FORRESTER®

FOR SECURITY & RISK PROFESSIONALS

Artificial Intelligence Will Revolutionize Cybersecurity

But Security Leaders Must View All Vendor Claims With Skepticism

by Chase Cunningham, Joseph Blankenship, and Mike Gualtieri

September 14, 2017

Why Read This Report

Security vendors are inundating CISOs with products purporting to use artificial intelligence to dramatically improve the accuracy and speed of both threat detection and response. However, much of this messaging is confusing, even misleading. How do you know fact from fiction from enthusiastic marketing? S&R pros should read this report to understand what is really possible with AI today to take cybersecurity efforts to the next level.

Key Takeaways

AI Helps, But It Isn't 2058

There are two types of artificial intelligence: pure and pragmatic. Pure AI is the science fiction you have seen in Star Trek and Ex Machina. Forget about that. Concentrate instead on the building block technologies of pragmatic AI — which enterprises use now for all manner of applications, including cybersecurity.

AI Is Not A Silver Bullet . . . But It Is A Bullet

Security analysts struggle to keep up with new and emerging threats as well as the deluge of alerts and events they must analyze and respond to every day. AI building blocks like machine learning and natural language processing can provide security pros with insights about current and future threats.

Human Knowledge Still Reigns And Must Cooperate With AI

Tony Stark, AKA Iron Man, is like you — human. But he is mega-augmented with technology. That's the way to think about cybersecurity and AI. AI will give you insights, but it is still your experience and knowledge that will protect the enterprise.

Artificial Intelligence Will Revolutionize Cybersecurity

But Security Leaders Must View All Vendor Claims With Skepticism

by Chase Cunningham, Joseph Blankenship, and Mike Gualtieri
with Stephanie Balaouras, Bill Barringham, and Peggy Dostie

September 14, 2017

Table Of Contents

Related Research Documents

Counteract Cyberattacks With Security Analytics

TechRadar™: Artificial Intelligence Technologies, Q1 2017

Vendor Landscape: Security User Behavior Analytics (SUBA)

FORRESTER®

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
+1 617-613-6000 | Fax: +1 617-613-5000 | forrester.com

AI Enhances The Scale, Speed, And Accuracy Of Security Operations

AI is here. Embrace it, but don't be fooled by it. There are two types of AI. Pure AI is the science fiction version: true intelligence that's indistinguishable from, or even superior to, human intelligence in all respects. That's not what we're talking about, and we're still a long way away from it. Moreover, AI is not one universal technology. Rather, it's composed of technology building blocks such as machine learning (ML) that, individually or in combination, are advanced enough to add some intelligence to applications that can lead to significant business or operational transformation. This is pragmatic AI.

Security professionals can use the building blocks of pragmatic AI today to: 1) predict and adapt to future threats; 2) identify, prioritize, and remediate existing vulnerabilities; and 3) detect and stop cyberattacks in progress — at a scale and speed that is simply not possible with human analysis and manual processes. In fact, it's not just security pros who see the potential; according to our data, 34% of global data and analytics decision makers whose enterprise firm is planning to use or is currently using AI, do so to mitigate security risks (see Figure 1). The scale and speed is particularly important today because of the need to:

- **Analyze massive volumes of data.** Security analysts in the security operations center (SOC) find themselves drowning in a flood of alerts and other events. Properly analyzing and prioritizing these alerts is time consuming and frustrating. Couple this with compliance requirements that require security analysts to store and analyze all log data concerning attack activities and it becomes clear that no human, or even team of humans, can hope to keep up. AI in cyberspace operations can help address these issues and keep your teams ahead of the curve by rapidly analyzing stored data and notifying human analysts with meaningful alerts when it discovers anomalous activity.

