Security Gets Smart with AI

G.W. Ray Davidson

SANS

A SANS Survey

Security Gets Smart with AI

Written by **G.W. Ray Davidson, PhD,** and **Barbara Filkins**

March 2019

Sponsored by:
Cylance

Analyst Program

Executive Summary

The concept of artificial intelligence (AI) has been with us since the term was coined for the Dartmouth Summer Research Project on Artificial Intelligence in 1956. Today, while general AI strives for full cognitive abilities, there is a narrower scope—this more well-defined AI is the domain of machine learning and other algorithm-driven solutions and is the point where cybersecurity has embraced AI.

Cybersecurity professionals are always on the lookout for tools to help them to deal with threats and attacks. As more organizations adopt a policy of continuous monitoring, security teams find themselves with voluminous quantities of monitoring and other operational data. At the same time, computing power and the science of "data science" has progressed to the point where we can use machines to exhaustively sift the data to detect patterns and then use the patterns to create predictions, which can be tested to create more data, and so on. Given the ready availability of resources—as well as difficult problems and the largely unknown limitations of AI—there is considerable interest in the subject.

However, a lack of clarity remains around AI in general. How do security professionals view AI in terms of its maturity and fundamental capabilities? How are they utilizing AI? What key technologies underlie AI implementation? What risks and barriers are holding back wider adoption? In light of the increasing interest and the lack of clarity, SANS surveyed the information security community in late 2018 to access and characterize perceptions of AI's capabilities and potential to create this report.

Recognizing that communication is the basis of better security, we summarize the results in this paper to facilitate communication among developers, providers and users of AI technology in the information security community.

Key Findings

- **85%** see AI as an enabler, with 67% believing that traditional tools will remain in force.
- **46%** consider AI-based security solutions as "maturing," although lack of technology maturity is a leading barrier to adoption.
- Top three technologies considered part of an AI-enabled solution:
  - Predictive analysis (**76%**)
  - Deep learning platform (**74%**)
  - Machine learning platform (**73%**)
- Top three uses for AI:
  - Cyber defense (**75%**)
  - Malware prevention (**71%**)
  - Advanced threat detection/ prevention (**69%**)

Setting the Context: Demographics

SANS directed this survey at professionals working or active in cybersecurity, and involved with or interested in the use of AI for improving the security posture of their organization. Entities of all sizes are represented, as shown in Figure 1. More than 60% represents organizations with 5,000 or fewer workforce members, weighting this survey toward small- to medium-sized businesses.

What is the size of the workforce at your organization, including employees, contractors and consultants?

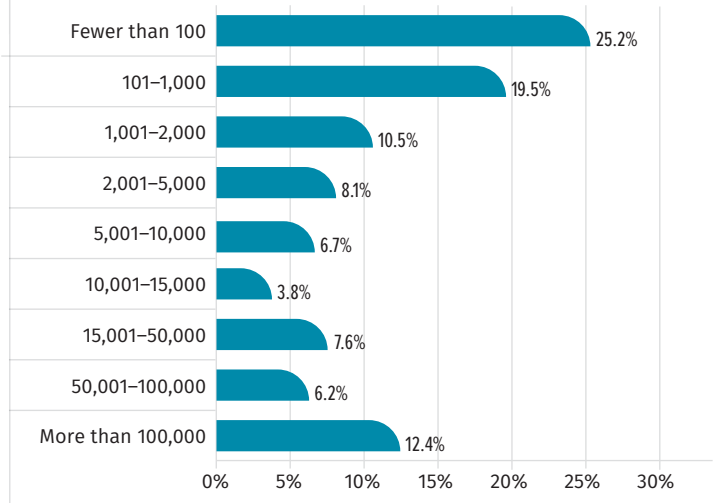

Figure 1. Size of Respondent Organizations

The top five industries include both suppliers, technology and service firms, whose interest is to enhance a core competency through AI, as well as verticals focused on protecting the organization and its intellectual property, allowing it to concentrate on its core business:

1. Technology (17%)
2. Cybersecurity service provider (14%)
3. Banking or finance (11%)
4. Education – beyond Grade 12 (8%)
5. Government – federal (7%)

Traditionally, management and staff have different perspectives on technology—staff uses the technology, while management is responsible for budget and resources as well as ensuring successful project completion. To evaluate whether perceptions around AI vary by role, we categorized respondent roles according to whether they were management (including C-suite positions) or technical staff, and by security or non-security (including IT, compliance and business-related roles). See Table 1.

Table 1. Role Categories[1]

	Security	Non-security
Management (including C-suite)	10.0%	13.8%
Staff (analysts and admins)	31.4%	29.1%
Total	**41.4%**	**42.9%**

Defining and Confining Artificial Intelligence

Artificial intelligence is a broad term, implying the capability to reason and "think" as a human, while exhibiting the following key characteristics of human intelligence:[2]

- Altering behavior based on past experiences, e.g., when encountering new and unseen situations (learning)
- Engaging in the encoding, storage and retrieval of past experiences (memory)
- Drawing logical conclusions and generalizing/deriving rules based on sample data (reasoning and abstraction)
- Systematically coming up with possible solutions and deriving the best answer to a problem (problem-solving)
- Generating multiple solutions to a given problem (divergent thinking)

In the pursuit of this lofty goal, researchers have developed numerous supporting technologies, including digital image processing, natural language processing, neural networks and so on. The price/performance ratio of computing power has continued to decrease, and companies have continued to automate and instrument their operations, continuously accelerating the growth in generation and availability of data (especially machine-generated data). This has, in turn, supported the development of a new field—data science—and new advances in underlying algorithms and methods of analyzing and interpreting the new flood of data.

[1] Categories do not sum to 100%. We included only those respondents who selected a specific title from the answer options provided; for this table we excluded the "Other" category.

