Cybersecurity Analytics

Clinton Firth
EY Cyber Leader
Africa, India & Middle East

EY
Building a better
working world

What is Artificial Intelligence?

It is a discipline

Machine learning
(supervised and unsupervised)

Predictive analytics

Key components for AI

70% Data

20% Processing
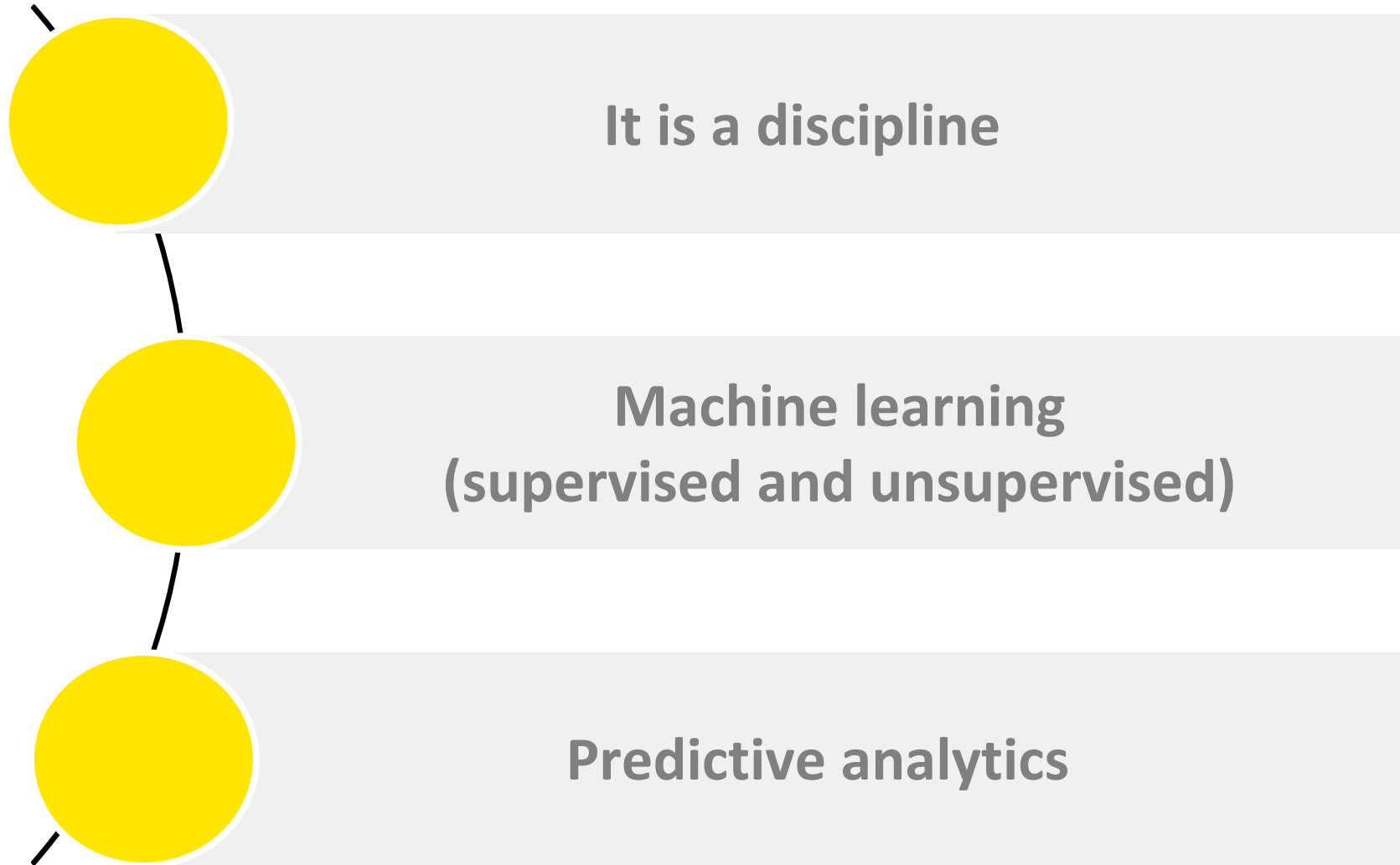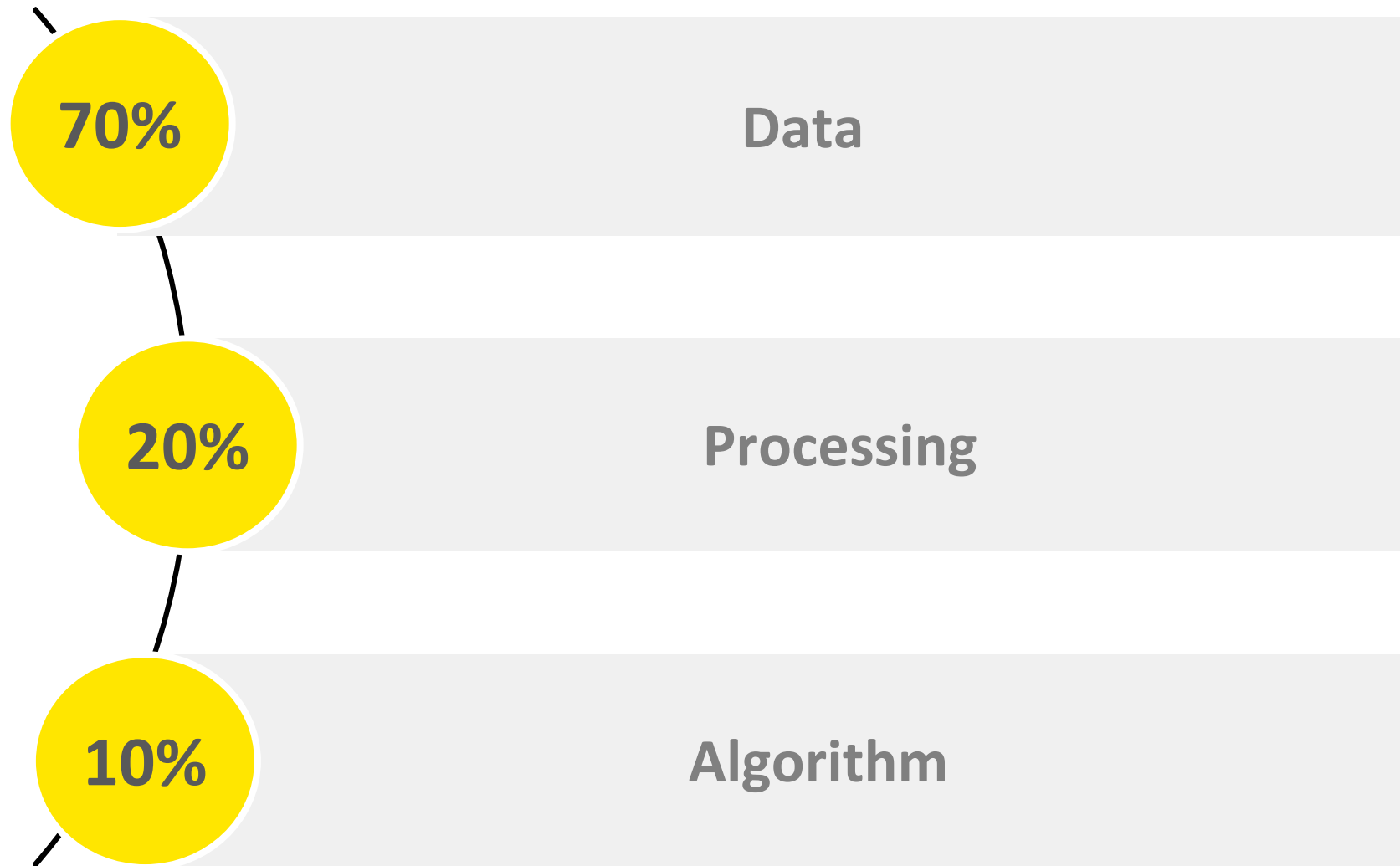
10% Algorithm

