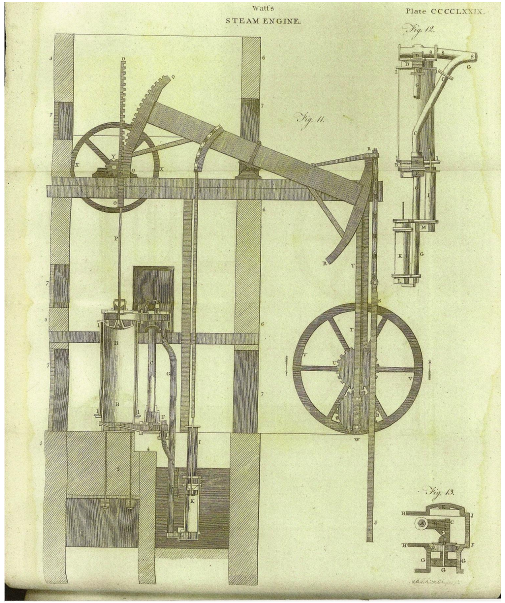
AI You Ready? Attention. Go!

Raphaël Marichez
Chief Security Officer, Southern Europe
Palo Alto Networks

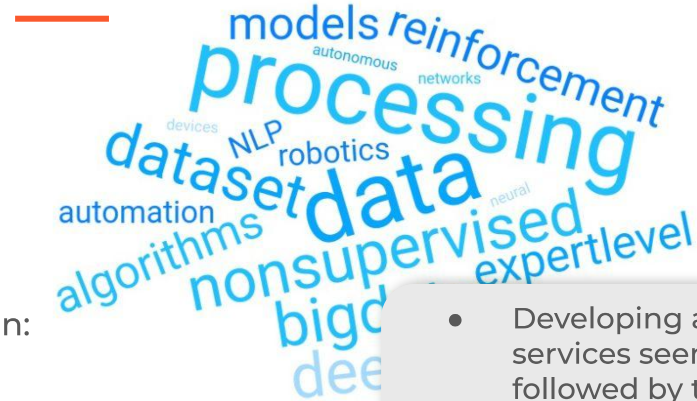
2021

Is there a common understanding of AI?

- Watt steam engine (around 1775)
  - Watt linked a steam regulator valve to a centrifugal governor to roughly correct the speed with a feedback loop
  - These improvements allowed the steam engine to replace the water wheel and horses as the main sources of power for British industry, thereby freeing it from geographical constraints and becoming *one of the main drivers in the Industrial Revolution.*



Is there a common understanding of AI?



- Artificial Intelligence relies on:
  - The availability of quality, consistent and accessible **data**
  - **Models, schemas** (eg. image classifier)

    And
  - **Algorithms** (processes) leading to a decision, a parameter change, a switch between models...)
  - And of course traditional infrastructure

- Developing applications using cloud-based AI services seems the most popular at present (48%), followed by those choosing to buy applications with ML built in (25%).
- More than half (51%) of enterprises have indicated that their current infrastructure cannot scale to meet future AI workload demands without considerable changes.
- Public cloud being the most popular core execution venue (over enterprise DC and 3dr-party DC)

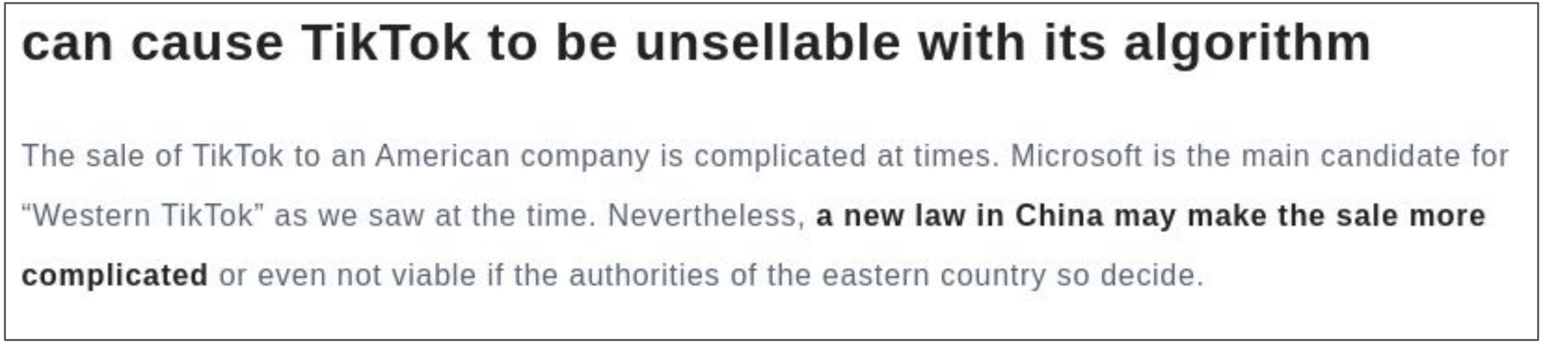
(S&P Global Market Intelligence, 2020)

From IaaS, SaaS to "*AI as a Service*"

AI is the key of the future of technology

If you want to enter by the door, you need to own the key, not the door itself.

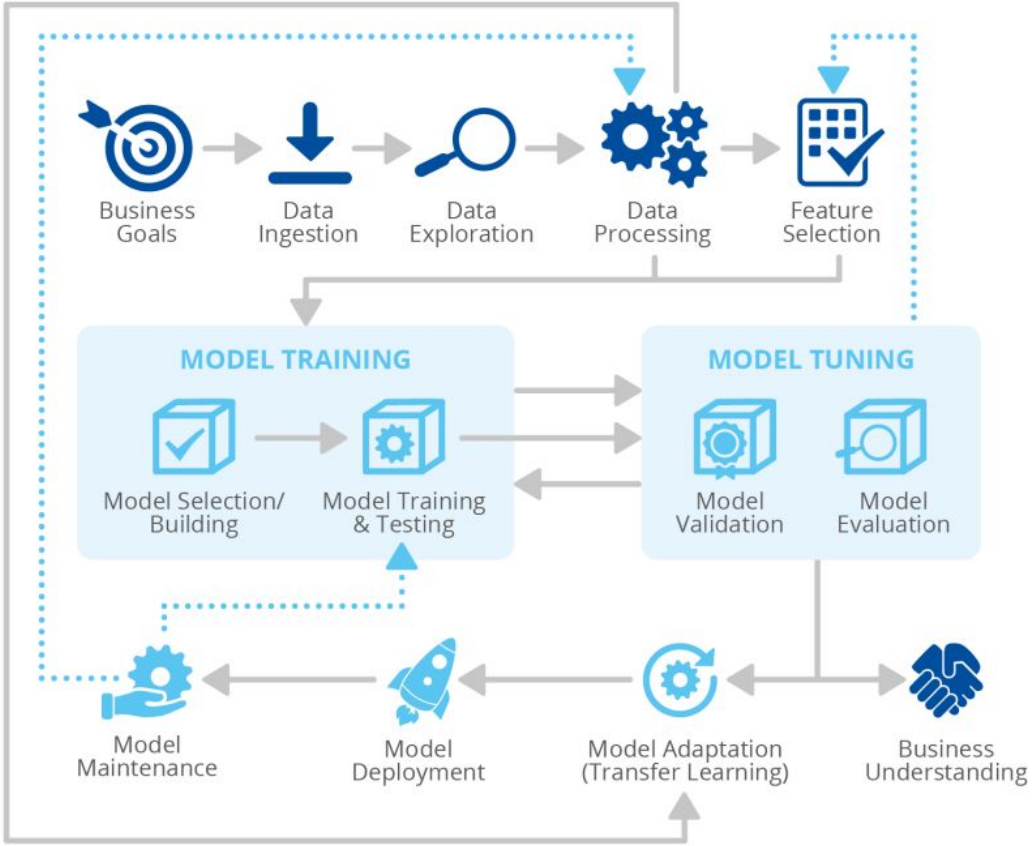
(August 2020)



