There's No Intelligence in AI Without Security Data Management

RSA Conference
Moscone Briefing Center

Wednesday, March 6th
11:10 – 11:30 a.m.

sas
THE POWER TO KNOW.

Scott Mongeau

Data Scientist Cybersecurity

+31 (0)68 370 3097

scott.mongeau@sas.com

Scott Allen Mongeau

SARK7

NYENRODE BUSINESS UNIVERSITEIT

sas

40+

Years of BUSINESS ANALYTICS

14,000

SAS employees worldwide

93 of the top 100 companies on the FORTUNE GLOBAL 500 LIST

World's #1 LARGEST privately held software company

80,000+

Customer sites in 148 countries

23%

Annual reinvestment in R&D

sas THE POWER TO KNOW.

Data Management Challenges = AI Challenges

Hang Seng 29022.29 0.21% ▲ U.S. 10 Yr 2/32 Yield 2.711% ▲ Crude Oil 56.12 -0.78% ▼

THE WALL STREET JOURNAL.

U.S. Edition ▾ | March 5, 2019 | Print Edition | Video

CIO JOURNAL

AI Efforts at Large Companies May Be Hindered by Poor Quality Data

Executives surveyed by PwC said cleaning up their data would lead to big cost savings and revenue gains

https://www.wsj.com/articles/ai-efforts-at-large-companies-may-be-hindered-by-poor-quality-data-11551741634

WIRED

Gear | Business | Politics

No AI until the data is fixed

Artificial Intelligence is changing business forever. But is the data powering it good enough?

22 Feb 2019

—

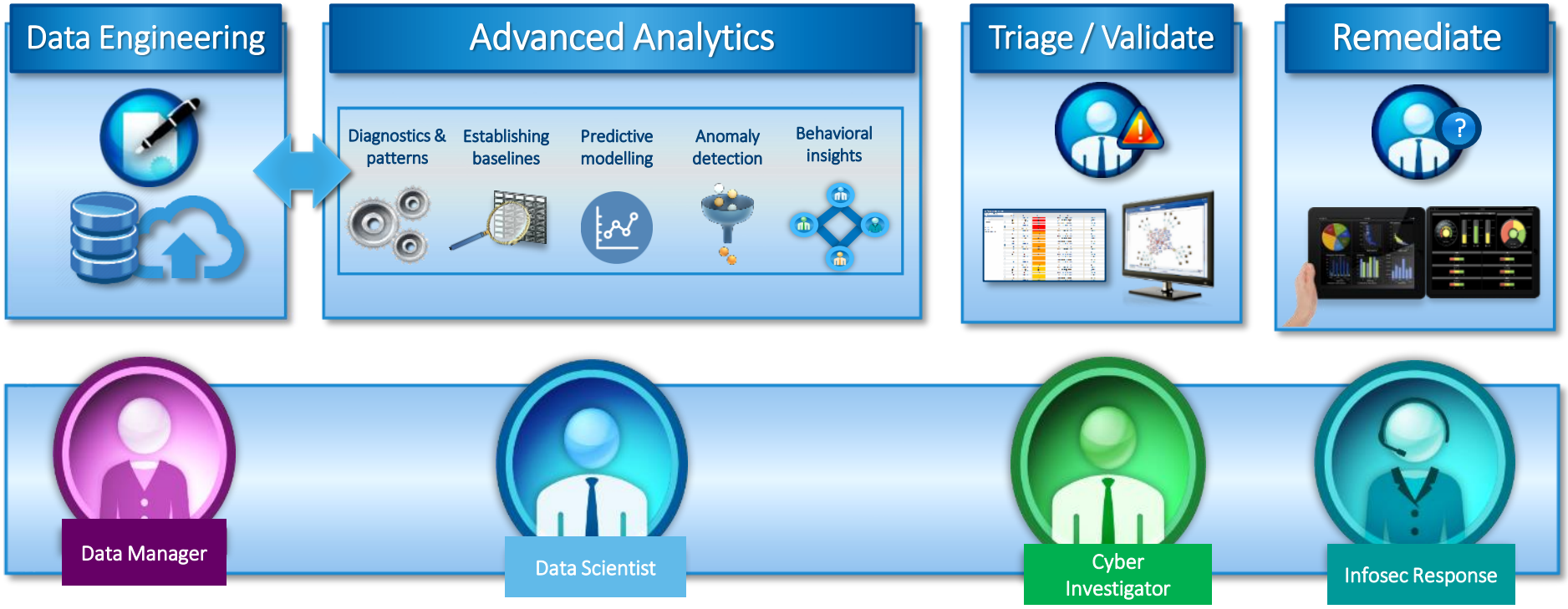
In partnership with **ACCENTURE**

https://www.wired.co.uk/article/no-ai-until-the-data-is-fixed

sas

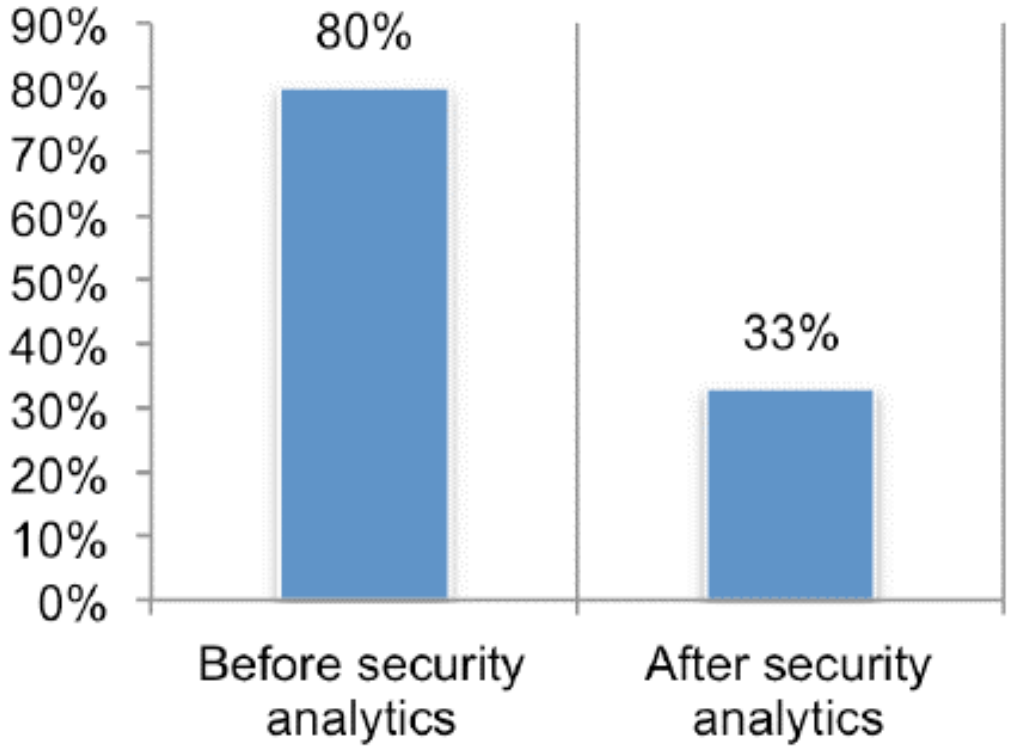
Why AI for Cyber?

Cybersecurity AI as-a-process



https://www.sas.com/en_us/whitepapers/ponemon-how-security-analytics-improves-cybersecurity-defenses-108679.html

Level of difficulty in reducing false alerts



Survey of 621 global IT security practitioners

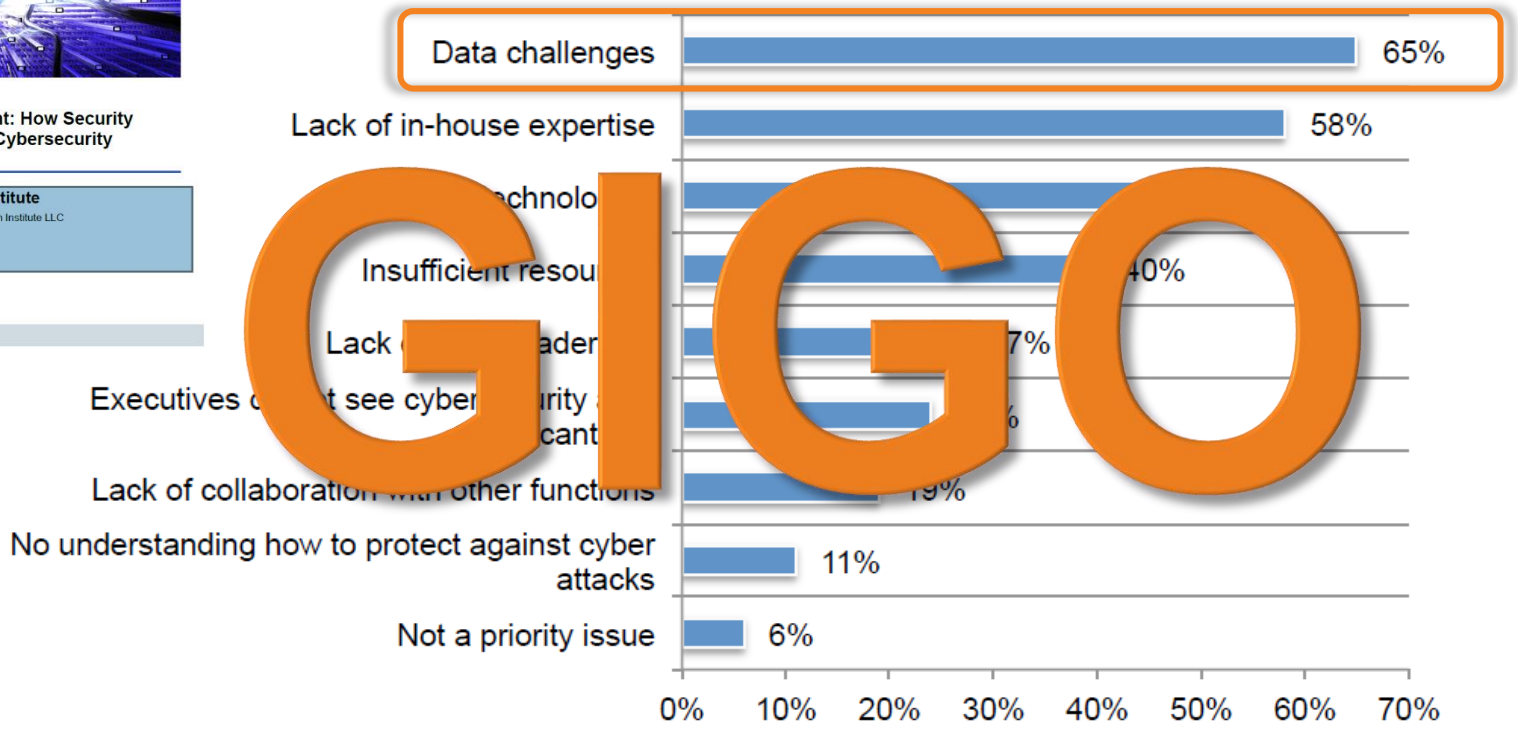
Sounds great!
What's the problem?