[2] "Artificial Intelligence Framework: A Visual Introduction to Machine Learning and AI," https://towardsdatascience.com/artificial-intelligence-framework-a-visual-introduction-to-machine-learning-and-ai-d7e36b304f87

While the umbrella of AI encompasses many different technologies, the interest in AI for cybersecurity is being driven by the newfound wealth of data and analysis methods—to mix a metaphor, researchers are searching for new ways to find and use the needles in the data lakes, using machine learning and predictive analytical techniques. Using these technologies, computers can be trained to accomplish specific tasks by processing large amounts of data and recognizing patterns in the data.

An Actionable Framework for AI-Enabled Security

We first asked respondents what they considered to be the basic capabilities of an AI-enabled solution for cybersecurity. According to one classification schema, an AI-based platform should exhibit the following capabilities, working in concert:

- **Discover—**Learn from data without upfront human intervention.
- **Predict—**Provide alternatives about what will happen in the future, based on the understanding of the data sets through intelligent discovery.
- **Justify—**Explain actions taken, making outcomes recognizable and believable. Note: There is a distinction between *transparency* and *justification*. The former tells the user what algorithms and parameters were used; justification provides the why.
- **Act—**Provide the user experience for the intelligent applications in the business process, seeing new data and automatically executing the loop of discover–predict–justify on a frequency that makes sense for that business process. For example, network monitoring may be continuous process with loops measured in minutes or even seconds.
- **Learn—**Detect and react as data evolves.

The responses present a clear picture: Respondents are looking for technology that will extend the current knowledge base and make it more dynamic, asking questions about previously "unknown unknowns" to learn, discover and predict new things. Only half as many respondents are interested in having the system support existing processes, either by acting upon the supported process or workflow (e.g., automation playbook) and consequently lowering the need for insight or justification in how the machine took action. See Figure 2.

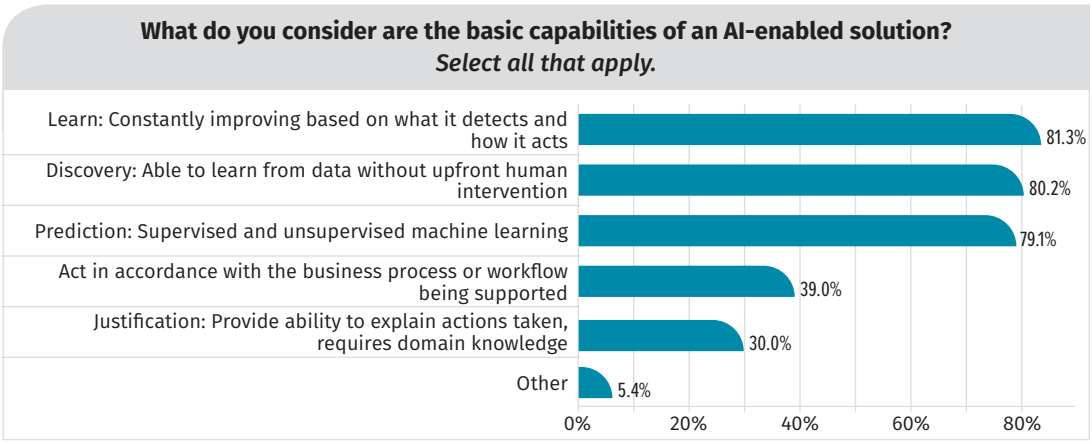

Figure 2. Perceptions of AI Capabilities

[3] "Five Components That Artificial Intelligence Must Have to Succeed," www.bloomberg.com/professional/blog/five-components-artificial-intelligence-must-succeed

Interestingly, however, staff (both security and non-security) place greater emphasis on the justification component than their management counterparts, possibly because they see the importance of transparency and trust in evaluation AI-based action as a means to improve security performance. See Figure 3.

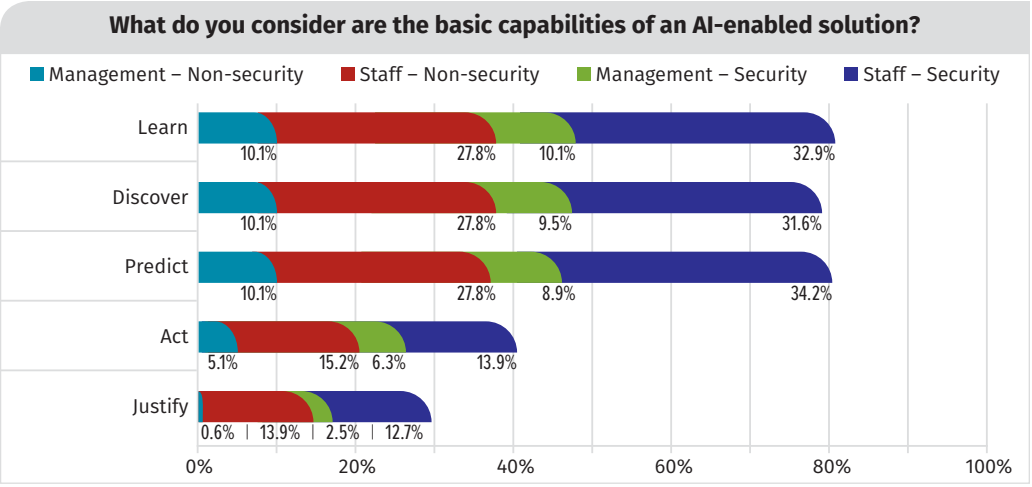

Figure 3. Relative Importance of All Steps by Role

The industry of respondents also had an effect on considerations of the basic capabilities of the AI solution. Technology organizations value learning, likely because of an emphasis on product improvements needed to support other industries. The banking and finance vertical, on the other hand, is most concerned with discover and predict—not surprising given the need for rapid response related to the detection of fraud. See Figure 4.

