exabeam

HOW MACHINE LEARNING IS CHANGING THE SOC

MAY 2019

Presented by Richard Cassidy – Sr. Director Security Strategy

#whoami – Richard Cassidy

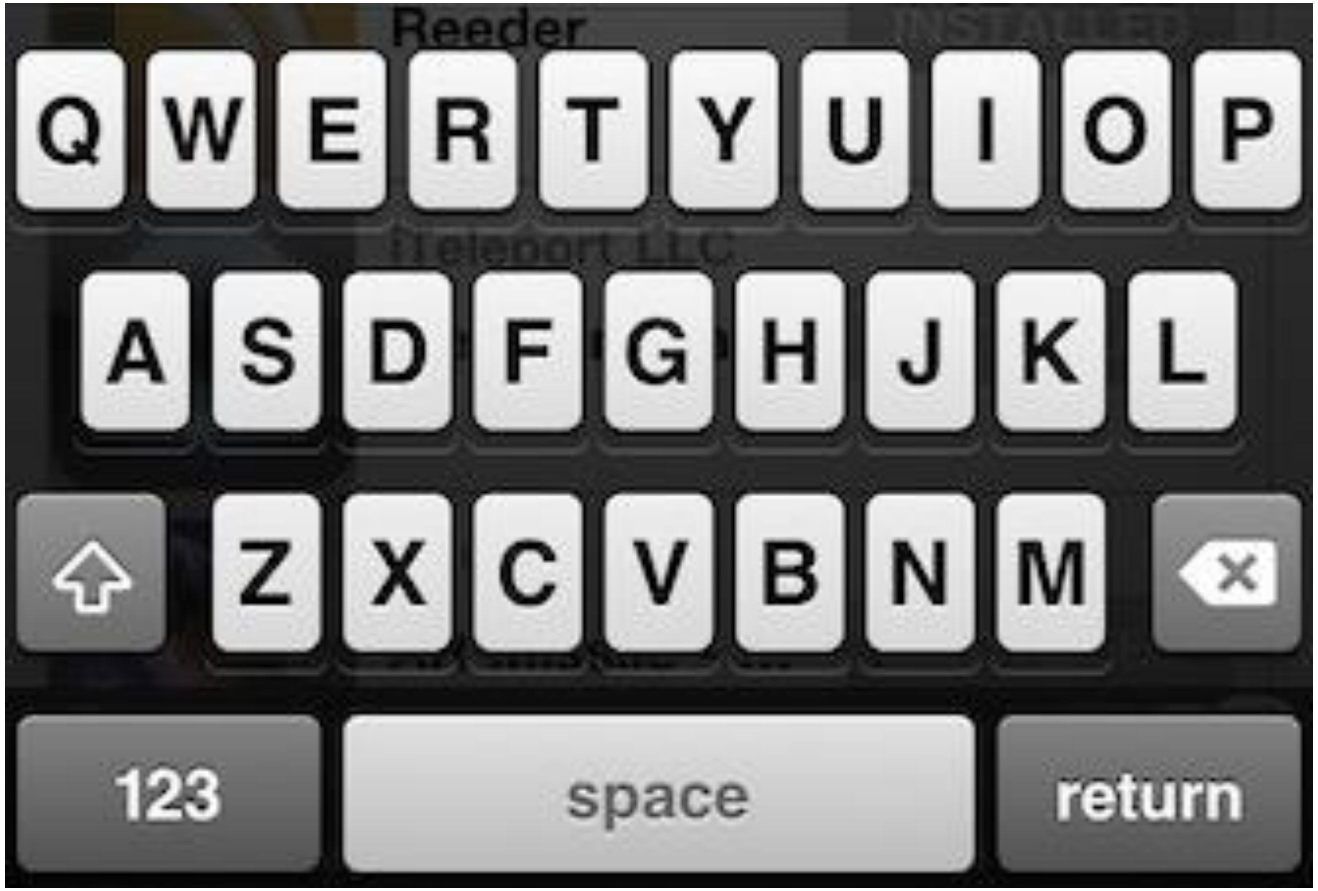
- Sr. Director Security Strategy - EMEA
- 19.25yrs Industry hands-on, in Cyber Security, Cloud & Services Technologies
  - Netscreen (Juniper), Fortinet, Virtual Computer (Citrix), Forescout, Alert Logic, Cybereason, Synack & Exabeam
- Industry Thought Leader
  - 100’s of Industry Press Contributions & Educational Columns
- Managed/Managing Global Cyber Security Projects
  - Government, Military, Finance, Manufacturing, Retail, Gaming
- SOC, Threat Intelligence & Security Services Delivery