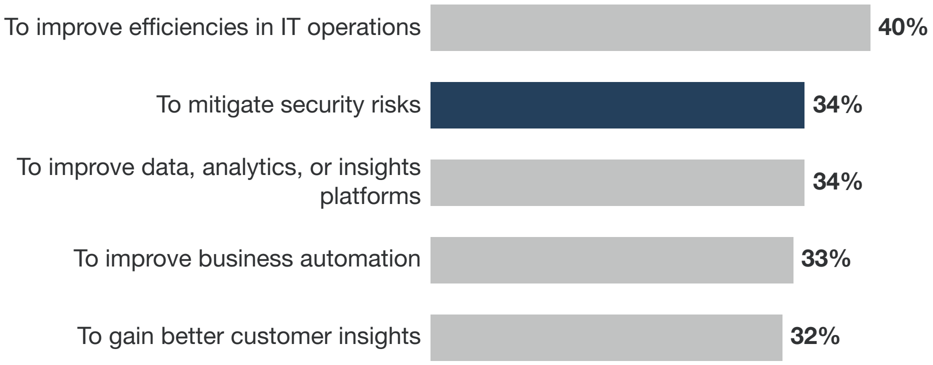
- **Address the cybersecurity skills gap.** According to multiple studies, there will be between 1.5 million and 3.5 million open positions in cybersecurity by 2021.[1] You simply can't hire enough immediately qualified and trained personnel to fully address your organization's skills gap. AI can help with this.[2] By augmenting and empowering the current workforce in an operational context, using AI in these capacities can greatly increase your current workforce's skill sets and empower them to do more with less.

- **Constantly adapt to evolving threats and attack patterns.** While they are considered a new or developing technology in the security space, some technical AI tools can enable better detection and increased knowledge of threats. These capabilities can enhance or even augment current rules or signature-based solutions, and in some cases replace them, as they're more focused on using proven scientific approaches to finding the threat needles in the haystack of data that security teams are presented with today.

- **Limit the customer and business impact of cyberattacks and breaches.** When the inevitable breach happens, the key is to be able to react quickly and with maximum efficacy. Easier said than done, because the menu of remediations at your disposal is not always obvious, and the cure may be more painful than the disease. AI solutions can optimize the best response by analyzing what actions have worked in the past combined with expert rules by security pros.

FIGURE 1 Firms Plan To Use AI To Mitigate Security Risks

Top five use cases/application scenarios firms are planning to use or are currently using artificial intelligence technologies for



Base: 911 global data and analytics decision makers whose firm is planning to use/currently using artificial intelligence technologies (1,000+ employees)

Note: Multiple responses accepted.

Source: Forrester Data Global Business Technographics® Data And Analytics Survey, 2017

Each AI Building Block Brings Security Benefits

No security vendor today has an end-to-end AI solution — one that senses, thinks, and acts — but some vendors are incorporating one or more of the key building blocks (see Figure 2).

- **Biometrics.** Biometrics solutions authenticate individuals based on unique physical characteristics, such as facial, voice, iris, and fingerprint, or behavioral characteristics, such as typing speed, mouse movements, and touchscreen interactions. Moreover, they can authenticate individuals — employees, partners, customers, etc. — at scale across multiple digital channels.[3] Behavioral biometrics provide for continuous authentication of users and uses machine learning to perform risk scoring.[4] Biometrics can help dramatically reduce fraud rates and improve security posture by stopping cyberattacks using stolen credentials. The data generated from these solutions can also feed other security analytics solutions to more quickly and accurately detect anomalous user behavior.[5]
- **Natural language processing.** Natural language processing (NLP) technology reads and understands human-generated text. People can say the darndest things, and what they say may be indicative of threats or improper activities.[6] Email spam filters are the obvious use case, but NLP also has the potential to detect phishing schemes and other threats by analyzing free-form text. NLP is also useful to security analysts conducting investigations and research. For example, IBM Watson for Cybersecurity has natural language understanding that allows it to understand written text, and analysts can query it much as they would a human analyst. Security analysts will be able to use Watson like an extra analyst that can answer questions and make recommendations.