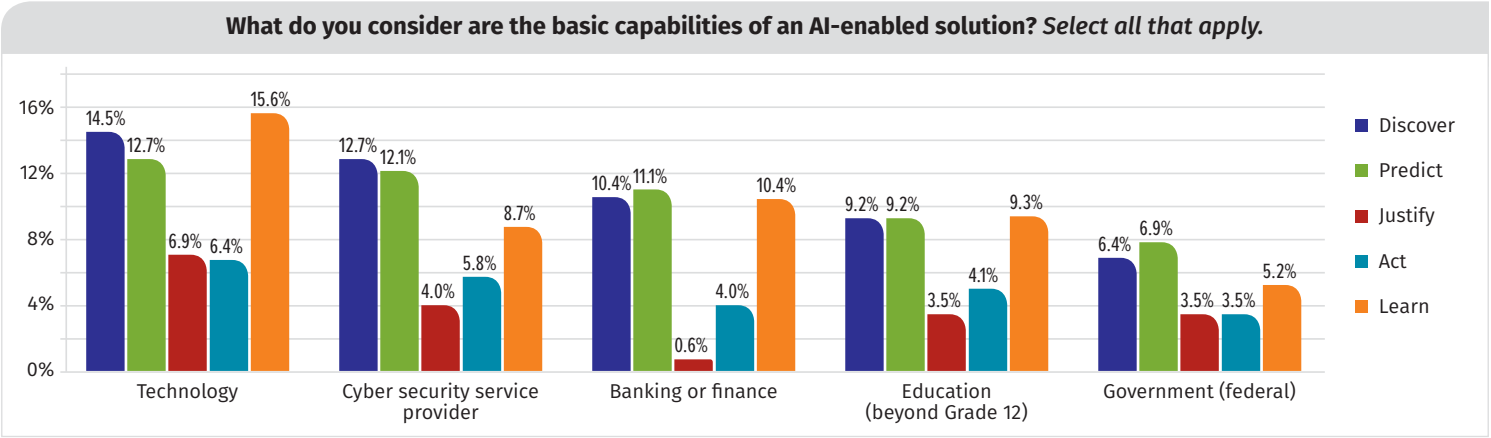

Figure 4. Emphasis on AI Capabilities by Top 5 Industries

Moving from the Abstract to the Practical

Respondents overwhelmingly selected predictive analytics, deep learning platforms and machine learning platforms as the top three technologies to consider as part of an AI-enabled solution for security. See Figure 5 on the next page.

Note: The top three technologies are also the leading technologies that have been implemented by the 35% of respondents who have actually implemented an AI platform.

As we examine perceptions of the technology, it almost seems that respondents are more likely to associate a technology with AI if it is difficult to describe easily. Statistical learning technologies like image processing, speech recognition and natural language processing (NLP) are less likely to be included in the perception of AI, and these are also technologies with which many are familiar, at least in concept. This unfortunate

tendency, known as the AI Effect, can be expressed as "AI is (only) what hasn't been done yet."[4]

Respondents were more likely to perceive AI as including machine-associated deep learning techniques such as neural network learning and predictive analytics—technologies that are able to ingest and analyze more data than ever to characterize a problem space, investigate that space virtually and provide unexpected information—as part of artificial intelligence. This is consistent with the AI Effect as well—researching "unknown unknowns" is well within any perceived definition of AI.

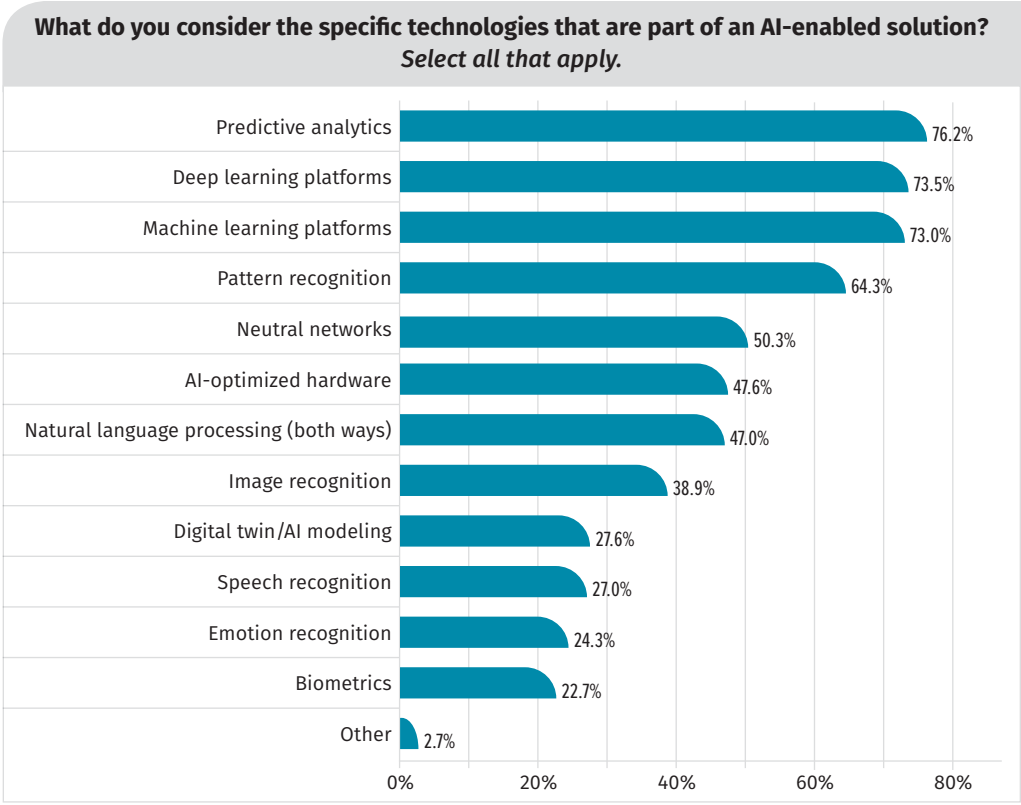

Figure 5. Perceptions of AI-Related Technologies

Systems That Learn: A Threat Detection Example

Machine learning is an AI discipline and a key driver behind many of the practical advances associated with this field. Let's take a quick look at a sample threat detection system enabled by AI/ML and some of the implementation concerns.