can cause TikTok to be unsellable with its algorithm

The sale of TikTok to an American company is complicated at times. Microsoft is the main candidate for “Western TikTok” as we saw at the time. Nevertheless, **a new law in China may make the sale more complicated** or even not viable if the authorities of the eastern country so decide.

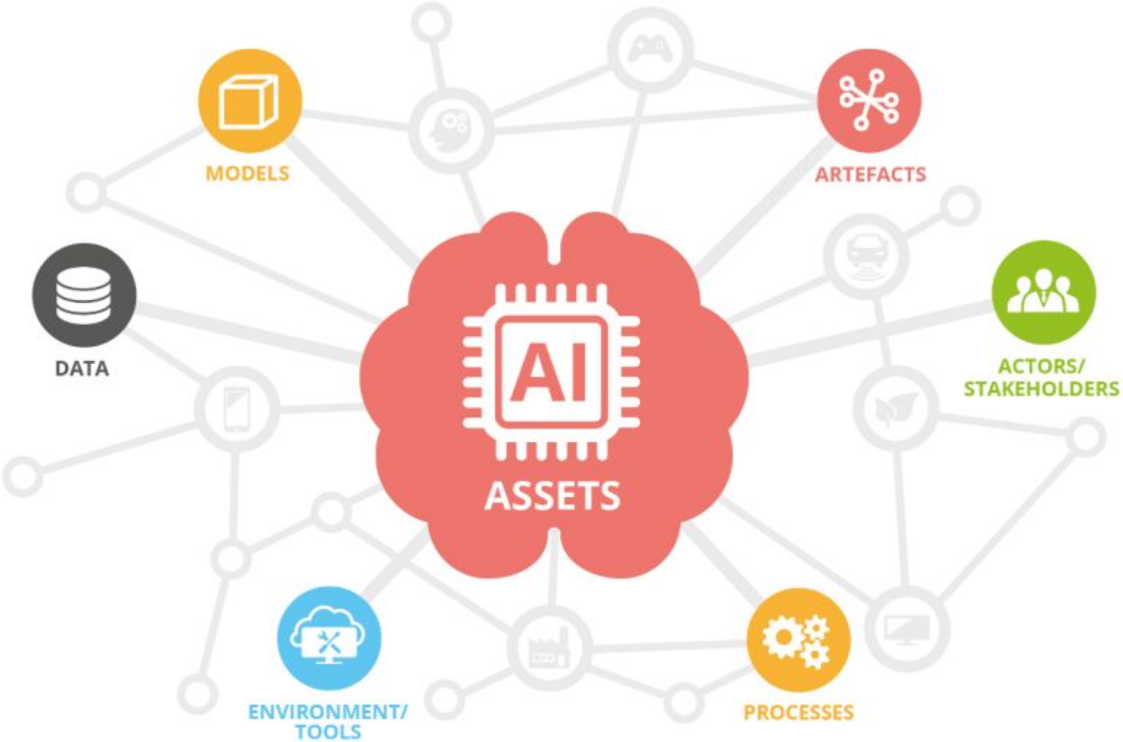
AI introduces new threat landscape



AI lifecycle

(by ENISA)

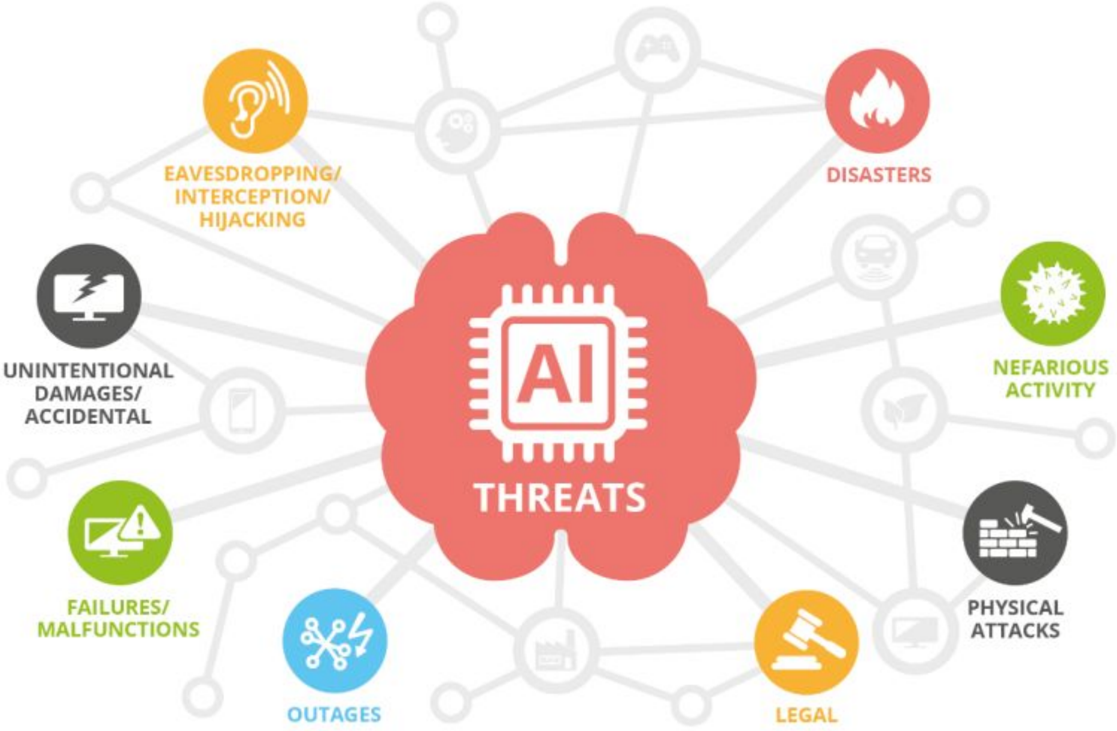
AI introduces new threat landscape



AI assets' caterogies

(by ENISA)

