The Future of AI and Cybersecurity

OCTOBER 30, 2019 | BEN BUCHANAN

The Cipher Brief

Ben Buchanan recently testified at a House Homeland Security Subcommittee Meeting on Preparing for the Future: An Assessment of Emerging Cyber Threats. He is a Senior Faculty Fellow, Center for Security and Emerging Technology, Mortara Center, Assistant Teaching Professor, Georgetown University.

Cybersecurity, already rife with challenges, is becoming even more complex with the rise in prominence of artificial intelligence systems. Where AI and cybersecurity intersect, society in general and national security in particular would be better served by greater attention to the potential pitfalls while they can still be addressed.

First and most significant is the cybersecurity of AI systems themselves. AI systems are just as likely to be susceptible to the kinds of software vulnerabilities that are present in other kinds of computer code. As we have seen for decades, hackers can exploit these vulnerabilities for their own ends. There is no reason to think that hackers will not try to do the same to AI systems, and there is no reason to think that they will not at times, succeed. This possibility is particularly worrying given the high stakes of some AI applications; it is not a reason to avoid using AI, but vigilance is imperative to preserve cybersecurity.

Yet to stop our analysis at just the traditional kinds of software vulnerabilities is to miss a great deal of the cybersecurity risk that AI systems pose. The neural network architecture that underpins a lot of modern AI is immensely powerful but presents a new class of cybersecurity risks that we are only beginning to uncover. We call this field adversarial learning.

Using adversarial learning, hackers can cause neural networks to make bizarre errors, causing systems that rely on those networks to fail or reveal confidential information. This is a field that requires a great deal more attention. My colleague Jason Matheny, the

former director of IARPA and the founding director of the Center for Security and Emerging Technology, estimates that only around 1% of AI research spending goes to security. That is simply far too low.

Second, AI can also change traditional offensive cyberattacks against regular computer systems. Modern hackers in many cases do not need artificial intelligence to achieve their ends. That said, I think it is noteworthy that some of the most potent cyberattacks we have seen–including Stuxnet, the 2016 blackout in Ukraine, and the 2017 attack known as NotPetya that caused at least ten billion dollars in damage–feature some forms of automated propagation and attack capability. I can imagine a world in which future cyber operations will use more sophisticated automated capabilities to achieve particular tasks, such as vulnerability discovery, target selection, command and control, and attack execution.

I suspect that such automation could offer significant upsides to sophisticated hackers faced with complex targets. In some respects, the possible upside to automation is higher in this area than in physical warfare; whether a plane is operated by a person or a human, the laws of physics still apply, but it is likely that automated cyber capabilities–if sophisticated enough–could operate much faster than their human-directed counterparts. I stress, however, that we have not seen this come to fruition yet.

This leads to the third area of analysis: the possibility that AI might help on cyber defense. This idea is also the subject of a lot of hype and a lot of venture capital investment. There seem to be discrete ways in which AI can indeed help secure computer systems, both in discovering vulnerabilities before hackers do and also in detecting the presence of malicious code. However, we must be careful not to let the hype outrun the reality on this front. In evaluating cybersecurity advances in this area, we should compare them to the baseline of technologies we already use–many of which already involve automation–and understand how, if at all, artificial intelligence improves our defenses.

These three areas–adversarial learning, cyber offense, and cyber defense–deserve a lot more attention, and quickly. Policymakers have begun to consider these important issues, including at a recent House Homeland Security Subcommittee hearing, but the

conversation must continue. During the early decades of the internet age, we raced ahead to deploy world-changing technology without always thinking through its security implications. In the age of AI, we should not repeat that mistake.

To get more involved in actionable conversations on the future of cybersecurity, join the Cyber Initiatives Group, hosted by The Cipher Brief. This community of dedicated cyber thinkers and leaders provides insights and shared knowledge of key cyber issues. The group hosts principals including General Michael Hayden, General Keith Alexander, FireEye CEO Kevin Mandia, Former Deputy Director of the NSA Rick Ledgett, and others, to bring our members briefings from public and private cyber leaders.

Join the CIG's next briefing on Monday, October 28 at 1p EST with DHS' CISA Director Christopher Krebs.

Read more national security insights, perspectives and analysis in The Cipher Brief.