ML tools can be extremely effective, but they produce vastly different results depending on the source and quality of data being analyzed. Specific domain knowledge related to security is needed to design a threat detection system using appropriate ML mathematical and statistical algorithms as opposed to another domain, such as clinical research or finance. A data scientist must apply security domain knowledge to identify primary and secondary sources of data, determine how to clean and transform acquired data, and select the best ML analytical method or algorithm for the problem.[5]

Primary data sources for this domain include network packets, and an ML-based analysis reveals otherwise invisible communication patterns from an attacker inside the network. Secondary sources are logs routinely collected from other devices, which may provide additional depth to the analysis but not direct evidence of activity because of the nature of their role in providing security defenses.

Machine learning describes a collection of algorithms and techniques used to design systems capable of acquiring and integrating knowledge automatically.[6] Machine learning (ML) can be *supervised*, inferring a function from labeled training data, or *unsupervised*, developing and then modifying the behavior of a model without knowing a previous model, through constantly analyzing available data. Supervised ML can identify a threat almost immediately without *a priori* knowledge of that threat or environment; unsupervised ML needs to learn the local context of what is "normal."

Deep learning is a subset of machine learning where artificial neural networks, algorithms inspired by the human brain, learn from large amounts of data.

Predictive analytics is the branch of the advanced analytics that is used to make predictions about unknown future events. Predictive analytics uses many techniques from data mining, statistics, modeling, machine learning and AI to analyze current data to make predictions.[7]

[4] *Gödel, Escher, Bach: An Eternal Golden Braid*, by Douglas Hofstadter (Random House, 1980), p. 601

[5, 6] "The Expanding Role of Data Analytics in Threat Detection," March 2015, www.sans.org/reading-room/whitepapers/analyst/expanding-role-data-analytics-threat-detection-36362, pp. 6–7

[7] www.predictiveanalyticstoday.com/what-is-predictive-analytics

A system capable of learning results is an automated threat detection and management system that continuously self-improves to be efficient and effective. Figure 6 illustrates the three main layers to the theoretical system.

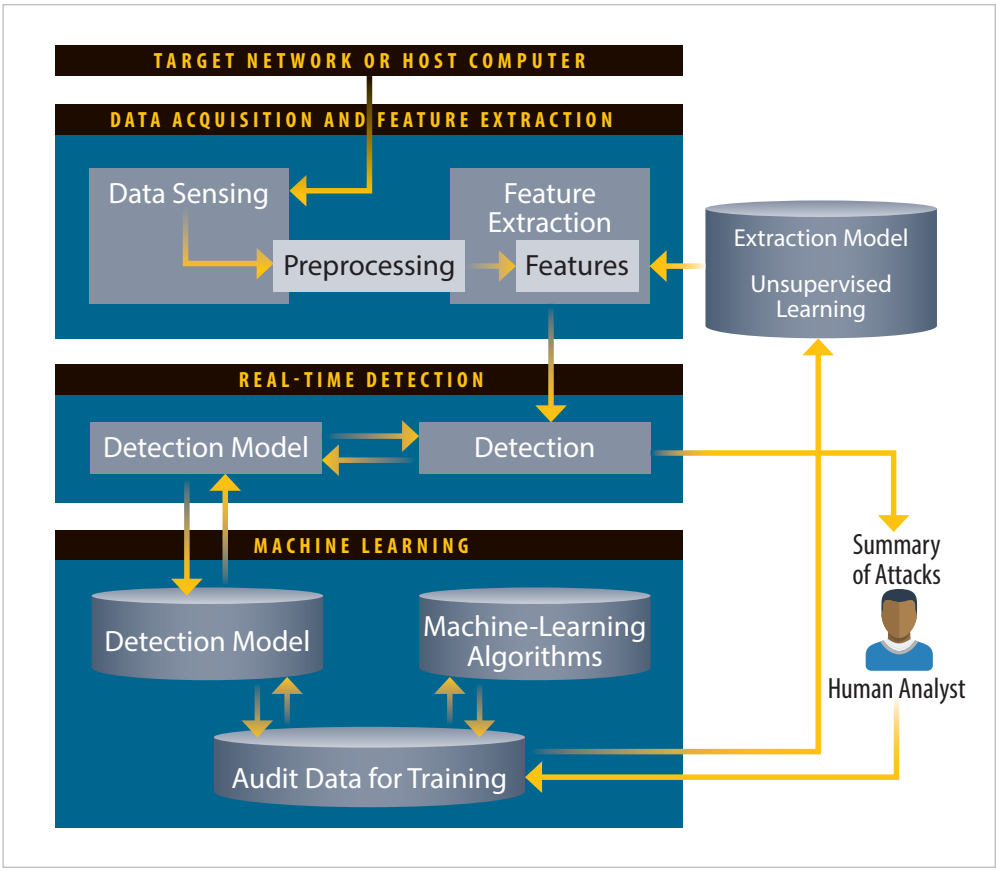

Figure 6. Threat Detection Reference Architecture Based on ML

1. **Data acquisition and feature extraction—**To be useful in threat identification, data—that is, network traffic or process execution events from various endpoints and host systems—must be detected and captured in ways that allow for analysis consistently over time and despite changes in the components producing the data being gathered. A feature extraction module is used to convert the raw data into feature vectors or datasets. This is a key step. Features that are improperly selected have a negative effect on the performance of detection models.
2. **Real-time detection—**This determines whether an observed pattern or a sequence of patterns is normal or abnormal, as it happens, based on the detection model and how the system has been trained.
3. **Machine learning—**This is the heart of the system, containing the various algorithms for anomaly detection, audit or training data (dynamically updated by either human analysts or the ML algorithms), and the actual behavior-based detection model. Selection of an appropriate algorithm includes several considerations, all of which can affect performance and the validity of the outcomes: accuracy, training time, use of linearity, the number of features/parameters and others.

Reaping the Benefits

From a security practitioner standpoint, 57% of respondents are using or planning to use security solutions that use AI, although only 35% have direct experience with such platforms. The range of applications is striking. The areas eliciting the largest response are oriented toward general defense and prevention—cyber defense, malware prevention and advanced threat detection/prevention. Other applications, often representing more specific use cases, showed a lower rate of adoption—roughly about half or less than the top three. See Figure 7.