AI introduces new threat landscape



AI Threat Taxonomy
(by ENISA)

AI introduces new threat landscape - some examples

- Nefarious activity / abuse
  - ACL / Group permissions inheritance ⇒ implicit **privilege escalation attacks**
  - **Adversarial examples** (perturbations imperceptible to the human eye) ⇒ impact on ML models
  - Insider attacks, hard-to-detect **parameter changes**
  - Limited, biased, erroneous or tampered **input dataset** (secrets, lack of understanding...)
  - Insert attacks on **training datasets** (eg. certain pixel pattern for a surveillance camera / image classifier)
  - **Data poisoning / tampering** (legitimate or illegitimate access) ⇒ Adversely affect AI operations, biases
  - Flawed or poisoned **schemas** or compromised cloud-based **models** (backdoor in libraries...)
  - **3rd parties models** backdoors or biases
  - Compromise of **data brokers** (poisoning via insertion, filtering) ⇒ Biases in the decision process
  - **DDoS** attacks (storage, CPU...)
  - **Timing attacks** (public interfaces) ⇒ loss of confidentiality
  - **ML model** confidentiality
  - Unauthorized access to data sets and data transfer process, or to models' code

AI introduces new threat landscape - some examples

- Legal or privacy concerns
    - **Unintentional data breaches** (personal data, models' code, weak encryption...)
    - Disclosure of **Personal Information** by correlation, profiling users, lack of randomization...
    - Lack of **data governance policies** (when personal data are processed)
    - Lack of data protection **compliance** of 3rd parties providing or processing data
    - **SLA breaches** with 3rd parties
    - **Vendor lock-in** (libraries, data storage...)

AI introduces new threat landscape - ENISA’s report summary

Nefarious activity/abuse
“intended actions that target ICT systems, infrastructure, and networks by means of malicious acts with the aim to either steal, alter, or destroy a specified target”.

Eavesdropping/Interception/ Hijacking
“actions aiming to listen, interrupt, or seize control of a third party communication without consent”.

Physical attacks
“actions which aim to destroy, expose, alter, disable, steal or gain unauthorised access to physical assets such as infrastructure, hardware, or interconnection”.

Unintentional Damage
“destruction, harm, or injury of property or persons and results in a failure or reduction in usefulness”.

Failures or malfunctions
“Partial or full insufficient functioning of an asset (hardware or software)”.

Outages
“unexpected disruptions of service or decrease in quality falling below a required level“.

Disaster
“a sudden accident or a natural catastrophe that causes great damage or loss of life”.

Legal
“legal actions of third parties (contracting or otherwise), in order to prohibit actions or compensate for loss based on applicable law”.

UNIT 42
BY PALO ALTO NETWORKS

Artificial intelligence for cybersecurity is not an option!

- Surface attack discovery
  With massive **cloud practices adoption**
  and **exponentially growing attack surface**
  (devices, softwares, data, connections, nodes...)

- **New era for the honeypots...**

- Unit 42 latest blog post (22 Nov. 2021)
  Observing Attacks Against Hundreds of Exposed Services in Public Clouds
  https://unit42.paloaltonetworks.com/

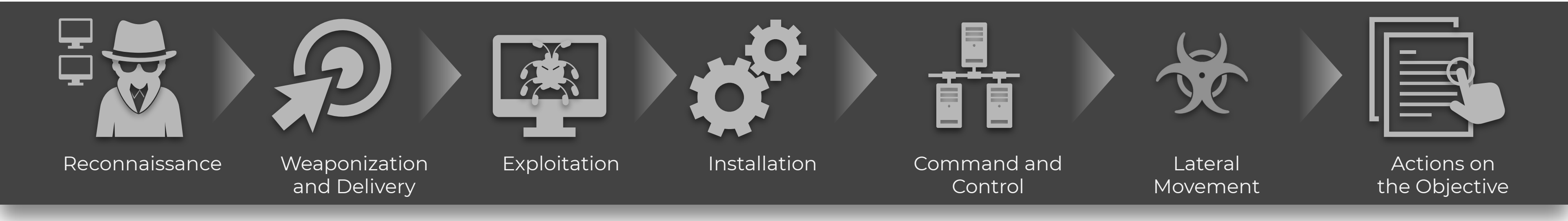
Notorious ransomware groups such as **REvil** and **Mespinoza** are known to exploit exposed services to gain initial access to victims' environments.

Using a honeypot infrastructure of 320 nodes deployed globally, **Unit 42** researchers aim to better understand the attacks against exposed services in public clouds.

80% of the 320 honeypots were compromised within 24 hours and all of the honeypots were compromised within a week

- On average, each SSH honeypot was compromised 26 times daily.
- One threat actor compromised 96% of our 80 Postgres honeypots globally within 30 seconds.
- 85% of the attacker IPs were observed only on a single day. Layer 3 IP-based firewalls are ineffective as attackers rarely reuse the same IPs to launch attacks. **A list of malicious IPs created today will likely become outdated tomorrow.**

AI-enabled attackers



Reconnaissance

Weaponization and Delivery

Exploitation

Installation

Command and Control

Lateral Movement

Actions on the Objective

Victim profiling

ML-supported fuzzing and exploitation

Build automatically a distribution network

Advanced ad-hoc evasion and self-defense techniques

Device clustering (e.g Ransomware)

Data clustering (e.g Espionage)

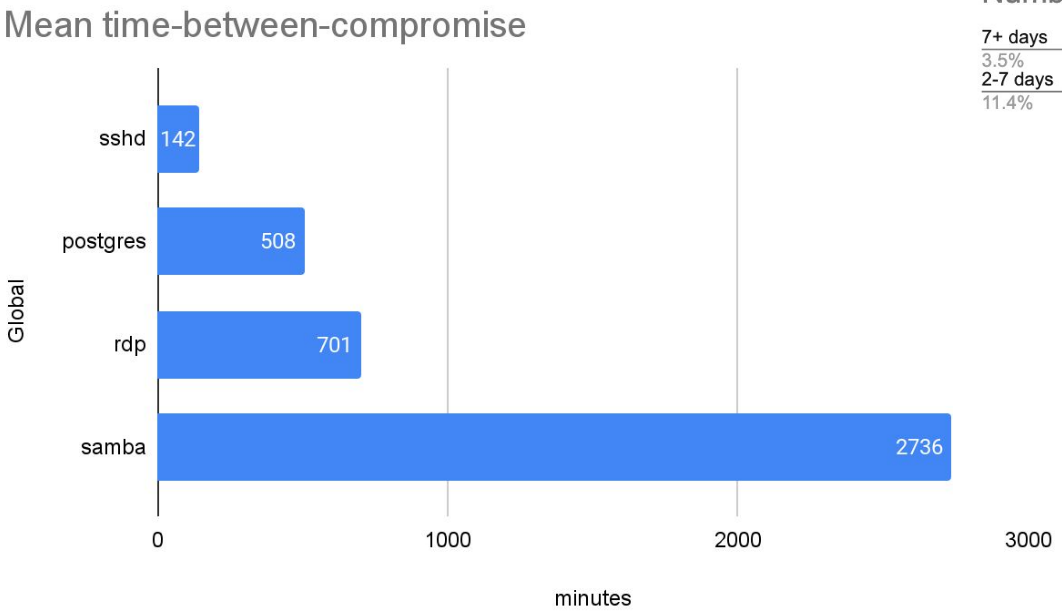
High Level of automation possible already today

