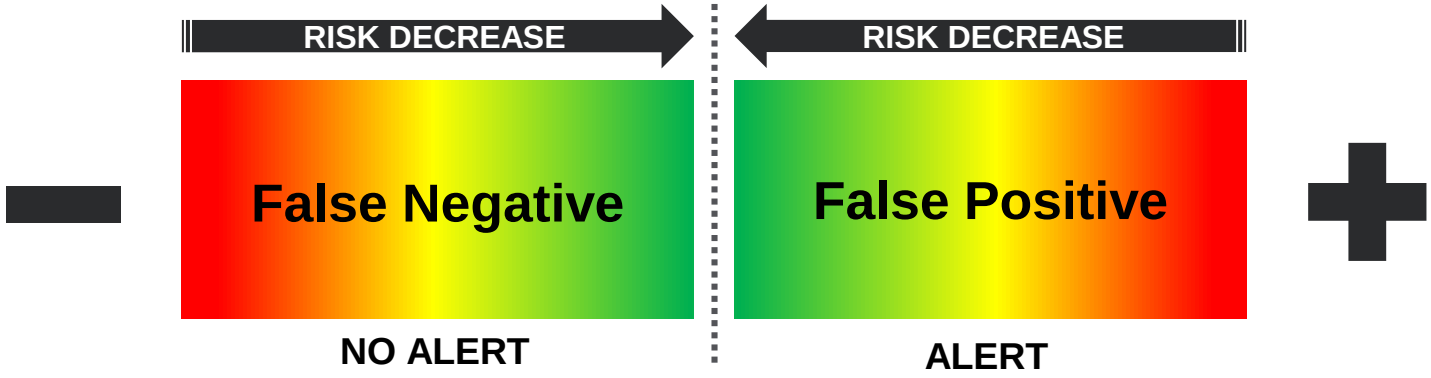
Time	Firewall	POS	Back Office	EPC
01:00:00	FW Event A	POS Event A	BO Event A	EPC Event A
01:01:00	FW Event B	POS Event B	BO Event A	EPC Event B
01:02:00	FW Event C	POS Event A	BO Event B	EPC Event C
01:03:00	FW Event D	POS Event C	BO Event C	EPC Event D

SIEM – the modern approach to dumping logs

CONEXXUS
solve forward

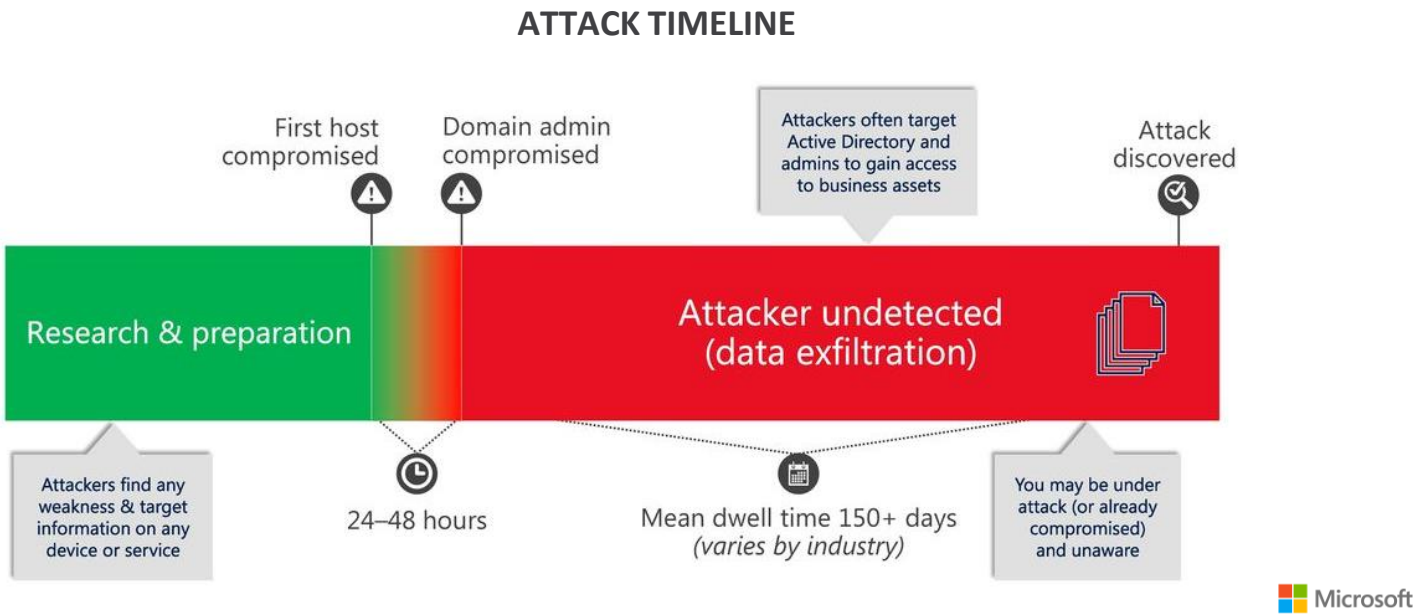
False Positives. Positively Negative.

1. A retail chain can produce thousands of alerts per year in a SIEM platform.
2. Standard application updates and processes can inadvertently trigger alerts.
3. Security teams do not always have an easy method for prioritizing alerts – no queuing methodology.

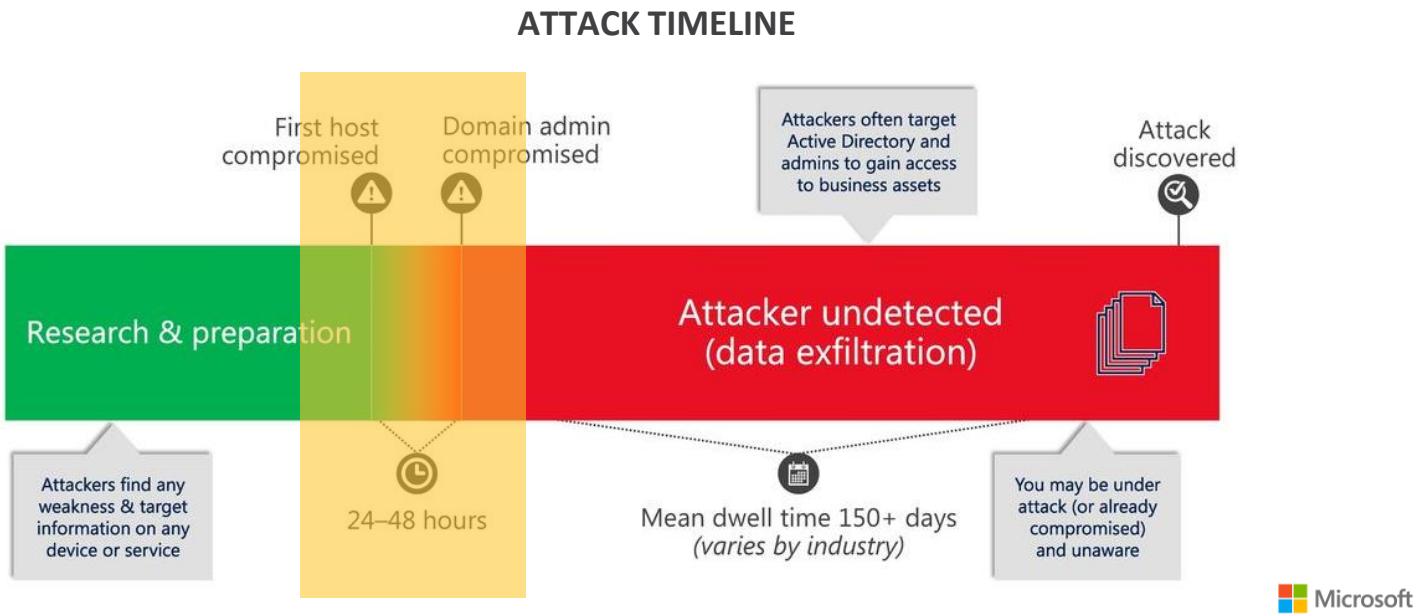


CONEXXUS
solve forward

Security teams need to review alerts that truly matter, or else it might be too late.



The 24-48 hour compromise window is all that is needed for initial exfiltration of data.



CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

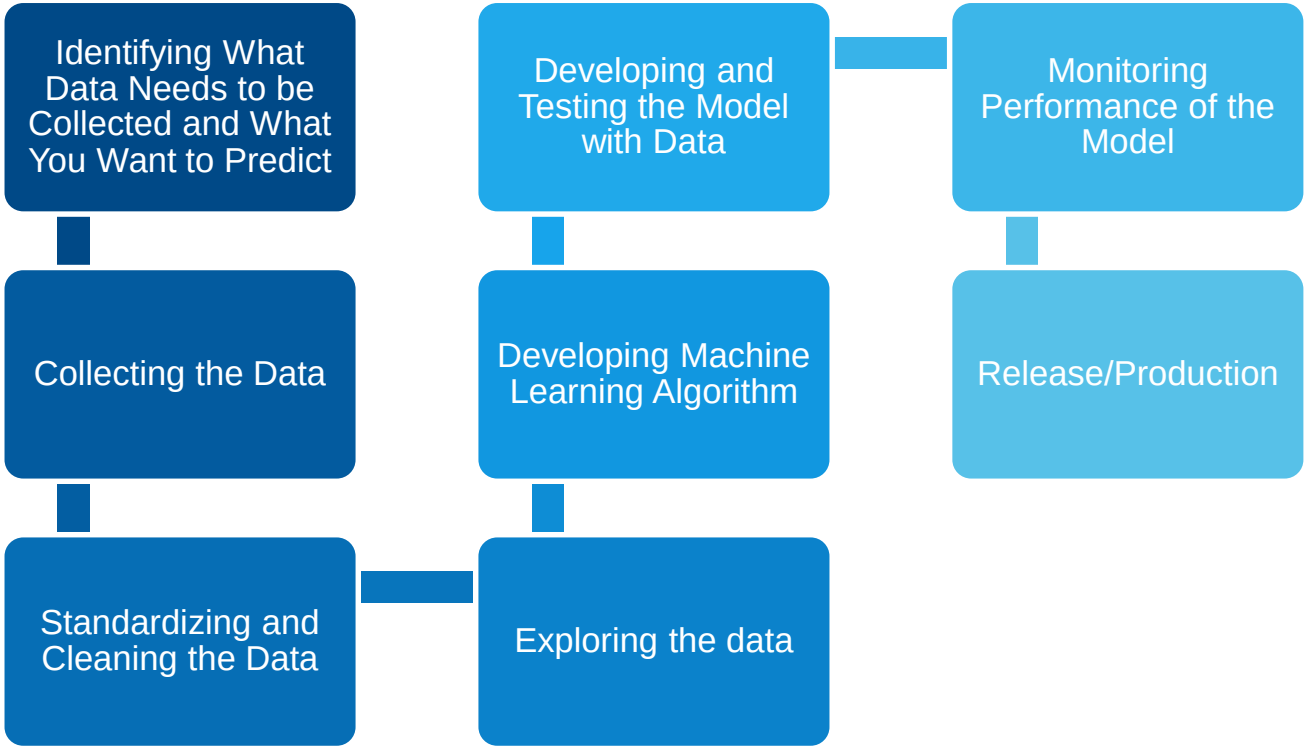
Aligning security with overall analytics strategy

Easing the burden of security management

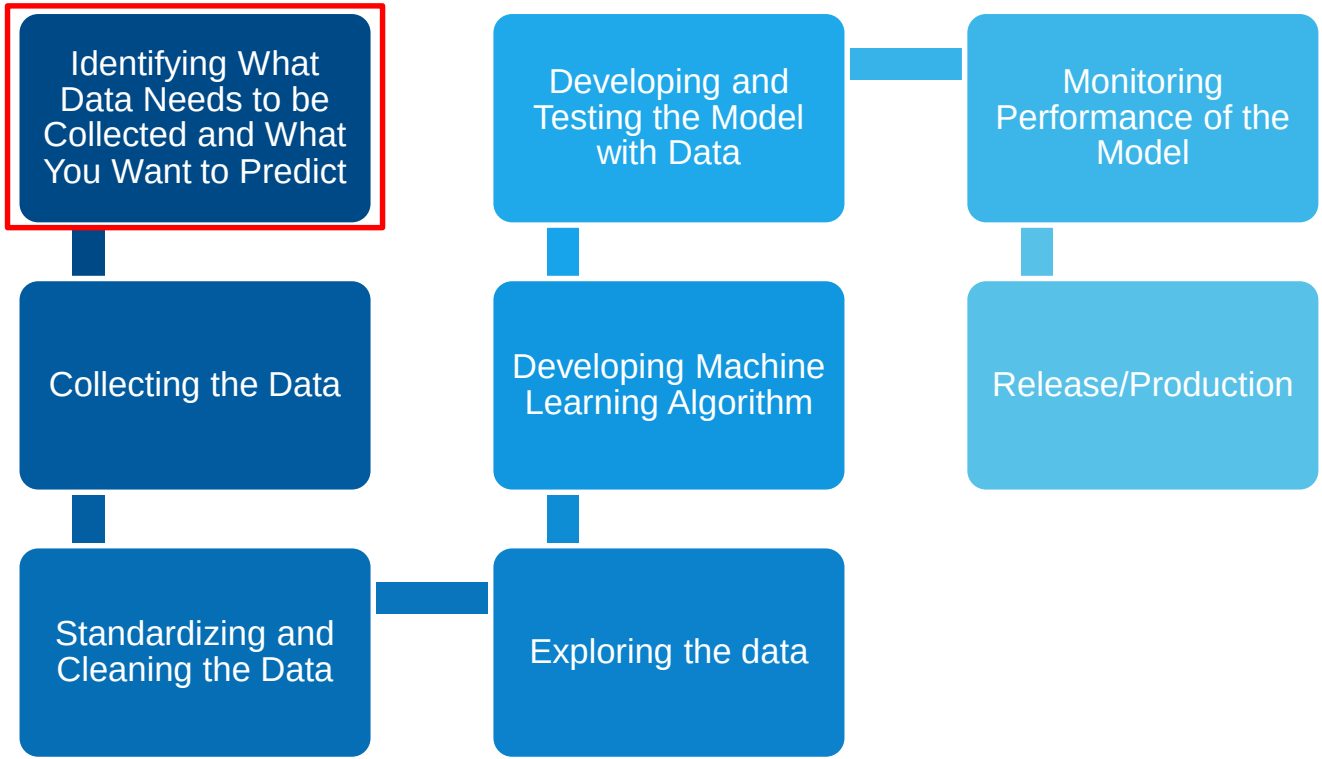
Grounding data science in reality

CONEXXUS
solve forward

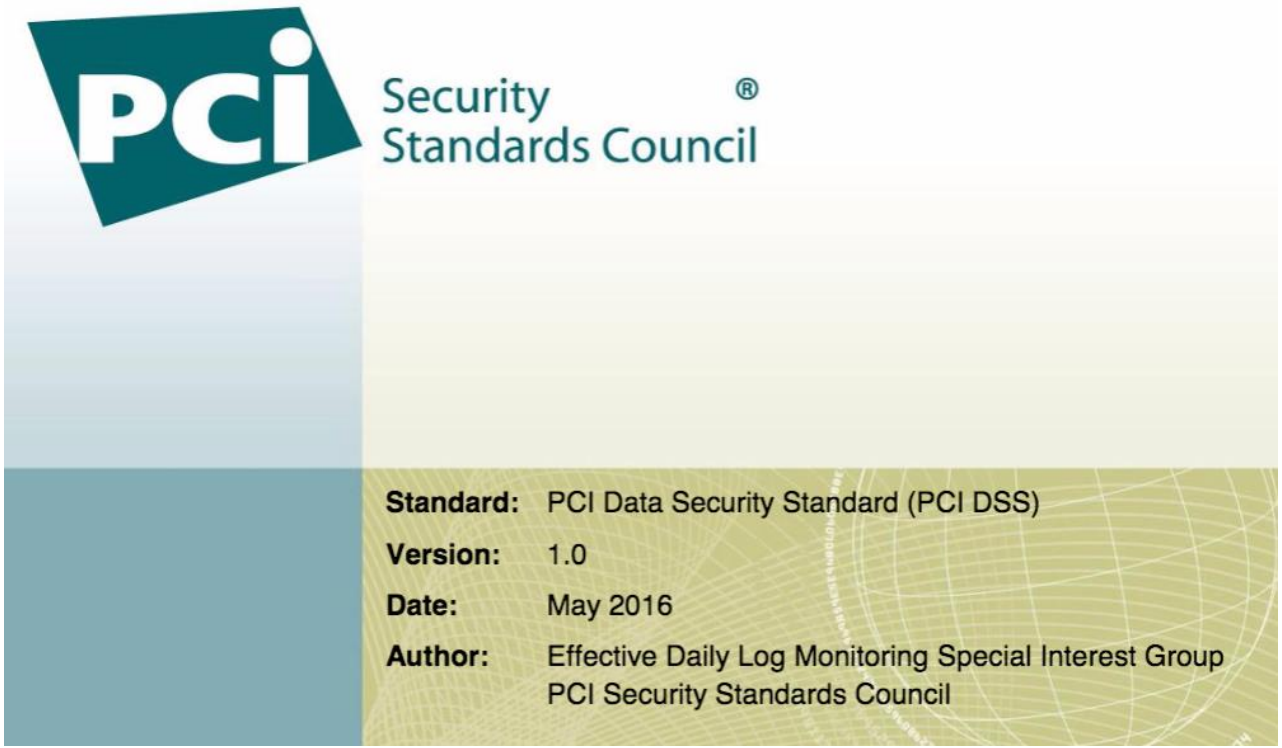
Data Science is a methodology.



Identify what data needs to be collected.