Challenges preventing successful use of cybersecurity analytics*

When Seconds Count: How Security Analytics Improves Cybersecurity Defenses

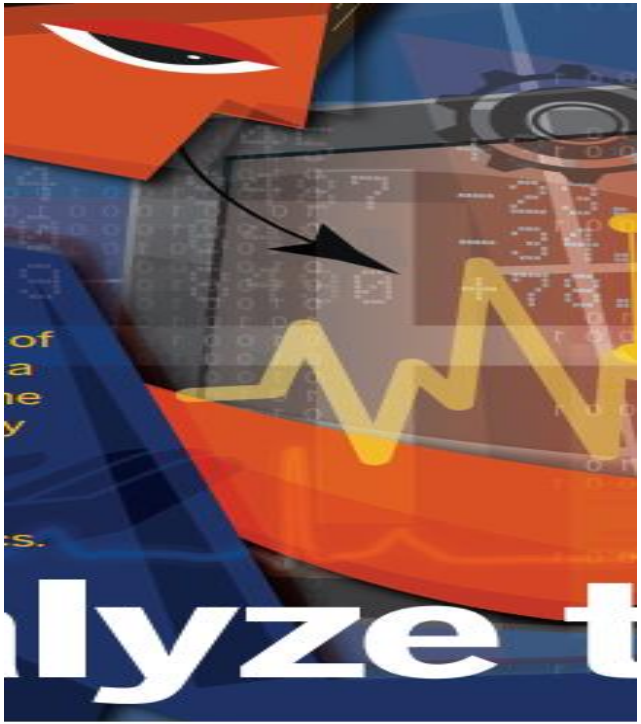
Sponsored by SAS Institute
Independently conducted by Ponemon Institute LLC
Publication Date: January 2017

Ponemon Institute® Research Report

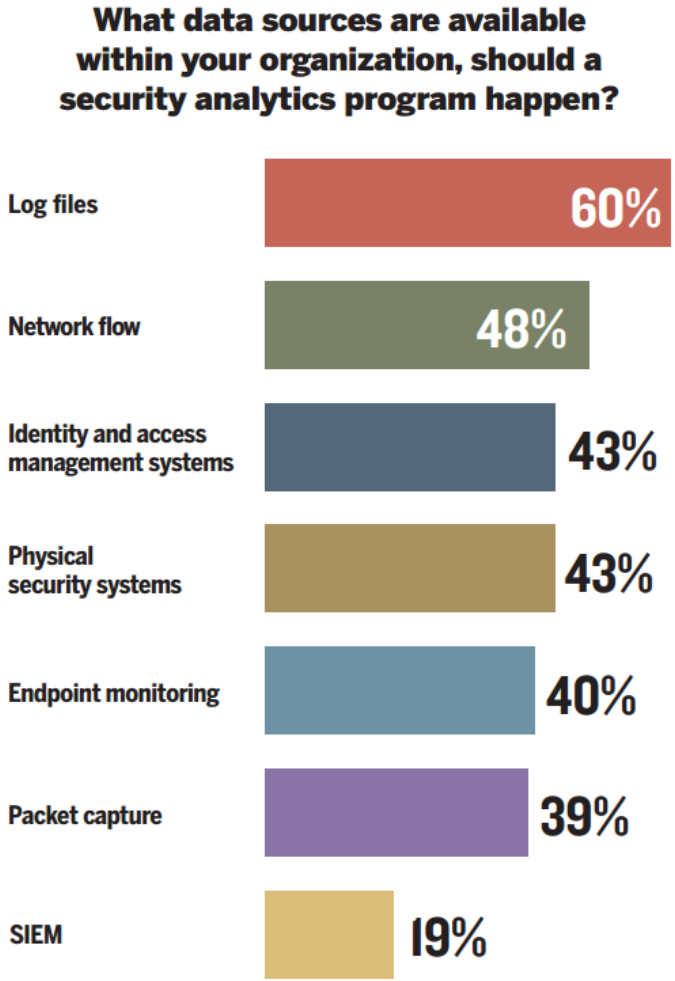


https://www.sas.com/en_us/whitepapers/ponemon-how-security-analytics-improves-cybersecurity-defenses-108679.html

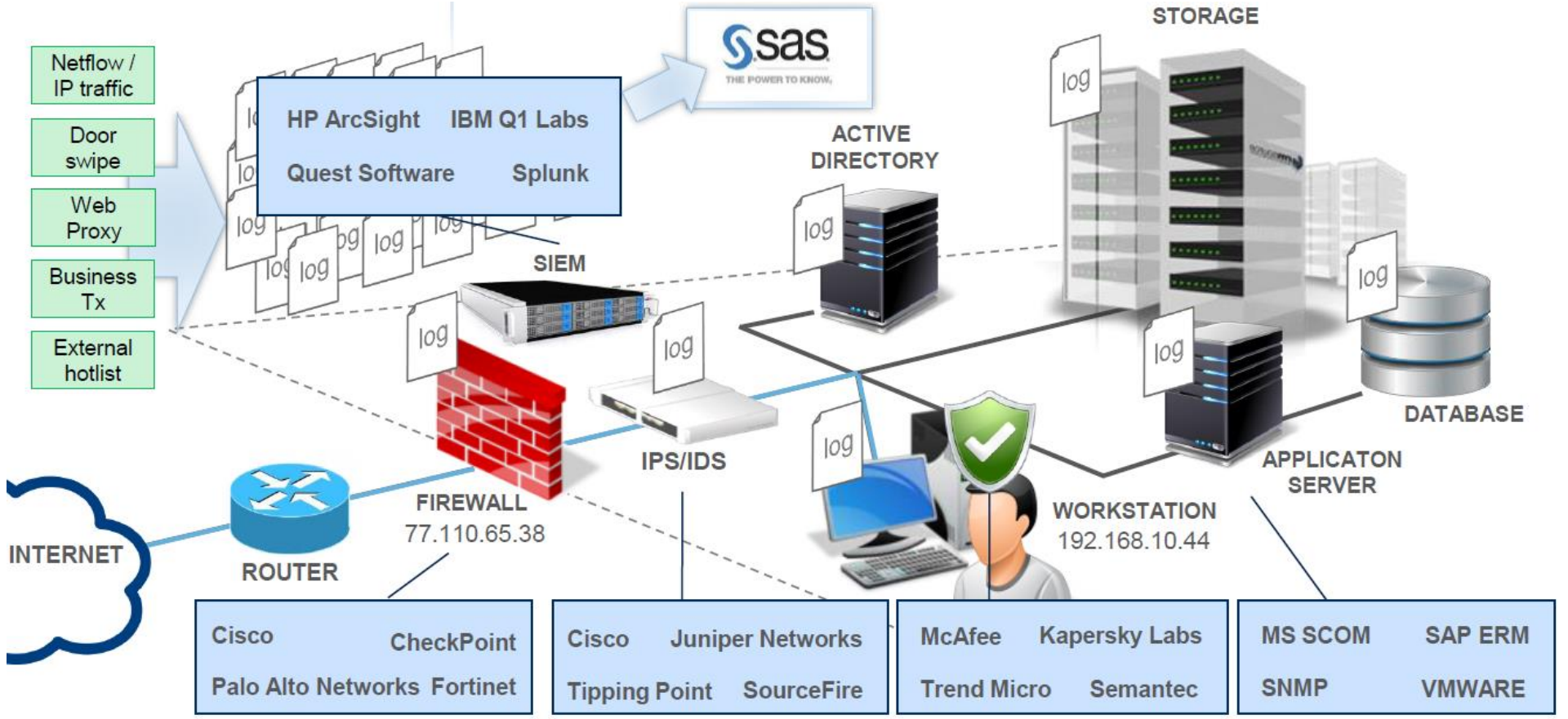
** Survey of 621 global IT security practitioners*



SOURCE
Security Brief Magazine. (2016). "Analyze This! Who's Implementing Security Analytics Now?" Available at https://www.sas.com/en_th/whitepapers/analyze-this-108217.html