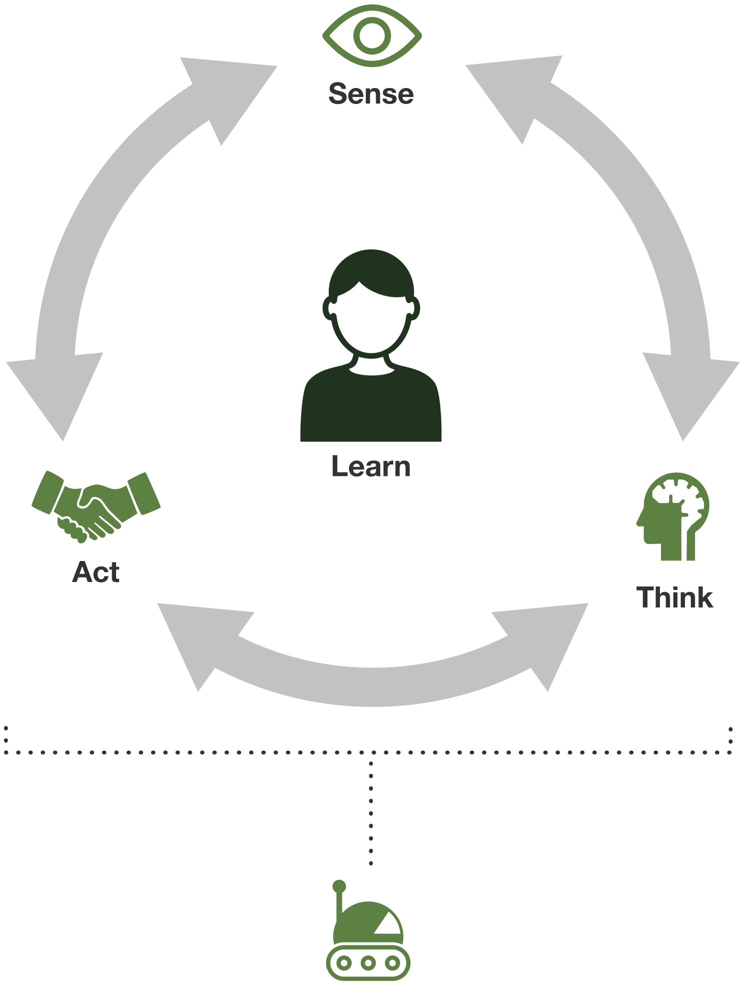
- **Machine learning.** Machine learning is composed of tools, techniques, and algorithms to analyze data that S&R pros and data scientists use to create predictive models or identify patterns in data. Machine learning is not a singular approach to analyzing data. There are dozens of specialized classes of algorithms that focus on specific problem domains. For example, some machine learning algorithms detect malicious file activity while others monitor users for unusual behavior. Cognitive search technology uses machine learning to identify recurring patterns in search results to make them increasingly relevant to analysts over time.[7]

- **Deep learning.** Deep learning is a branch of machine learning that specifically focuses on algorithms that construct artificial neural networks that are loosely inspired by how biological neural networks form in the brain.[8] Today, all the internet giants use it to analyze and predict online behavior, improve search, and label images. Other enterprises can experiment with deep learning to organize information and predict outcomes or to boost the accuracy of other AI building blocks, such as image analysis and speech recognition. In cybersecurity, security researchers use deep learning techniques to automate the mining of massive data sets for threats.

- **Security automation and orchestration (SAO).** SAO solutions employ AI building blocks to assist in the threat investigation and response process. Vendors like Demisto and Hexadite (recently acquired by Microsoft) use ML for event triage and to guide investigations.[9] The technologies are not yet mature enough for enterprises to rely on them to make decisions without human intervention, but they can assemble relevant context and inform human analysts about possible next steps. Over time, these solutions can be taught to take automated actions, based on previous outcomes to further accelerate operations.

- **Security analytics.** Security analytics solutions use ML to detect malicious behaviors. Security information management (SIM) platforms, long plagued by inefficient rulesets, use ML to reduce false positives and to detect activity missed by existing rules. SUBA tools detect unusual user behavior patterns, alerting analysts to suspicious user activity. Standalone security analytics tools use ML for threat detection and threat hunting by sifting through large quantities of security data.[10]

FIGURE 2 The Building Blocks Of AI For Cybersecurity



Sense
- Biometrics
- Text analytics and natural language processing
- Sensors and telemetry data