INSERT PICTURE HERE

exabeam

MACHINE LEARNING, STORY TIME

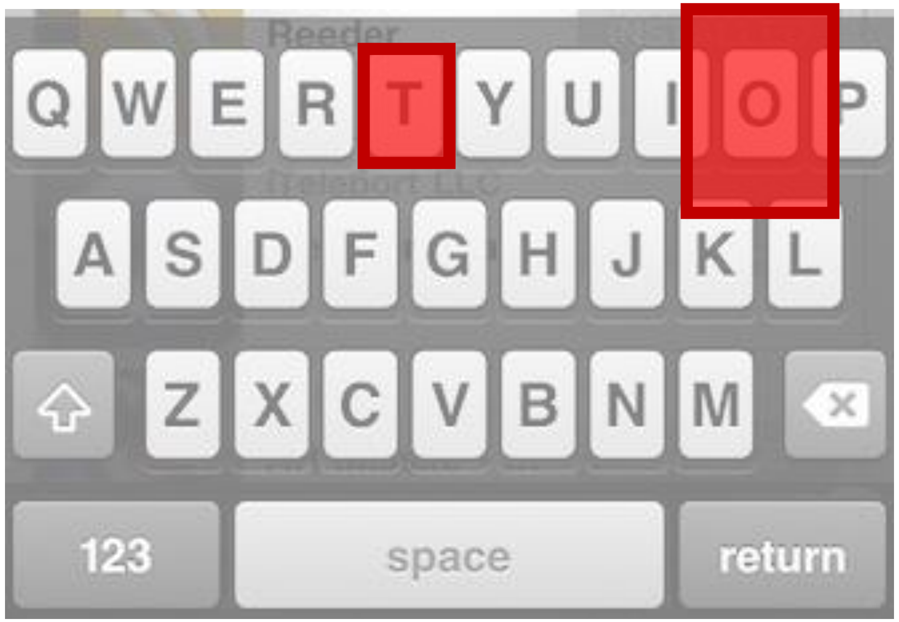
The First Touch Keyboard



The First Touch Keyboard

N-Gram

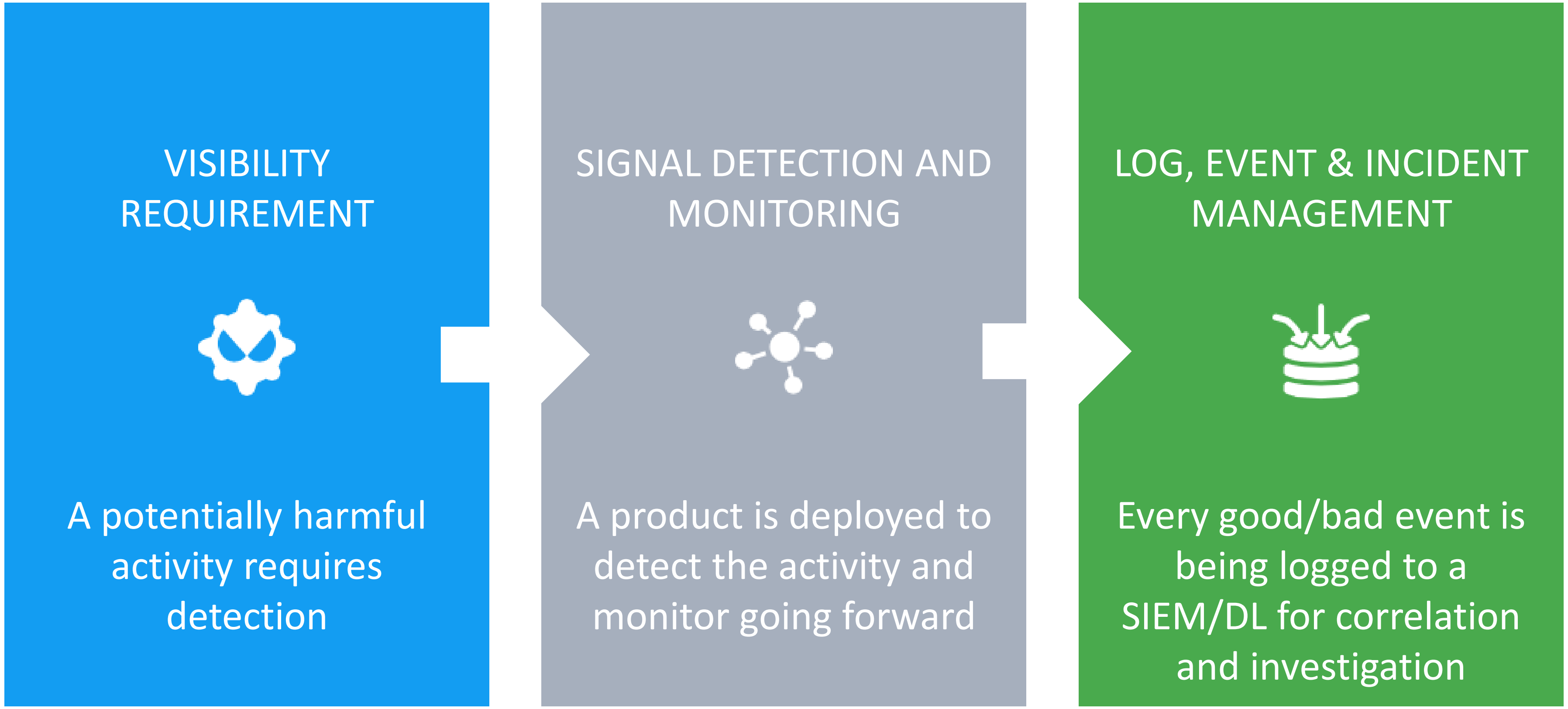
- W:E = 0.782
- **T:O = 0.851**
- L:O = 0.799
- C:R = 0.705
- M:Y = 0.913
- …



Source: https://www.youtube.com/watch?v=xxBc1c3uAJw

SECURITY MONITORING THROUGH LOGS

Today’s security monitoring best practices



VISIBILITY REQUIREMENT

A potentially harmful activity requires detection

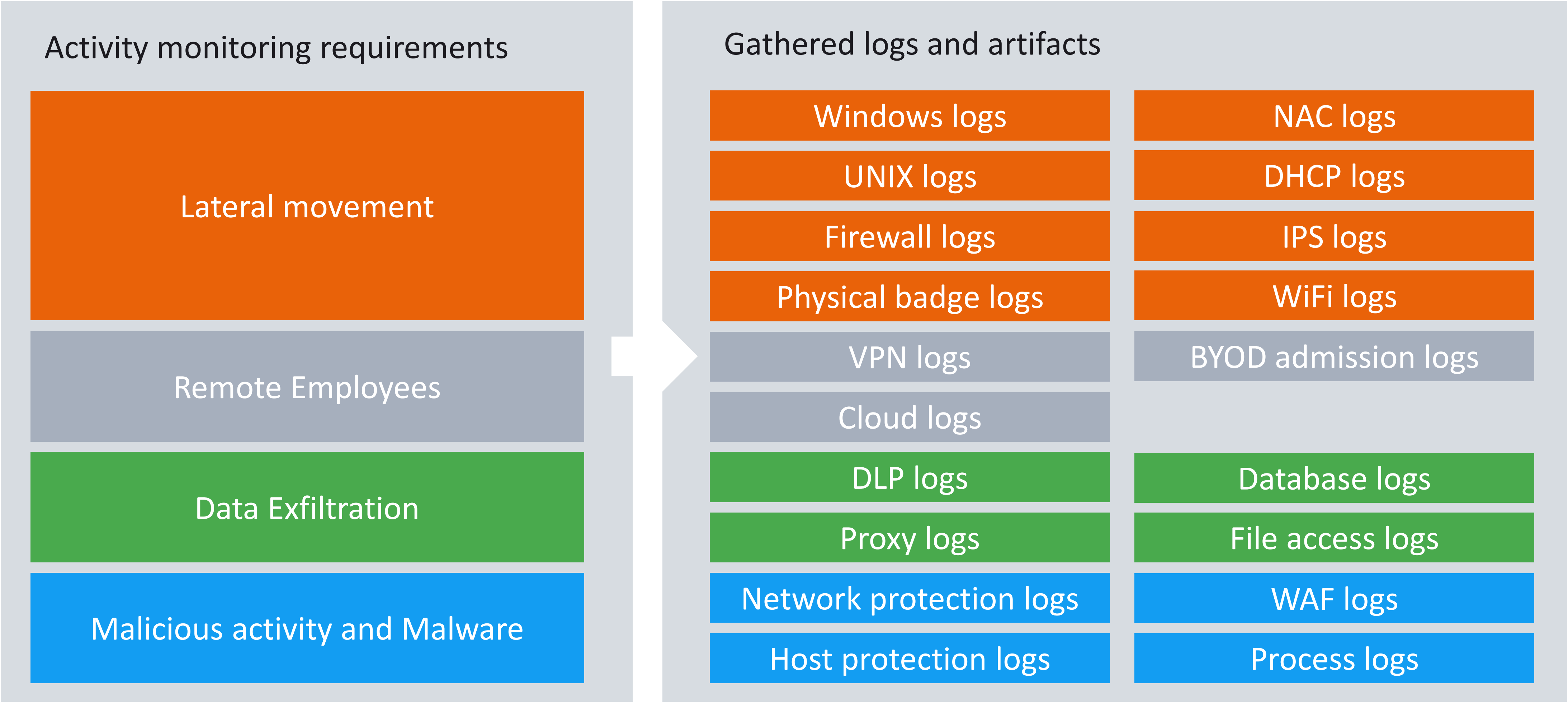
SIGNAL DETECTION AND MONITORING

A product is deployed to detect the activity and monitor going forward

LOG, EVENT & INCIDENT MANAGEMENT

Every good/bad event is being logged to a SIEM/DL for correlation and investigation