A full 85% of respondents see AI as an enabler for improved security, with the majority (67%) believing this will happen by augmenting, rather than replacing, traditional tools. Figure 8 shows the most important enabler as better identification of unknown threats. Respondents feel that AI can improve the identification of unknown threats, a response consistent with the fact that 69% of respondents use or plan to use AI to help address advanced threat detection and prevention. See Figure 8.

Interestingly, however, this viewpoint varies by industry. Technology placed greater emphasis on improved times between infection and remediation as opposed to better identification of unknown threats. Banking and

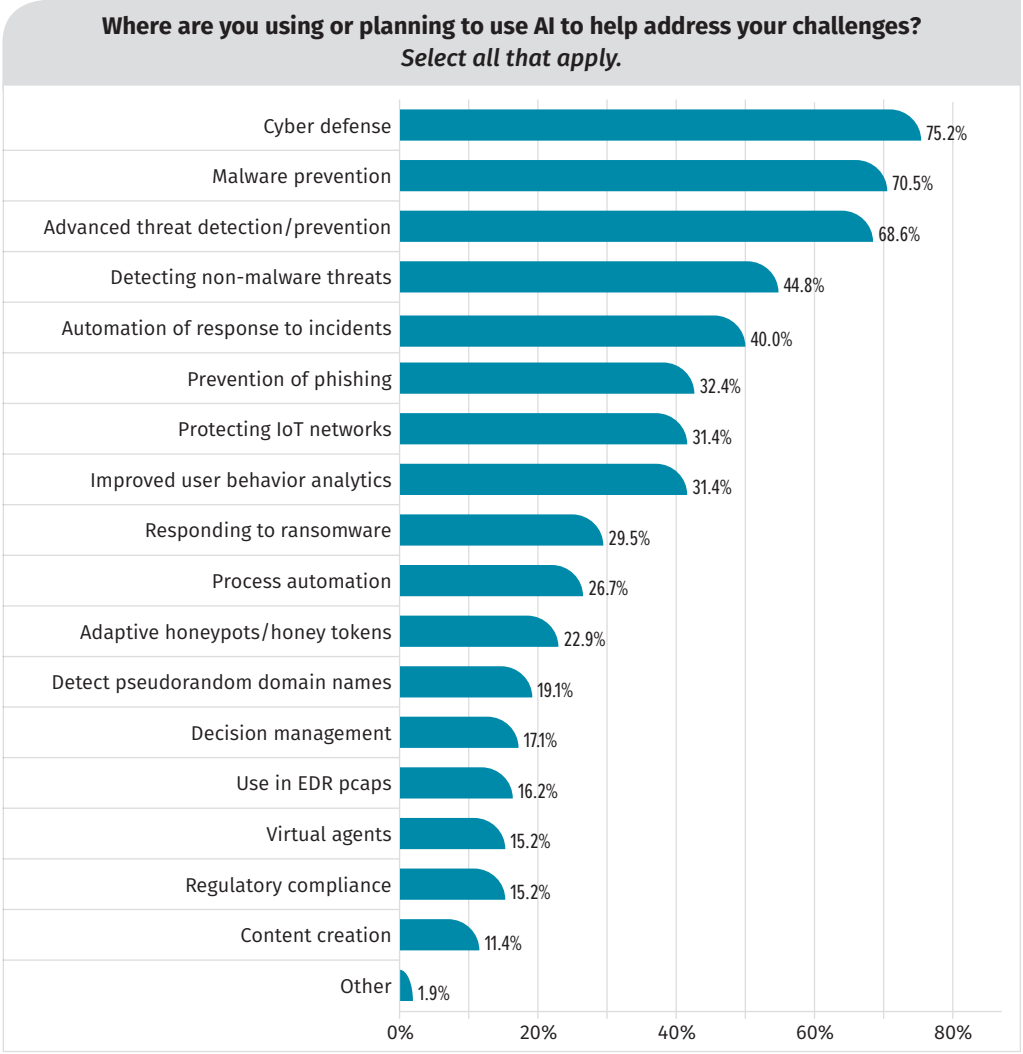

Figure 7. Planned Security Applications of AI Technologies

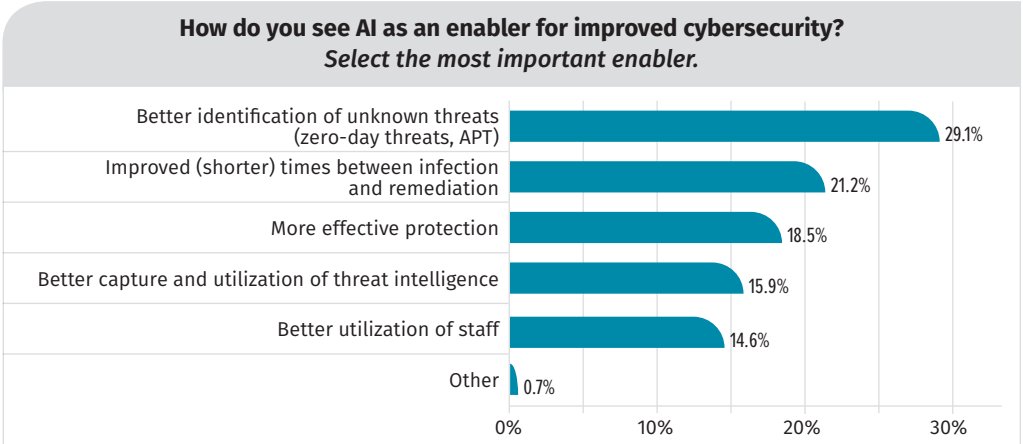

Figure 8. AI as an Enabler

finance were overwhelmingly concerned with better identification of unknown threats as opposed to the use of AI in other cases. Bottom line: How an organization perceives the best use of AI in improving its security posture will inevitably be driven by the mission and existing capabilities of that organization. See Figure 9.