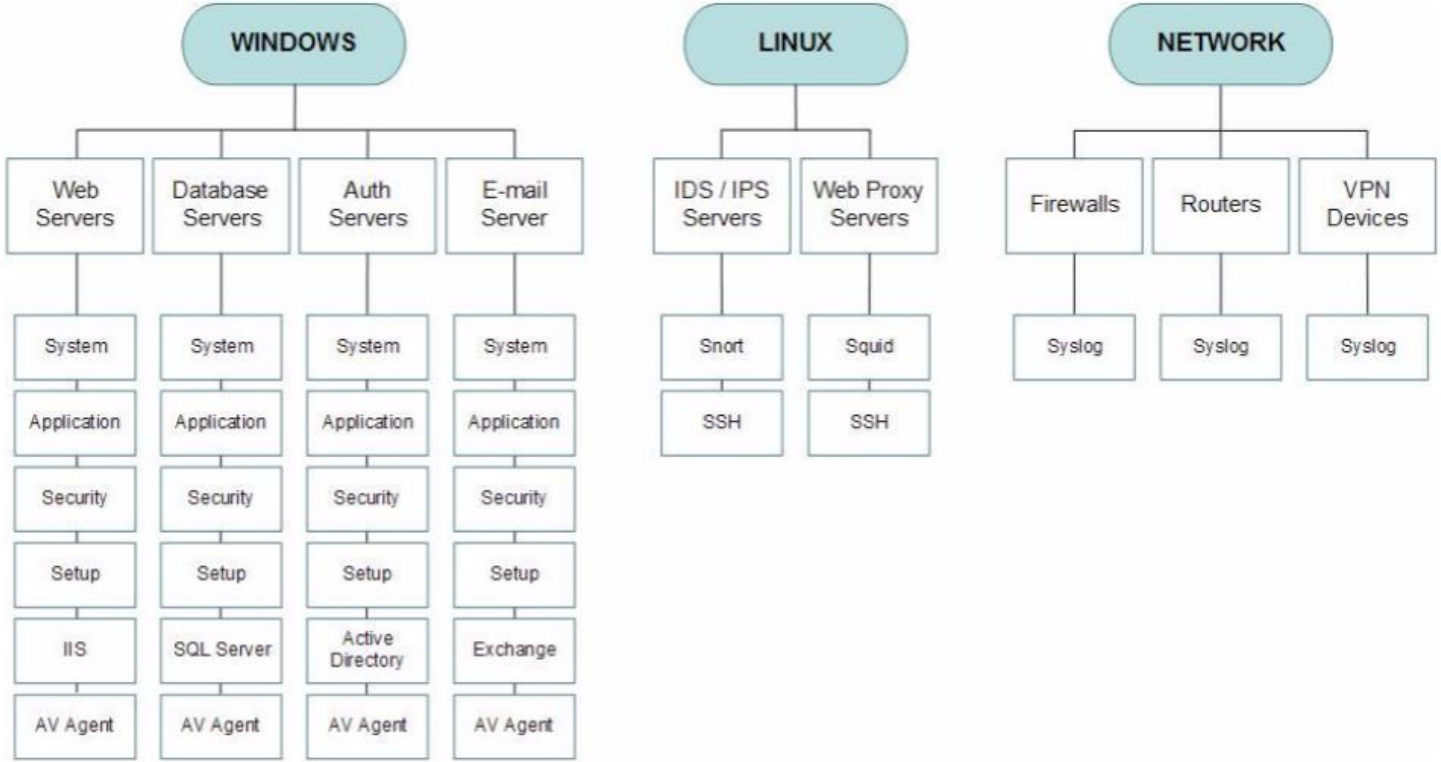


Identify what data needs to be collected.



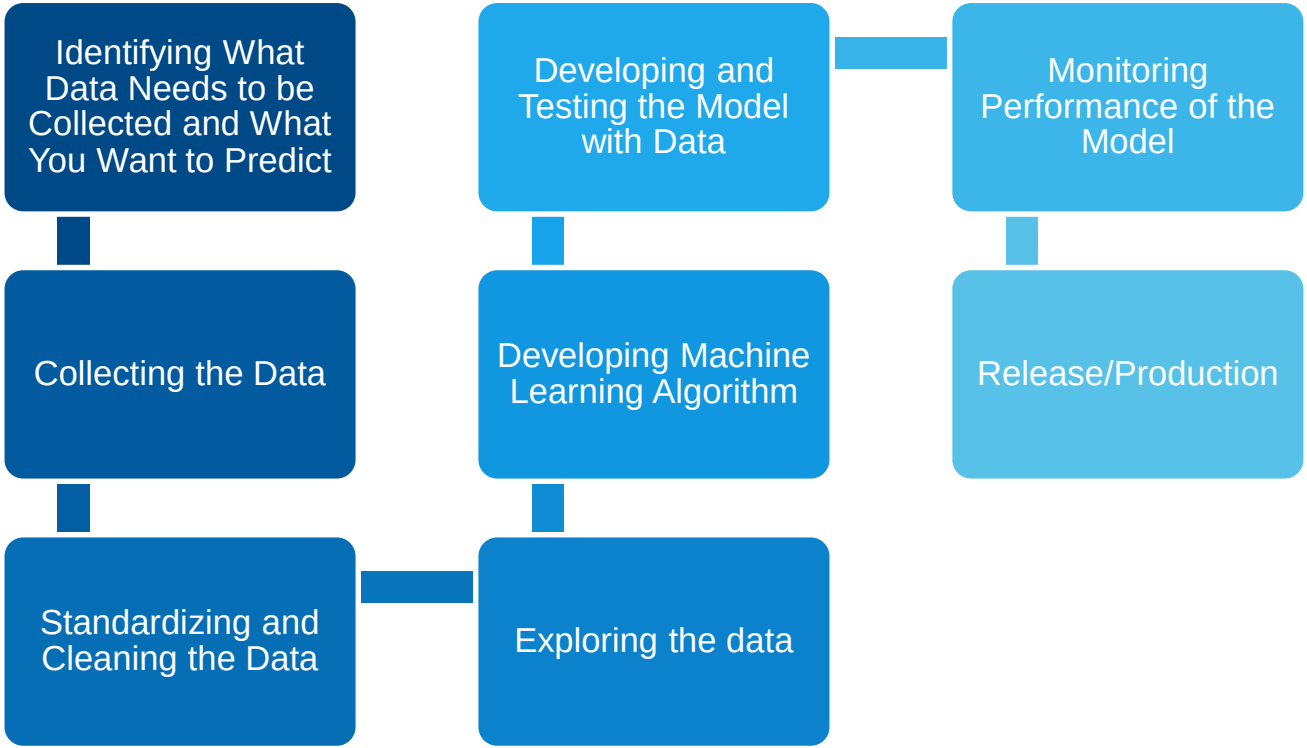
CONEXXUS
solve forward

Identify what data needs to be collected.



