White Paper

HardenStance

AI in Cyber Security: Filtering out the Noise

By Patrick Donegan, Principal Analyst, HardenStance

Sponsored by:

FORTINET JUNIPER NETWORKS kpn NOKIA

February 2019

HardenStance

"Trusted Research, Analysis and Insight in IT & Telecom Security"

Executive Summary

- Artificial Intelligence (AI) has proven potential to improve threat detection and enable critical, time-saving automation in cyber security operations.
- Many security teams are already reaping the 'low-hanging fruit' benefits of automated machine-learning that's embedded in their cyber security infrastructure. Early adopters of advanced AI use cases in End Point Protection (EPP), Network Traffic Analysis (NTA) and Security Orchestration, Automation and Response (SOAR) report significant challenges in learning how best to exploit AI's potential.
- AI can help security teams respond better to ambiguity and evolve to a more probabilistic model of security operations. But controlled introduction is needed to mitigate the very real new risks that AI creates.
- While ease-of-use and risk-related barriers to deep AI integration remain, security teams will need to rely on the expertise of vendors, MSPs and MSSPs.

Introduction and Definitions

The most frequently heard perspectives on Artificial Intelligence (AI) in cyber security tend to revolve around one of two themes. The first assumes that AI is in the midst of one of its periodic hype-cycles. The term 'AI' is grossly mis-used; most of what's being sold as 'AI' is nothing of the kind. Worse, in this view, AI carries more risk than opportunity. Hence headlines from leading media outlets like these from August 2018:

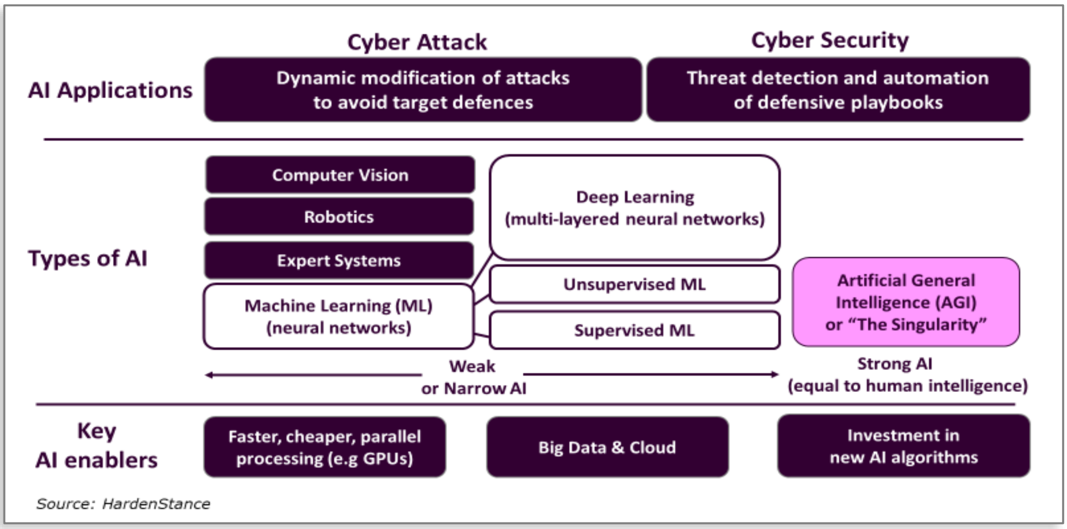
- "AI Security Hype Putting Businesses at Risk" (Computer Weekly)
- "AI for Cyber Security is a Hot New Thing and a Dangerous Gamble" (MIT Review)
- "Is the Hype Around AI Deceiving Cyber Security Professionals?" (Silicon Republic)

This White Paper aims to eliminate the many false positives and negatives that cloud much of the reality of AI adoption in cyber security to date.

The second perspective, advanced most often by a subset among prominent new entrant vendors in advanced threat detection, is that AI is the answer to all our problems and the silver bullet we need is already here. This White paper's starting point is that neither perspective is right. It aims to eliminate the many false positives and negatives that cloud much of the reality of AI adoption in cyber security.

The paper seeks to filter out this noise and provide as accurate an understanding as possible of how AI can help cyber security defenders today with minimal risk. A representation of the family of types of AI and AI nomenclature is depicted in **Figure 1**. The key technology enablers that have triggered the recent and marked upward spike

Figure 1: Varieties of AI and Cyber Applications



Cyber Attack
Cyber Security
AI Applications
Dynamic modification of attacks to avoid target defences
Threat detection and automation of defensive playbooks
Types of AI
Computer Vision
Robotics
Expert Systems
Machine Learning (ML) (neural networks)
Deep Learning (multi-layered neural networks)
Unsupervised ML
Supervised ML
Artificial General Intelligence (AGI) or "The Singularity"
Weak or Narrow AI
Strong AI (equal to human intelligence)
Key AI enablers
Faster, cheaper, parallel processing (e.g GPUs)
Big Data & Cloud
Investment in new AI algorithms

Source: HardenStance

in investment in AI are also shown. Also shown are the main ways that AI can be put to use - by cyber attackers as well as defenders.