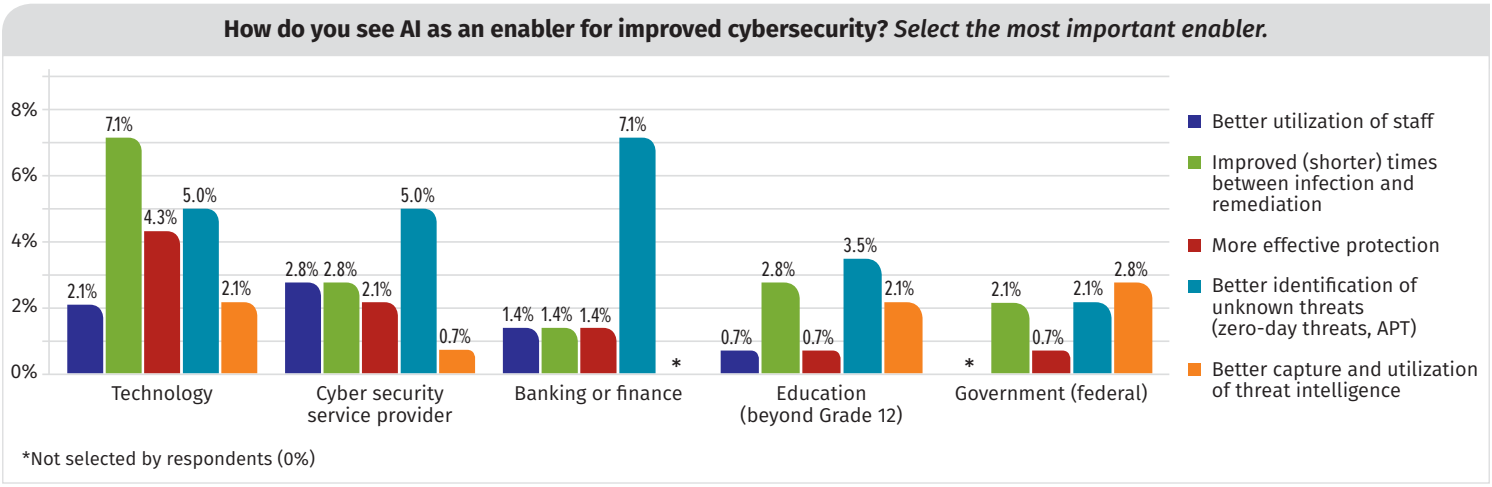

Figure 9. AI as an Enabler by Top 5 Industry

Still the New Kid on the Block

The majority of respondents view AI-based security solutions as a maturing market. See Figure 10.

In general, security staff have more confidence in the maturity of AI-based solutions than does their management. See Figure 11.

This viewpoint discrepancy suggests potential challenges related to funding and resources for implementation of AI-based security solutions.

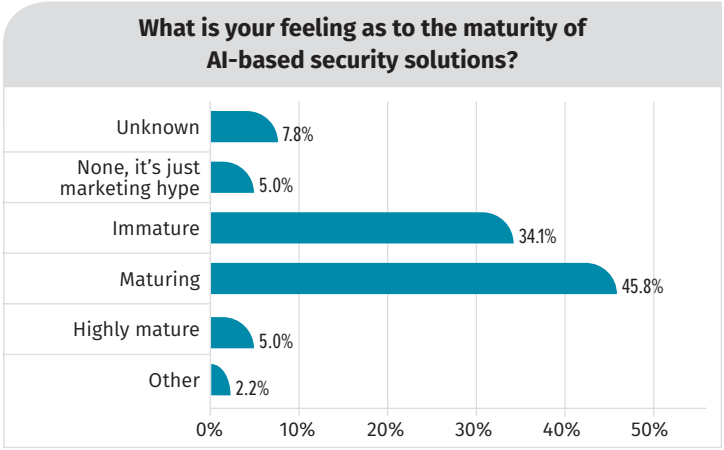

Figure 10. Maturity of AI-Based Security Solution

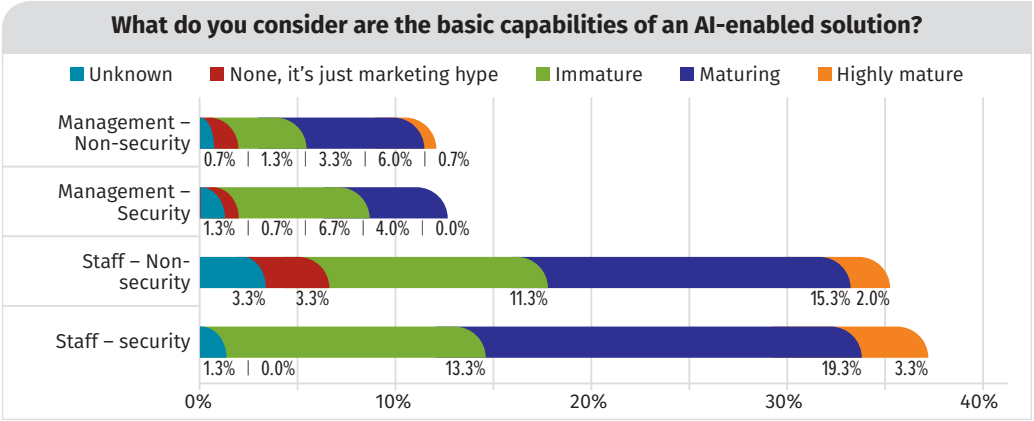

Figure 11. AI Maturity Levels, Based on Roles

Risks: It's Really All About the Data

Respondents ranked the risks behind AI, listed from highest to lowest:

1. Loss of privacy due to the amount and types of data that needs to be consumed
2. Over-reliance on a single, master algorithm
3. Not understanding the limitations of the algorithms used
4. Inadequate protection of data and metadata used by the AI platform
5. Improperly or inadequately trained solutions
6. Lack of visibility into decisions reached through AI
7. Selection of the wrong algorithms for the problem being solved

Data Governance for AI

Consider implementing a data governance program for your AI platform, asking some key questions to help define your approach. Trusting outcomes from an AI platform depends upon trusting the integrity and validity of the source data being consumed and the metadata and algorithms used. Ask:

- What protections are needed for source protection? How do you ensure its integrity and validity?
- What protections are needed for metadata? How do you verify there has been no tampering?
- How do you trust the outcomes from the platform, especially if you determine quality or trust issues with your source data and metadata, and you don't fully understand the algorithms?