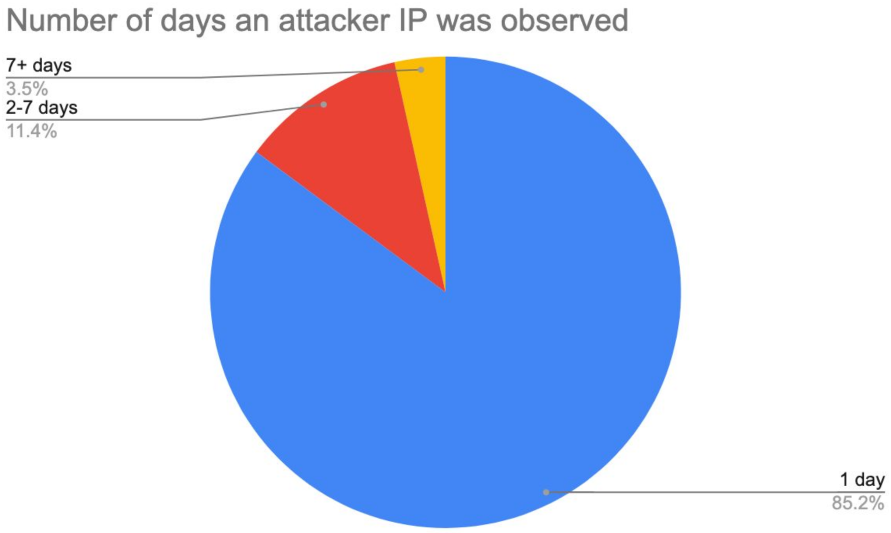
➕

Mature business model with several functions provided “as a service”

Artificial intelligence for cybersecurity is not an option!

UNIT 42 BY PALO ALTO NETWORKS



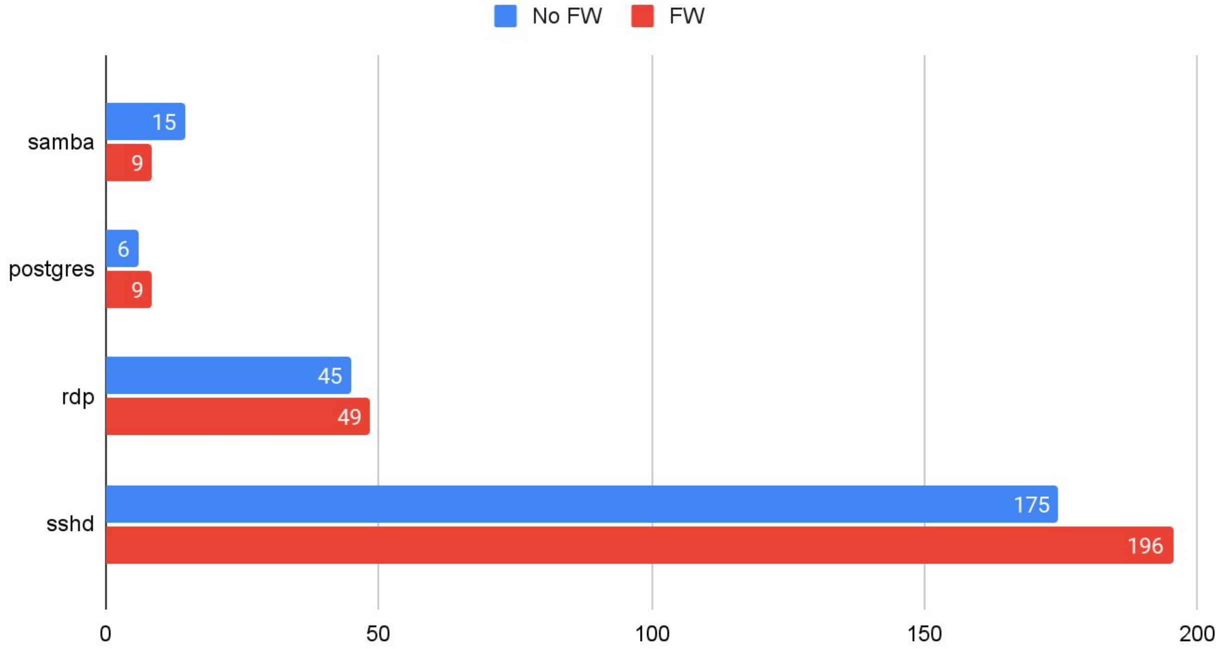


Artificial intelligence for cybersecurity is not an option!

Unit 42 applied firewall policies to block IPs from known network scanners. The firewall policy blocks the IPs that have been scanning a specific application daily in the past 30 days.

Blocking known scanner IPs is ineffective in mitigating attacks.

Number of attackers in 30 days per honeypot



Artificial intelligence for cybersecurity is not an option!

- Cyber threat intelligence management

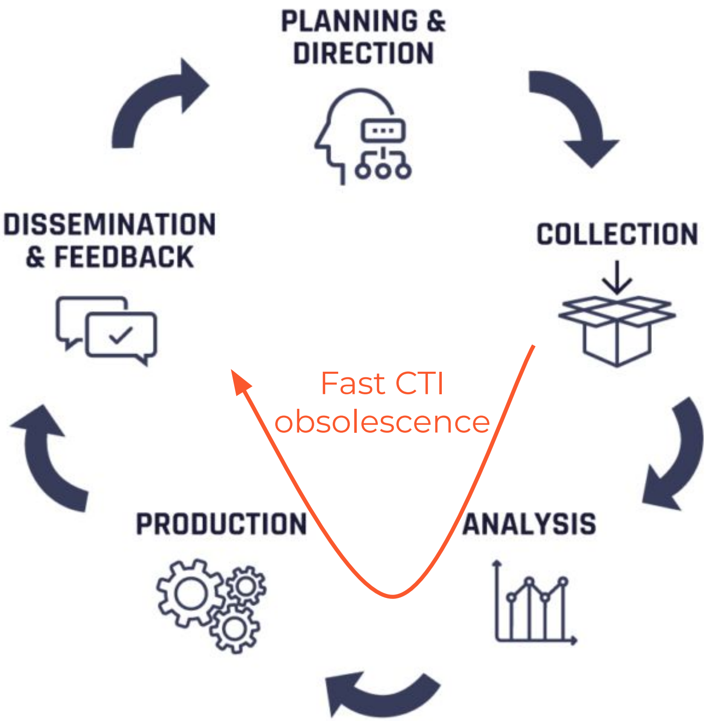
Challenge:
Prevent your collected CTI from being obsolete before dissemination across your infrastructure

- AI is a game changer for the offensive part

- The attacker isn't subject to national regulations

He can practice:

- Deception (aversary learning data)
- Low-noise attacks (undetectable by human eyes)
- Try, fail and try again (while ***we*** can't!)