Throughout the paper an AI algorithm is assumed to be one that classifies data according to probabilistic reasoning. The exact way it derives outcomes isn't pre-determined by human programming. As shown in **Figure 1**, this broad definition includes newer unsupervised and deep learning techniques. It also embraces traditional supervised – human-trained – machine learning that has been widely deployed for many years.

Some additional definitions are as follows:

Artificial General Intelligence (AGI), which is also called 'The Singularity', is a state of technology maturity in which machines possess the same cognitive powers as humans. This is widely assumed to be ten to twenty years away at the very least.

Augmented Intelligence: AI is an enhanced data analytics tool for augmenting the intelligence that is available to assist humans in accomplishing repetitive or complex tasks. AI is therefore a key enabler of augmented intelligence.

AI-assisted and AI-driven: The term 'AI-assisted' is used throughout this paper to describe products or services comprised of several different types of analytics algorithms, of which one or more AI instances is a subset. The term 'AI-driven' is used to describe products and services that are largely or entirely made up of AI algorithms.

The Potential of AI in Cyber Security

This paper's definition of AI embraces newer unsupervised and deep learning techniques as well as supervised machine learning that has been widely deployed for many years.

The growing volume and sophistication of cyber threats is overwhelming cyber security teams. The volume reduces the time available to do anything but fight fires. The sophistication is driving a change in the security operations model from one rooted in determinism (something is either a threat or it isn't) to one that is increasingly reliant on the interpretation of ambiguity and is driven by probabilism (this looks unusual and it may be threatening, let's take a look).

Designed and operated well, AI significantly improves detection rates through automation. It also generates high quality probabilistic findings that can help explain whether complex patterns that appear threatening are in fact malicious or benign. This is unique to AI – other algorithms in the security infrastructure can't do this.

The other core value of AI is that it saves time. And time is the security team's most precious asset. The faster an organization can prevent, detect and remediate both known and unknown threats, the stronger its cyber security stance.

Today, some of the load of manually trawling through huge volumes of known threats is already offloaded onto several security controls. Even the subset that requires manual investigation still tends to be overwhelming, however. AI can take on more of this load in three main ways:

- **Spotting outliers and minor modifications** to known threats that other security controls miss;
- **Flagging anomalies to learned patterns** of communications and application behaviour that are trusted across an organization's unique infrastructure.
- **Enabling the business case for automation in security operations**.

A simplified view of where and how AI can be applied in the cyber security infrastructure is depicted in **Figure 2** on page 5. As the following sections show, machine learning is already quite widely deployed in the background of IT security, assisting the detection efficacy of some security products. In many of these use cases the AI algorithms don't require much customization or any AI expertise on the user's part.

Artificial Intelligence in the Hands of Threat Actors

The Defense Advanced Research Projects Agency (DARPA) Cyber Grand Challenge of August 2016 provided an unnerving glimpse into the future of cyber security, one which we can already see starting to take shape today via the increasing role of machine automation in generating cyber attacks.

The Grand Challenge pitched eight autonomous hacking machines – defined as Cyber Reasoning Systems - against one another in a "Capture the Flag" contest to see which one of them would fare best at spotting unknown software vulnerabilities. The winner of the $2 million first prize was 'Mayhem', built by ForAllSecure, a Carnegie Mellon University startup.

As the Grand Challenge recognized, there is as much opportunity for attackers in leveraging AI as there is for defenders. Among the most attractive use cases for attackers leveraging AI is to improve lure content for social engineering attacks including by making chat bots sound more human.

A November 16th 2017 Wall Street Journal headline announced: "First AI-Powered Cyber Attacks are Detected." In the last year some well-known, brand- name, cyber security vendors have echoed this, stating categorically that they themselves have also seen the first AI-driven attacks. Most vendors have not committed themselves in this way, due in large part to the difficulty of attribution. A really high degree of confidence that a given attack was AI-driven requires reliable confirmation all the way to the source. That includes validation from a credible government security agency that has managed to capture and reverse-engineer the exact machine used.

Attribution is very difficult – if it hasn't happened already, it will

From the perspective of most security teams this "did they or didn't they use AI?" question largely misses the point. The more important reality is far simpler: AI takes the cyber security arms race to the next level. Whether the first AI-assisted attacks have or have not been launched yet is much less important than the inevitability that they will be – and probably soon.

Initially we should expect AI-assisted or AI-driven modules in initial breach and lateral movement phases. This will evolve across the kill chain until defenders face highly intelligent, fully automated, hacking machines. Hence for defenders, plotting a strategy for how best to take advantage of AI becomes mandatory rather than optional. You don't bring a knife to a machine fight; you don't bring a dumb machine to a smart machine fight.

Learning how to get the best out of AI engines is proving to be a steep learning curve.

The End Point Protection (EPP), Endpoint Detection and Response (EDR) and Network Traffic Analysis (NTA) use cases that command so much AI-related mindshare in cyber security today are a lot more disruptive of the traditional security model. Security teams are used to dealing with ambiguity in data. They're used to security tools that arrive at binary decisions (some of which prove to be wrong). They are not so used to security tools that can offer brilliant insight with a high degree of confidence when presented with one complex problem but then offer a high level of ambiguity when presented with the next problem. This is how AI engines tend to present their findings. Learning how to get the best out of them and apply that to each use-case – knowing when to trust an AI instance's conclusions and when not to – is proving to be a steep learning curve.