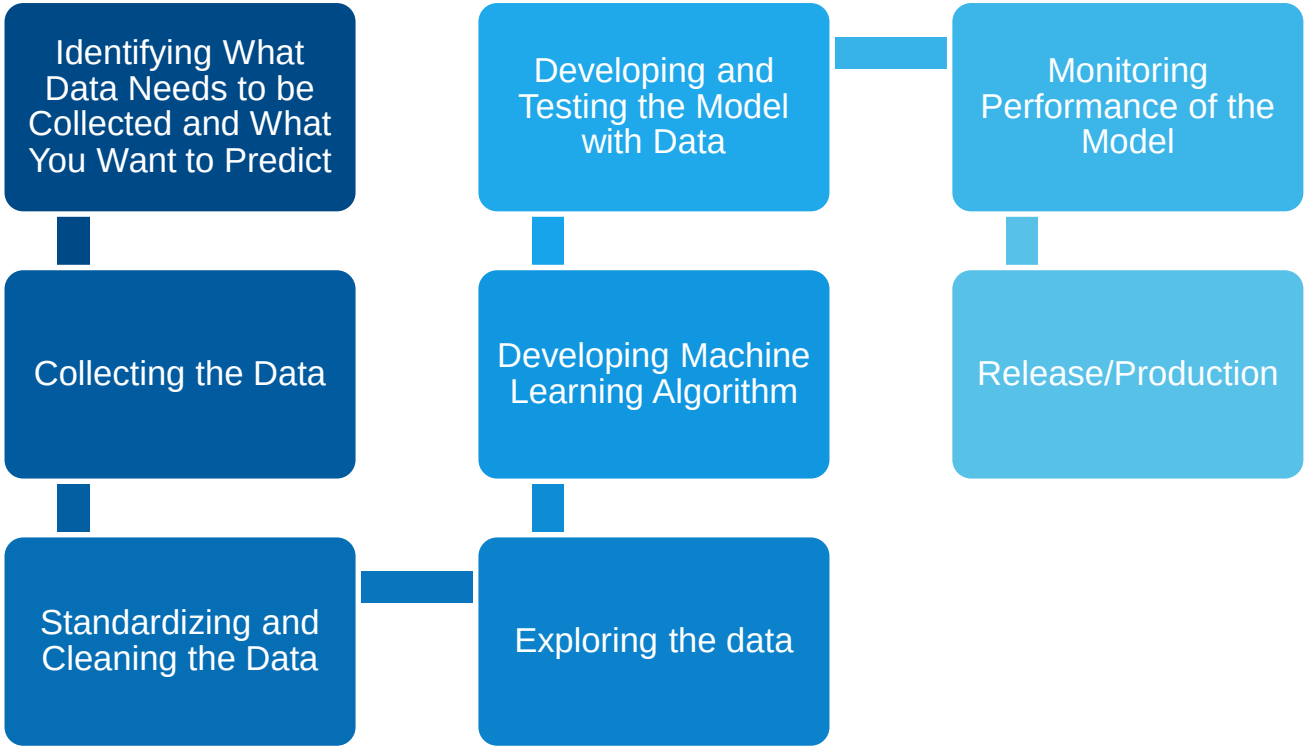
CONEXXUS
solve forward

We are trying to classify and predict malicious behavior



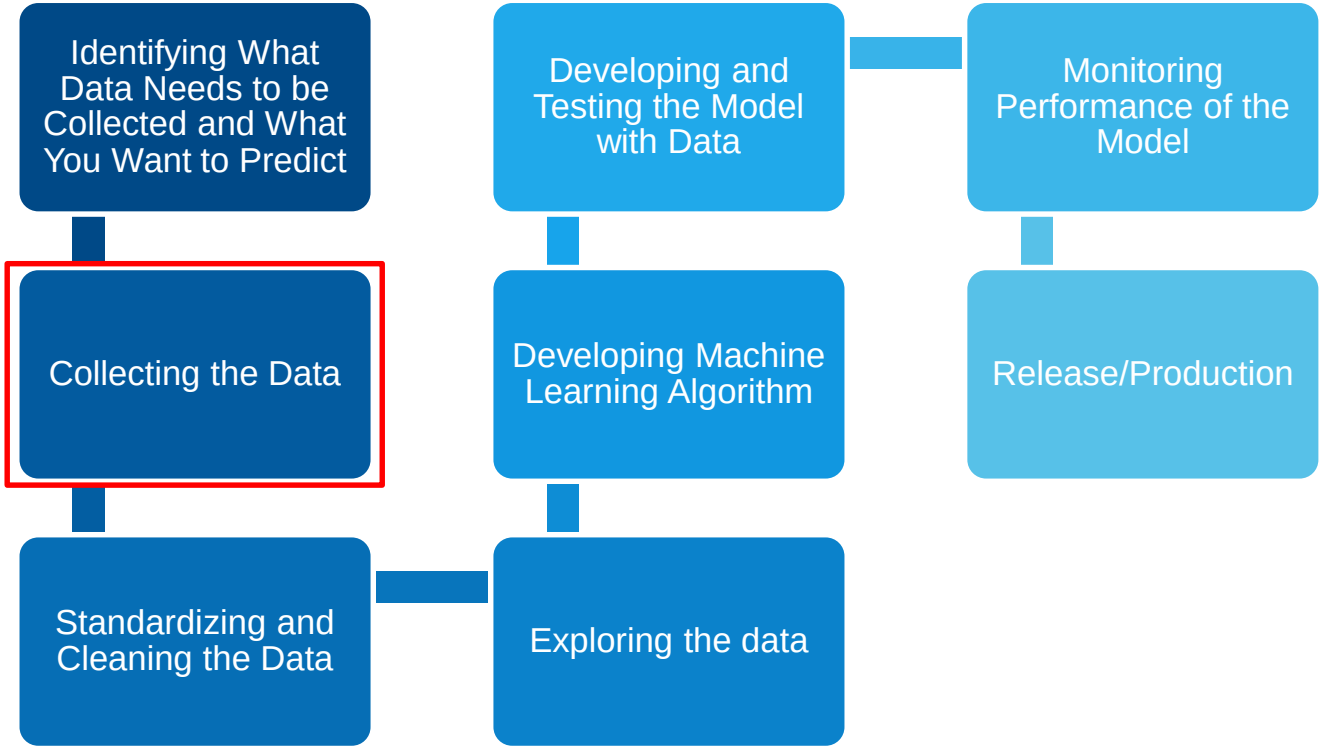
CONEXXUS
solve forward

We are also trying to classify and predict what is NOT malicious behavior

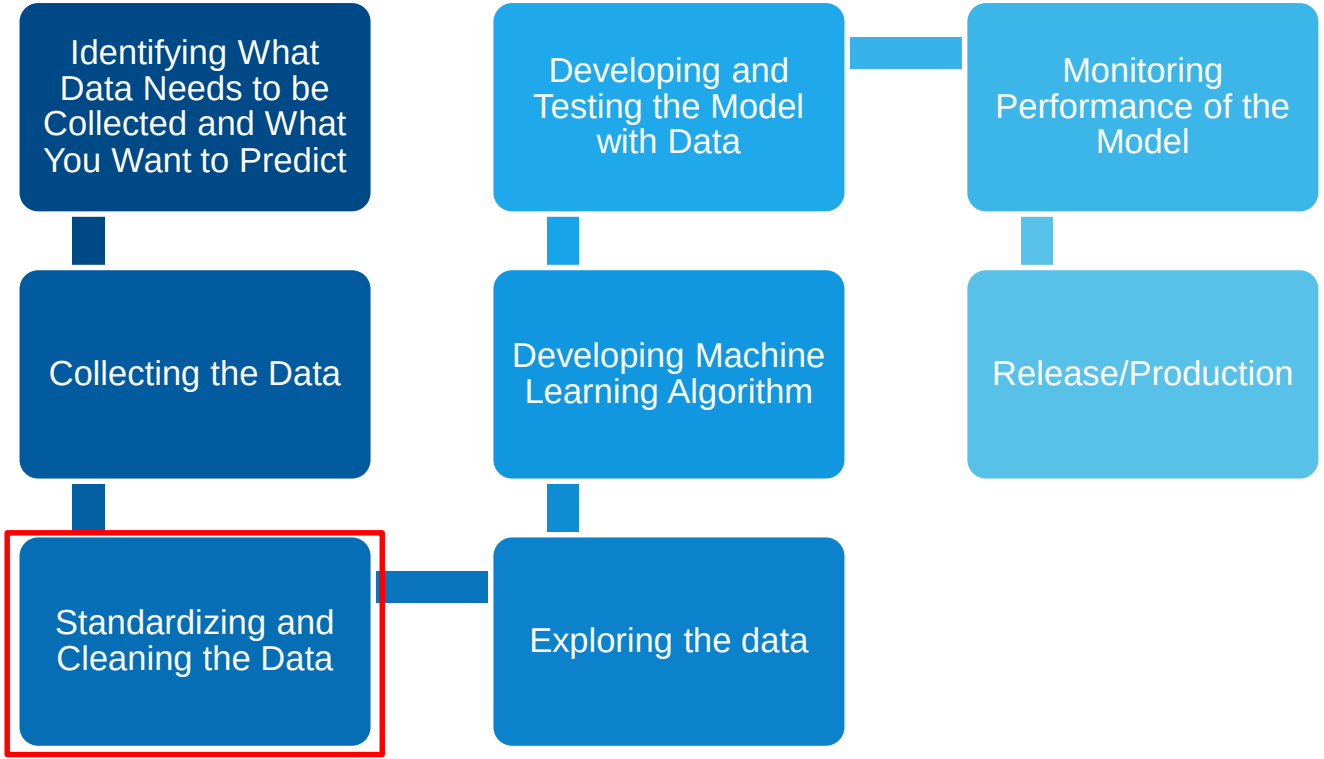


CONEXXUS
solve forward

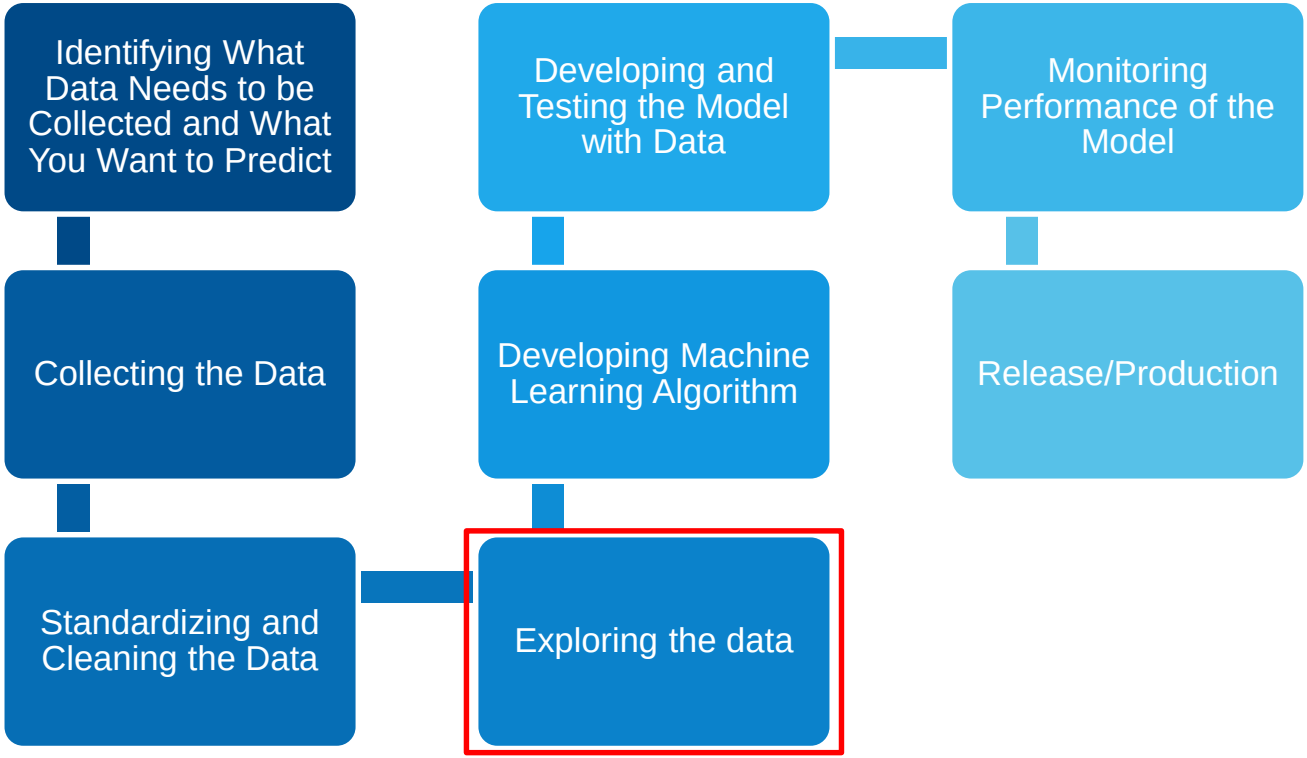
Data must be sent, received, and stored.



Data must be in a format that is usable.



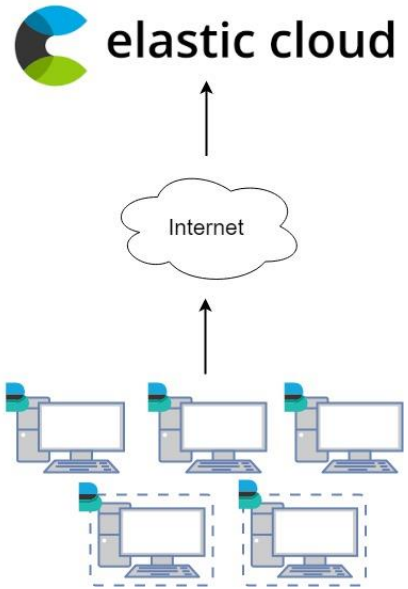
Data must be explored before it can be understood and applied.