A few requirements, lots of log feeds



Activity monitoring requirements

Lateral movement

Remote Employees

Data Exfiltration

Malicious activity and Malware

Gathered logs and artifacts

Windows logs

NAC logs

UNIX logs

DHCP logs

Firewall logs

IPS logs

Physical badge logs

WiFi logs

VPN logs

BYOD admission logs

Cloud logs

DLP logs

Database logs

Proxy logs

File access logs

Network protection logs

WAF logs

Host protection logs

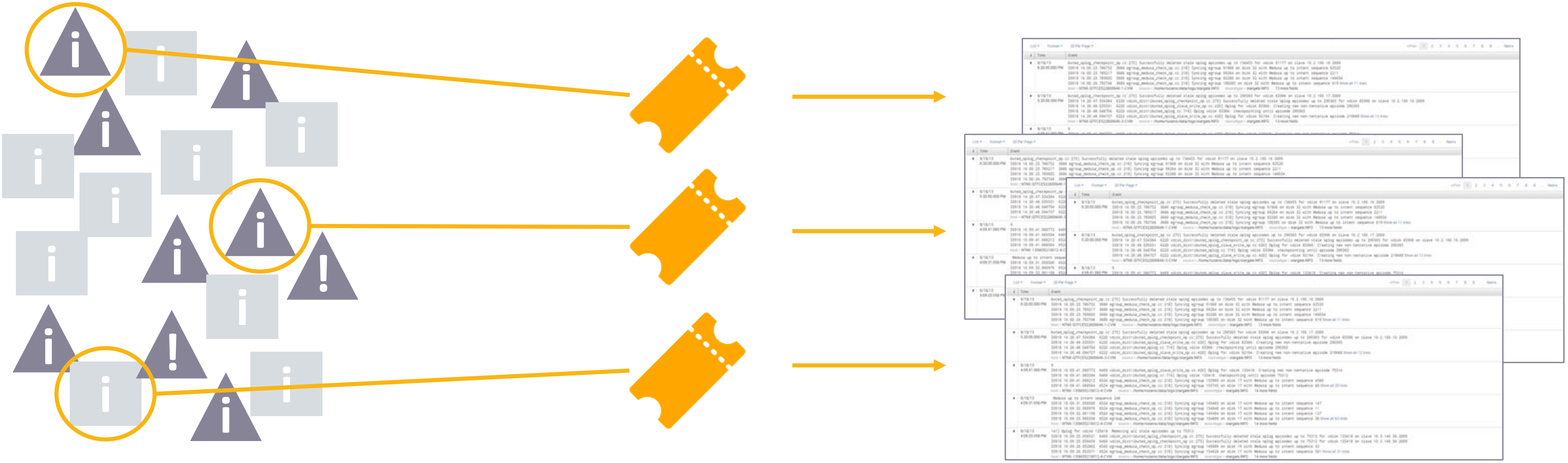
Process logs

Analyst workflow

Stash of Logs

Correlation Rules and Alerts Create Incidents

Case assembly and investigation through log search



exabeam

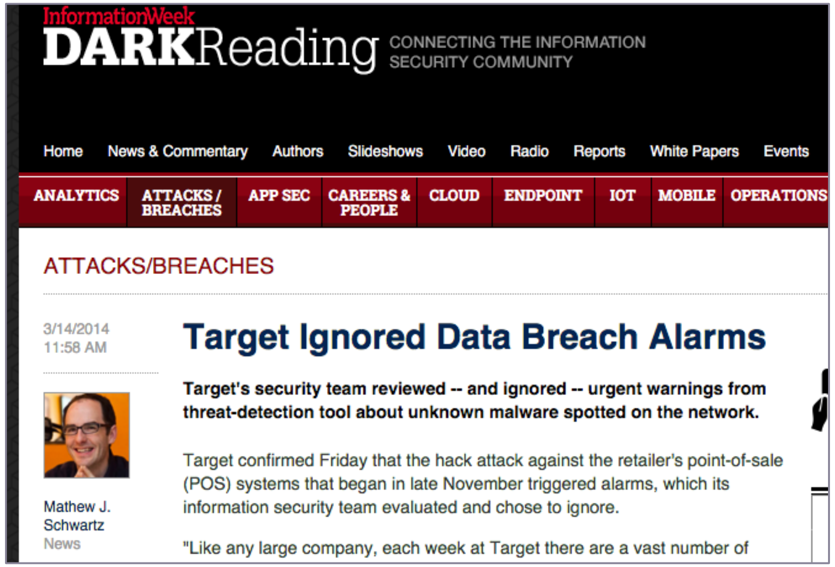
THE LEGACY APPROACH IS BROKEN

Familiar incidents In The Press

The Target breach, you are only human…

- ~Thanksgiving/Christmas 2013, 40m records of credit and debit card numbers were stolen using POS Malware at Target
- FireEye sent alerts of the then-unknown malware but were wrongfully interpreted and ignored.
- From DarkReading's interview with Target:

> "**Based on their interpretation and evaluation of that activity, the team determined that it did not warrant immediate follow up,**" she said. "With the benefit of hindsight, we are investigating whether, **if different judgments had been made, the outcome may have been different."**



ATTACKS/BREACHES

3/14/2014
11:58 AM

Mathew J. Schwartz
News

Target Ignored Data Breach Alarms

Target's security team reviewed -- and ignored -- urgent warnings from threat-detection tool about unknown malware spotted on the network.

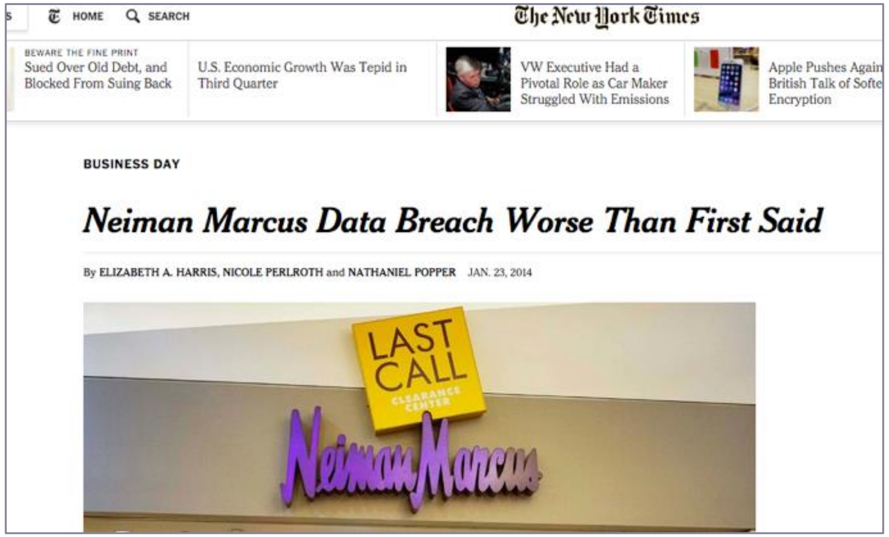
Target confirmed Friday that the hack attack against the retailer's point-of-sale (POS) systems that began in late November triggered alarms, which its information security team evaluated and chose to ignore.

"Like any large company, each week at Target there are a vast number of

Source: http://www.darkreading.com/attacks-and-breaches/target-ignored-data-breach-alarms/d/d-id/1127712
Source: http://www.scmagazine.com/target-did-not-respond-to-fireeye-security-alerts-prior-to-breach-according-to-report/article/338201/
Source: http://www.bloomberg.com/bw/articles/2014-03-13/target-missed-alarms-in-epic-hack-of-credit-card-data

Neiman Marcus… needle in the needle-stack

- ~1.1m Credit cards information exposed (NYT, Jan 13, 2014)
- Industry Averages
  - The average enterprise, logs **~160m-200m events** a day
  - The average enterprise logs up to **150k security events** a day
- Neiman Marcus had **60k security alert events** per day, yet suffered from a 3 month breach. (Damballa State of Infections Report 2014)
- Those are just security alerts, numbers exclude noteworthy infrastructure events