Many data sources… increasing data volumes



High false alerts… slow investigation processes

sas

Data, Data Everywhere…

Data Challenges

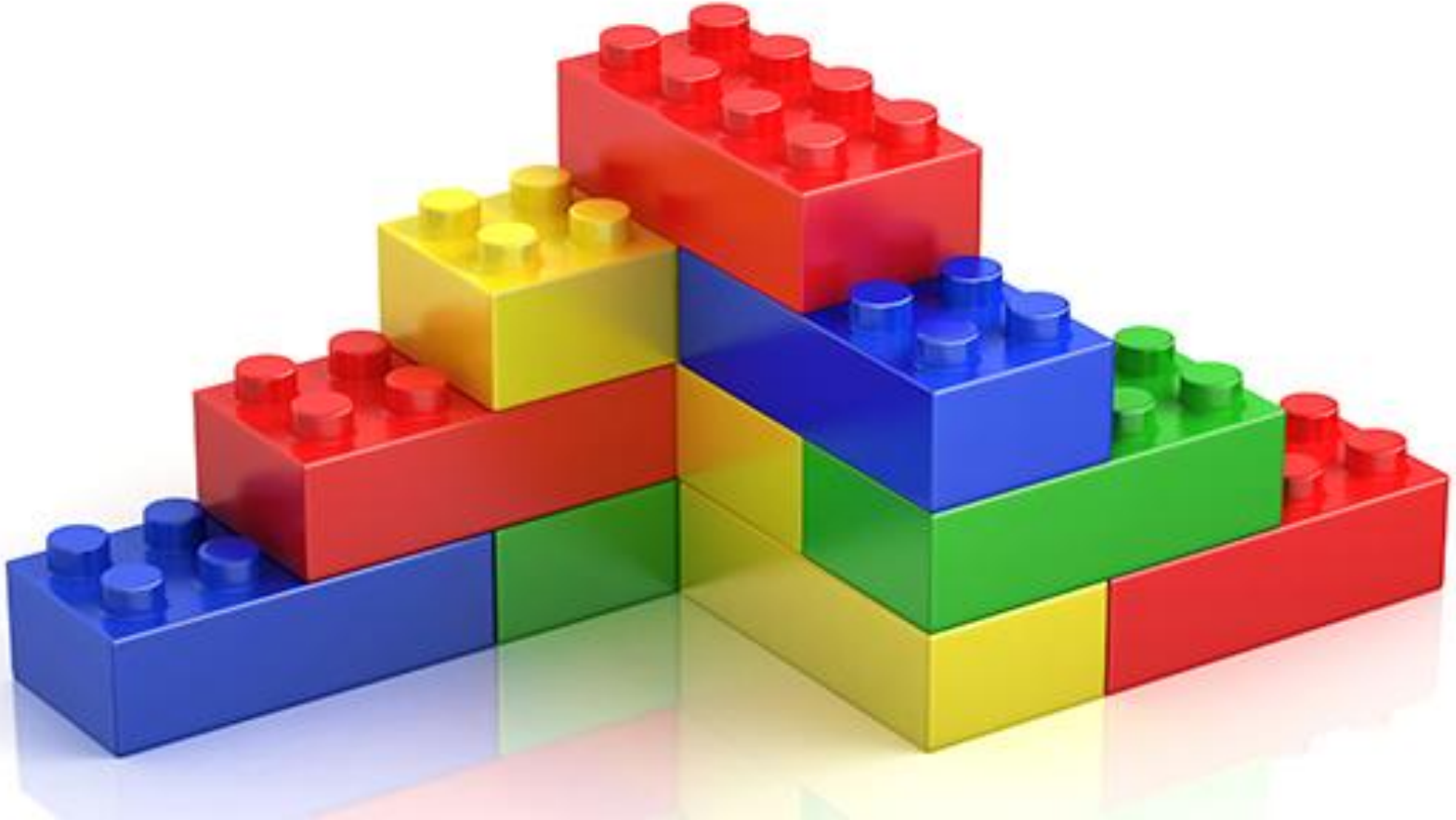
IP address

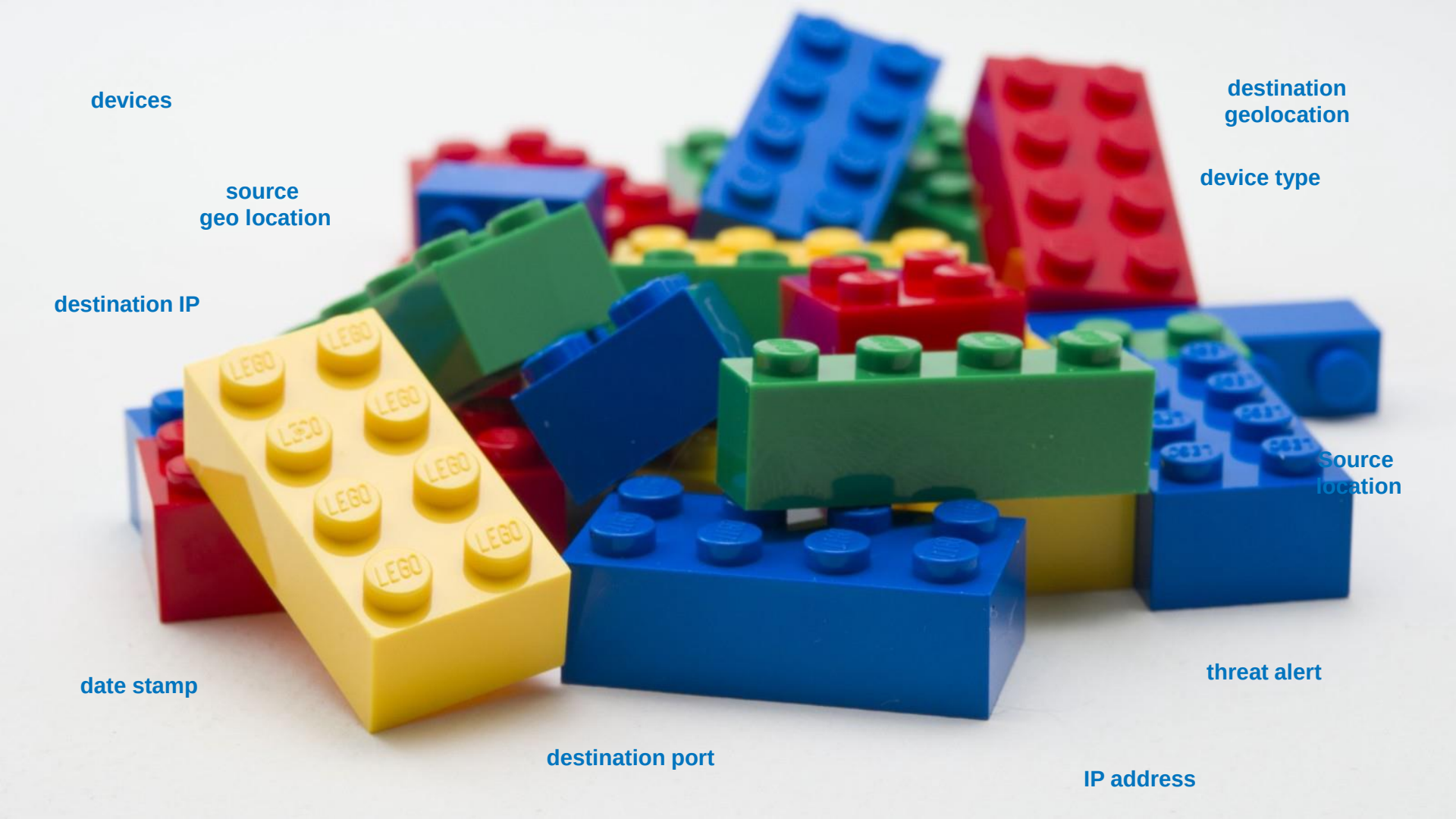
time stamp

userid

destination port

log file




devices
source
geo location
destination IP
date stamp
destination port
IP address
threat alert
Source
location
device type
destination
geolocation