The top four risks all involve either data protection or concerns about appropriateness of the algorithms. One respondent summed up the concern nicely: "AI could be compromised and act in [an] improper way such as giving wrong decisions [if the data is compromised in any way]."

Barriers: AI Coming of Age

When asked about barriers to implementation, 35% pointed to the lack of maturity in the AI technology. Another 27% pointed to the lack of time and skilled resources needed to implement an AI-based solution, closely followed by 24% citing what is often the leading barrier: lack of management commitment and budget. See Figure 12.

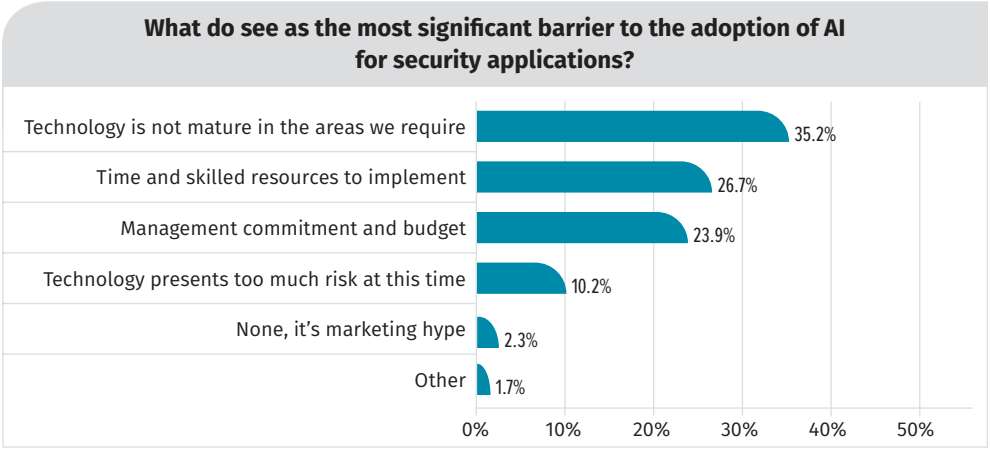

Figure 12. Barriers to the Adoption of AI for Security Applications

Survey results demonstrate that technical staff has more confidence in the maturity of AI-based security solutions. If you are an analyst or security architect looking to implement AI, be aware—management may look for quantifiable returns, given the perceived risks related to maturity. Be sure to address legal and policy issues as well as the human factors, transparency and trust.

Putting It All Together: Operationalizing

As you plan AI-based security projects, consider the following:

1. **Know your use cases.** Understand your current use cases—threat detection, malware prevention, etc. Even if not designed for a real-time environment, these use cases can serve as a foundation for the use of AI/ML for operational situations, providing the domain knowledge needed to guide your solution.

   Remember to keep your security experts in the loop. Don't be dissuaded by marketing hype—AI for cybersecurity still has a long way to go to match the cognitive capabilities of the human analyst.

2. **Understand the data and its limitations.** Don't underestimate the effect of data. Your data leads to metadata (the features extracted from the data), and your AI algorithms base outcomes on that metadata. Bad data can lead to bad metadata and, consequently, the wrong result. When developing data protection and governance strategies, pay specific attention to your metadata. It's easier to keep bad data out than to clean it out once it is in.

   Also, determine which data sources are appropriate for your use case. Some software applications analyze raw network data to spot an irregularity, while others focus on user/asset/entity behavior to detect patterns that deviate from normal.

3. **Establish transparency and trust.** The machine provides an output, but how can the user trust that the machine made the right decision or recommended the correct action? An AI platform needs to provide transparency in its decision-making process by allowing insight into the ML algorithm (without having the user be a trained data scientist or mathematician), what data was important for its reaching a particular conclusion, and actions that may be needed to fine-tune or adjust the platform to achieve the best outcomes. This understanding will become increasingly important as findings in AI applications become admitted as legal evidence.

   Users also need training. How did the machine identify a spearfishing threat? What were the clues and anomalies that led to the suggestion of a threat? How does one evaluate whether the process is correct or incorrect?

4. **Don't underestimate the time to train the AI platform.** For as long as ML has existed, training has been the most lengthy and cumbersome part of the implementation, although solutions have now been developed that permit the software to train itself autonomously, at least in part. But it still takes time to

Implementation Advice

Implementers—You are undertaking a big project. Prepare to work closely with a project team to determine an appropriate pilot environment. Don't shortcut your implementation; it will take time to show results.

Technology/platform providers—Explain your technology in more detail than you normally would, even for a technical audience. AI is based on math and abstract algorithms—your challenge will be to describe how AI supports the user in the cybersecurity domain.

prepare the data from which to train the platform. You might have years of data for the system to learn from. To create a system that delivers results in seconds, you might need weeks to prepare the data and additional time to train the platform.

5. **Make intelligent and informed decisions about the technology.** Understand your requirements and your infrastructure. For example, do you want a cloud-based solution?

   Undertake a procurement process that smokes out the false claims from the vendor. Whether planning your cybersecurity technology road map or acquiring individual technologies, be cautious about buying into solutions that boast AI or ML capabilities without some level of due diligence around product claims.

Conclusion

Continuing advances in AI-enabled tools point toward a future in which security goes from a reactive, forensic operation to an adaptive—and predictive—discipline, greatly reducing the risks of advanced threats. Emerging solutions bring true promise for holistic, intelligence-driven security; reduced time needed for threat detection and incident response; real-time alerts of anomalous behavior; and increased efficacy of existing investments and human analysts.