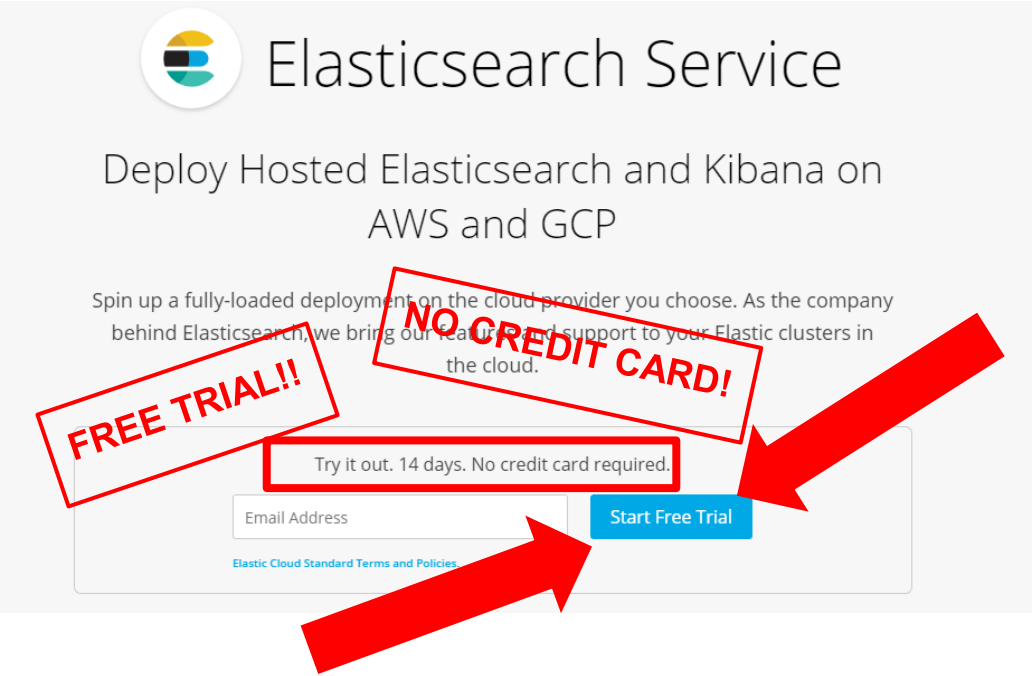
Let’s explore!

ELK Stack by Elastic – an open source analytics toolkit

Methodology for demonstration



- Visualize live data to form connections and patterns
- Automated Scripts
- Real-life workflow



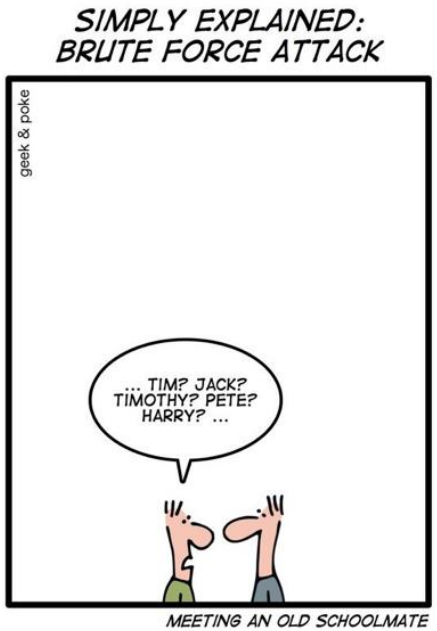
CONEXXUS
solve forward

Attack #1 – brute force password attack

Objective: Attempt to gain access to a user account / fi' share by guessing the password for a known username

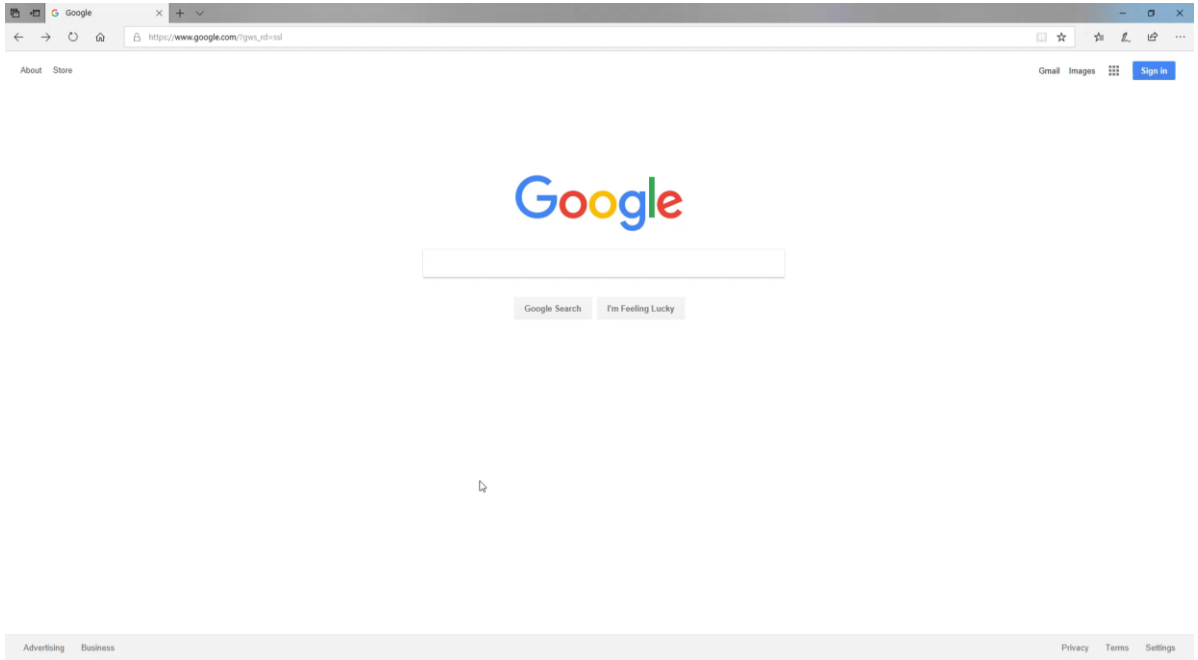
Tips -

- Research the environment
- DON’T GET CAUGHT
    - Consider common thresholds for security alerting



Google = the hacker's primary analytics tool

...need to find username, let's Google it!



CONEXXUS
solve forward

PCI Jiu Jitsu – using the rules against you

PCI Standards for accounts?

- Lock account after 6 failed attempts
- Disable account for 30 minutes

8.1.6 Limit repeated access attempts by locking out the user ID after not more than six attempts.	**8.1.6.a** For a sample of system components, inspect system configuration settings to verify that authentication parameters are set to require that user accounts be locked out after not more than six invalid logon attempts.
	8.1.6.b ***Additional testing procedure for service provider assessments only*:** Review internal processes and customer/user documentation, and observe implemented processes to verify that non-consumer customer user accounts are temporarily locked-out after not more than six invalid access attempts.
8.1.7 Set the lockout duration to a minimum of 30 minutes or until an administrator enables the user ID.	**8.1.7** For a sample of system components, inspect system configuration settings to verify that password parameters are set to require that once a user account is locked out, it remains locked for a minimum of 30 minutes or until a system administrator resets the account.

Windows defaults?

- NONE! (Please make sure this isn't you..)
  - We will assume PCI standards

Policy	Security Setting
Account lockout duration	Not Applicable
Account lockout threshold	0 invalid logon attempts
Reset account lockout counter after	Not Applicable

Policy	Security Setting
Enforce password history	0 passwords remembered
Maximum password age	42 days
Minimum password age	0 days
Minimum password length	0 characters
Password must meet complexity requirements	Disabled
Store passwords using reversible encryption	Disabled

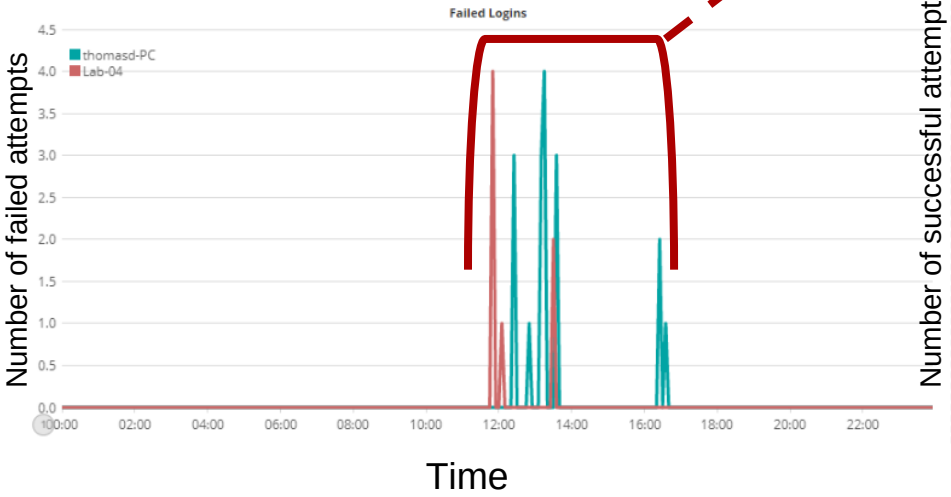
CONEXXUS
solve forward

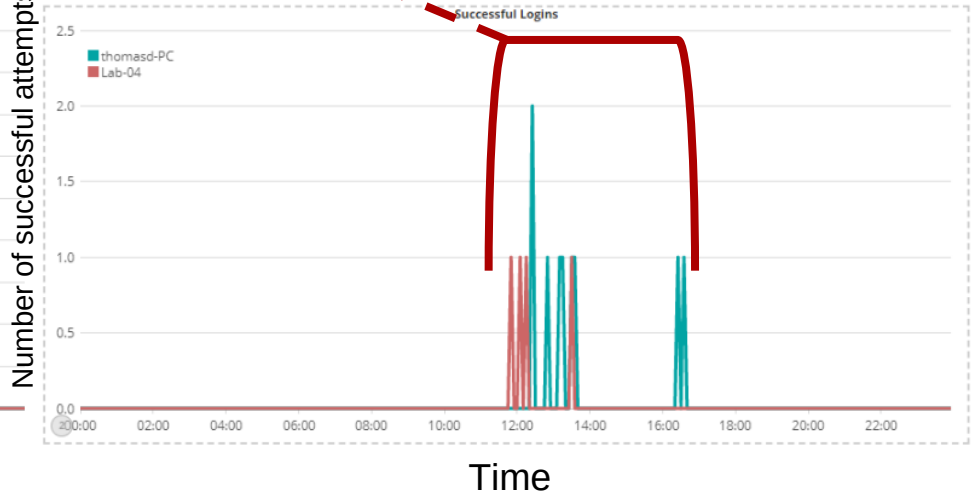
Correlating events between endpoints helps identify human behavior.