Think
- Machine learning
- Deep learning

Act
- Security automation and orchestration

Sense, think, *and* act
- AI-enhanced security analytics platforms

Six Ways To Scale Security Ops With The Power Of Machine Learning

ML can automatically identify suspicious patterns faster in network and user behavior that appear anomalous, and it allows for classification and grouping of activities, events, and data points that are useful in analysis and investigations. There is a variety of ML techniques, and how a vendor supports them will govern the value of the solution (see Figure 3). Utilization of these techniques includes:

1. **Thresholds to detect anomalies.** The most common capability to predict threats is to set thresholds on continuous metrics coming from infrastructure and application monitors and logs. If a metric rises or falls through a set threshold value, then it indicates a potential threat. For example, if a slew of password changes occur at the same time, it may indicate a breach of the sign-on directory.

Advantage: A threshold is very simple to configure.

Disadvantage: It may detect situations after the fact rather than ahead of time.

2. **Built-in rules using vendors' years of expertise.** The most advanced security vendors don't just provide simple threshold rules or dashboards for real-time monitoring. They also provide more complex built-in rules that already understand the vagaries of the applications and infrastructure components they're monitoring. Built-rules can automatically raise alerts based on this internal knowledge of the monitored threat surfaces.

   Advantage: Built-in rules require little setup. Built-in rules codify vendor expertise with other customers.

   Disadvantage: Rules may not exist for all threat surfaces, and rules may be based on outdated information.

3. **Customizable rules to let security pros apply their years of experience.** Built-in rules are extremely useful at predicting well-understood threats, but enterprises have their own unique, complex combinations of software and systems. Security pros have gotten to know the "personalities" of their threat surfaces and, over time, understand cause and effect that leads to problems. Some vendor solutions allow security pros to adjust rules or define a more complex set of rules that goes beyond just thresholds to if/then expressions.

   Advantage: Security pros can codify their expertise within the solutions.

   Disadvantage: Security pros may create rules based on theories rather than concrete data.

4. **Built-in models that go beyond rules to address complex relationships.** The key difference between rules and models is that rules are created by humans (security pros) and models are created by machine learning algorithms that analyze historical data. Algorithms can analyze more historical data more quickly and thoroughly than humans, finding complex, nuanced relationships that humans are likely to miss.

   Advantage: Models are created by machine learning algorithms that analyze historical security data, yielding better predictions that improve over time.

   Disadvantage: Models require more data science knowledge to tune and maintain.

5. **Built-in models that can learn the peculiarities of an enterprise's threat surface.** Machine learning can analyze large amounts of monitoring and log data to create predictive models solely based on historical data peppered with incidents. This learning is not perfect and can result in false positives and false negatives. This is particularly true if the incidents the models are trying to predict don't occur frequently: Machine learning relies on a history of several or many incidents, so if a problem occurs only a few times a year, machine learning models probably can't predict it until a year or more has passed. False positives and false negatives can often be reduced by adding custom rules to explicitly reject or accept what is obvious to security pros.

Advantage: Predictive models are based on actual data collected from the infrastructure and analyzed by machine learning algorithms.

Disadvantage: False positives and false negatives are often problems with predictive models generated by machine learning.

6. **External, importable models that let a community of enterprises share knowledge.** As the market for AI security matures, we can envision a community, either open source or vendor-specific, where enterprises can share AI models. The upside of such a community is that enterprises can gain predictive capabilities that they would not otherwise share. Vendors could also enable communities by sharing a framework for infrastructure component providers to offer their own AI models that any enterprise could use. The advantage to the component vendors is that they will offer more reliable, or at least more fixable, components.

   Advantage: Enterprises can share and reuse security AI models.

   Disadvantage: Community models may vary widely in efficacy and applicability to specific enterprises.