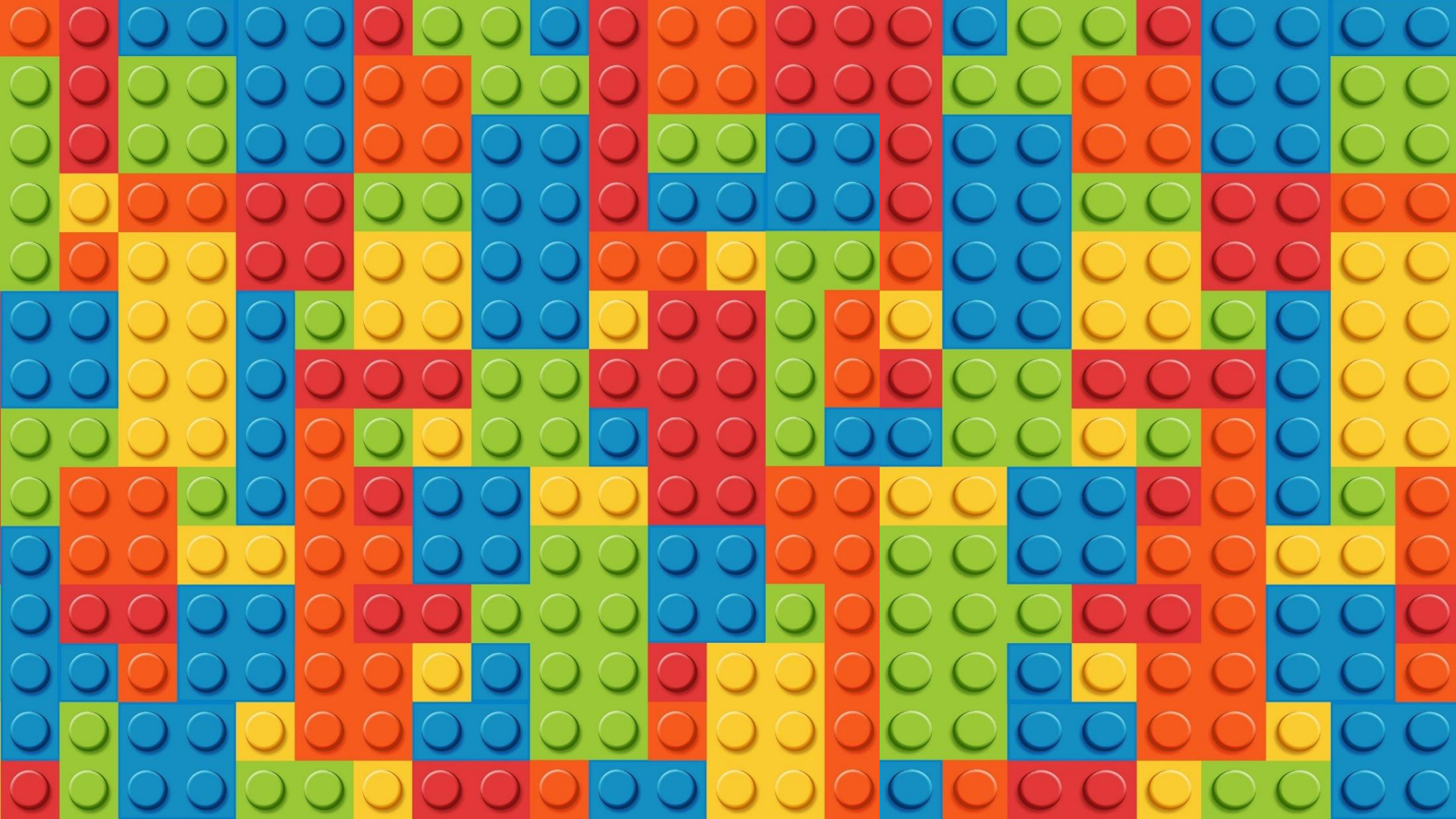
Time Series

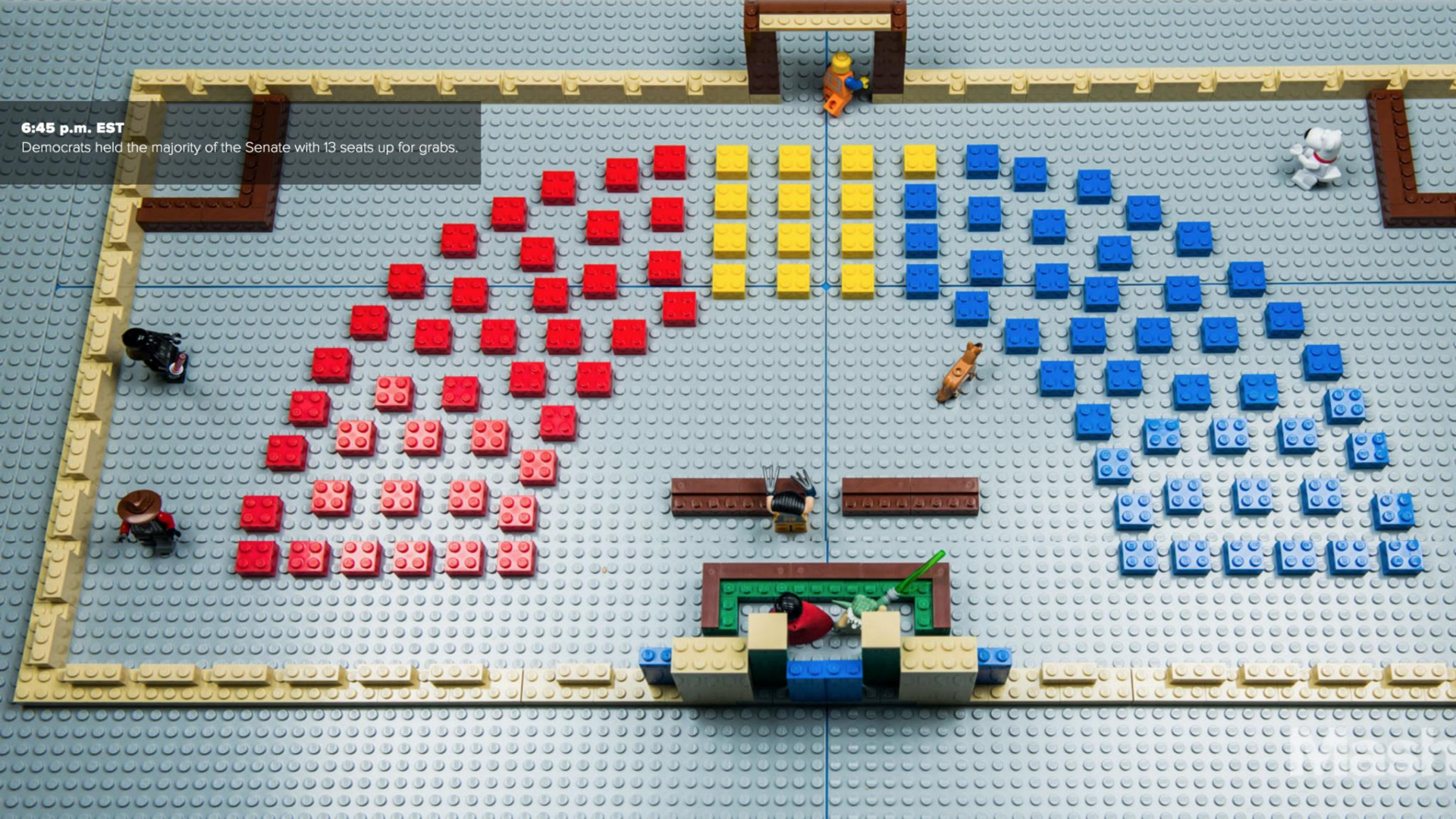
Irregular and Complex Events

WANTED!
Data, Every Which Way

Data Delivery


behavioral
profile




6:45 p.m. EST
Democrats held the majority of the Senate with 13 seats up for grabs.