Algorithm example – internal attack

Probability that email is malicious (p_1)

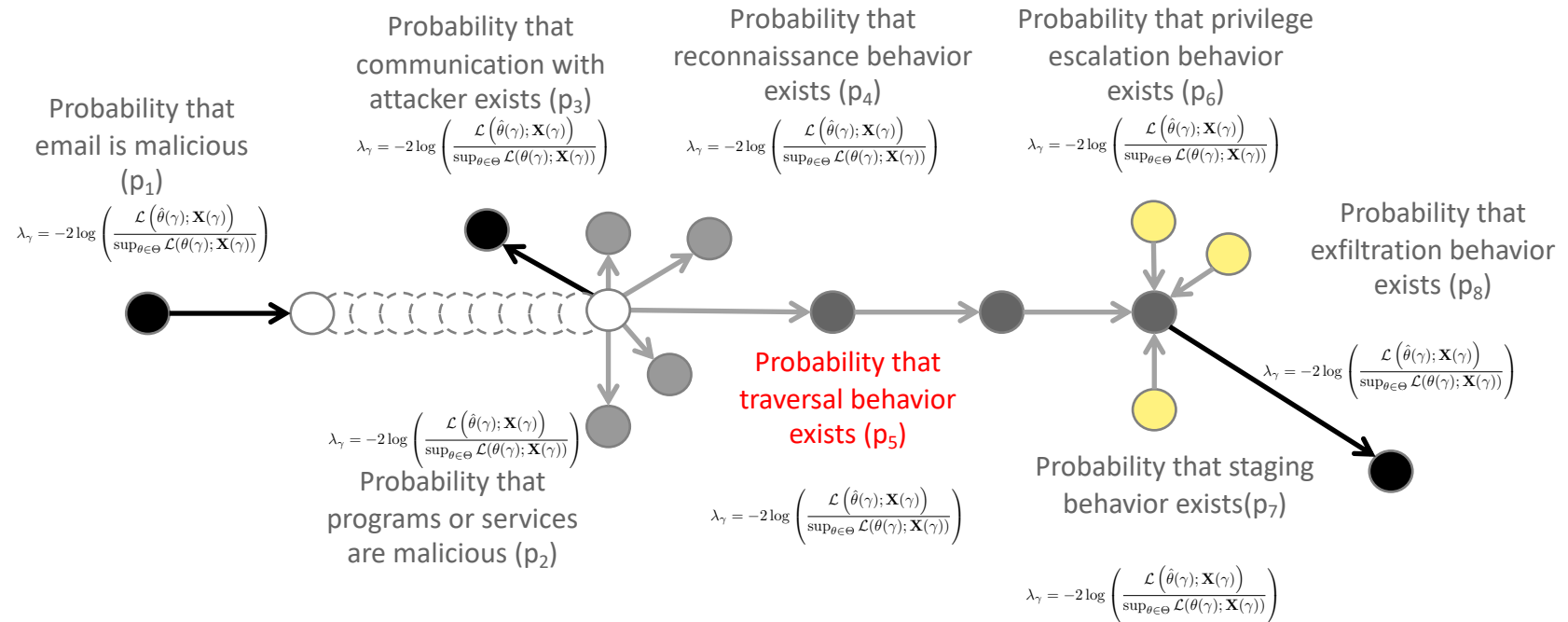$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

Probability that communication with attacker exists (p_3)

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

Probability that reconnaissance behavior exists (p_4)

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

Probability that privilege escalation behavior exists (p_6)

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

Probability that exfiltration behavior exists (p_8)

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

Probability that programs or services are malicious (p_2)

Probability that traversal behavior exists (p_5)