FIGURE 3 Applying ML Techniques To Cybersecurity

Technique	Description	Cybersecurity use case
Classification	This is a method of comparing an unknown data point against a larger data set that has a variety of known characteristics. The more data available with previously identified entries, the faster the new classification will likely be. However, if the comparative data has not been accurately classified or contains invalid data points, the new classification will be inaccurate and will grow in its inaccuracy as the machine "learns" from that data set.	Applied to cybersecurity, this is the ability of a solution to take an unknown, seemingly benign data point — a file sample, email, or log entry — and compare it to a large data set (such as data lake or other data repository) of previously identified data and have the system say this is malicious. For many current security solutions that use ML for malware categorization and identification systems, this is the current state of the art.
Clustering	With clustering, the goal is to find data points that naturally appear similar in nature. To date, this technique has been used in natural language processing (NLP) and genetic or medical research as the data therein lends itself well to clustering. In cybersecurity, clustering is possible but challenging because of the wide variety of data possibilities present in items like log entries, text alerts, and file samples. This makes it difficult for an ML algorithm to cluster items correctly without wide variances in accuracy and without significant human-assisted contouring of the data. However, when cybersecurity vendors target clustering for a specific use case, it can be very effective at identifying anomalies.	In cybersecurity, security teams use clustering in an attempt to identify network attacks by analyzing alert outliers visually plotted on a clustering map. In a modern SOC that has a network analysis and visibility (NAV) solution employing this technique, a data map would display a graphical representation of attack data and normal data. The places where the outliers had large variances from the mean would be the alerts or areas of focus for the SOC team. This helps the SOC quickly and easily identify likely areas of concern because they are visually displayed as blips that are well outside of the clusters of normal data. Clustering is also a technique used in security user behavior analysis (SUBA) solutions that attempt to highlight anomalous user activity.

FIGURE 3 Applying ML Techniques To Cybersecurity (Cont.)

Technique	Description	Cybersecurity use case
Regression	With regression, the goal is to measure the statistical relationship between variables based on historical data or a training data set. The variable attempting to be predicted is the dependent variable. The variables that have an impact on the dependent variable are the independent variables. In cybersecurity, for example, vendors use regression analysis to determine which factors have the most impact on the determination of whether something is malware or malicious activity.	This particular ML technique is best exemplified by security solutions such as Cylance and Trend Micro's Deep Learning system. These solutions looking at a variety of sample data and a variety of other data (e.g., a classification) and then compare specific mathematical measurements of the data as they relate to a standard mean value. Data points measured as being too far from the standard mean value are likely the ones that merit further analysis and would be used to either prompt a SOC operator to dive deeper on that alert or file sample, or the system might automatically conduct some further analysis on its own to determine the cause of the variances.

Deep Learning Enhances Machine Learning

With the massive volumes of data now so abundant in cybersecurity (logs, logs, and more logs, plus feeds and inputs from other security solutions as external threat intelligence feeds), security teams are inundated with data that leads to paralysis by analysis. Deep learning (DL) techniques within ML tooling are focused on mining those vast data repositories to look for specific indications of threat activity and anomalies that are indicative of compromise actions. Below are a few current use cases that show the benefits from applying DL techniques:

- **Automated data mining for indicators of compromise.** Security researchers are already using DL techniques to automate data mining for compromise indicators stored in log data that resulted from domain hijacking. Researchers in Singapore fed their systems over 1 billion log entries across an infected subnet, and then they unleashed a DL algorithm to seek out indications of anomalies and domain hijacking. The DL algorithm found those indications and infections in minutes instead of days or weeks and was nearly 100% effective during this testing.[11]
- **Automated classification of malicious activity — unsupervised.** DL learns "on its own." The DL algorithms and neural networks that are in use for a variety of purposes today are capable of discerning what appears to be anomalous and then classifying that indicator as malicious. The more often this type of action occurs, the more solutions the algorithm will learn, and it will improve intuitively.

AI Should Empower Human Analysts, Not Replace Them

With pure AI, the machine itself has a basic consciousness. On its own, it can walk through a problem or ask “How will cybercriminals target my organization?” It would understand the concept of whatever one asks it, process the question, formulate a response, determine the efficacy of that response, and then “think” on how good or bad the response was relative to the question. With pure AI, security leadership could remove human operators from the SOC and replace them with a machine that can do their job as well or better. Pure AI is years, likely even decades, away, especially in cybersecurity. However, in the short term, security teams can use solutions that use the building blocks of practical AI together with solutions that facilitate automation and orchestration to build a SOC that can keep up with the scale, speed, and adaptability of today’s threats. Here’s what AI can do in the short term:

