Center for Cyber
& Homeland Security
THE GEORGE WASHINGTON UNIVERSITY

Trends in Technology and Digital Security

Artificial Intelligence for Cybersecurity: Technological and Ethical Implications

Issue Brief 2 in a Series
Based on Fall 2017 Symposium Proceedings

Panelists
Michael Brett - QxBranch
George Duchak - Department of Defense
Anup Ghosh - Sophos
Kristin Sharp - New America Foundation

Panel Moderator
Frank J. Cilluffo

Panel Rapporteur
Sharon L. Cardash

This publication is the exclusive work product of the Center for Cyber & Homeland Security. It was made possible thanks to the financial support of Razor's Edge Ventures and Raytheon Company.

Center for Cyber
& Homeland Security

THE GEORGE WASHINGTON UNIVERSITY

Issue Brief Series on Trends in Technology and Digital Security
Artificial Intelligence for Cybersecurity: Technological and Ethical Implications

On September 14, 2017, CCHS convened a Symposium on Trends in Technology and Digital Security. Four panels addressed emerging threats and their implications for security policy, with a focus on digital infrastructure protection and anticipatory analysis. In a series of Issue Briefs, CCHS shares the findings and recommendations that emerged from the Symposium, primarily on a not-for-attribution basis. This second Brief in the series addresses Artificial Intelligence for Cybersecurity: Technological and Ethical Implications.

Emerging Technologies: Workforce Impact

The impact of emerging technologies on the workforce—including automation and artificial intelligence (AI)—has been explored using the technique of scenario-planning. A 120-participant study brought together technologists, traditional industry leaders, policymakers, and cultural exemplars to think through what they are seeing in AI, its impacts on the way these participants and their companies/organizations are functioning, and to explore a variety of potential futures. In total, the Commission considered 43 different future scenarios sited ten to twenty years in the future, and these boiled down to four categories:

The first is a future of less work, done in a non-traditional form, where we farm out “scut work” to machines, and humans focus on care and craft work. In this scenario, people are using their expertise to train, coach, and educate others. In the care industry, they are working specifically to develop products that are human-based only, and that are appealing to people based on their tactile nature. The second future is a world in which fewer people have jobs, so they are more competitive. Here, corporations would take a larger role in helping the displaced in their communities, and would be an overarching force in the types of jobs that people could have. The third future is a world, called “the contingent world,” in which most jobs would be project-based, and most people would have a variety of income sources, a sort of “souped-up” and extended version of the on-demand economy that we see today. And fourth is a potential world in which almost everyone is augmented in some way by technologies, and every action, surface, and technology is interconnected.

Three insights were common across all of these worlds. First, the “one-and-done” model of education that we see and use today—where a student focuses on one thing in high school or college and expects to use that knowledge for the whole rest of his/her career—is done. People will need to constantly retrain and acquire new skills and new information, and use them in different ways throughout their careers. Second, most emerging opportunities will be self-motivated and individually-driven, so people will have to take much more responsibility in finding, creating, and training for the kinds of work they want to do. Third, with respect to this self-driven and self-led feature, the nature of jobs in our economy is such that probably 20-30% of new jobs are contingent in some way. Yet most people when surveyed or polled in discussions say that their highest priority in a job is some sense of stability and predictability; they are looking for an ability to forward plan their income, to have a sense of benefits, to know that they will be able to have a job going forward. That is not something that employers are providing as much as they used to, so we will have to explore new systems for creating the kind of stability that lets workers be successful.

Cybersecurity is a particularly interesting example field because there are hundreds of thousands of open jobs now that go unfilled. Estimates for the undersupply of talent are anywhere from 30,000 in Virginia, to 300,000 nationally. With the cybersecurity industry as a job category expanding dramatically, and expected to expand dramatically over the course of the next ten or twenty years, we need to think through ways that we can identify, train, and get new types of people into the pipeline—meaning non-traditional students, and those with non-traditional backgrounds, either academic or demographic.