As will be shown, deploying a vendor's AI in any one security domain can have demonstrable value. For many early adopters that have invested in EPP, EDR or NTA over the last couple of years, this is their entrée into the emerging world in which AI instances are available to consult as trusted assistants in the context of a specific security domain or security product silo.

Longer term, far greater value lies in being able to extract AI learnings as well as apply them throughout the security infrastructure. The ultimate goal in cyber security is to take today's manually-driven and time-consuming defensive playbooks and automate them on a par with the way attack playbooks are automated as characterized by the cyber security kill chain. This is where the third primary use case – AI-assisted and AI-driven Security Orchestration Automation and Response (SOAR) platforms – comes in.

It's debatable whether fully automated, AI-driven, security operations is even achievable within twenty years. We should expect that a subset of targeted cyber threats will continue to be so sophisticated – and unravelling and remediating them will remain so complex – that even reducing human involvement to a bare minimum is still years away.

Most security teams have yet to take even the first step in integrating AI deeply into their operations. Even the world's most capable teams are still only starting out on this journey. But there is no doubt that the momentum is under way now. It starts with AI supporting humans. If it's done securely, it can evolve towards humans supporting AI.

The rest of this section looks at the rate of adoption of AI in the three primary protection, detection and response use cases in cyber security. It describes the common real-world experience of early adopters in each use case. It also reflects on the impact of these real-world experiences on the perceptions and realities of AI as cyber security tool.

In recent years several cyber security vendors have augmented their back-end analytics with machine learning.

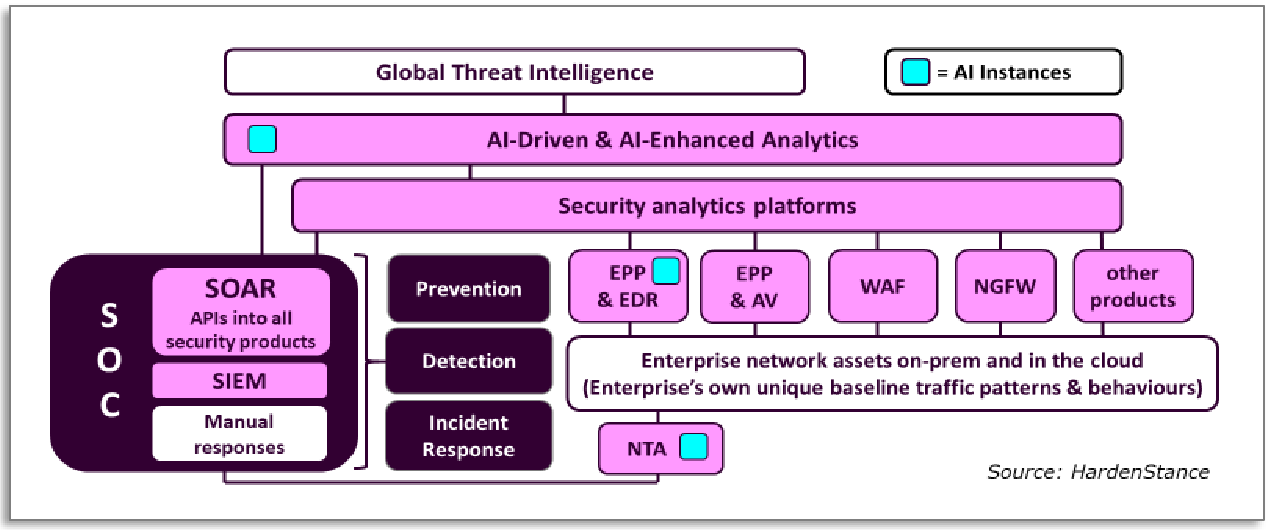
AI-assisted Prevention: The Low-Hanging Fruit

It's a reasonable estimate that hundreds of thousands of enterprises are already deriving some incremental protection from the use of supervised machine learning algorithms in their cyber security architecture. Many of the most popular preventative security tools like Web Application Firewalls (WAFs), Next Generation Firewalls (NGFWs), spam filters and antivirus (AV) products rely on back-end analytics applied to threat data in the cloud to drive their findings. In recent years, several vendors in these and other product spaces have augmented their back end analytics with machine learning in this way.

Many common threat prevention use cases are relatively straightforward. As quickly as adversaries automate tweaks in their attacks, the machine learning algorithm adjusts and automatically generates a signature for the new variant. Spotting polymorphic malware is one use case. AI-assisted use cases like this are the low-hanging fruit of AI in cyber security. Nowadays, some vendors attribute as much as 10% of the total threats they detect to a new machine learning algorithm added to their back end analytics.

Since the tasks assigned to the AI instance are fairly simple in cases like this, users need little or no AI expertise. Today, many users of AI-assisted products don't even know the algorithm is there. That's largely because until recently, vendors of AI-assisted products tended to see machine learning as an incremental evolution in their capabilities rather

Figure 2: AI Options in Cyber Security Operations


Global Threat Intelligence
■ = AI Instances
AI-Driven & AI-Enhanced Analytics
Security analytics platforms
S O C
SOAR
APIs into all security products
SIEM
Manual responses
Prevention
Detection
Incident Response
EPP & EDR
EPP & AV
WAF
NGFW
other products
Enterprise network assets on-prem and in the cloud
(Enterprise's own unique baseline traffic patterns & behaviours)
NTA

Source: HardenStance

than something to try and capture AI mindshare with. And because it's relatively easy to trust an AI instance for these relatively straightforward use cases, it's also proved relatively easy for users to allow the recommendations it generates to be automated.