BEWARE THE FINE PRINT
Sued Over Old Debt, and Blocked From Suing Back

U.S. Economic Growth Was Tepid in Third Quarter

VW Executive Had a Pivotal Role as Car Maker Struggled With Emissions

Apple Pushes Again British Talk of Softe Encryption

BUSINESS DAY

Neiman Marcus Data Breach Worse Than First Said

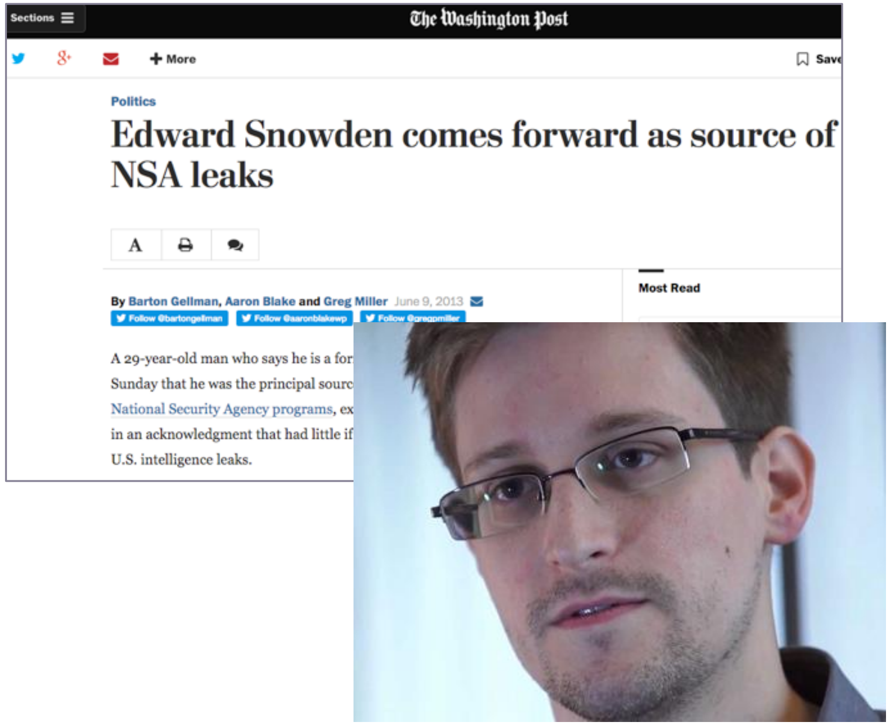
By ELIZABETH A. HARRIS, NICOLE PERLROTH and NATHANIEL POPPER JAN. 23, 2014

Impossible Signal/Noise Ratio

Source: http://www.nytimes.com/2014/01/24/business/neiman-marcus-breach-affected-1-1-million-cards.html
Source: https://www.damballa.com/downloads/r_pubs/Damballa_Q114_State_of_Infections_Report.pdf

Snowden… in those we trust.

- Highly privileged and trusted user with access rights to sensitive information
- Creates the mother of all data leaks
- Noteworthy
  - Changes his behavior over time
  - Avoids stepping in any traps
  - No malware, only credentials – mostly his own
  - Appears to be just like any other trusted insider user



Politics

Edward Snowden comes forward as source of NSA leaks

By Barton Gellman, Aaron Blake and Greg Miller June 9, 2013

A 29-year-old man who says he is a for
Sunday that he was the principal sourc
National Security Agency programs, ex
in an acknowledgment that had little if
U.S. intelligence leaks.

To an analyst, he appears just like anyone else

Source: https://www.washingtonpost.com/politics/intelligence-leaders-push-back-on-leakers-media/2013/06/09/fff80160-d122-11e2-a73e-826d299ff459_story.html

exabeam

THE MODERN SOC?

Is this your SOC?

Alert fatigue results in missed incidents

Signal to Noise ratio is unmanageable

One user’s malicious activity, is another user’s standard

The Analyst World – Through Different Eyes

What The Board Thinks We Do

What The SOC Manager Thinks We Do

What We Actually Do!

exabeam

A NEW HOPE

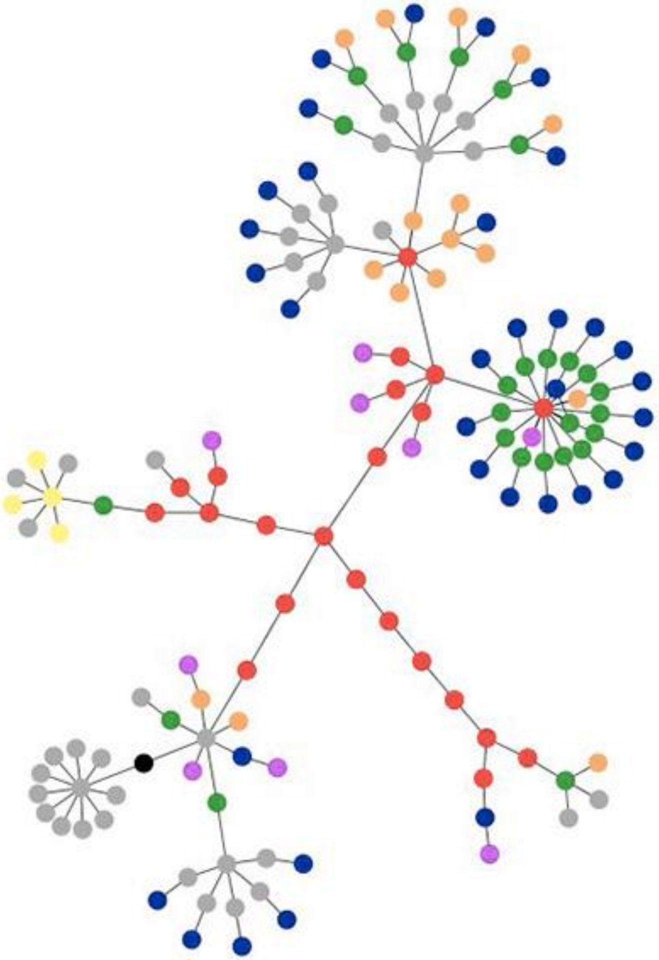
The connection graph

Stitching together user activities that cross accounts, devices, IPs and networks requires a new type of data structure:

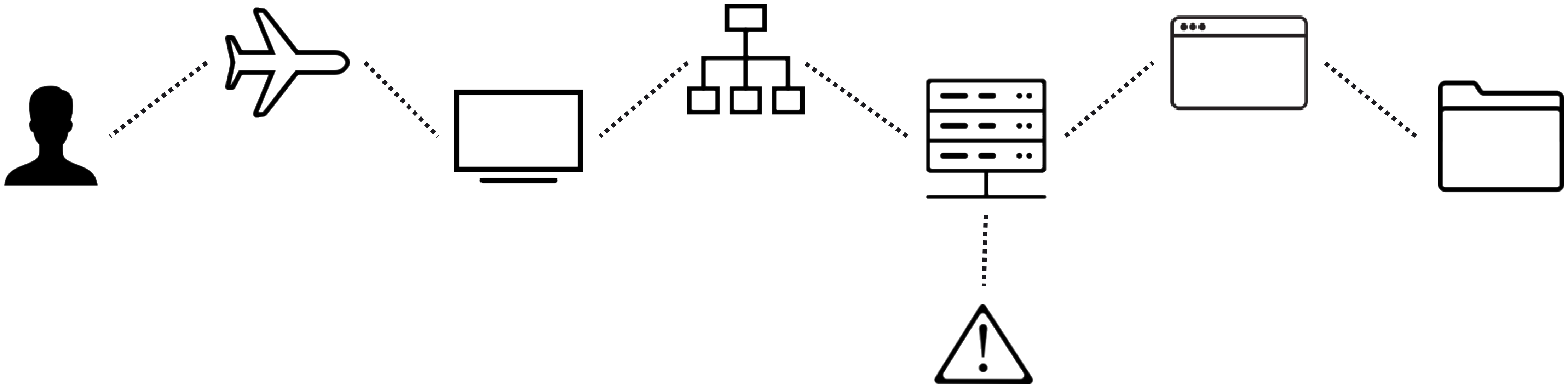
- **Integrates state changes** – so that the attackers stays visible as he changes accounts, IPs, across a session