A counter-argument is that since demand far outpaces the supply of qualified professionals with cybersecurity jobs, it is likely that AI software will replace these jobs out of necessity. For example, traditional security operations centers (SOC) are mostly staffed with tier one analysts staring at screens, looking for unusual events or detections of malicious activity. This activity is similar to physical security personnel monitoring video cameras for intruders. It is tedious for humans, but it is a problem really well-suited to machine learning. So, when we talk about unfulfilled demand for people in cybersecurity, you will see that software will begin to replace conventional and mundane cybersecurity jobs with techniques like machine learning for pattern recognition.

Deep-Learning Neural Networks: Battling Malware

Our cybersecurity defenses as a nation are ill-equipped to deal with the nation-state-level threats that we are being attacked by. Historically, the U.S. Government and defense sector classified nation-state attacks and other breaches on its networks, which meant the commercial sector was largely unaware of the critical details of these attacks and could not develop technology to protect against them. Accordingly, a goal of one Symposium participant—Anup Ghosh, Founder and CEO of Invincea, a Sophos company— was to develop defenses that did not need signatures of threats in order to defend against them. Ultimately that led Dr. Ghosh down the path of machine learning. Part of his company was a group that did Defense Advanced Research Projects Agency (DARPA) R&D, and participated in a program where the basic idea was to look at the whole corpora of malware and identify the core attributes of malware that you can learn. If you create a model, you can then detect variants of this. Note that a super-majority of malware is variations of previously released malware.

The company developed these techniques using deep-learning neural networks and it worked remarkably well—so well, that a larger commercial anti-virus firm bought the company because they understood that innovation in machine learning is critical to combating current and future threats. The acquisition will result in machine learning technology reaching a very broad market through its products. The target market of the acquired company is small- to mid-sized business, which is a market segment that is largely ignored by just about every startup and next-generation security company—yet 99% of all businesses are small- to mid-size. This is the soft underbelly of American companies that gets attacked, and has no protection. Now, this underserved segment will get state-of-the-art machine learning technology to defeat threats not previously known.

Quantum Computing: Applications and Implications

Another Symposium participant, Michael Brett, Chief Executive Officer of Q^{x}Branch, explained the work that his company does applying predictive analytic technologies to a range of commercial outcomes. Working with the financial, insurance, and technology sectors, the firm conducts pricing and risk analysis, and seeks better understanding of customer behavior, using a range of probabilistic and predictive analytics techniques.

Quantum computing is an emerging technology that the company considers to be a strategic long-term enabler of advanced predictive analytics. Quantum computing has attracted significant investment over the past five years and is rapidly gaining attention from enterprises with high computational analytics needs. Q^{x}Branch has been involved in the field for about four years, with access to early stage hardware; and is working with partners in the commercial sector to explore and validate applications and the potential impact of the technology.

Some of the organizations that this company is working with to help understand the impact of quantum computing are banks, pharmaceutical groups, and oil & gas companies—to look at the kind of problems that quantum computing can assist with solving. These are all data analytics problems that, within those industries, are very computationally intensive—or practically unsolvable using classical techniques. Quantum computers will allow us to efficiently solve quantum math; and that could help to unlock some new breakthrough applications.

Quantum computing is at a really interesting stage of its technology development, where the focus is transitioning from research labs and universities, into corporate R&D. We have recently seen major investments from Google, Microsoft, and IBM; they are investing significant R&D resources into their own capabilities plus a startup ecosystem. There are also significant efforts led by the U.S. government including the programs run by the Intelligence Advanced Research Projects Activity, the National Research Foundation, and NIST. Globally we are observing a major push forward in the maturity and the availability of prototype quantum computing hardware that companies are getting access to; and this enables us to start to explore and validate the applications relevant to both industrial and national security issues.

Placing Artificial Intelligence in Context

Artificial intelligence has been around for over sixty years. A lot of the algorithms that you are seeing now are not new, but are things that we can do now because of the confluence of different technologies—including the GPU, big data analytics, and the massive connectivity of the cloud, where you can actually take data, exploit it, and use it at scale. About ten years ago, Ray Kurzweil was quoted as saying that all companies are essentially information companies—largely because information and the automated handling of it, is foundational to a firm's operations. Ten years from now, one participant suggested, I think we will say that all companies will be AI companies—largely because now that you have the information, you have to do something with it: exploit it, try to extract value from it. And the way that we are doing that now is this intersection of big data analytics, improved hardware, and AI.