Artificial intelligence benefits for cybersecurity

These 5 items are aligned with Gartner's PR on "Security & Risk Management Summit Day 4 Highlights"

- **Infrastructure protection**

  Classify patterns (applications rather than network protocols, file contents rather than extensions…)
  Analyse system calls within binaries (static analysis or living sandboxes…)
  Correlate with IPs, hosting infrastructure, network activity to map interaction graphs
  Correlate across time, to classify known behaviours and alert on unseen behaviour

- **Identity & Access management**

  Identity profiling
  Entitlements (mainly in public cloud permissions graphs)

- **Risk management**

  Suggestions based on new software or 3rd parties
  Find and factorize similar controls or rules across different subsidiaries or infrastructures

- **Application and Data Security**

  Classify bad / good applications based on seen behaviour, and/or reputation
  Classify text/images patterns (mixed) to trigger DLP techniques

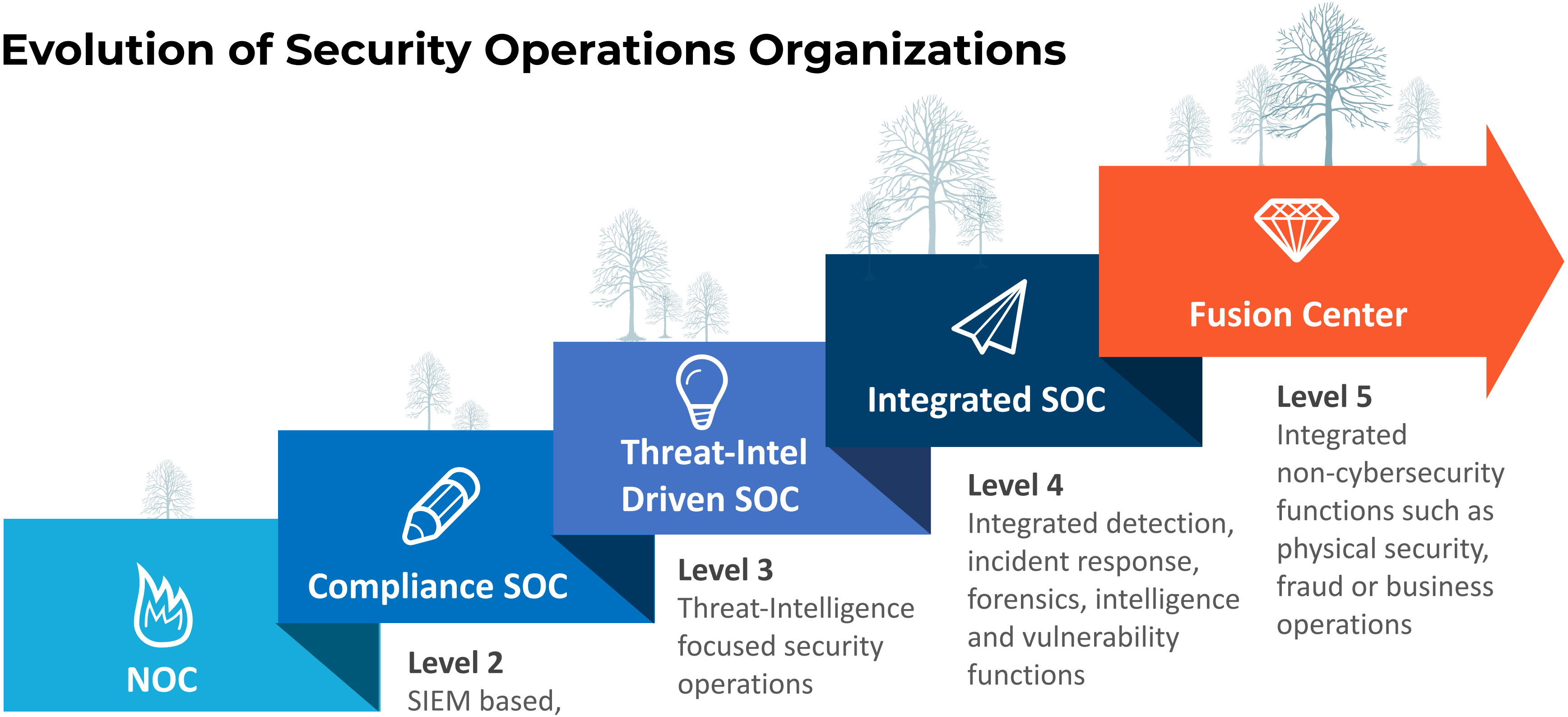
- **Security Operations**

  Most obvious use case (alert fatigue, repeatable low-value tasks…)

Artificial intelligence benefits for cybersecurity

- **General recommendations:**
  - AI is a collection of complementary techniques, it’s not magical, it’s not a product
  - Misunderstanding AI can cause AI malfunction and misuses!
  - AI doesn’t replace human. It changes what humans do (in a better way, we expect)
    - **Automation** replaces humans
    - Automation is **enabled** by AI with growing complexity infrastructures overwhelming humans capacity
  - Your AI should always be explainable. Measure it to improve or fix it.
  - Challenge your AI’s results with your field expertise.
  - AI value grows with good data input and supervision. Supervise it to improve or fix it.