- **Incorporates time** - to understand that C happened after B happened after A

- **Abstracts individual events** – so that the entire session can be queried



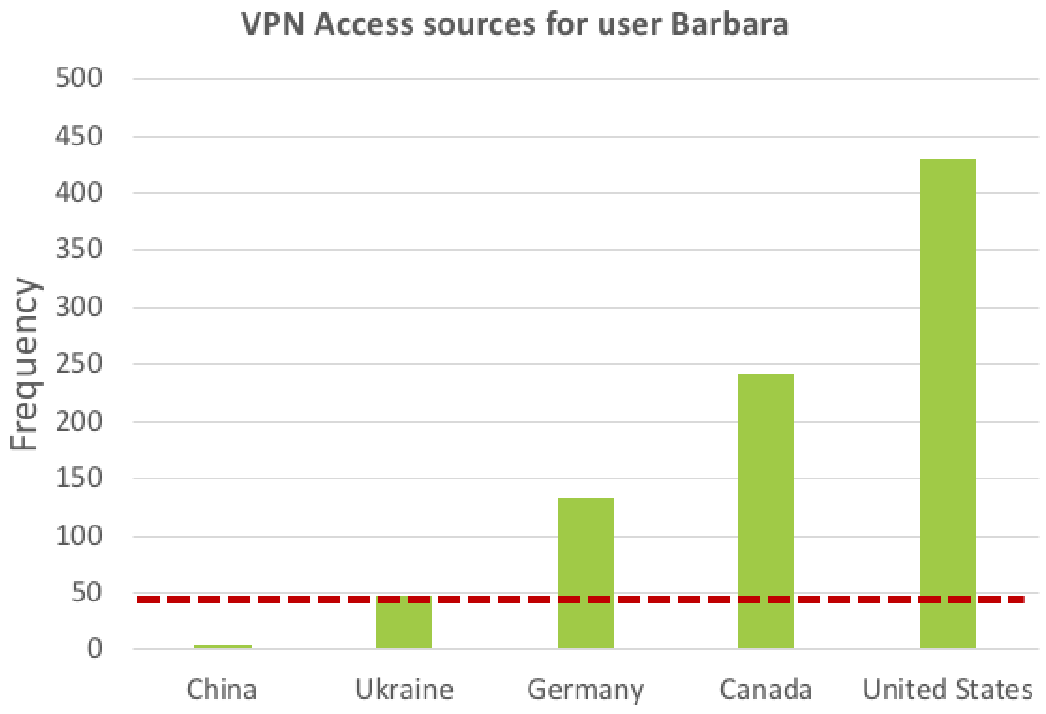
User activity session as a connected graph



APPLYING MACHINE LEARNING

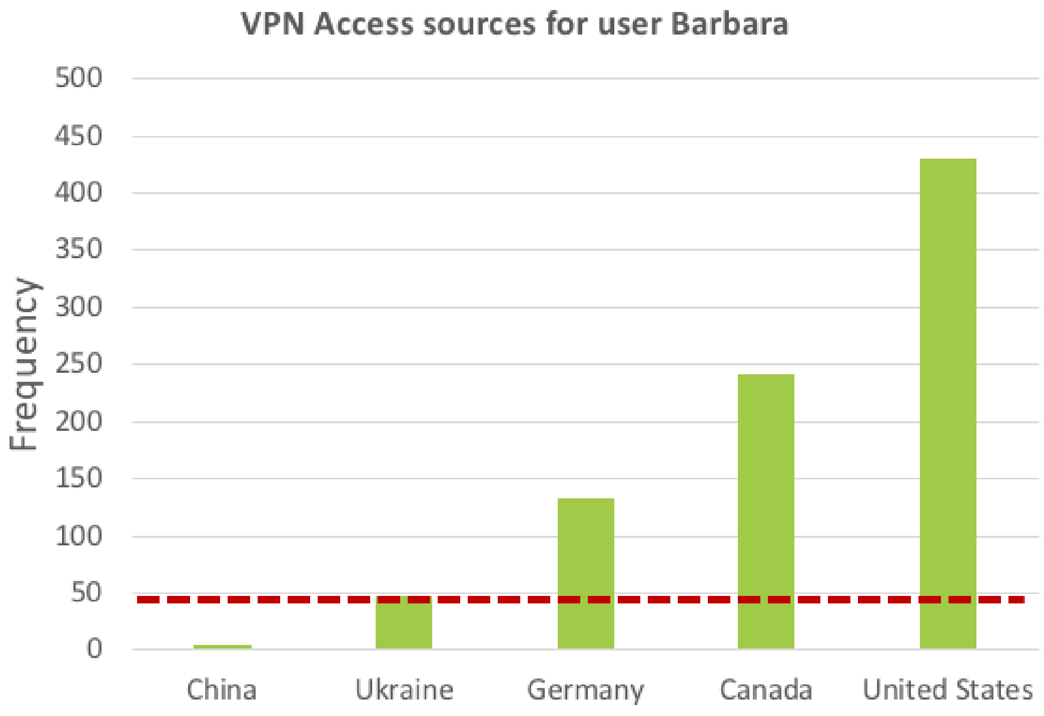
Learning a user’s behavior over time

User **Barbara** connected to VPN from **US**
User **Barbara** connected to VPN from **US**
User **Barbara** connected to VPN from **US**
User **Barbara** connected to VPN from **GR**
User **Barbara** connected to VPN from **GR**
..
..
User **Barbara** connected to VPN from **CN**

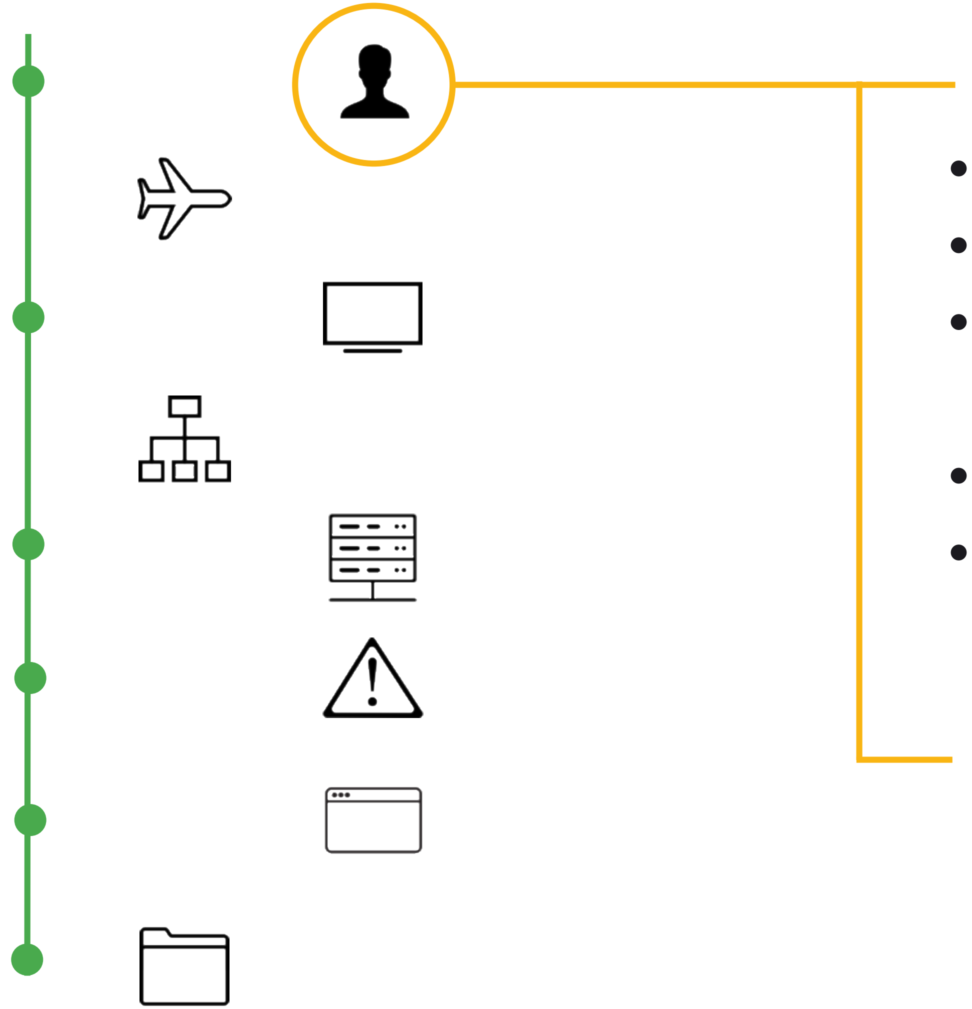


Let data speak for itself…

- Barbara regularly connects from **United States**
- It is abnormal for Barbara to connect from **China**
- Barbara never connected from **Brazil**