CONEXXUS
solve forward

Comparing login patterns can give insight into behavior.

Correlation between failed and successful login from legitimate user.





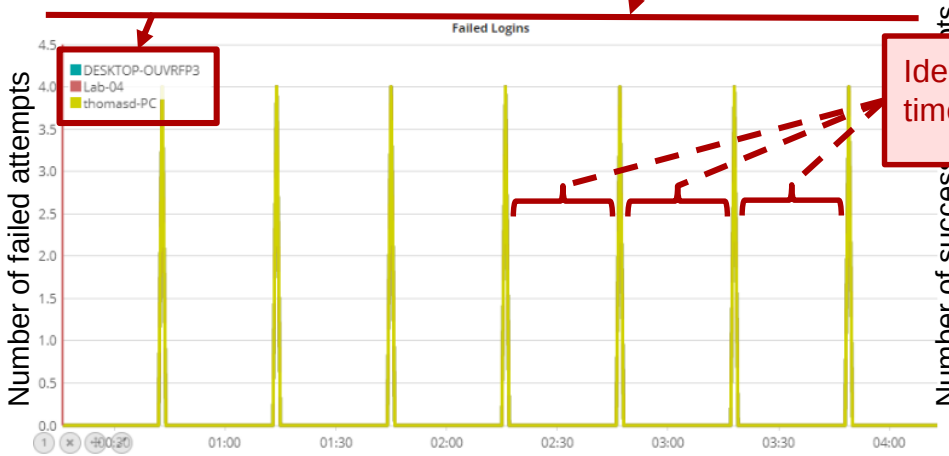
CONEXXUS
solve forward

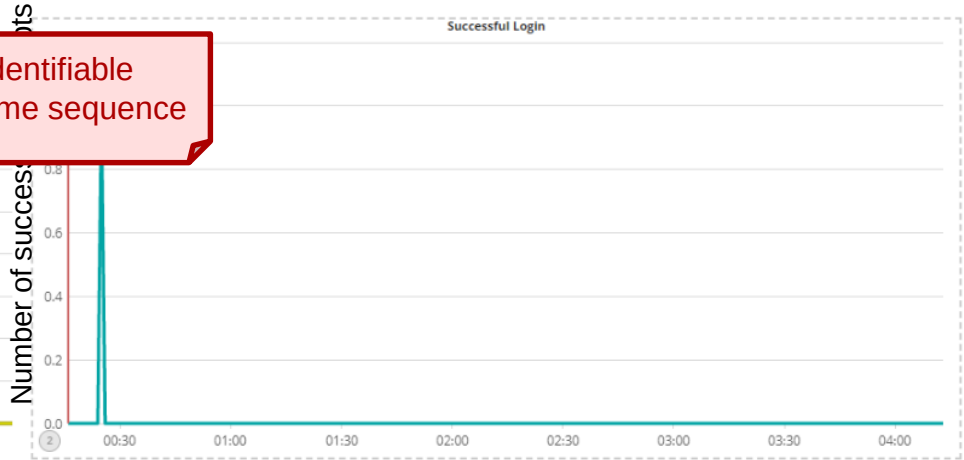
Comparing login patterns can give insight into behavior.

Multiple machines being targeted

User Access Threshold
Amount of failed logins before an alert is triggered

Identifiable time sequence

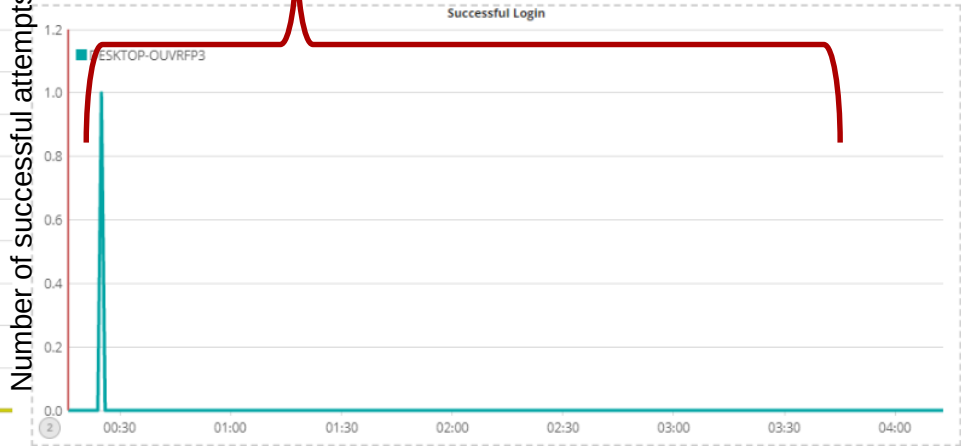




CONEXXUS
solve forward

Comparing login patterns can give insight into behavior.

No correlation between failed and successful logins



CONEXXUS
solve forward

Log correlation helps visualize malicious behavior.

BEFORE

- Common thresholds
- Isolated alerts
- No visualization

AFTER

- Behavioral insights
- Aggregated alerts
- Visualization

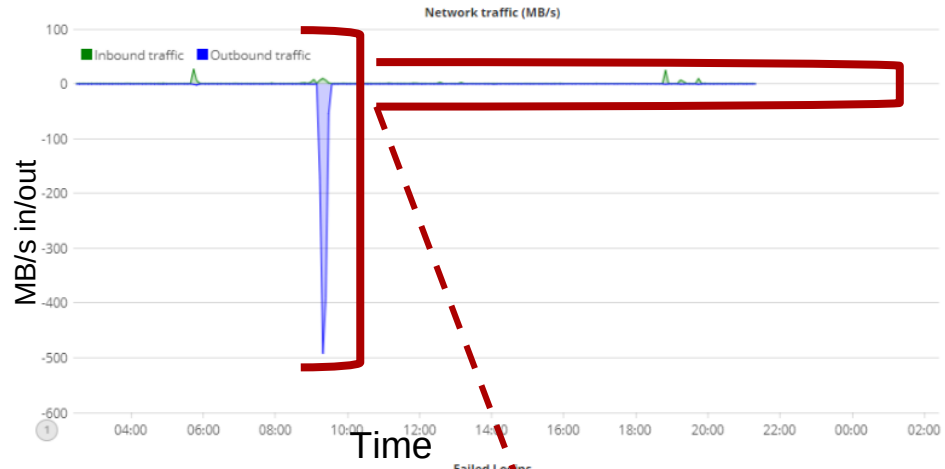
Demo - Correlating events from multiple data sets helps identify and confirm anomalies in file integrity monitoring.

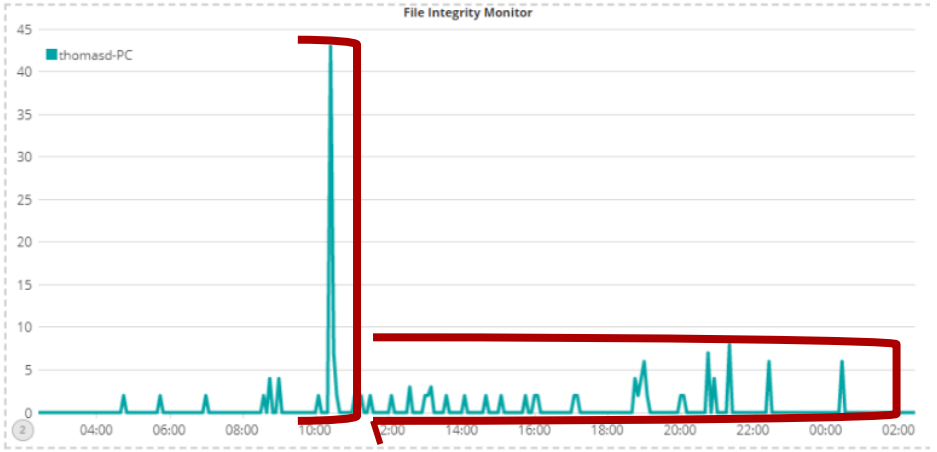
CONEXXUS
solve forward

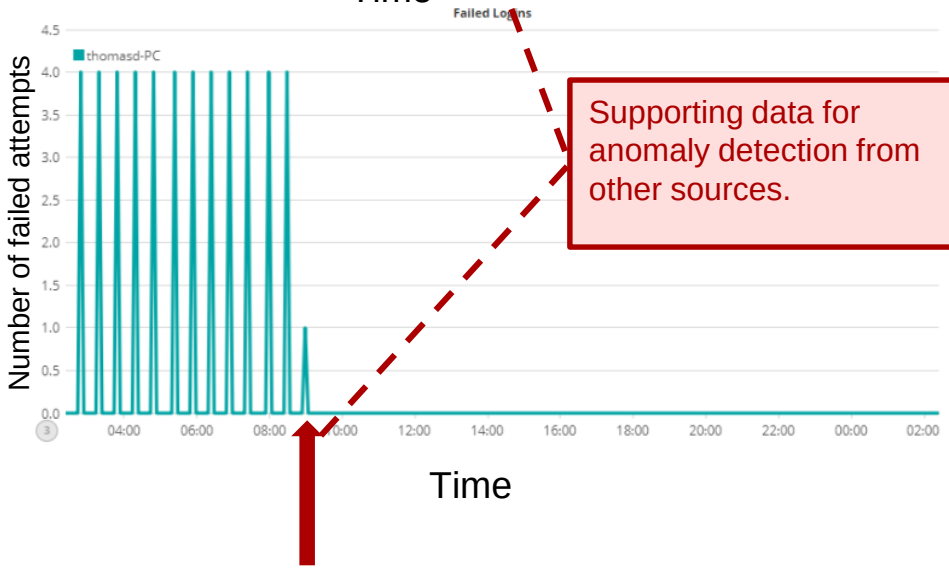
Change in pattern correlates with compromise of system.