Evolution of Security Operations Organizations



NOC

Compliance SOC

Threat-Intel Driven SOC

Integrated SOC

Fusion Center

Level 1
Perimeter focused operations

Level 2
SIEM based, Policy driven operations and static playbooks

Level 3
Threat-Intelligence focused security operations

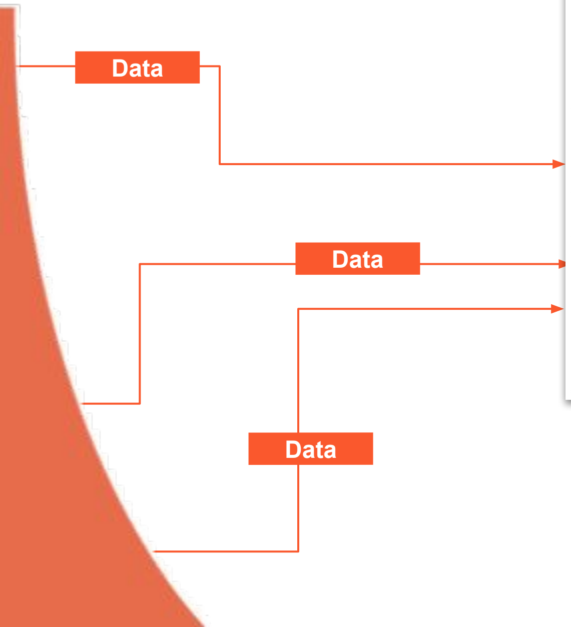
Level 4
Integrated detection, incident response, forensics, intelligence and vulnerability functions

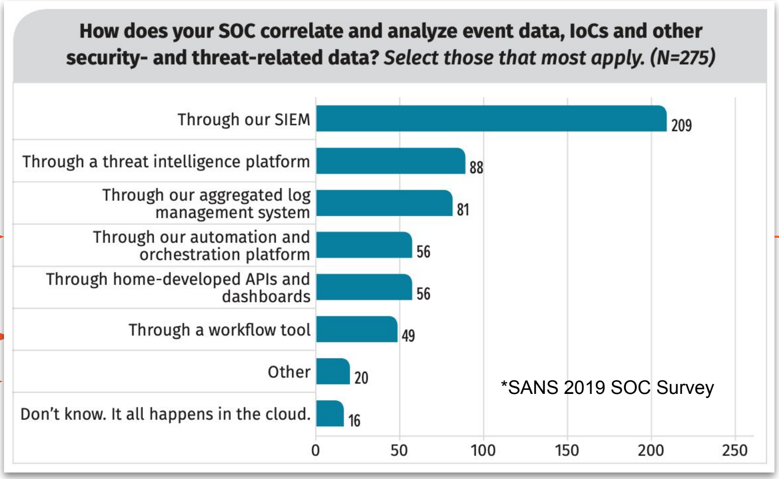
Level 5
Integrated non-cybersecurity functions such as physical security, fraud or business operations

Legacy organization structure

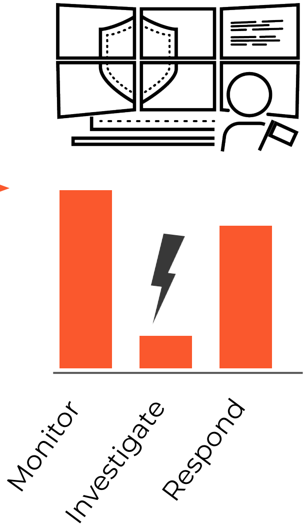
“We ask the most inexperienced “Tier 1” analyst to distinguish between APTs and commodity threats. Does this work?”

SIEMs and our obsession for data

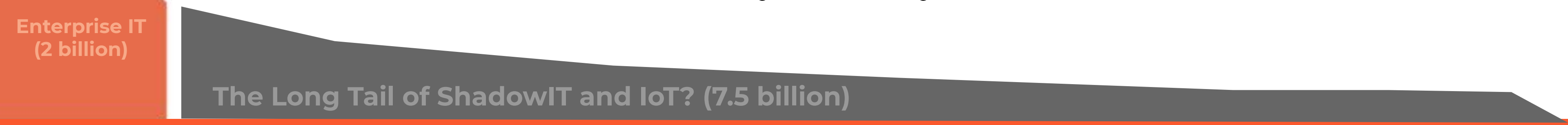



How does your SOC correlate and analyze event data, IoCs and other security- and threat-related data? ***Select those that most apply. (N=275)***

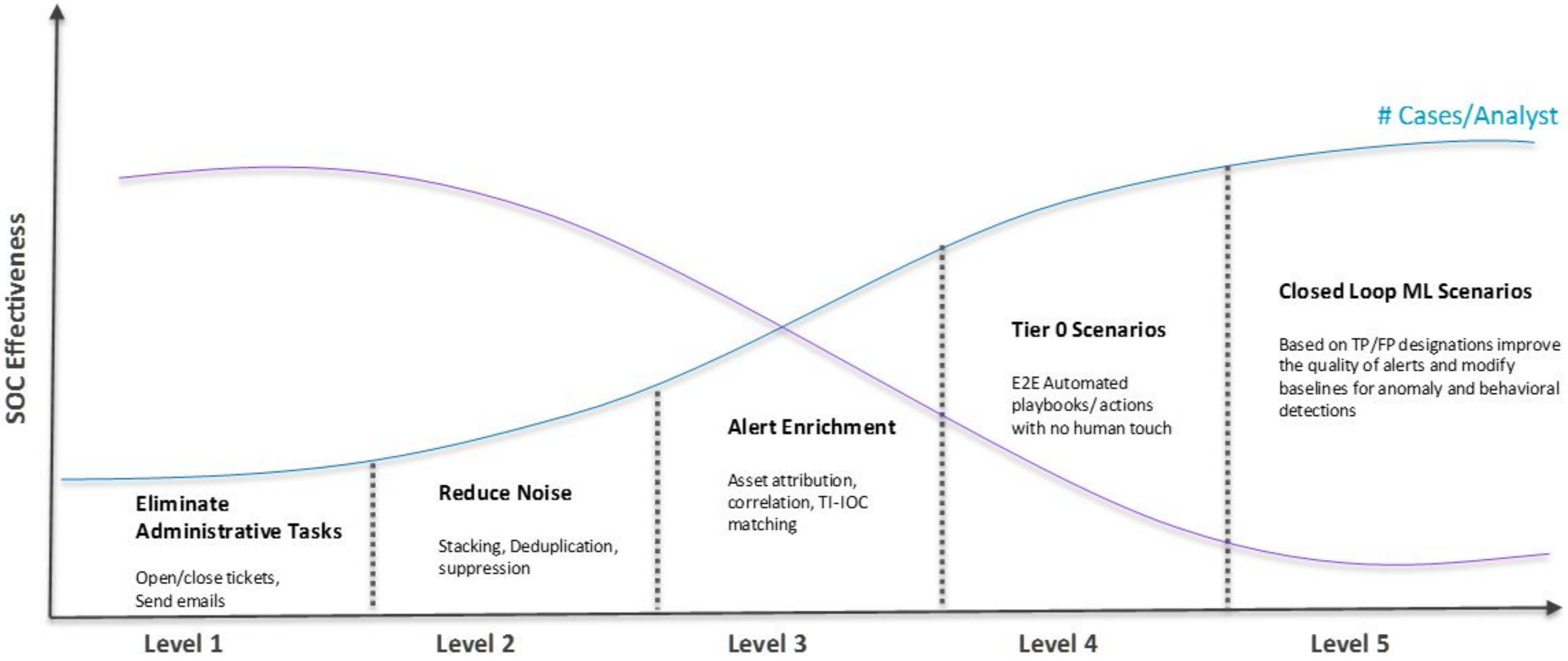
*SANS 2019 SOC Survey



1. Number of log sources != Maturity
2. Overwhelmed with low-fidelity data and false positives
3. Console burnout leading to ignored alerts
4. Fire-and-forget mitigation
5. Manual response = Lack prevention
6. No time to do investigations or hunting