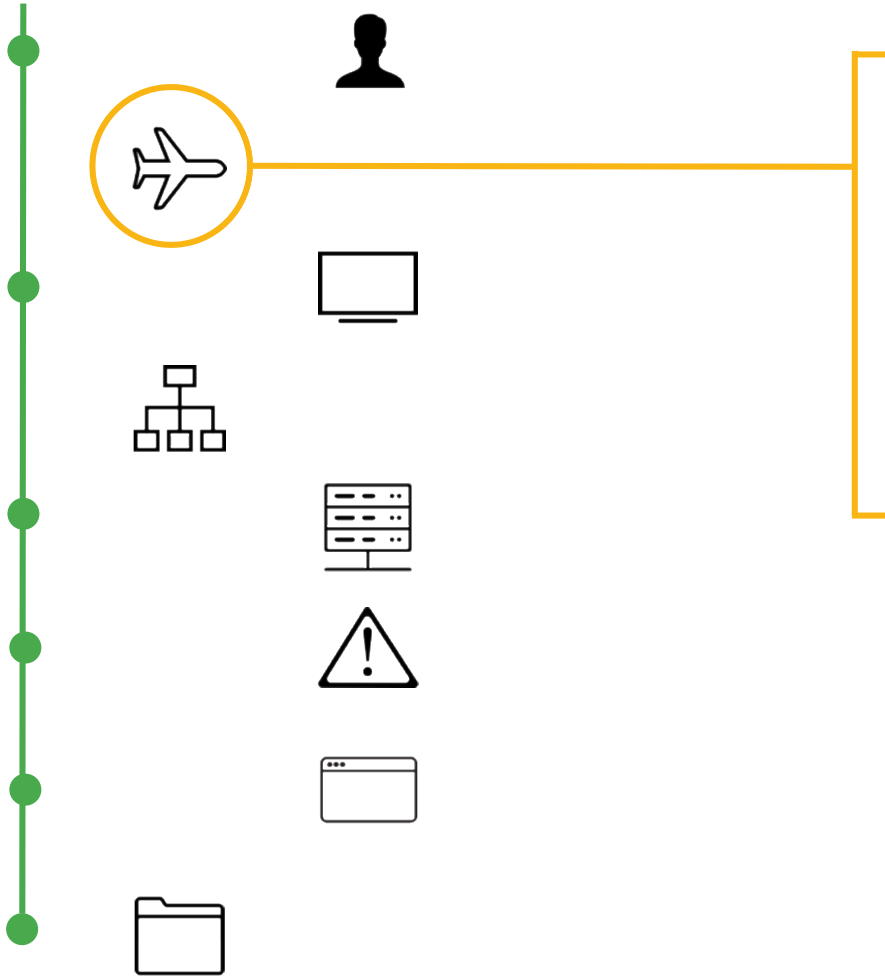


Applying machine learning to user behavior



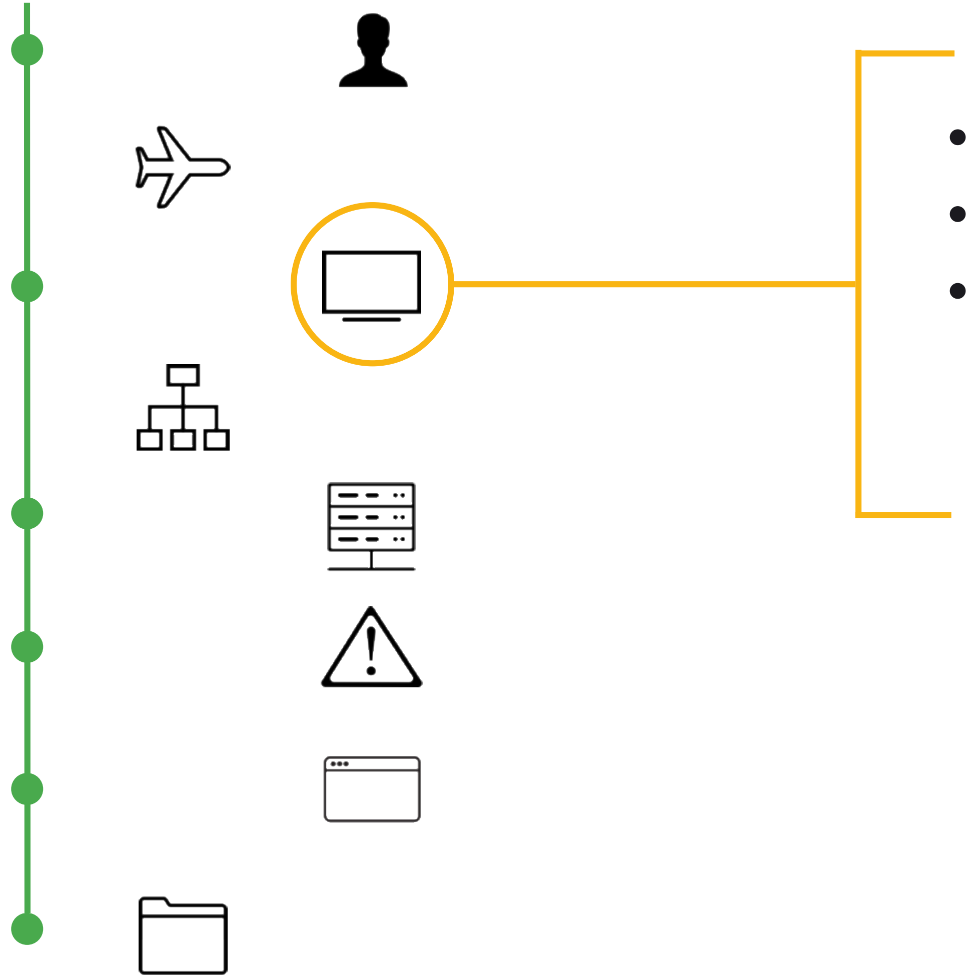
- Who is this user to the organization?
- What are this user’s peers
- Does this user have any special characteristics?
  - Is he an Executive? A privileged user perhaps ?
- When does the user usually login ?
- When did we last see this user ?