As a parting word, however, remember: What is good for the defender is also good for the attacker. At the 2016 Black Hat conference, John Seymour and Philip Tully gave a presentation on an experimental AI that sent simulated spearphishing links to Twitter users. The results, shown in Table 2, tip the advantage to the AI approach. This experiment is perfectly analogous to use of AI for cybercrime. Whether DDoS attacks, ransomware or some other kind of malware, a threat actor can use AI to spread the threats faster and target more vulnerable machines through automation.

Table 2. Man vs. Machine Two-Hour Bakeoff[8]

Actor Metric	Person	SNAP_R (AI-based Twitter Spearphishing Tool)
Total targets	~200	819
Tweets/minute	1.67	6.85
# Click-through (victims)	49	275
Observations	Manual copy/paste messages to different hashtags	Arbitrarily scalable with number of machines

[8] www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter.pdf

About the Authors

G.W. Ray Davidson, PhD, is the former dean of academic affairs for the SANS Technology Institute. He continues to serve as a mentor, subject matter expert and technical reviewer for the SANS Institute and holds several GIAC certifications. Ray started his career as a research scientist and subsequently led global security projects for a major pharmaceutical company. He has taught at the college level and cofounded a security startup. Ray is currently the program manager of the Michigan Cyber Civilian Corps.

Barbara Filkins, a senior SANS analyst, holds several SANS certifications, including the GSEC, GCIH, GCPM, GLEG and GICSP, the CISSP, and an MS in information security management from the SANS Technology Institute. She has done extensive work in system procurement, vendor selection and vendor negotiations as a systems engineering and infrastructure design consultant. Barbara focuses on issues related to automation—privacy, identity theft and exposure to fraud, plus the legal aspects of enforcing information security in today's mobile and cloud environments, particularly in the health and human services industry, with clients ranging from federal agencies to municipalities and commercial businesses.

Sponsor

SANS would like to thank this paper's sponsor:

BlackBerry | CYLANCE

Last Updated: August 5th, 2019

Upcoming SANS Training

Click here to view a list of all SANS Courses

SANS San Jose 2019	San Jose, CAUS	Aug 12, 2019 - Aug 17, 2019	Live Event
SANS Prague August 2019	Prague, CZ	Aug 12, 2019 - Aug 17, 2019	Live Event
SANS Minneapolis 2019	Minneapolis, MNUS	Aug 12, 2019 - Aug 17, 2019	Live Event
SANS MGT516 Beta Three 2019	Arlington, VAUS	Aug 19, 2019 - Aug 23, 2019	Live Event
SANS Amsterdam August 2019	Amsterdam, NL	Aug 19, 2019 - Aug 24, 2019	Live Event
SANS Chicago 2019	Chicago, ILUS	Aug 19, 2019 - Aug 24, 2019	Live Event
SANS Virginia Beach 2019	Virginia Beach, VAUS	Aug 19, 2019 - Aug 30, 2019	Live Event
SANS Tampa-Clearwater 2019	Clearwater, FLUS	Aug 25, 2019 - Aug 30, 2019	Live Event
SANS New York City 2019	New York, NYUS	Aug 25, 2019 - Aug 30, 2019	Live Event
SANS Hyderabad 2019	Hyderabad, IN	Aug 26, 2019 - Aug 31, 2019	Live Event
SANS Copenhagen August 2019	Copenhagen, DK	Aug 26, 2019 - Aug 31, 2019	Live Event
SANS Philippines 2019	Manila, PH	Sep 02, 2019 - Sep 07, 2019	Live Event
SANS Munich September 2019	Munich, DE	Sep 02, 2019 - Sep 07, 2019	Live Event
SANS Brussels September 2019	Brussels, BE	Sep 02, 2019 - Sep 07, 2019	Live Event
SANS Canberra Spring 2019	Canberra, AU	Sep 02, 2019 - Sep 21, 2019	Live Event
SANS Network Security 2019	Las Vegas, NVUS	Sep 09, 2019 - Sep 16, 2019	Live Event
SANS Oslo September 2019	Oslo, NO	Sep 09, 2019 - Sep 14, 2019	Live Event
SANS Dubai September 2019	Dubai, AE	Sep 14, 2019 - Sep 19, 2019	Live Event
SANS Raleigh 2019	Raleigh, NCUS	Sep 16, 2019 - Sep 21, 2019	Live Event
Oil & Gas Cybersecurity Summit & Training 2019	Houston, TXUS	Sep 16, 2019 - Sep 22, 2019	Live Event
SANS Rome September 2019	Rome, IT	Sep 16, 2019 - Sep 21, 2019	Live Event
SANS Paris September 2019	Paris, FR	Sep 16, 2019 - Sep 21, 2019	Live Event
SANS Bahrain September 2019	Manama, BH	Sep 21, 2019 - Sep 26, 2019	Live Event
SANS San Francisco Fall 2019	San Francisco, CAUS	Sep 23, 2019 - Sep 28, 2019	Live Event
SANS London September 2019	London, GB	Sep 23, 2019 - Sep 28, 2019	Live Event
SANS Dallas Fall 2019	Dallas, TXUS	Sep 23, 2019 - Sep 28, 2019	Live Event
SANS Kuwait September 2019	Salmiya, KW	Sep 28, 2019 - Oct 03, 2019	Live Event
SANS DFIR Europe Summit & Training 2019 - Prague Edition	Prague, CZ	Sep 30, 2019 - Oct 06, 2019	Live Event
SANS Tokyo Autumn 2019	Tokyo, JP	Sep 30, 2019 - Oct 12, 2019	Live Event
SANS Northern VA Fall- Reston 2019	Reston, VAUS	Sep 30, 2019 - Oct 05, 2019	Live Event
SANS Cardiff September 2019	Cardiff, GB	Sep 30, 2019 - Oct 05, 2019	Live Event
Threat Hunting & Incident Response Summit & Training 2019	New Orleans, LAUS	Sep 30, 2019 - Oct 07, 2019	Live Event
Supply Chain Cybersecurity Summit & Training 2019	OnlineVAUS	Aug 12, 2019 - Aug 19, 2019	Live Event
SANS OnDemand	Books & MP3s OnlyUS	Anytime	Self Paced