Enterprise IT (2 billion)
The Long Tail of ShadowIT and IoT? (7.5 billion)

The automation maturity model and automation journey



© Microsoft Cyber Defense https://www.microsoft.com/security/blog/2017/08/03/top-5-best-practices-to-automate-security-operations/

Example automation benefits for SecOps functions

Threat Intelligence

↓

Automated enterprise sweeps, countermeasures

Security Monitoring

↓

Faster triage and investigation

Incident Response

↓

Faster countermeasure deployment

Vulnerability management

↓

Accelerated patching

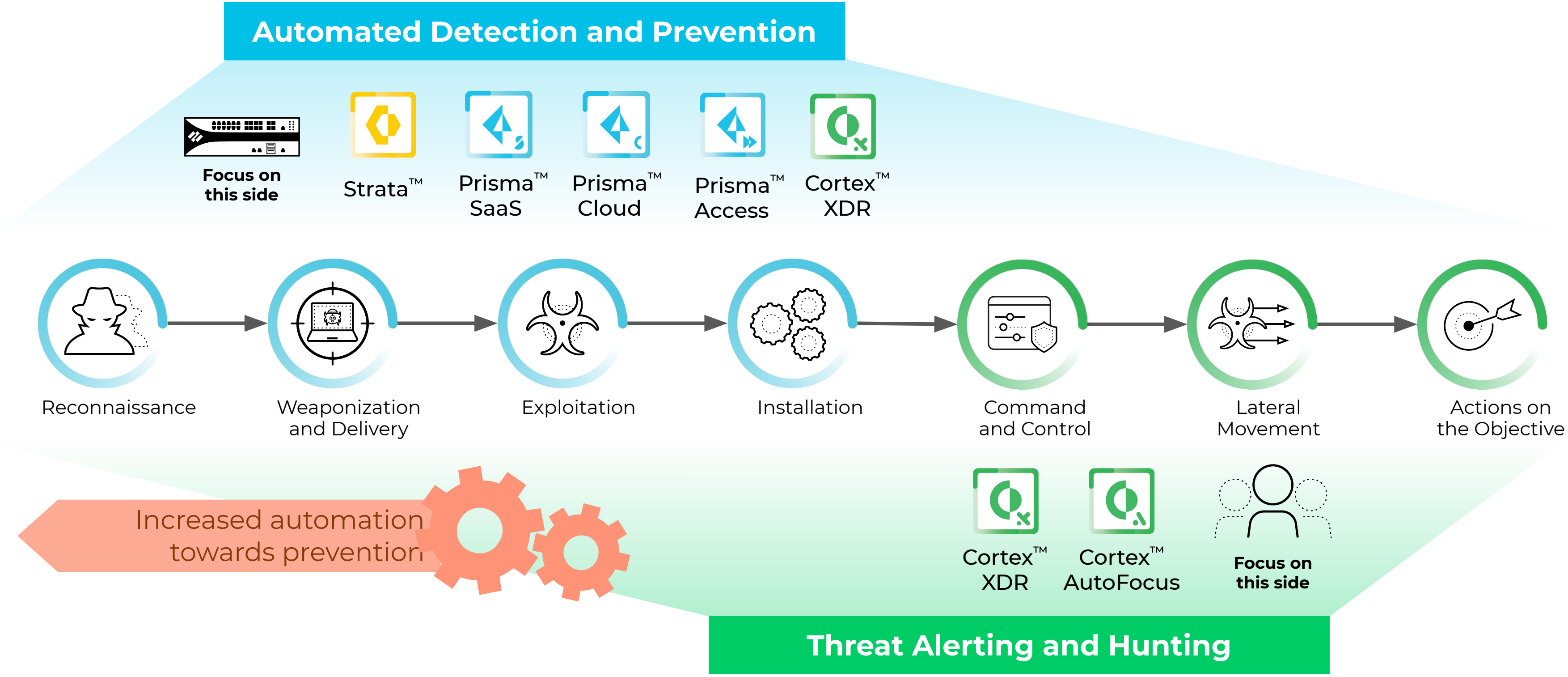
Red Team

↓

Continuous simulation of threats

More time to focus on what matters.

Focus People Effort on Right Side of Cyber Attack Lifecycle



Automated Detection and Prevention

Focus on this side

Strata™

Prisma™ SaaS

Prisma™ Cloud

Prisma™ Access

Cortex™ XDR

Reconnaissance → Weaponization and Delivery → Exploitation → Installation → Command and Control → Lateral Movement → Actions on the Objective