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$

Probability that staging behavior exists (p_7)

$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$
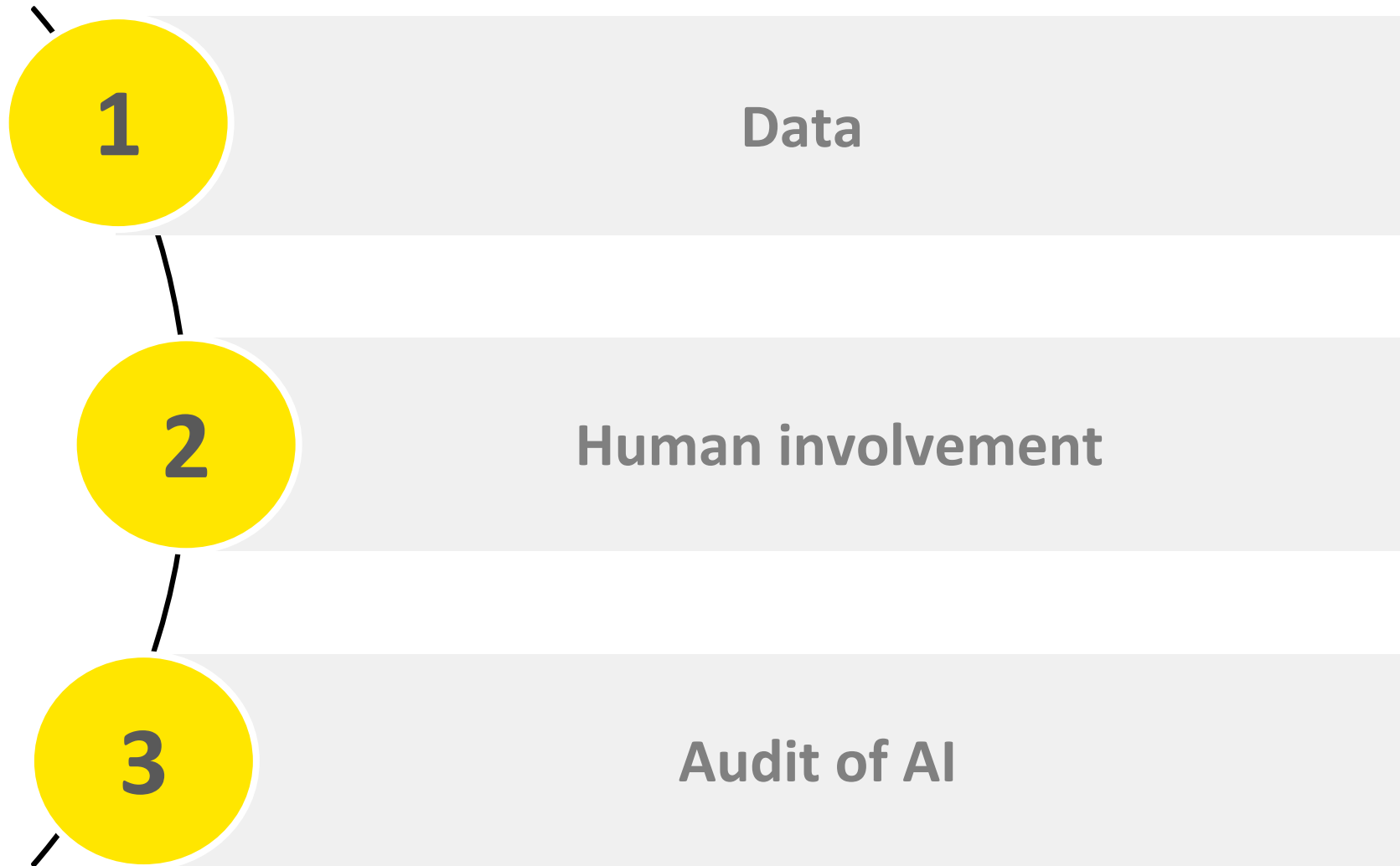
$$\lambda_\gamma = -2\log\left(\frac{\mathcal{L}\left(\hat{\theta}(\gamma); \mathbf{X}(\gamma)\right)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta(\gamma); \mathbf{X}(\gamma))}\right)$$

The overall probability of the above attack existing is a statistical combination of the individual anomaly scores.