The real-world experience of AI-assisted solutions is therefore making a significant, albeit small, contribution to security automation already. It hasn't been entirely painless either - many AI-assisted products are just as vulnerable to generating false positives as other products. But the experience of these initial use cases alone demonstrates a basic subset of some of the potential that AI offers security teams.

New Entrants in Advanced Threat Detection

The Ponemon Institute's "2018 State of Endpoint Security Risk" survey of 660 IT security professionals in the US provides valuable data on user experience with endpoint threat detection. It shows that 77% of respondents that suffered endpoint attacks that successfully compromised data assets and/or IT infrastructure said that the attack was a new or unknown zero-day attack. Just 19% said it was an existing or known attack.

Advanced threat detection uses AI to spot threats that are sophisticated enough to have evaded at least one set of other security controls.

With universal recognition that the security perimeter is permanently compromised and that more manpower alone cannot fix it, solutions that position AI as a tool for spotting complex threats inside the perimeter are pitched at a critical area of vulnerability.

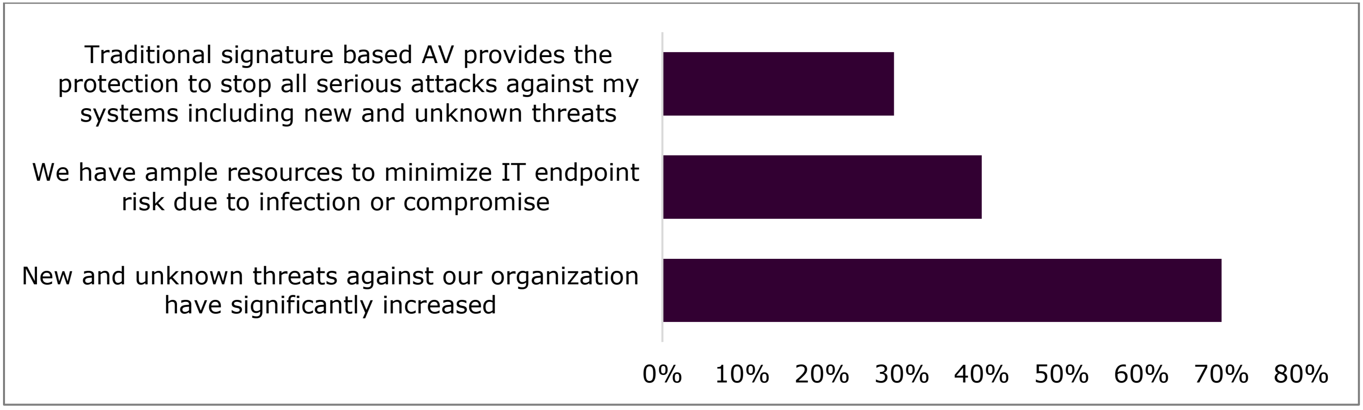
It's this space that AI-driven – and heavily AI-branded – new entrants in the EPP, EDR and NTA product spaces have entered. A number of these vendors state that they use unsupervised machine learning, which gives the AI greater autonomy to 'learn' independently of human inputs. With high levels of VC-funding and big marketing budgets, these firms have captured a lot of the AI-related mindshare in cyber security. They've also contributed to the creation of new product categories like threat hunting.

An Intersection between AI-assisted and AI-Driven Models

There is an intersection between the advanced threat detection models that AI-driven start-ups are pursuing and some AI-assisted approaches. For example, some traditional AV vendors have AI-assisted their products to pitch them into the EPP and EDR space. Equally, some AI-driven start-ups target AV now as well as advanced threat detection.

Advanced threat detection uses AI to spot threats – including dormant or heavily disguised threats – that are sophisticated enough to have evaded at least one other set of security controls. The complexity in identifying them sometimes lies in the fact that the exact same patterns and behaviours can indicate high risk in one unique enterprise environment but low or no risk in another. As shown on the next page, research by HardenStance, together with new and credible survey data, allows for a number of high level conclusions to be drawn about the day to day experience of security teams that have been early adopters of advance threat detection products.

Figure 3: Perceptions of Endpoint Security Risk



Source: The Ponemon Institute's "2018 State of Endpoint Security Risk" survey of IT security professionals

Ponemon's '2018 State of Endpoint Security Risk' survey yields these additional findings:

- **30% of mostly medium and large organizations are using EDR.** This is according to the survey's US-based respondents so it can be assumed that EDR adoption elsewhere and among smaller organizations is a lot lower.
- **Training takes time**. Only 23% of the survey's respondents who have deployed EDR said they had fully deployed it within three months or less.
- **False positives and false negatives are still too high.** In the case of EDR "a high number of false positives and IT Security Alerts" is the biggest challenge early adopters have faced according to 58% of the survey's respondents.