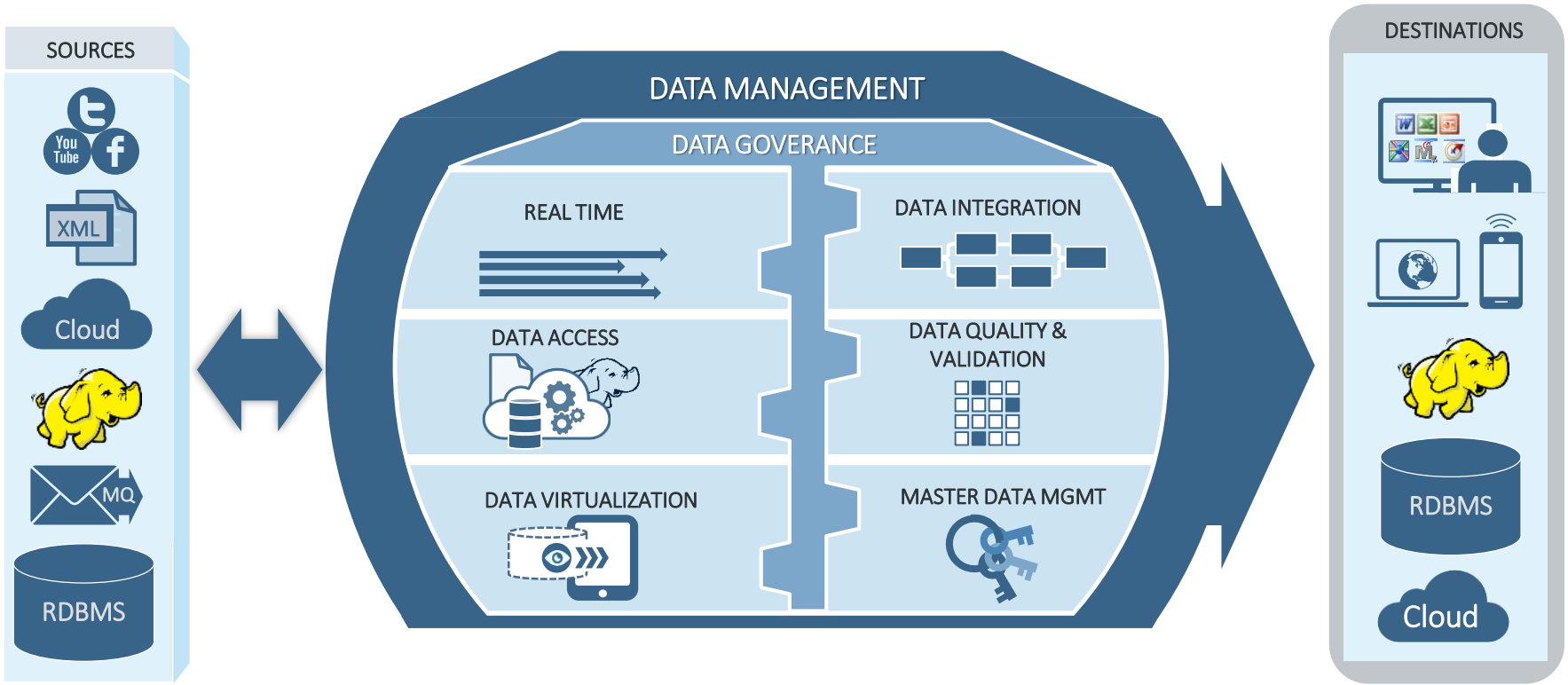
AI Models = Active Data Vehicles

Data… Delivered

Data Engineering


POLICE

Data Ops: Fusion, Quality and Delivery



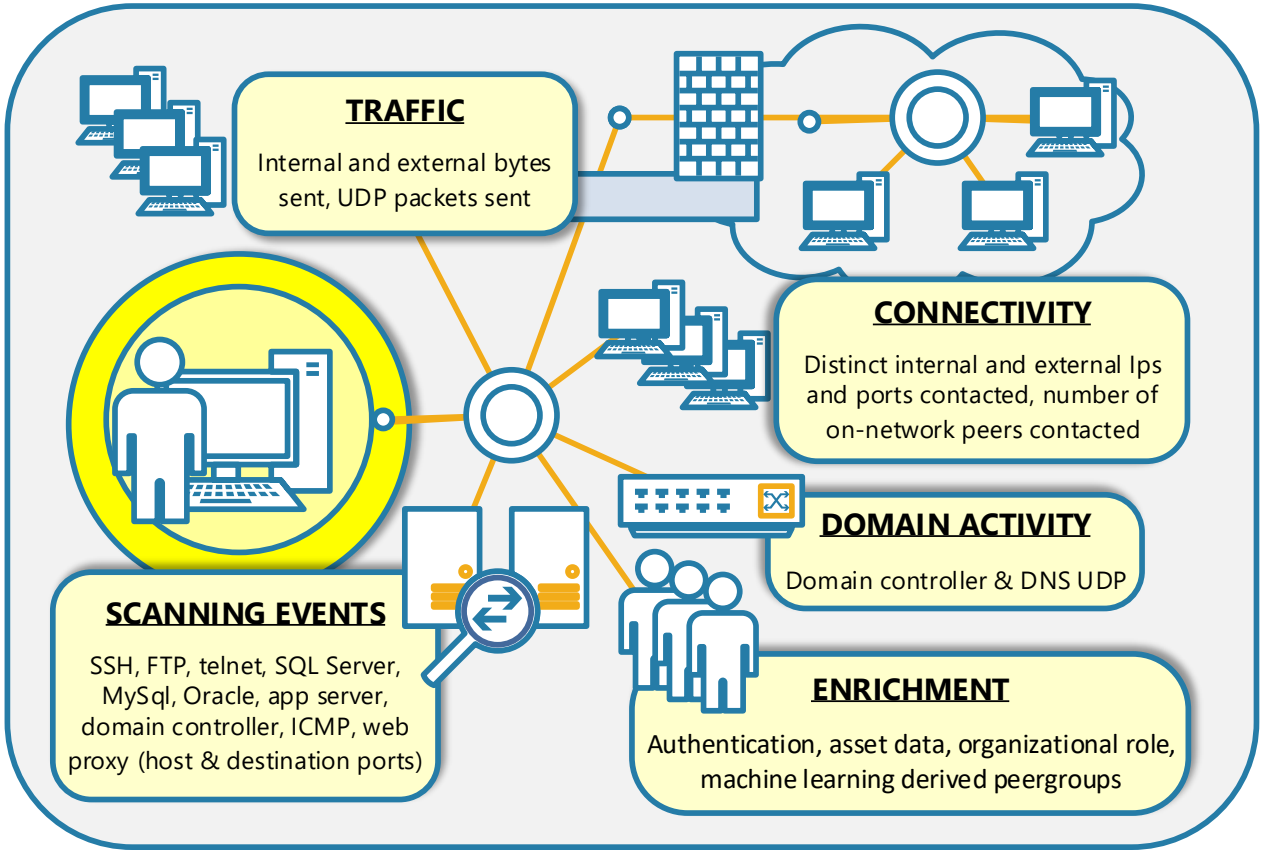
"Organizing data is a critical first step in figuring out what data means"

Larry Alton, Information Management Feb 14th, 2019

Whitepaper: A Comprehensive Approach to Big Data Governance, Data Management and Analytics

http://sww.sas.com/cosmos/a/cosmos-images/107968_0718.pdf

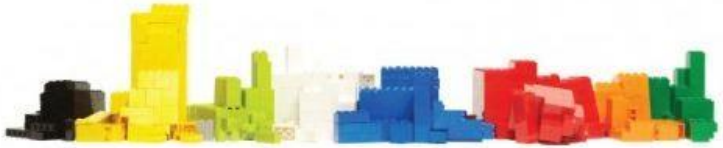
Stitching Together Cyber Events



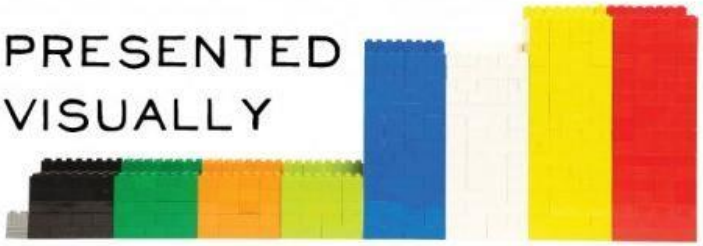
DATA

SORTED

ARRANGED



PRESENTED VISUALLY



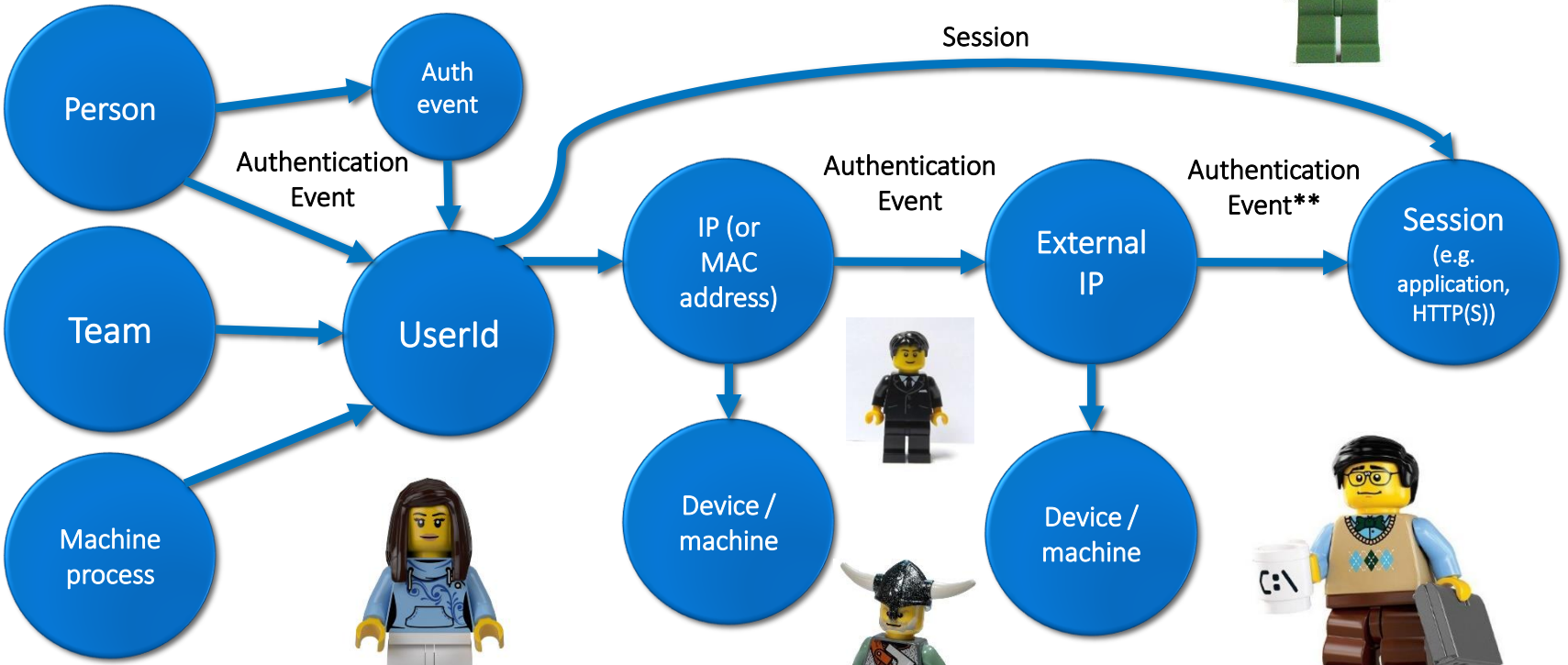
Self-Service Visual Analytics

Data Quality and Context

Blending, Cleansing, Shaping

[Feature Engineering]

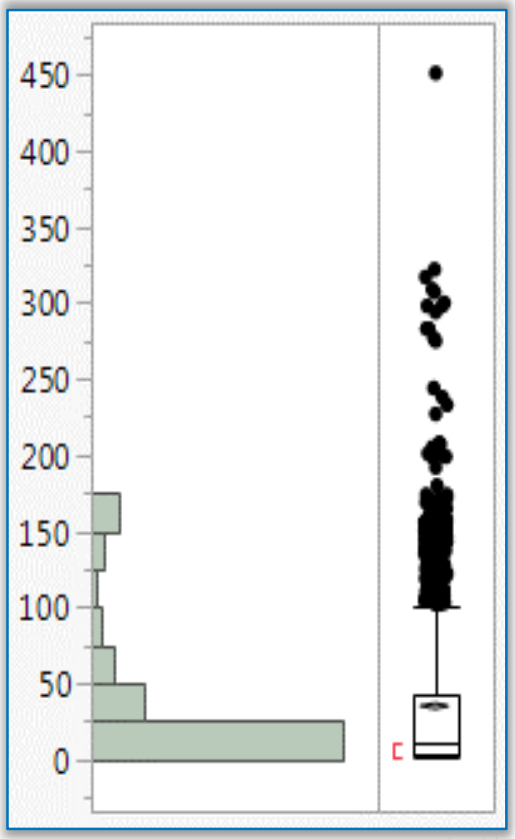
What is a User, Anyway?