CONEXXUS
solve forward

Not all FIM is created equal.





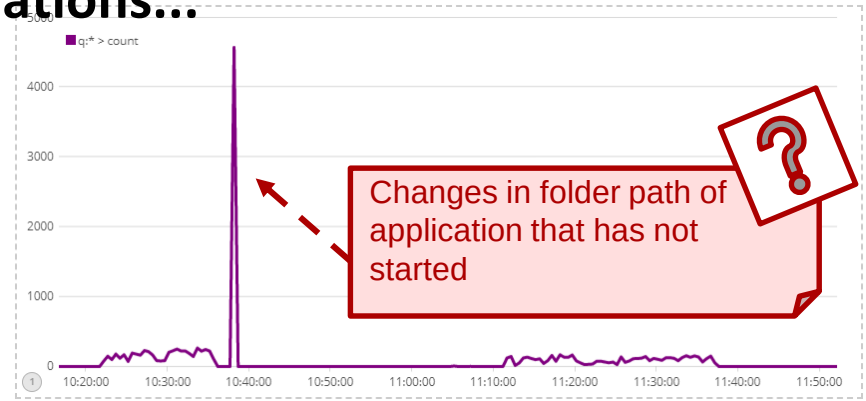


Supporting data for anomaly detection from other sources.

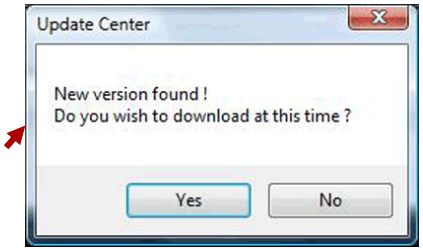
File change anomaly

CONEXXUS
solve forward

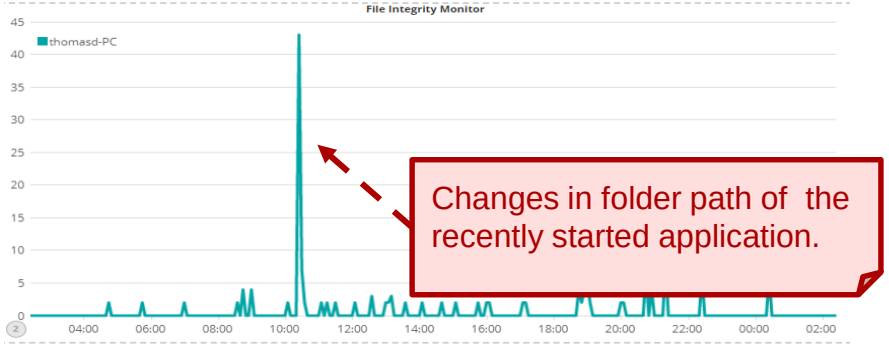
Other FIM considerations...



vs.



Logging the start of an application



CONEXXUS
solve forward

Without proper correlation to filter false positives or accurately assign priority, FIM will remain ”noisy.”

BEFORE

- No context to changes
- Frequent alerts
- First come first serve alerting

AFTER

- Better understanding of changes
- Prioritization
- Fewer false positives

PCI rules for file integrity monitoring are the absolute baseline. Weekly review of alerts might be too late!

- Fewer false positives
- Faster response time
- Greater insights into machine behavior

11.5 Deploy a change-detection mechanism (for example, file-integrity monitoring tools) to alert personnel to unauthorized modification (including changes, additions, and deletions) of critical system files, configuration files, or content files; and configure the software to perform critical file comparisons at least weekly.

CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

Aligning security with overall analytics strategy

Easing the burden of security management

Grounding data science in reality

CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

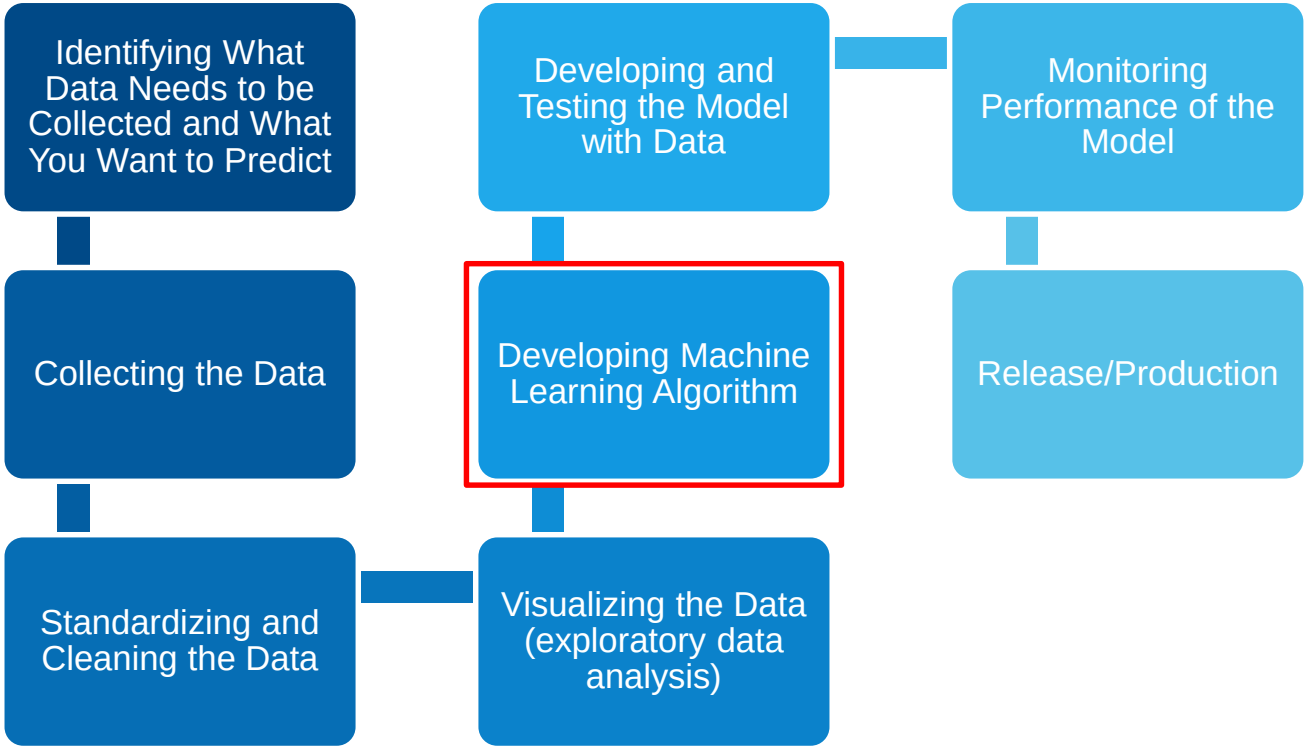
Aligning security with overall analytics strategy