HardenStance's perspective on these data points is that they reflect the new nuance-driven operations paradigm that AI introduces. Even though they can suffer from many of their decisions to block or allow proving to be wrong, security teams are familiar with the traditional security model in which an AV tool at least makes a binary decision one way or another. With advanced threat detection products, users are taking time to adapt to a model in which their findings are ambiguous.

According to HardenStance's research, which is also echoed in the Ponemon survey data, six to twelve months training in EPP, EDR and NTA products is often needed. Moreover, the challenge can sometimes be complicated by vendors being unable or unwilling to commit to recommending that a high level of trust in their product's findings should be triggered at or above a specific score outcome.

Polarization among key industry players on the objectivity of testing of products that are mostly closely associated with AI is harmful.

In addition, the following are representative highlights of early adopters' experiences:

- **Some vendors are causing confusion with messaging that compares their own use of what they call "Advanced AI" with competitors that are "just" using machine learning or not using "real AI".** While not especially harmful, this is neither illuminating nor helpful. Some of the best CISOs continue to complain about lack of transparency in many of these products. KPN CISO, Jaya Baloo, for example, states: "it's still hard for me to poke through the marketing and find out what some of these products really do. Honestly, it can be hard to tell."
- **Independent testing frameworks aren't serving users well**. The architectures of many advanced threat detection products differ greatly, for example as regards their interaction with enterprise environments on-premises and in the cloud, as well as with their own cloud. Their security efficacy varies greatly in different use cases and environments – more so than is typical for other security products. The industry is used to legal battles between independent test houses and vendors over comparative test-results. But the September 2018 anti-trust suit filed by NSS Labs against three vendors and the sixty-member-strong Anti Malware Testing Standards Organization (AMTSO) has taken this to a new extreme. Such polarization among key industry players on the objectivity of testing methodologies for products that are most closely associated with AI is harmful. It clouds understanding of the value of AI - in endpoint security and more broadly too.
- **Data scientists with expertise in training AI systems are in very short supply.** Colleges and universities have only started offering AI-focused degree courses in the last two or three years**.** The availability of qualified people to train AI algorithms in cyber security applications - and train analysts to use them – directly impacts the value security teams are able to extract. Training out an AI-generated false positive tends to have more than one single effect, for example. People without data science backgrounds struggle with this. Hence to integrate AI deeply into their operations, security teams will need to rely on AI expertise that's accumulated and delivered as a service by their vendors; Managed Service Providers (MSPs); Managed Security Service Providers (MSSPs) and other specialists.

Figure 4: Currently Deployed AI Use Cases in Cyber Security

Use Case	Examples of AI-supported products & services	AI-assisted or AI-driven	AI expertise needed to operate	Potential impact on overall security posture	Share of AI learnings used to drive automated action as of today
Protection	WAF, NGFW,	Assisted	Low	Low	High
Detection	EPP, EDR, NTA	Assisted or Driven	Medium	Medium	Low
Response	SOAR	Assisted or Driven	High	High	Low

Source: HardenStance

- **There's a long way to go with automation**. Anecdotal feedback suggests that progress in automating defences based on AI-driven EPP, EDR and NTA findings is still in its early stages. This reflects the initial lack of familiarity with the trustworthiness of AI-driven findings in these use cases.

Proof-points in the EPP/EDR and NTA Space

Some AI-driven and AI-assisted advanced threat detection products are clearly delivering value to some customers in live production networks:

- Many new entrant vendors in this space can cite several customer testimonials stating that they have spotted high-risk threats that other products missed.
- One vendor has been widely feted for leading the investigation into the hack of the Democratic National Committee (DNC). Another has been celebrated for spotting how hackers exfiltrated data from a casino via a fish tank in a Las Vegas hotel lobby.
- Most of the world's largest MSSPs have either integrated some of these AI-driven vendors' products into their own managed services portfolio or are actively re-selling them (or both). These companies have some of the very best product testing environments. Products that have proved themselves in these exacting test environments are offering added value (although many smaller, less advanced organizations aren't necessarily skilled enough to extract that value).

Many of these vendors can cite several customer testimonials stating that they have indeed spotted high risk threats that other products missed.

AI in Incident Response (IR)

According to the Ponemon Institute's March 2018 'Third Annual Study on the Cyber Resilient Organization' surveying 2,848 IT and IT security professionals around the world, 77% of organizations don't have a formal cyber security incident response (IR) plan applied consistently across the organization.

This is a powerful data point. It shows that whereas investment in prevention and detection are generally considered table stakes in cyber security, IR is still considered something of a high-end capability that only a small minority of businesses are currently investing in (though many more find they have to make an emergency call to summon IR providers when hit by a big breach).

Today, most IR work consists of human analysts leveraging security tools to carry out investigations and implementing remediation actions leveraging vast amounts of event logs and other data available to them. It can take months to plough through this data to understand exactly what has occurred. During that time the adversary may well remain embedded in the network while the organization's costs continue mounting

Since the costs to an organization following a major incident are higher the longer it takes to respond and remediate it, what IR teams want above all else is to reduce the time this takes through automation. That requires reducing the time it takes to execute responses from weeks and months to days or hours through defence playbook automation. But that also requires reducing the time taken to understand what has happened – and to decide what to do – by a similar amount.