Increased automation towards prevention

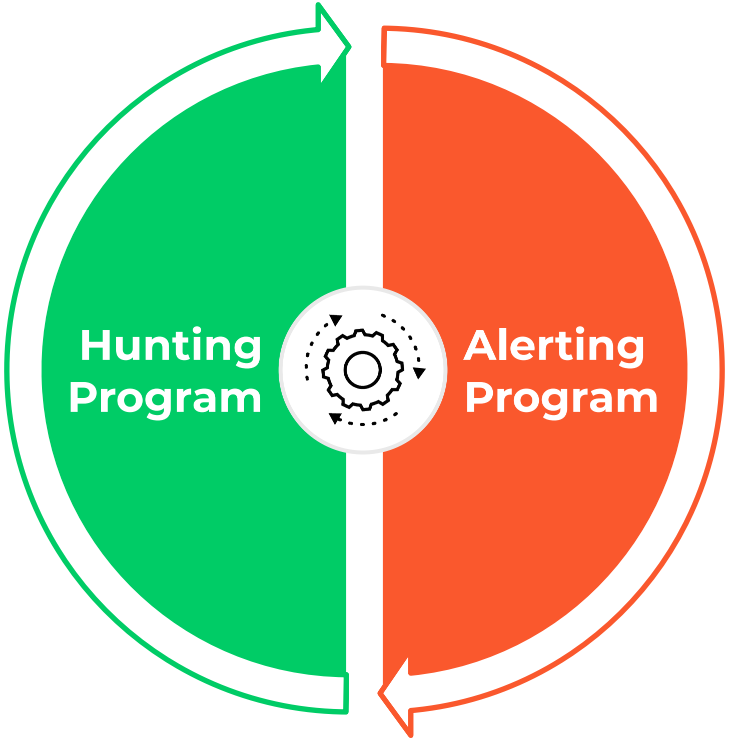
Cortex™ XDR

Cortex™ AutoFocus

Focus on this side

Threat Alerting and Hunting

Continuous Improvement of Alerts and Hunting



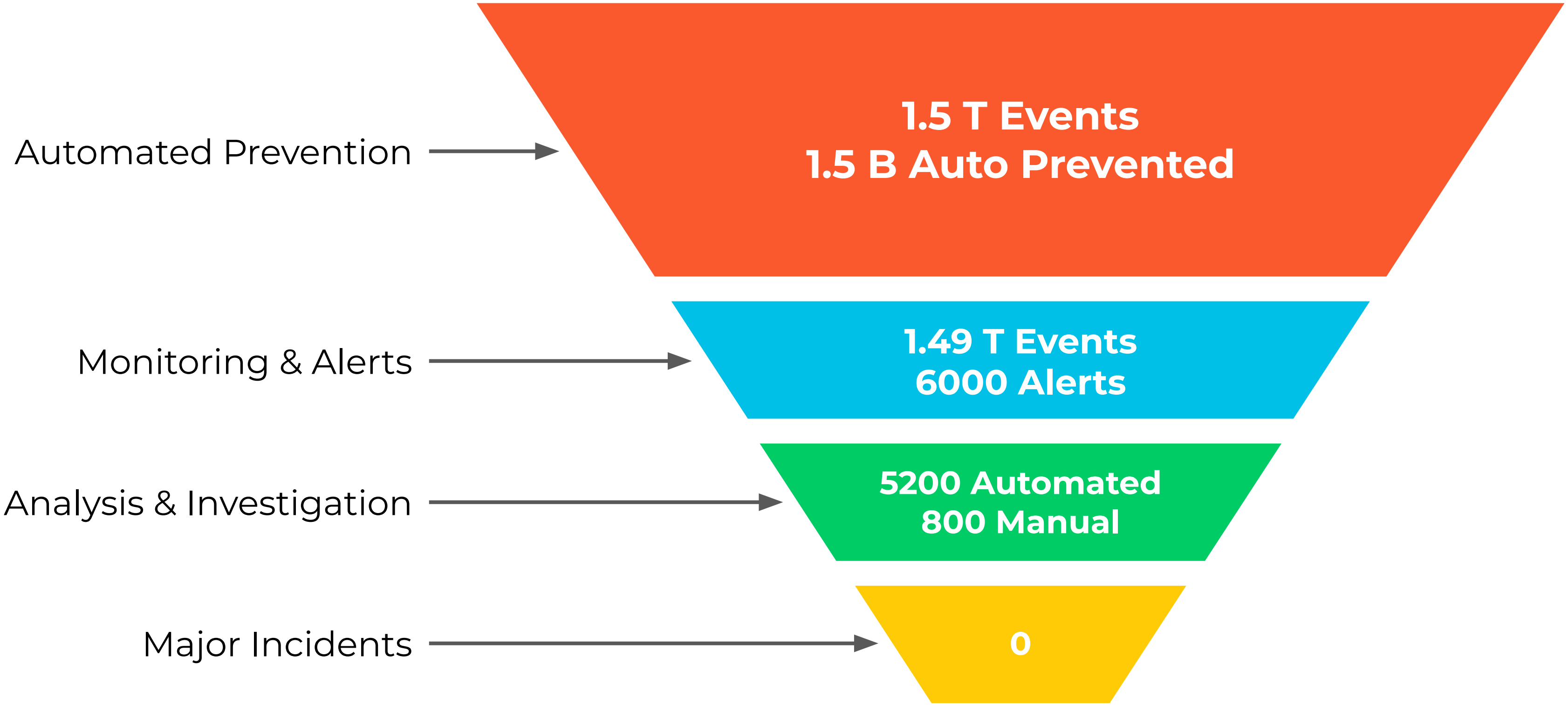
Dedicated hunt time

Move high fidelity Hunts to Alerting

Alerts must be 90%+ True Positive

Move low fidelity Alerts to Hunting

Log and Alert Volume (90 days)



Automated Prevention

1.5 T Events
1.5 B Auto Prevented

Monitoring & Alerts

1.49 T Events
6000 Alerts

Analysis & Investigation

5200 Automated
800 Manual

Major Incidents

0

Internally, Cortex XSOAR does the work of ~9 virtual FTEs

Automation Type	Count	Analyst Hours Saved
Enrich Alerts	1090	635.8 hours
De-duplicate alerts	7,783	648.6 hours
Ask user for more details (Email/Slack)	308	128.3 hours
Request re-image with IT (open Service-Now ticket)	5	2.1 hours
Coordinate password reset	4	1.7 hours
GCP Remediation	33	16.5 hours
Other Jobs*	*	29.8 hours

Repetitive tedious SOC work that nobody wants to do…

Total hours saved in 1 month:

1,463

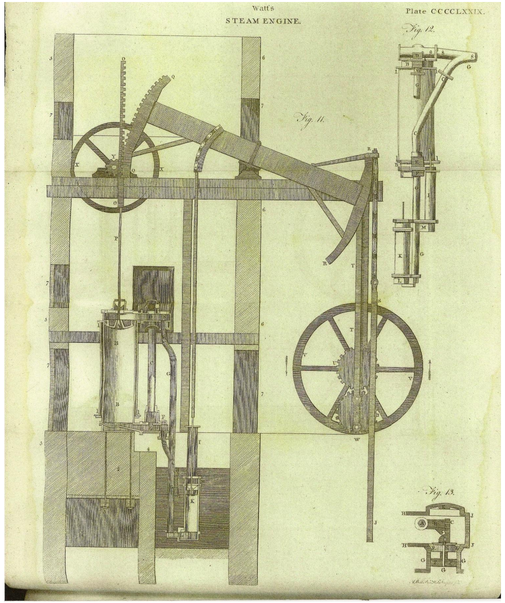
XSOAR automates the workload of 9.1 FTEs.

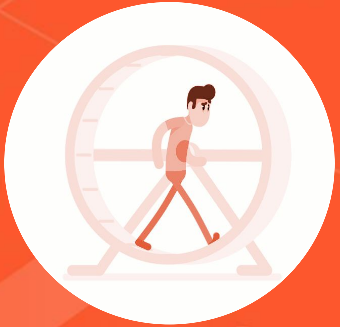
*PhishMe metrics, RSS feed job, content update job, hunting assignments and metrics, daily monitoring ticket creation

Conclusion

- Thank you!

- We presented:
  - Root components of AI,
  - Its subsequent (new) threats
  - AI as an opportunity for the attacker
    So why it’s not something you can deny
  - AI as an opportunity for the defender
  - Operational use case with SOC automation





Question?

rmarichez@paloaltonetworks.com