sas

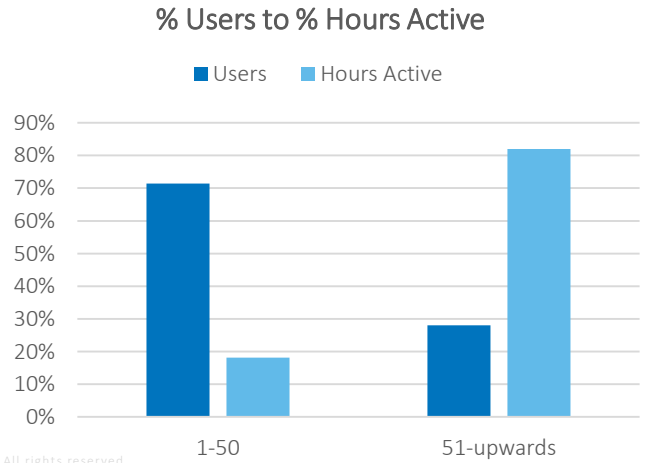
Feature Selection / Extraction

Understanding Network Behavioral Patterns



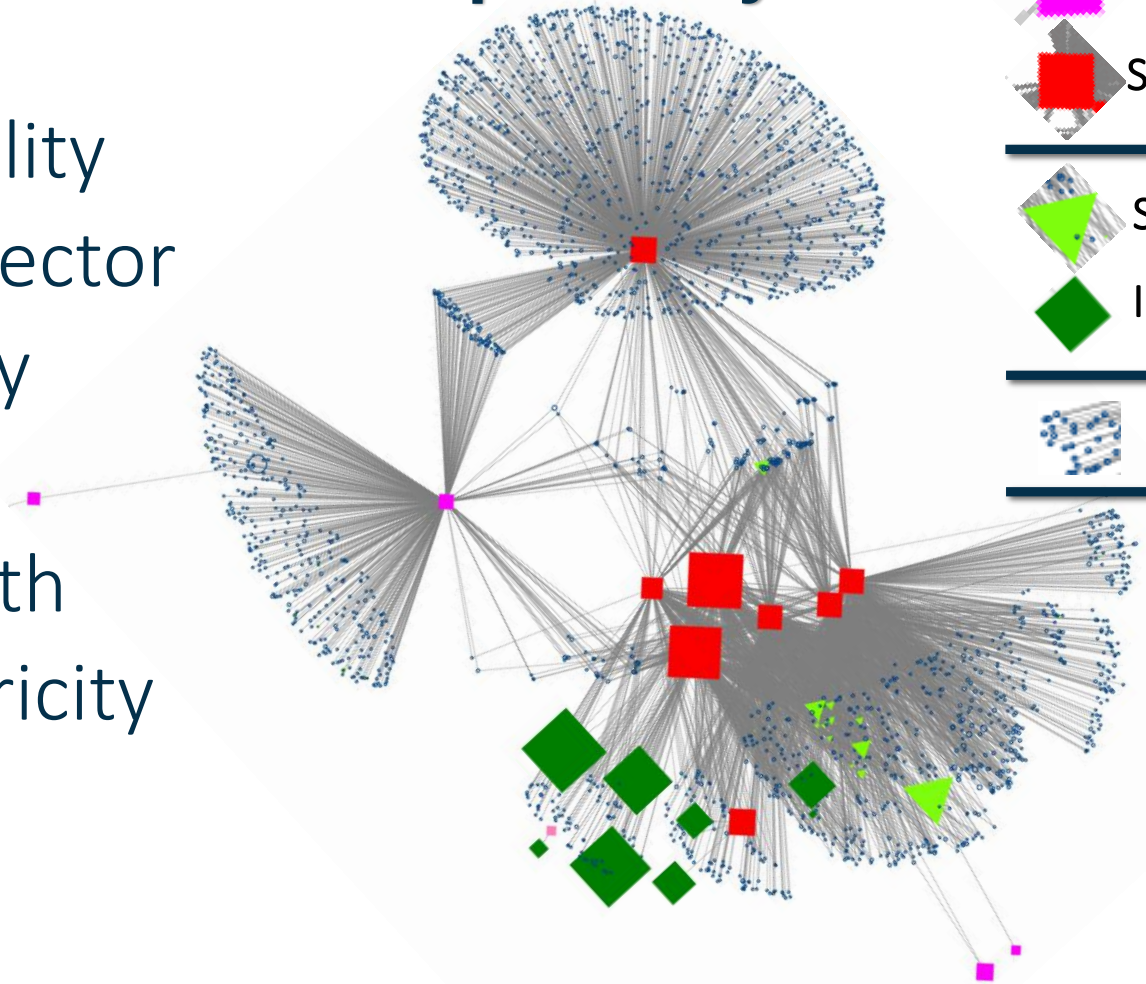
Pareto Principle

- **80/20%** pattern in network-usage
- *Outliers*: multiple devices 24 hours online
- High correlation: hrs online and breadth of activities
- Pattern observed across multiple networks



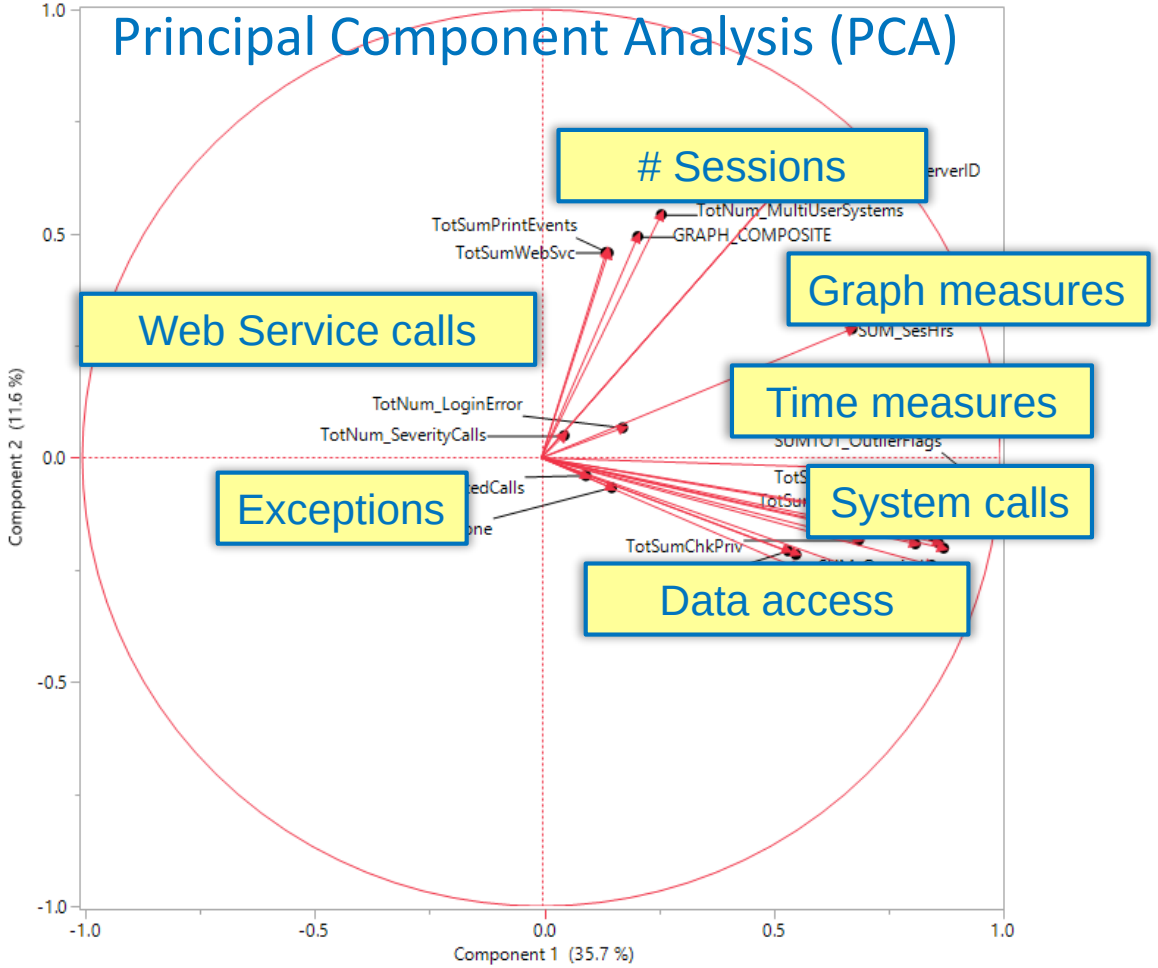
Network Graph Analytics

- Centrality
- Eigenvector
- Density
- Reach
- Strength
- Recopricity



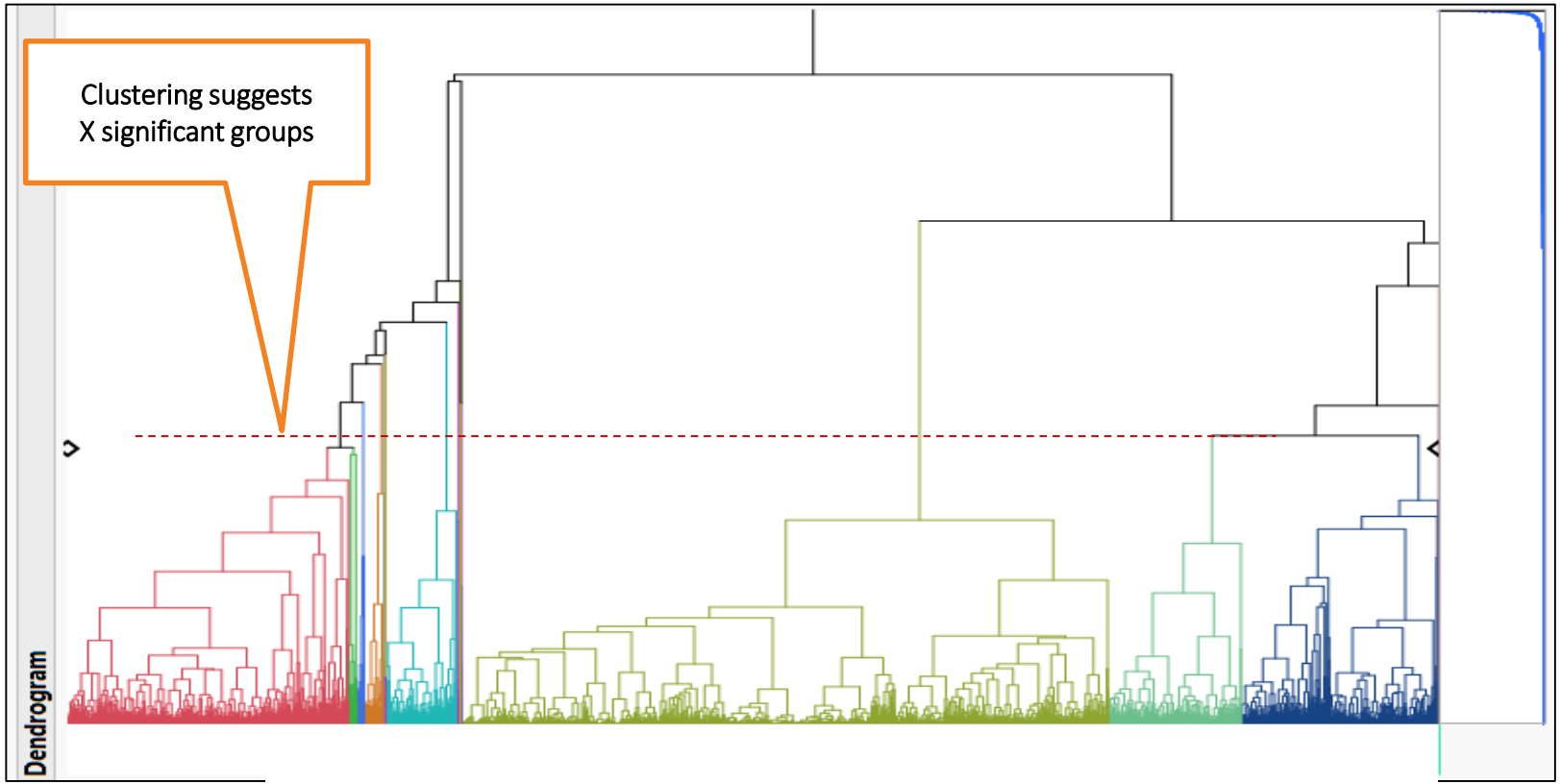
System Manager
System Host
System-User
Interface-User
Human Users