Number of SIEM use-cases per category	Intelligence Gathering	Initial Exploitation		Command & Control		Privilege Escalation		Data Exfiltration		
	Background Research	**Initial Attack**	**Establish Foothold**	**Enable Persistence**	**Enterprise Recon**	**Move Laterally**	**Escalate Privilege**	**Gather & Encrypt Data**	**Steal Data**	**Maintain Presence**

Cyber Kill Chain →

Example attack vectors (MITRE ATT&CK)	Pre Attack	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Execution	Collection	Exfiltration	Command and Control
	Recon, Information Gathering	Browser Extensions	Exploitation of Vulnerability	Hidden Files and Directories	Hooking	Network Service Scanning	Remote Desktop Protocol	InstallUtil	Data Staged	Data Compressed	Custom Command and Control Protocol

Sample Attack Detection Counter-measures	• Threat intel feeds • Dark web crawlers • Deep web crawlers • Social media crawlers	• Intrusion detection/ prevention system • Perimeter log analysis • Sandbox detonator	• Network IOC monitoring • Endpoint IOC monitoring	• Network IOC monitoring • Endpoint IOC and logs monitoring	• Net flow anomalies analysis	• Net flow anomalies analysis (point to point)	• Network IOC monitoring • Endpoint IOC and logs monitoring • File integrity monitoring	• Net flow anomalies analysis	• Network IOC monitoring	• Network IOC monitoring • Endpoint IOC and logs monitoring

Primary Detection Method	Intelligence, Monitoring Services	Signatures, Application Whitelisting	Signatures, behaviors, application whitelisting		Signatures, behaviors	Anomalies and behaviors	Anomalies and behaviors	Anomalies and behaviors	Anomalies and behaviors	Anomalies and behaviors, app monitoring

Vendor logos by stage: iSIGHT PARTNERS, tenable, digital shadows, CROWDSTRIKE, NORSE | Check Point SOFTWARE TECHNOLOGIES LTD., CISCO, paloalto NETWORKS, FireEye | CROWDSTRIKE, CARBON BLACK, Fidelis Cybersecurity, FireEye | splunk>, QRadar, ArcSight, LogRhythm | DARKTRACE, CISCO | CYBERARK, tripwire, BeyondTrust | RSA SECURITY, LogRhythm, SOLERA | websense, Symantec Data Loss Prevention, McAfee DLP Endpoint | splunk>, QRadar, ArcSight, LogRhythm

Attack Time Horizon

Traditional Detection Shortfall	Hours to Months	Minutes to Hours	Months to Years	**170 days** mean time from original breach until organizations detect a cyber attack
	Cost of Breach to Organization 20% ≥		**Cost of Breach 80% ≤**	

Unsupervised machine learning + AI models and algorithm based. Lightweight full ecosystem visibility to fill capability voids in signature and behavioural based traditional cyber detection and prevention

EY Analytics	• AI based dark web crawler			• Not rule based	• Aggregated lateral movement		• Lightweight, not PCAP	
	EY CyberEYE	EY Malware Web based Phishing	EY Los Alamos Credential analytics / EY Los Alamos Sandoxing	EY Los Alamos PathScan®	EY Los Alamos PathScan®	EY Los Alamos Credential analytics / EY ID compromise	EY PathScan®	EY Exfiltration

EY Cyber Analytics is focused on improving efficacy of analysts through automation and advanced data science to reduce time from breach to detection

Key issues / considerations for AI

1 Data

2 Human involvement

3 Audit of AI

Thank you

EY | Assurance | Tax | Transactions | Advisory