Applying machine learning to user behavior



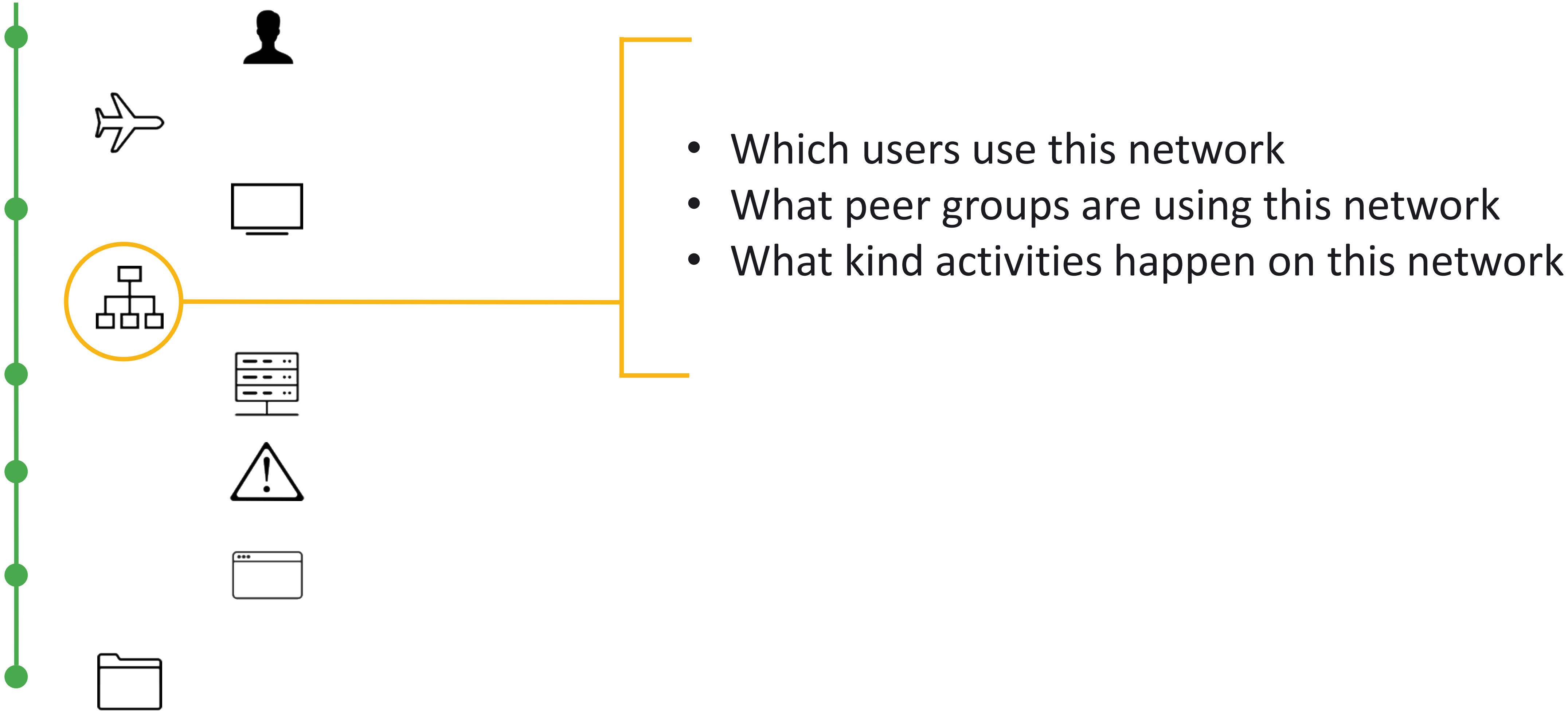
- Where is the user connecting from ? ISPs, States, Countries, etc…
- When does the user usually login
- Is the source IP known as good/bad IP?

Applying machine learning to user behavior



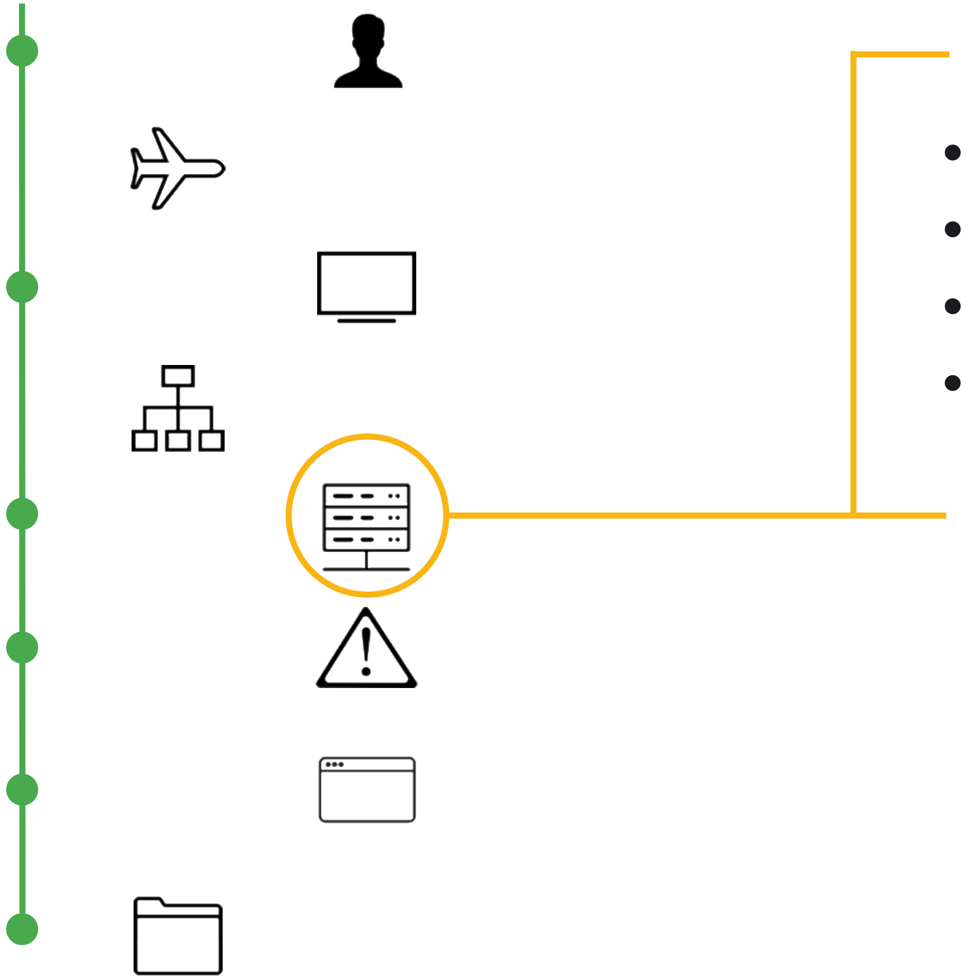
- Who used this asset before ?
- What kind of users login to this asset
- What actions are normally performed on this asset

Applying machine learning to user behavior



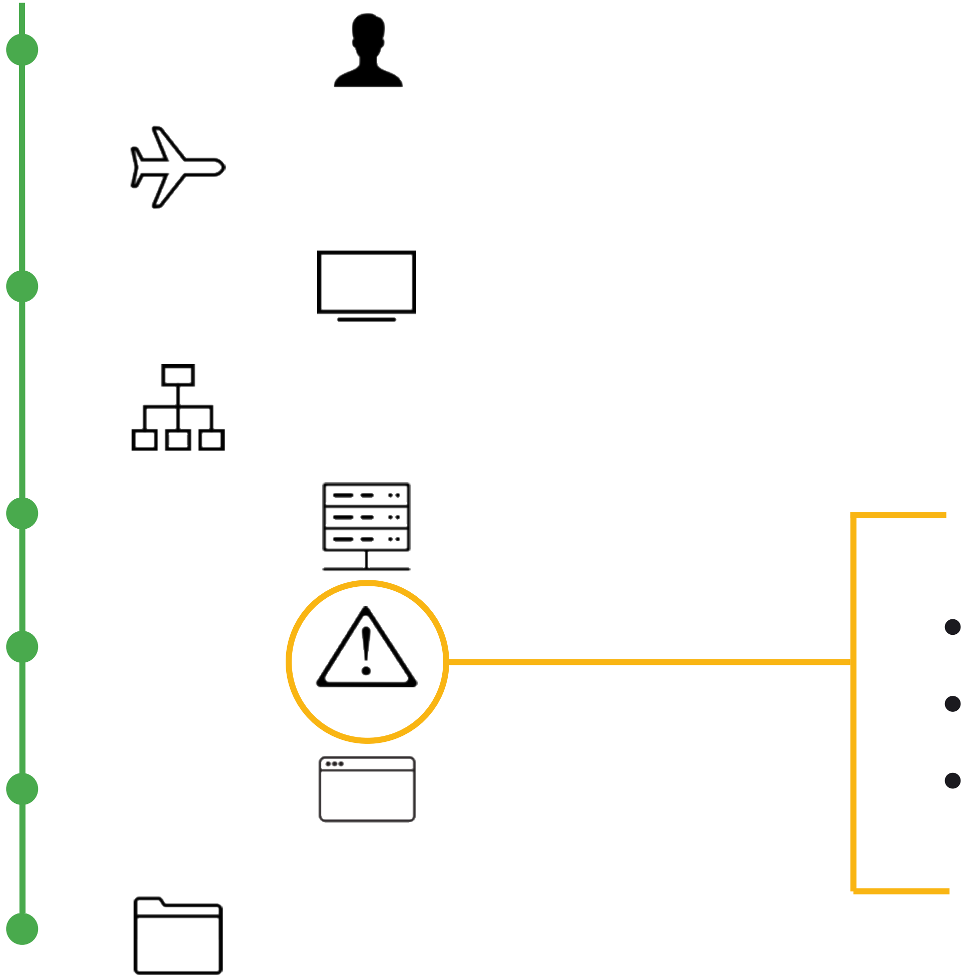
- Which users use this network
- What peer groups are using this network
- What kind activities happen on this network