Dimensionality Reduction



Unsupervised Machine Learning: Cluster Analysis

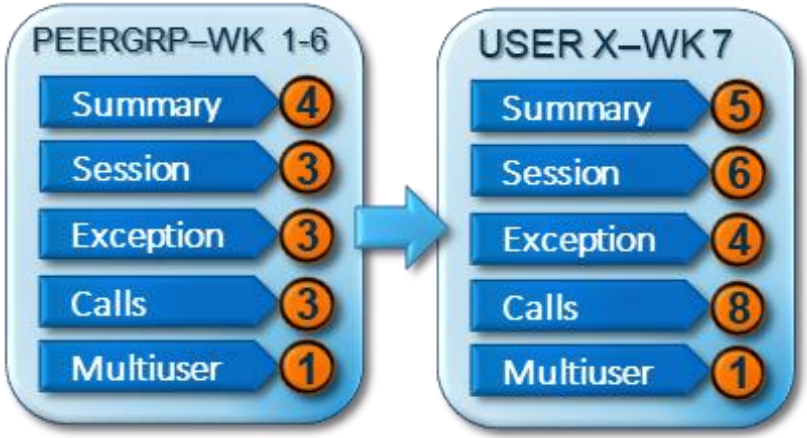
Extracting Statistically Self-Similar Groups



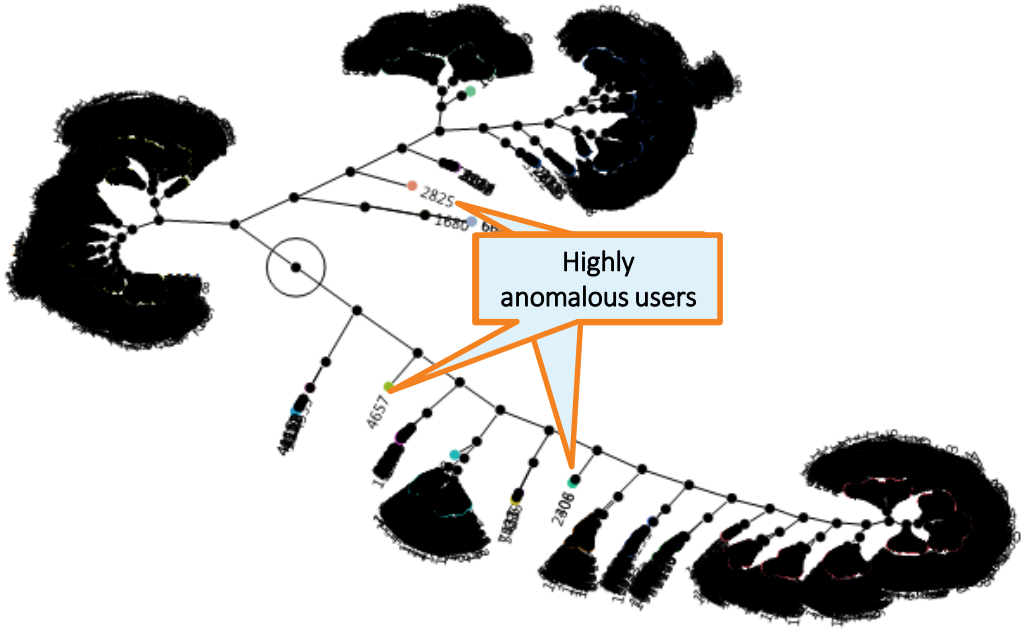
Unsupervised Machine Learning: Cluster Analysis

Statistical Baselining for ‘Normal’ versus ‘Abnormal’

USER DEVIATION FROM PEERGROUP



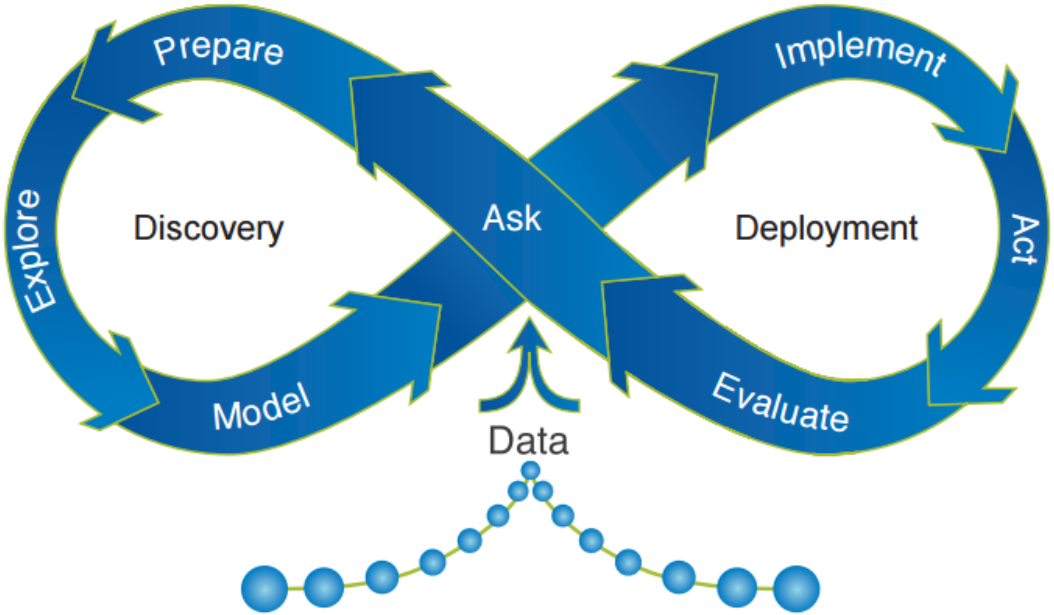
USER MULTIVARIATE ANOMALIES



Analytics Lifecycle

Raw Data — Feature Selection — **Features** — Feature Engineering — **Modeling** — **Insights**

Exploration and Discovery



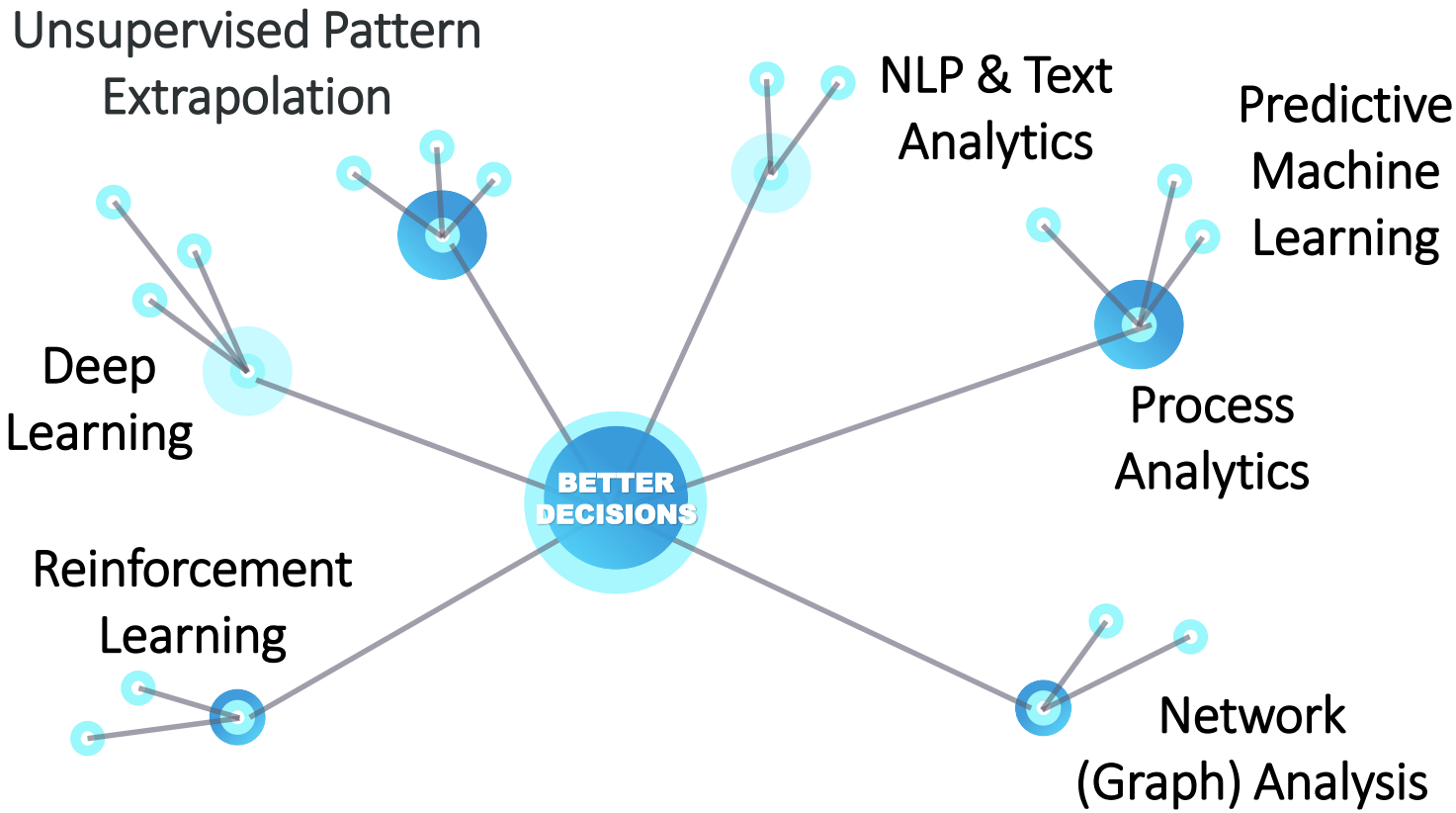
Detection and Prediction

SAS Institute. (2016). "Managing the Analytical Life Cycle for Decisions at Scale." Available at https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/manage-analytical-life-cycle-continuous-innovation-106179.pdf