SOAR platforms are becoming increasingly important in terms of response automation. These are designed to leverage open APIs to reach into an organization's entire security operations, across all prevention, detection and response domains. They are designed to dynamically adjust workflows to optimize an organization's security posture. Many big security vendors have invested in SOAR start-ups in recent years. There are several independent start-ups too.

SOAR platforms have stand-alone value with or without AI. But integrating AI into the IR investigation and decision making process - and feeding those outcomes into the SOAR environment for automated response - is where big wins in improved security efficacy and time-saving are available to security teams.

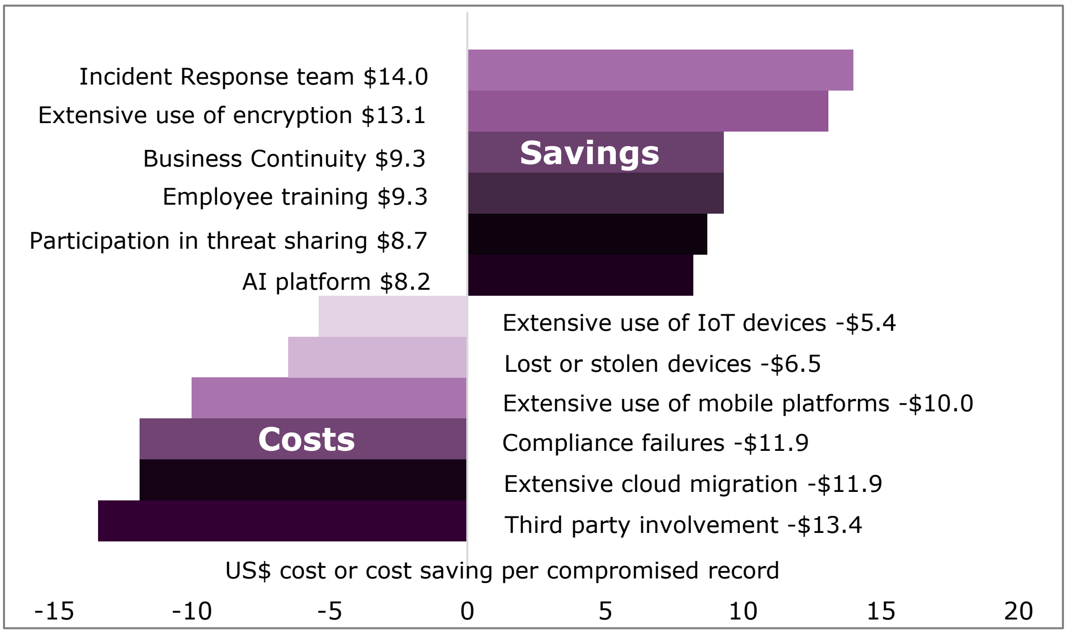
Despite it being early days, there is nevertheless some empirical evidence that use of AI is starting to prove its worth in Incident Response.

It's clear from both anecdotal and empirical evidence that it's only a small subset of organizations that have begun automating their IR environment. The Ponemon Institute's "2018 Cost of a Data Breach" survey shows just 15% of respondents stating that security automation is "fully deployed" in their organizations. The typical pattern is for IR processes to be comprised of several steps – often several dozens of steps – and for a subset to be automated selectively as and when it's considered safe to do so.

Despite it being early days, there is nevertheless some empirical evidence that use of AI is already starting to prove its worth in protecting organizations against the impact of data breaches. As shown in **Figure 5,** having an AI platform came in sixth behind having an IR team and extensive use of encryption among factors considered key by users to saving on the costs of a data breach. On average, the Ponemon survey respondents said that having an AI platform saves $8.20 cents per compromised record.

This is an AI use case which should increasingly be prioritized. The goal that the cyber security community should be targeting by combining AI and SOAR platforms is a level of defensive playbook automation that's on a par with attack playbook automation as characterized by the cyber security kill chain. The industry is a very long way from achieving that target today. However, it's clear that AI can have a key role to play here – and its value is already starting to be recognized.

Figure 5: Factors that most impact the per capita cost of a data breach



Source: The Ponemon Institute's July 2018 "Cost of a Data Breach Study"

Nine Key Challenges and Recommendations

This chapter draws on the experience of early adopters to point to the following key challenges and recommendations for security teams that are planning for, or just beginning, to extend AI into their environment. They span key issues relating to ease of use; risk management and compliance.

It's worth bearing in mind that while some of these challenges are relevant to adversaries as well as defenders, adversaries are free to navigate the AI adoption curve without the legal and ethical constraints that defenders are bound by.

It's also worth considering that most organizations – even many large ones – don't have the expertise to navigate these challenges by themselves. Acquiring that expertise will be very challenging for all but the most well-resourced organizations. There is therefore a huge opportunity here for vendors, Managed Service Providers (MSPs), Managed Security Service Providers (MSSPs) and other third party players to accumulate expertise in the application of AI to cyber security and sell that as a managed service over the coming years.

The 9 key challenges and recommendations around deploying AI in cyber security operations are as follows:

1. Protecting the Integrity of AI Instances

Conventional security controls as well as emerging detection and repair methodologies can protect the integrity of AI instances.