Easing the burden of security management

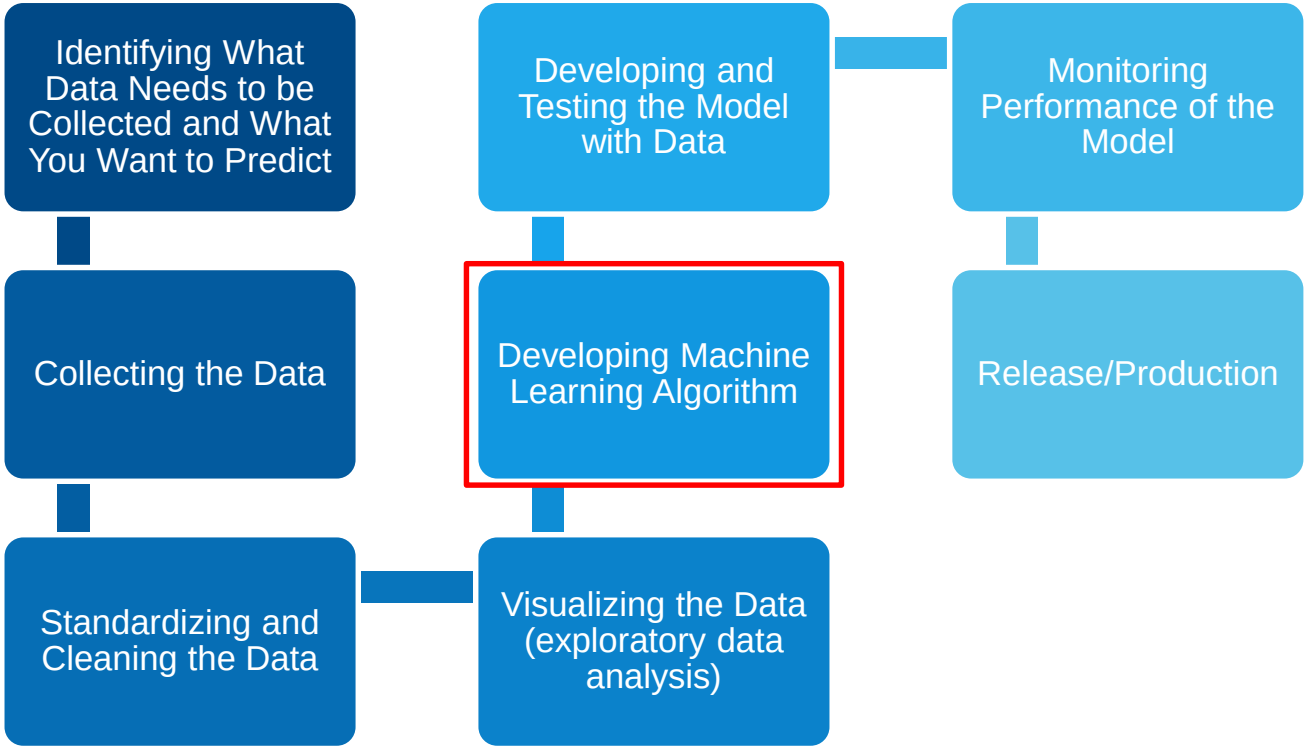
Grounding data science in reality

CONEXXUS
solve forward

Apply insights gained during exploration.



You must train before you can learn.



Applying insights from our exploration.

X = multiple login attempts over regular time intervals in 2 or more machines with a sudden stoppage in login attempts after 1 successful login

Y = specific pattern of file change activity

- A pair of X and Y, (X1,Y1), should be assigned a higher probability and fall into the queue as a higher priority alert. This ensures that more likely signs of malicious behavior are being addressed first.
- Whereas the specific pattern of file change activity may represent only a 5% chance of being a sign of malicious behavior, the conditions of X AND Y being met could represent a 33% chance of malicious activity.

Applying insights from our exploration.

X = multiple login attempts over regular time intervals in 2 or more machines with a sudden stoppage in login attempts after 1 successful login

Y = specific pattern of file change activity

- When the alert is examined, the information collected in the SIEM has already established a potential relationship between X and Y. Hence, the investigation and forensic process is dramatically shortened.
- The fingerprint is readily available to see; no searching required.

Penetration testing that truly helps you “learn”…

- With data science, penetration testing can be even more valuable to security operations by feeding real data that truly represents malicious behavior.
- A SIEM program that leverages data science does not simply close tickets. Every alert is assigned a TRUE or FALSE value, which helps to train and refine models produced by the algorithm.

Analytics tools give you access to expert communities

- Hackers share knowledge and collaborate. Security teams too often operate in isolation. Knowledge exists in silos.
- Communities formed through analytics tools provide access to libraries of pre-built models, forums, data sets, open source code, and expert advice.
- Users of log management tools often share machine learning plugins.

elastic

graylog

splunk>

sumologic

CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

Aligning security with overall analytics strategy

Easing the burden of security management

Grounding data science in reality

CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

Aligning security with overall analytics strategy

Easing the burden of security management

Grounding data science in reality

CONEXXUS
solve forward

Example: door counters – how marketing innovation can also support security efforts.

- Modern people traffic counters are able to measure inbound and outbound foot traffic, thereby providing an accurate count of overall store traffic at a given point in time.
- Many people counter models are able to export data in .csv formats for further analysis. Log analysis platforms can serve as an ideal tool for collecting and analyzing this data – with the potential to implement alerting.
- While data from people/traffic counters is typically used to measure the success of marketing efforts or the conversion rate of foot traffic, it can also be used for security information and event management.

Example: door counters – how marketing innovation can also support security efforts.

Time	Transaction ID	Transaction Amount
15:05:00 PM	120	$10.50
15:06:30 PM	125	$6.85
16:07:55 PM	279	$22.10
17:03:00 PM	400	$31.10

Example: door counters – how marketing innovation can also support security efforts.

Time Period	Total Foot Traffic
3:00 – 3:59 PM	123
4:00 – 4:59 PM	59
5:00 – 5:59 PM	28
6:00 – 6:59 PM	93

Example: door counters – how marketing innovation can also support security efforts.

Time Period	Total Foot Traffic	Transaction Volume
3:00 – 3:59 PM	123	61
4:00 – 4:59 PM	59	28
5:00 – 5:59 PM	73	37
6:00 – 6:59	81	80

Example: door counters – how marketing innovation can also support security efforts.

Time	Firewall	POS	Back Office	Foot Traffic
01:00:00	FW Event A	POS Event A	BO Event A	18
01:01:00	FW Event B	POS Event B	BO Event A	20
01:02:00	FW Event C	POS Event A	BO Event B	20
01:03:00	FW Event D	POS Event C	BO Event C	18

"Hey John, we detected an unauthorized USB drive with a malicious file at 3AM. Our data shows only six people in the store at that time. Please contact the store manager and see if you can learn anything more."

CONEXXUS
solve forward

Example: door counters – how marketing innovation can also support security efforts.

Time	Firewall	POS	Back Office	Conversion Rate
01:00:00	FW Event A	POS Event A	BO Event A	49%
01:01:00	FW Event B	POS Event B	BO Event A	63%
01:02:00	FW Event C	POS Event A	BO Event B	20%
01:03:00	FW Event D	POS Event C	BO Event C	18%

Example: door counters – how marketing innovation can also support security efforts.

Time	Firewall	POS	Back Office	Conversion Rate
01:00:00	FW Event A	POS Event A	BO Event A	49%
02:00:00	FW Event B	POS Event B	BO Event A	63%
03:00:00	FW Event C	POS Event A	BO Event B	20%
04:00:00	FW Event D	POS Event C	BO Event C	18%

Example: door counters – how marketing innovation can also support security efforts.

Time	Firewall	POS	Back Office	Conversion Rate
01:00:00	FW Event A	POS Event A	BO Event A	49%
02:00:00	FW Event B	POS Event B	BO Event A	63%
03:00:00	FW Event C	POS Event A	BO Event B	20%
04:00:00	FW Event D	POS Event C	BO Event C	18%

Outage, maintenance, security incident, etc.

Business impact.

CONEXXUS
solve forward

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

Aligning security with overall analytics strategy

Easing the burden of security management

Grounding data science in reality

CONEXXUS
solve forward

New concepts are fun to learn about, but is this really feasible?

EXPLORE Security teams can get started by simply downloading and exploring tools. Widely shared libraries and data sets can help you get started. Take small steps forward.

PARTNER Managed security service providers should increasingly bring a rigorous data science approach to their operations; always test them for their knowledge.

UNDERSTAND As the Tesla example shows, business operators need to understand high-level concepts related to attack patterns and methods of detection.

Can we leverage data science to improve retail data security?

Greater intelligence and accuracy

Aligning security with overall analytics strategy

Easing the burden of security management

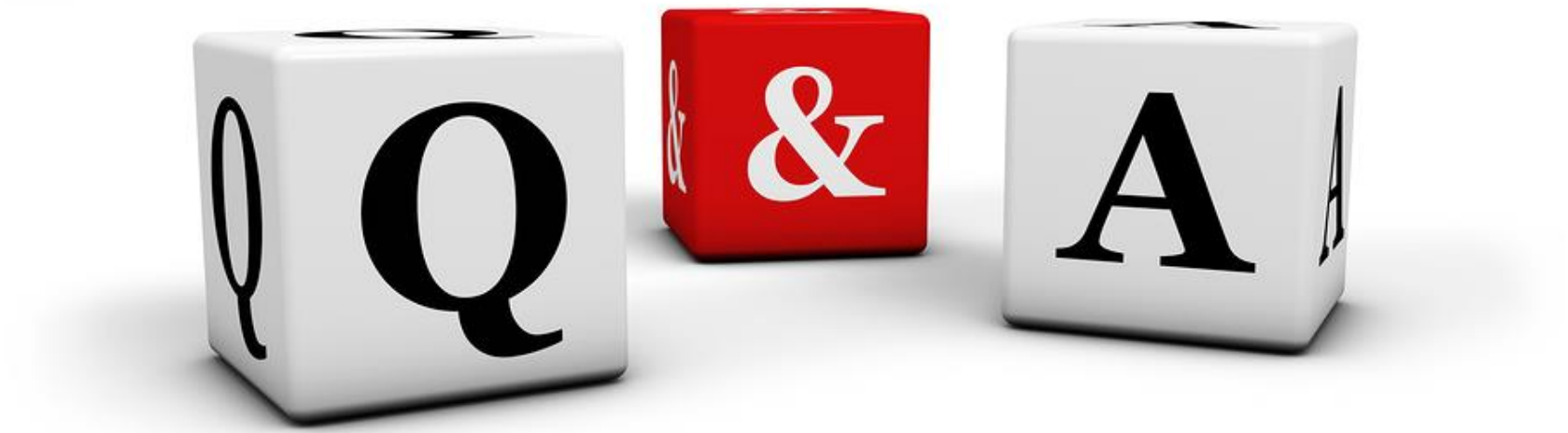
Grounding data science in reality

CONEXXUS
solve forward

Speakers

Ashwin Swamy
Director, Omega ATC
ashwin.swamy@omegaatc.com

Thomas Duncan
Security Strategist, Omega ATC
thomas.duncan@omegaatc.com



CONEXXUS
solve forward

- Website: www.conexxus.org
- Email: info@conexxus.org
- LinkedIn Profile: Conexxus.org
- Follow us on Twitter: @Conexxusonline