- **Improve detection and accelerate investigation.** Vendors are already combining NLP and ML techniques to mine data streams, reports, and data repositories for indicators of malicious activities that are going unnoticed by human analyst teams. These types of applications are taking the brunt of the detective work off of security analysts and seeking out specific threat indicators that teams can use to enhance defensive countermeasures across an enterprise.
- **Speed decision making and automate response.** SAO tools are already using practical AI for some decision making, although many users prefer to keep humans in the loop. These tools use playbooks and vectored threat response systems to prompt users toward specific decisions and actions and can automate response.[12] As SOC and MSSP offerings continue to grow and sprawl, these types of solutions will be used to augment and optimize current SOC staff and elevate junior analysts to act more like senior decision makers.

Key Questions For Differentiating Practical AI Security Vendors

These technologies are powerful and have the potential to totally change cybersecurity and SOC operations as we know them. However, you must be cautious of current vendors’ claims regarding AI. What we have today are some powerful AI building blocks and enhanced ML techniques with focused use cases. Don’t believe vendors that boast about “total AI”; we are years before any security technology can start replacing humans entirely with AI decision makers. What’s realistic today and for the foreseeable future is the ability of security solutions to incorporate AI building blocks to empower, enhance, and augment analysts and responders, not replace them with robots. Now that you understand how to use different AI building blocks to enhance your SOC, you can better evaluate security vendors purporting to have AI or ML solutions: At a minimum, ask vendors the following questions:

- How long does it take to begin recognizing suspicious patterns? How does the solution adapt to completely novel attacks? For behavioral analysis, how long does it take to establish a baseline?
- Can you provide me with specific use cases that apply to my environment and industry?
- How can your solution help me address skills shortage in the SOC? To what extent can it accelerate the productiveness and effectiveness of both junior and experienced SOC analysts?

- Do you have metrics on the improvement in capabilities (e.g., detection, analysis, prioritization, investigations, response) that your solution offers from others that have implemented this system?

Recommendations

Start Using AI Building Block Technologies Sooner Rather Than Later

Whether your team or group is ready for it, AI and ML will become a major capability in cybersecurity, and soon. The time to start figuring out how and where AI and ML technologies and techniques can augment and empower your security team is now. Embrace this technology slowly, but with an open mind and a focused plan on specifically where you think you can get the most benefit from these innovations. Specifically, we recommend that security pros:

- **Create a blueprint for an AI-driven or intelligent SOC.** Use a well-crafted plan and a focused strategy to begin the baby steps toward growing and using AI- and ML-enabled tools and capabilities. In cybersecurity today, these capabilities are well suited to enhancing and augmenting your security team and SOC. If you approach these techniques and technologies slowly and with a focus on empowering analysis and response efforts, your team will see tangible benefits.

- **Define and map your processes first.** Examine your operations and decide where you will rely on human resources and where you will augment. What are your biggest pain points due to lack of scalable human resources or overwhelming alert volume? Those are the areas that are immediately open to growth and optimization from current AI and ML approaches. If your team has a firm command and control of the specific needs and processes that are currently in place, then you have the starting blocks that are needed to use AI technologies and ML.

- **Identify and define your data sources and data types.** If your organization can't define the totality of the data that it seeks to use for an ML- or AI-related project, the benefits of these solutions will be missed. If you have large amounts of very structured data, then ML certainly applies. Work to define and clearly identify your specific security data, and focus on using those well-controlled data points and streams for early AI and ML applications. The more targeted your data use strategy is, the more powerful your uses of AI and ML will be.

- **Evaluate vendors for how they fit into the vision.** AI is not magic. And no vendor has *The Terminator*'s Skynet of the cyberfuture. Make your vendors prove that their tools and capabilities truly empower and augment your security needs. A solution may sound cool and look nifty, but that doesn't mean it really enhances your security program.

- **Know what you're expecting before you start asking for demos or POCs.** Have very specific requirements for how AI technologies or ML techniques will empower your analysts and decision makers, and focus on using your defined data to test those systems. You should be driving these POCs; don't let the vendor's visuals and interesting technical applications control these evaluations.