Applying machine learning to user behavior



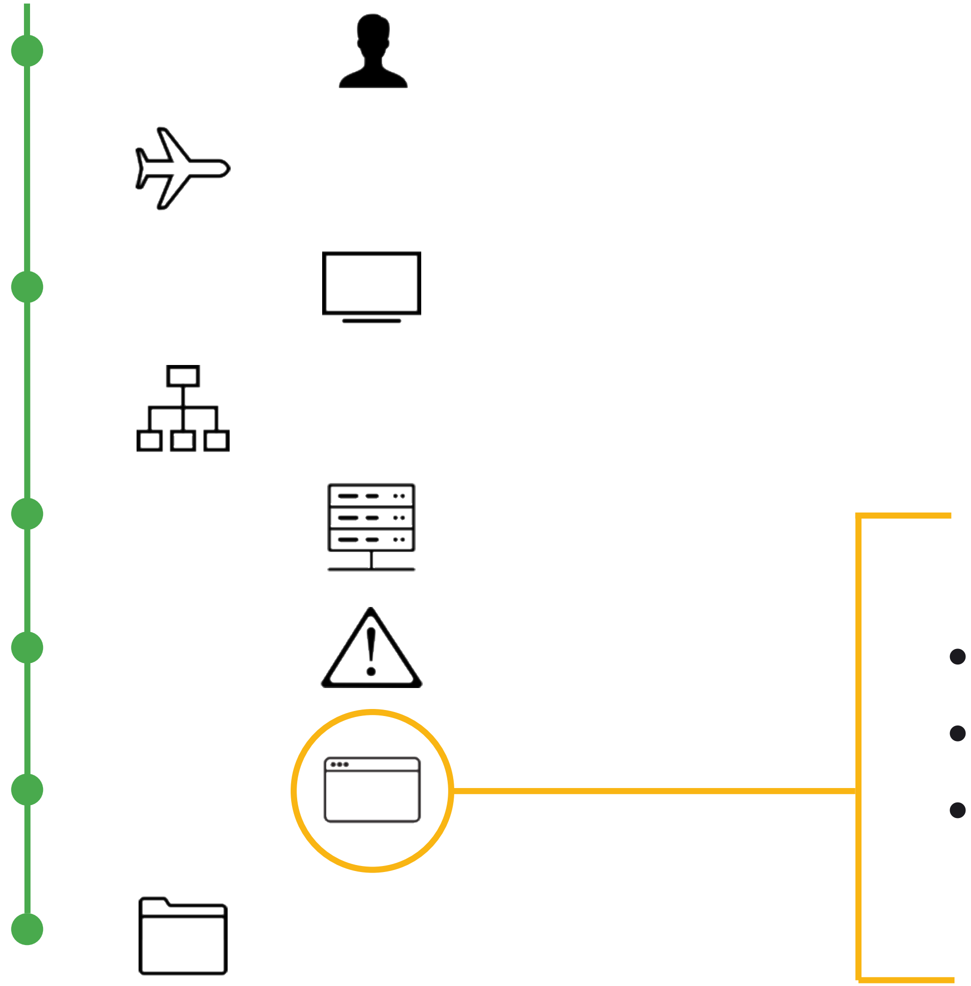
- Which users normally use this server
- Which users are the administrators to this server
- What applications run on this server
- How do users normally access this server

Applying machine learning to user behavior



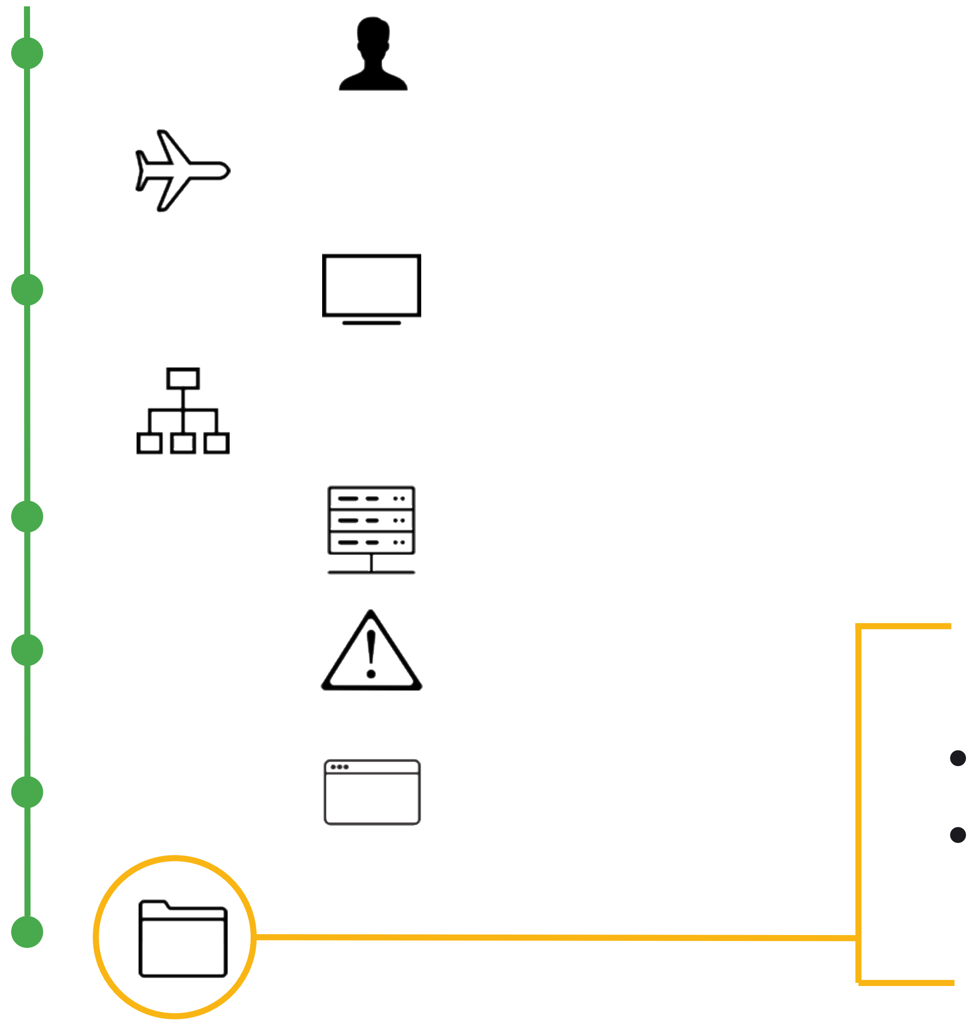
- Is this alert a common alert in the organization
- Has this alert ever fired before on this asset
- Has this alert ever fired before for this user

Applying machine learning to user behavior



- Which users access this application
- What hosts hold this application
- What arguments are used with this application

Applying machine learning to user behavior



- Who accesses this file normally
- Where was this file is accessed from

exabeam

ML IN PRACTICE

User and Entity Behavior Analytics

So those alerts again…

“Legitimate User VPN session out-of-hours from CN”

Would this event even register on the analysts radar?

“Another alert has fired; malware on host X”

How many analysts would dismiss this ?

“DB access by HR User, Table copied”

Where would an analyst even begin here?

Lets try this again, with ML functions…

- User **Barbara** has
  - Abnormally logged in using **VPN** from **China**
  - Is accessing networks she **never accessed** before
  - No one in her **peer group** uses this server
  - Normally **only reads this file** and does not edit
  - An alert has fired for **malware**
  - **First time** this malware is seen **in this company**

exabeam

WWW.EXABEAM.COM