An AI algorithm can be hacked, just like any other. Poisoning attacks on training data can corrupt the learning process and the final trained model. Adversarial samples can be crafted to fool the AI and cause it to mis-classify. Autonomous cars have been tricked like this into mis-interpreting traffic signs just by minor alterations to their appearance.

The prospect of a security tool making corrupted recommendations – or worse, automating the implementation of those recommendations – is among the greatest risks that AI presents to security teams. Conventional security controls as well as emerging detection and repair methodologies can protect against this. Security teams should also plan for algorithm diversity rather be over-reliant on a master AI algorithm.

2. Ensuring the Quality of Training Data

AI can only deliver valuable insights in cyber security if it is trained with a lot of correctly labelled good and/or malicious samples. This is the 'garbage in/garbage out' principle. Some estimates suggest that the amount of data needed for unsupervised machine learning is a hundred times the volume collected today by a SIEM.

Where relevant to the use case, security teams should grill vendors on how much data they've used to train their systems and what those sources of data are. Where relevant, users should be able to train systems with their own data. Greater threat intelligence sharing would also help here – at the level of individual companies; government agencies and specific industry sectors. The work of the Cyber Threat Alliance is a leading example here that users should welcome and support. As AI in cyber security matures, the data and threat data lakes of larger players will confer advantages. Players that can accumulate data across the cyber security infrastructure will have a potential advantage over those whose product footprint is confined to just one domain.

3. Ensuring Interpretability of AI Findings

Transparency is a huge topic in AI. Some amount is critical to enable AI to be accountable to those humans that use it – and those that are impacted by it. In cyber security there are self-evident limitations on the extent to which algorithm designers should be required to expose exactly how they work. Proprietary IPR needs to be protected – not just from competitors but from attackers too.

The key requirement for AI in cyber security is interpretability. Security analysts need to be able to understand at a high level how an AI instance has arrived at its conclusions.

That's 'how' in the sense of what data inputs have been correlated and in what way; not 'how' in the sense of what techniques have been used to arrive at a decision to correlate them.

This is absolutely critical. A security analyst's career hinges on making consistently good calls in their analysis and recommendations. To be able to trust an AI's conclusions, it's imperative that analysts have enough information about the methodology used to be able to gauge the validity and relevance of its findings. If the analyst can't properly interpret an AI instance's findings, the correct thing to do is ignore them.

4. Enabling Ease of Use

Interpretability is one component of ease-of-use but there are others. Younger security analysts, in particular, prefer visual and linguistic aids to reading data sheets and inputting commands when interacting with machines, including AI-driven machines. Being able to interact with, and query, AI engines using natural language (including language translation) and visual prompts will be increasingly important in driving usage.

Most of the experience of AI in cyber security to date is that offloading tasks onto AI-assisted or AI-driven machines creates an opportunity to upskill security analysts.

5. Reducing Design Debt

Design debt or technical debt is a well-known phenomenon in programming. It's the long term cost of cutting corners, doing things too quickly or cheaply at the outset which has to be paid further down the line. Technical debt in applying AI to cyber security is increased risk exposure. One of the keys to avoiding it is for the AI strategy to be driven by qualified data scientists supported by strong cyber security domain expertise.

6. Gauging the AI's Autonomy

Autonomy is another one of AI's double-edged swords. A high level of autonomy can identify patterns and correlations much faster than humans can. But giving an AI instance a lot of "freedom" to learn autonomously creates a risk of unwanted consequences. With some advanced threat detection vendors now marketing unsupervised learning algorithms, it can be difficult to gauge the autonomy of a given vendor's solution, and hence the level of risk associated with deployment. One option could be industry standards that generate certification or labelling of AI algorithms assigning them a score that benchmarks the level of autonomy that they're capable of.

7. Overcoming Human Resistance to AI

The fear of job losses arising from AI is a very real issue in many industries. Due to huge shortages of human analysts and the growing sophistication of many threats, cyber security isn't one of them. Much of the experience of AI in cyber security today, especially as deployed in EPP, EDR, NTA and IR use cases, is that offloading tasks onto AI-assisted or AI-driven machines creates an opportunity to upskill security analysts and support their progress from Tier 1 to Tier 2 or Tier 2 to Tier 3 level.

Where AI systems are effectively deployed, they can lead to greater job satisfaction by equipping analysts with better consultative tools to get the job done. Organizations should point to customer testimonials verifying this – for example from leading SOAR and IR as well as EPP/EDR and NTA vendors. Some of these customer testimonials point to the value of advanced AI tools for attracting new talent as well as reducing churn within the security team.

8. Achieving Transparency in Product Testing

The introduction of AI into cyber security products can create new challenges in product testing. In the EPP/EDR space, for example, it's important to test for a product's dependence on sandboxing when the AI encounters malware that it can't classify. If the design default is to always forward unclassifiable samples to a sandbox, many early adopters have found that this can have a huge impact on the CPU.

As previously stated, there are too few trusted independent test frameworks available for medium and small enterprise vendors organizations to draw on. There is also too much polarization within the industry about optimal approaches to third party testing.