Supplemental Material

Survey Methodology

The Forrester Data Global Business Technographics® Data And Analytics Survey, 2017, was fielded between February and April 2017. This online survey included 3,378 respondents in Australia, Brazil, Canada, China, France, Germany, India, New Zealand, the UK, and the US from companies with 100 or more employees.

Endnotes

[1] Source: Surendra Dhote, "Cybersecurity Ventures predicts there will be 3.5 million cybersecurity job openings by 2021," Electronics Media, June 10, 2017 (https://www.electronicsmedia.info/2017/06/10/cybersecurity-ventures-predicts-will-3-5-million-cybersecurity-job-openings-2021/); Steve Morgan, "1 million cybersecurity job openings in 2017," CSO, January 6, 2017 (http://www.csoonline.com/article/3155324/it-careers/1-million-cybersecurity-job-openings-in-2017.html); and Julie Peeler and Angela Messer, "(ISC)^2 Study: Workforce Shortfall Due to Hiring

Difficulties Despite Rising Salaries, Increased Budgets and High Job Satisfaction Rate," (ISC)^2 Blog, April 17, 2015 (http://blog.isc2.org/isc2_blog/2015/04/isc-study-workforce-shortfall-due-to-hiring-difficulties-despite-rising-salaries-increased-budgets-a.html).

[2] See the Forrester report "Quick Take: Your Next Security Analyst Could Be A Computer."

[3] See the Forrester report "TechRadar™: Biometric Authentication, Q1 2017."

[4] See the Forrester report "Vendor Landscape: Behavioral Biometrics."

[5] See the Forrester report "Vendor Landscape: Security User Behavior Analytics (SUBA)."

[6] Source: Eamon Javers, "You won't believe what gets an email flagged at Goldman: CNBC has the list," CNBC, June 16, 2016 (https://www.cnbc.com/2016/06/15/you-wont-believe-what-gets-an-email-flagged-at-goldman-cnbc-has-the-list.html).

[7] Source: Mike Gualtieri, "Cognitive Search Is The AI Version Of Enterprise Search," Forrester Blog, June 12, 2017 (https://go.forrester.com/blogs/17-06-12-cognitive_search_is_the_ai_version_of_enterprise_search/).

[8] See the Forrester report "Deep Learning: An AI Revolution Started For Courageous Enterprises."

[9] SAO solutions orchestrate security processes and automate activities to accelerate security operations. SAO can also act as an advisor, helping less experienced analysts complete investigations and elevate their skills. See the Forrester report "Breakout Vendors: Security Automation And Orchestration (SAO)."

[10] For more information about security analytics tools, see the Forrester report "Vendor Landscape: Security Analytics (SA)."

[11] Source: Vrizlynn L. L. Thing, "IEEE 802.11 Network Anomaly Detection and Attack Classification: A Deep Learning Approach," IEEE Xplore, March 19, 2017 (http://ieeexplore.ieee.org/document/7925567/).

[12] See the Forrester report "Rules Of Engagement: A Call To Action To Automate Breach Response."

FORRESTER®

CHALLENGE THINKING. LEAD CHANGE.

We work with business and technology leaders to develop customer-obsessed strategies that drive growth.

PRODUCTS AND SERVICES

- Core research and tools
- Data and analytics
- Peer collaboration
- Analyst engagement
- Consulting
- Events

Forrester's research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE

Marketing & Strategy Professionals

CMO

B2B Marketing

B2C Marketing

Customer Experience

Customer Insights

eBusiness & Channel Strategy

Technology Management Professionals

CIO

Application Development & Delivery

Enterprise Architecture

Infrastructure & Operations

› Security & Risk

Sourcing & Vendor Management

Technology Industry Professionals

Analyst Relations

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.

Forrester Research (Nasdaq: FORR) is one of the most influential research and advisory firms in the world. We work with business and technology leaders to develop customer-obsessed strategies that drive growth. Through proprietary research, data, custom consulting, exclusive executive peer groups, and events, the Forrester experience is about a singular and powerful purpose: to challenge the thinking of our clients to help them lead change in their organizations. For more information, visit forrester.com.