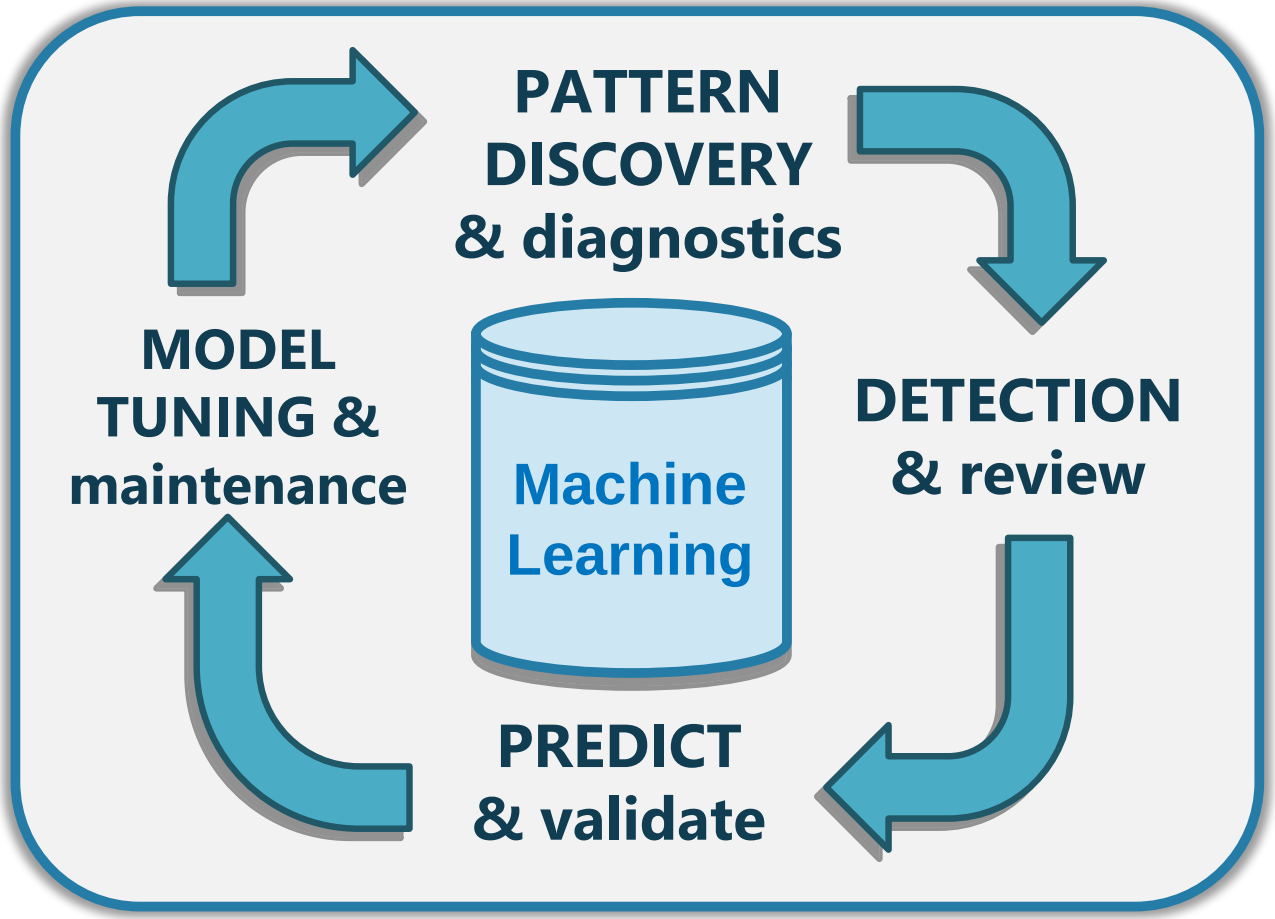
Towards Data Driven AI

Data Drives the Analytics Lifecycle

Advanced Analytics Toolkit



AI for Cybersecurity



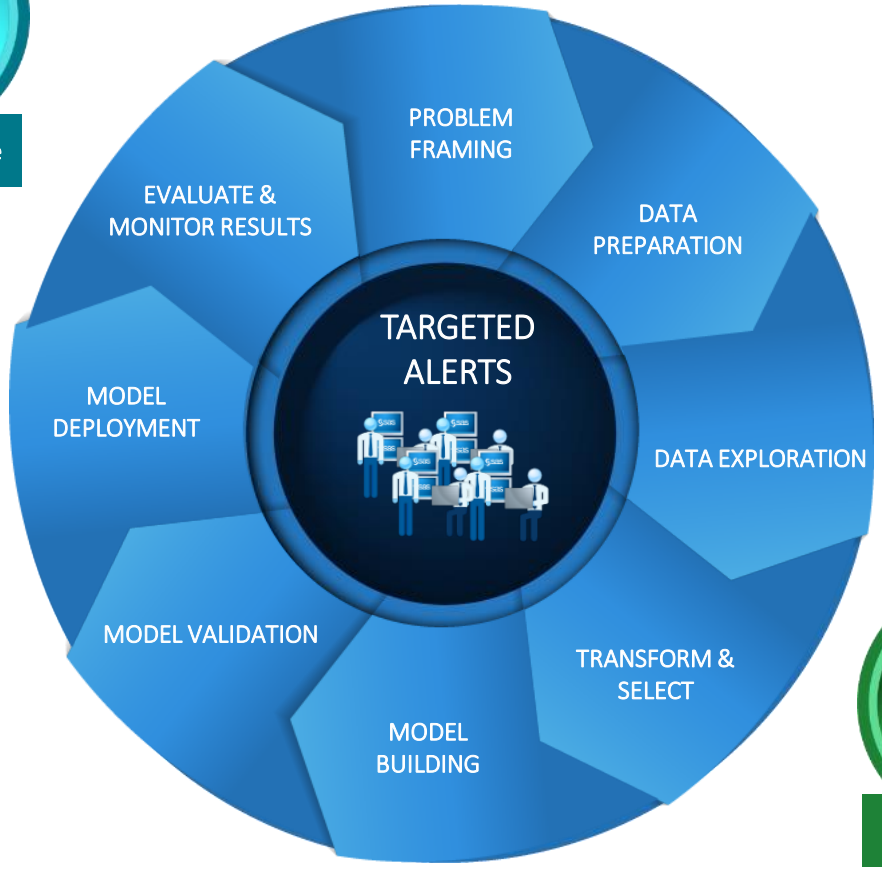
Cyber Analytics Life Cycle

InfoSec Response

Data Manager

Pattern Detection

Exploration and Insights



Data Scientist

Cyber Investigator

SOURCE SAS Institute. (2016). "Managing the Analytical Life Cycle for Decisions at Scale." Available at https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/manage-analytical-life-cycle-continuous-innovation-106179.pdf

sas

Where do we go from here?

Data Management for Cybersecurity AI

Facilitating fast and big data **ingestion** in **real time** or **batch**

Integrate, aggregate & enrich siloed, unstructured & streaming data

Advanced analytics and **machine learning** to extract patterns, predict, and optimize

Deliver focused alerts to dashboards, tickets, SIEM, and/or case management

Want to Know More?

SAS whitepaper '*Data Management for Artificial Intelligence*'

www.sas.com/en_us/whitepapers/data-management-artificial-intelligence-109860.html

SAS Cybersecurity Solution (SCS)

www.sas.com/en_us/software/cybersecurity.html

Scott Allen Mongeau
Data Scientist - Cybersecurity

scott.mongeau@sas.com

Scott Mongeau

REFERENCES

References

- Aggarwal, C. (2013). “Outlier Analysis.” Springer. http://www.springer.com/la/book/9781461463955
- Kirchhoff, C., Upton, D., and Winnefeld, Jr., Admiral J. A. (2015 October 7). “Defending Your Networks: Lessons from the Pentagon.” Harvard Business Review. Available at https://www.sas.com/en_us/whitepapers/hbr-defending-your-networks-108030.html
- Longitude Research. (2014). “Cyberrisk in banking.” Available at https://www.sas.com/content/dam/SAS/bp_de/doc/studie/ff-st-longitude-research-cyberrisk-in-banking-2316865.pdf
- Ponemon Institute. (2017). “When Seconds Count: How Security Analytics Improves Cybersecurity Defenses.” Available at https://www.sas.com/en_us/whitepapers/ponemon-how-security-analytics-improves-cybersecurity-defenses-108679.html
- SANS Institute. (2015). “2015 Analytics and Intelligence Survey.” Available at https://www.sas.com/en_us/whitepapers/sans-analytics-intelligence-survey-108031.html
- SANS Institute. (2016). “Using Analytics to Predict Future Attacks and Breaches.” Available at https://www.sas.com/en_us/whitepapers/sans-using-analytics-to-predict-future-attacks-breaches-108130.html
- SAS Institute. (2016). “Managing the Analytical Life Cycle for Decisions at Scale.” Available at https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/manage-analytical-life-cycle-continuous-innovation-106179.pdf
- SAS Institute. (2017). “SAS Cybersecurity: Counter cyberattacks with your information advantage.” Available at https://www.sas.com/en_us/software/fraud-security-intelligence/cybersecurity-solutions.html
- SAS Institute. (2019). “Data Management for Artificial Intelligence.” Available at www.sas.com/en_us/whitepapers/data-management-artificial-intelligence-109860.html
- Security Brief Magazine. (2016). “Analyze This! Who’s Implementing Security Analytics Now?” Available at https://www.sas.com/en_th/whitepapers/analyze-this-108217.html
- UBM. (2016). “Dark Reading: Close the Detection Deficit with Security Analytics.” Available at https://www.sas.com/en_us/whitepapers/close-detection-deficit-with-security-analytics-108280.html