Greater industry collaboration is increasingly needed to arrive at greater transparency, consensus and standardization around testing methodologies. NetSecOPEN and AMTSO are positive examples of many security vendors partnering within an organization to do that. NetSecOPEN is even submitting its work for ratification to Standards Development Organizations (SDOs) like the IETF. To avoid vendors dominating these organizations, new approaches are needed to incentivize users to participate in them.

9. Regulatory Compliance

There is both risk and opportunity in AI from a compliance perspective. Poorly controlled AI runs the risk of breaching data protection regulations like GDPR. By the same token AI can also be extremely useful in helping locate, connect and present the data that's required by those same regulations a lot faster than manual efforts would.

Poorly controlled AI runs the risk of breaching data protection regulations like GDPR.

Also, many of the world's largest and most valuable data lakes belong to global commercial entities. These include the big cloud providers, some of whom are in the security analytics business. Whilst security teams have an opportunity to use these organizations' data lakes as part of an AI strategy, there's also risk attached if usage requires dropping their own data into that cloud provider's cloud.

Here, security teams should look to workarounds like multi-partite encryption. This ensures that a host company can only access a guest customer's data with its permission, using shared encryption keys that the guest customer manages.

Figure 6: 9 Key AI Challenges and Recommendations

Challenge	Description	Solution
Protecting AI integrity	Vulnerability of AI itself to attack, leading to corruption of its findings.	Detection and repair methodologies.
Ensuring quality data	Garbage in, garbage out.	Data scientists. Vendor transparency. Threat intel sharing. Scale.
Interpretability of AI findings	AI's findings can't be trusted without understanding its high level methodology.	Vendor transparency.
Ease of Use	Analysts need to enjoy using AI	Visual and language aids.
Reducing design debt	Exposure to future software vulnerabilities arising from focus on near-term gains.	A long term, risk-aware strategy driven by data science.
Autonomy of AI algorithm	Need for an understanding of what limits there are to the AI's autonomy.	Industry standards and certification.
Human resistance	Security analysts can feel threatened with job insecurity by AI-assisted machines.	Communication around AI's limitations and upskilling impact.
Transparency in product testing	High quality, transparent frameworks for testing AI-assisted security products.	Industry standards and more user involvement in creating standards.
Regulatory Compliance	Protection of privacy when applying AI to data.	Alignment with GDPR and use of appropriate encryption techniques.

Conclusion

The binary opinions that are most commonly held regarding the application of AI to cyber security are unhelpful to security teams. The experience of early adopters demonstrates that there is both tremendous opportunity and significant risk the more deeply AI is integrated into cyber security operations. Arriving at an understanding of both and creating a safe framework for exploiting AI's clear potential will take time and will evolve over a number of years. Security teams that haven't yet embarked on this journey need to do so as this is where both attack and defence playbooks are headed.

Cyber security leaders among vendors, MSPs, MSSPs, test houses and other stakeholders need to work together to make the best use of limited resources, including in people that have both data science expertise and cyber security domain knowledge. While pursuing commercial goals in a highly competitive market place, industry leaders also need to look for ways to rise above commercial rivalries and help users arrive at a balanced perspective of opportunities and risks. They need to make AI capabilities more understandable and usable to users. ■

About The Sponsors

About Fortinet

Fortinet (NASDAQ: FTNT) secures the largest enterprise, service provider, and government organizations around the world. Fortinet empowers its customers with intelligent, seamless protection across the expanding attack surface and the power to take on ever-increasing performance requirements of the borderless network - today and into the future. Only the Fortinet Security Fabric architecture can deliver security features without compromise to address the most critical security challenges, whether in networked, application, cloud or mobile environments. Fortinet ranks #1 in the most security appliances shipped world-wide and more than 375,000 customers trust Fortinet to protect their businesses. Learn more at https://www.fortinet.com.

About Juniper Networks

Juniper Networks simplifies the complexities of networking with products, solutions and services in the cloud era to transform the way we connect, work and live. We remove the traditional constraints of networking to enable our customers and partners to deliver automated, scalable and secure networks that connect the world. Additional information can be found at Juniper Networks (www.juniper.net) or connect with Juniper Networks on Twitter, LinkedIn and Facebook.

About KPN

KPN is the leading telecommunications and information & communications technology (ICT) service provider in the Netherlands. KPN offers a broad portfolio of services to the business (SoHo, SME, large & corporate enterprise), consumer and wholesale market, varying from fixed and mobile telephony, fixed and mobile internet, and TV to a wide range of ICT services, such as cloud, workspace, internet of things and security. KPN makes life more free, fun and easy by connecting people. KPN is passionate about offering secure, reliable and future-proof networks and services, enabling people to be connected anytime, anywhere, whilst at the same time creating a prosperous and more sustainable world. More information can be found at www.kpn.com

About Nokia

We create the technology to connect the world. Powered by the research and innovation of Nokia Bell Labs, we serve communications service providers, governments, large enterprises and consumers, with the industry's most complete, end-to-end portfolio of products, services and licensing. We adhere to the highest ethical business standards as we create technology with social purpose, quality and integrity. Nokia is enabling the infrastructure for 5G and the Internet of Things to transform the human experience. nokia.com

About HardenStance

HardenStance provides trusted research, analysis and insight in IT and telecom security. www.hardenstance.com