On September 13, 2017, the Deputy Secretary of Defense, Mr. Shanahan, signed off on a memo saying that the Department of Defense is going to accelerate its movement to the cloud. That was largely motivated for purposes of exploiting data, using and applying AI to a lot of our problems in defense. Data, specifically training data, is the feedstock of AI. General training data, available to everyone, is often used to train AI algorithms. Democratized training data gives no one a competitive advantage. Algorithms are rarely a discriminator, but the training stock, data, often is. In the Department of Defense (DoD), probably 99%-plus of the data that it collects is dark, that is, never exploited. It just sits someplace, waiting for daylight. The movement to cloud is to try to get this data to be exploited. Our competitive advantage in DoD is the data we collect, with national technical means or otherwise. This is data that is not in the public domain—which gives us a competitive advantage in whatever AI algorithms that we have developed. The very premise of AI is the ability to learn from the data that is continuously collected.

As we discuss AI for cybersecurity, we should also talk about the cybersecurity of AI. We need to protect our models and data from manipulation. A canonical example of an image classifier, panda in this case, will result in the panda being classified as a gibbon with the introduction of a small perturbation in the training data. These "adversarial" examples show us that even simple modern algorithms, for both supervised and reinforcement learning, can act in surprising and unintended ways.

From a commercial standpoint, companies are the most competitive where they are able to bring a unique data set to that opportunity. The world is pretty flat when it comes to algorithms and machine learning techniques; but the world is not flat when it comes to access to unique and well-curated data. One of the competitive advantages that a company has is curating datasets that it owns (that are proprietary to it), that match somehow its industry partners' internal dataset. For example, a bank that has lots of transaction data; a company being able to match that with some other commercially available data that creates a force multiplier effect, is where a company is able to compete and obtain a win over others.

Is data a strategic advantage? Yes, in the Department of Defense; but less so in the commercial sector. You need really good data scientists, but besides your training data, it depends on your ability to execute that as a product. The industry group, Cyber Threat Alliance, believes that threat intelligence should be a public good; and that companies should feed on the ability to execute. On the other hand, one of the challenges with a common dataset is that you do develop blind spots to emerging threats. Machine learning is fallible to training on homogeneous data, so, to develop a really robust model you need a very diverse training set as well. Moreover, on the question of whether datasets are a strategic advantage that may shift, as we see improvements in unsupervised learning. In the future you might gain advantages from simulations or from simulated environments, more than you would/will from datasets.

Will AI benefit the attacker or the defender more? AI offers both promise and peril. It will be used for both offense and defense. It is too early to say for certain which side will have the advantage. Cybersecurity firms are using AI and machine learning to prevent attacks, and attackers are using AI to craft and respond to these defenses. At this stage, the technology is democratized—both parties have access to AI technology, and either side can use it. The discriminator, however, will be in the AI system that can learn and adapt the fastest. For example, we can use machine learning to write tweets that people will click on, for phishing. Or, we can use machine learning to write vulnerabilities in software that vendors can use to patch, and that adversaries can use to exploit. It is not a question of if or when; it is already happening. So, it is another continual evolution of technology for the good and the bad, at once.

Similarly, quantum computing has many great, beneficial applications; and then it also has some applications that are going to be very complicated for the United States and for the world. Consider breaking public key encryption: whether we would want to accelerate or decelerate the breaking of RSA encryption is really complicated, and there does not appear to be agreement about that, even within government. The encryption-breaking aspect of quantum computing, solving Shor's algorithm, is what seems to get all the press. Another use, in the era of big data, is using quantum computing for database search employing Grover's algorithm. Grover's algorithm searches for an entry in an unordered database with a polynomial speed advantage over the best classical algorithms. In the Department of Defense, this speed advantage can be a competitive advantage when optimizing command and control systems across air, land, sea, space, and cyber domains, for an optimized course of action.

About Us

The Center for Cyber & Homeland Security (CCHS) at the George Washington University is a nonpartisan "think and do" tank whose mission is to carry out policy-relevant research and analysis on homeland security, counterterrorism, and cybersecurity issues.

Website http://cchs.gwu.edu **Email** cchs@email.gwu.edu **Twitter** @gwcchs