What Corrodes Cyber, Infects its Offspring: Unlearned Lessons for Emerging Technologies

Dr. Chris C. Demchak

Making the Self-Evident Obvious[1]

What infects the cyberspace societal substrate also infects its technological offspring. Whatever relies on the current shoddy, insecure cyberspace substrate inherits its vulnerabilities. This great silent unlearned lesson among technologists, promoters, government officials, and ignorant or optimistic users seems self-evident and yet is repeatedly unlearned. It would seem self-evident that, unless the underlying substrate is transformed to be securable, any new technology built on those insecure cyber foundations will, in turn, fall prey to the same assaults. That adversary and criminal campaigns to poison data, corrupt algorithms, and 'p0wn' development processes are fairly predictable is logically obvious for AI systems as well as quantum, robotics, autonomous systems, synthetic biology, and any other emerging technologies. They all rely on the highly corruptible, existing cyberspace substrate and inherit its attack surfaces in addition to new ones of their own.

This logic and lesson of inter-generational sharing of vulnerabilities is clearly not obvious to the bulk of the information technology (IT) community responsible for developing the programs, collating the data, or using the results, tools, and products in established or emerging technologies. Even the advanced and most respected policy schools feeding the top ranks of the national policy elite did not perceive the logic despite having the president of the US declare in 2009 that cybersecurity is "one of the most serious economic and national security challenges we face,"[2] As late as the mid-2010s, most major master's programs in public affairs, political policy, international relations, business administration, law, or criminal justice did not yet include cybersecurity courses (or even cybersecurity components as part of their existing curricula) for their students.[3]

All the ideas in this essay are solely those of the author, and do not reflect the positions or policies of the U.S. Government, U.S. Navy, or the U.S. Naval War College.

With degrees in engineering, economics, and comparative complex organization systems/political science, **Dr. Chris C. Demchak** is Grace Hopper Chair of Cyber Security and Senior Cyber Scholar, Cyber Innovation Policy Institute, U.S. Naval War College. In publications and current research on cyberspace as a global, insecure, complex, conflict-prone "substrate," Dr. Demchak takes a socio-technical-economic systems approach to comparative institutional evolution with emerging technologies, adversaries' cyber/AI/ML campaigns, virtual wargaming for strategic/organizational learning, and regional/national/enterprise-wide resilience against complex systems surprise. Her manuscripts in progress are "Cyber Westphalia: States, Great Systems Conflict, and Collective Resilience" and "Cyber Commands: Organizing for Cybered Great Systems Conflict."

As further evidence of how little learning about computer vulnerabilities has penetrated the national zeitgeist, a myriad of revealed cyber-attacks over the past ten years has taken most of the past decade to spur broader and more concrete systemic action across the IT capital goods industry. For most of the past three decades, only occasional or niche admiration of this ubiquitous insecurity problem pockmarked the thinking across computer science professionals, Internet promoters, IT capital goods enterprises, and political decision-makers. Only recently have we seen the development of the "DevSecOps" (development security operations) concept across the information systems community, with the goal of securing software during its development.[4] However, it has yet to solidify as a defined set of basic processes embedded in the design, development, and implementation of new technology (software and hardware).[5]

Even further behind in learning this lesson are the various communities of emerging technologies. Especially noteworthy is artificial intelligence, arguably the most widely applied of the major emerging technologies. Although it has been around in either a software or an algorithmic form since the 1950s,[6] only since 2010[7] have some developers begun to warn about the cybersecurity threats to AI products. Moreover, only since 2017 – as though these vulnerabilities were something newly noticed and in need of explanation – has that warning become more vigorous.[8] Over the same period, however, according to one AI safety company's CEO, vulnerabilities have increased five thousand percent.[9]

The situation is only going to get worse, the longer the lesson about how a shoddy cyberspace parent infects its emerging technology offspring is ignored. *"Those attacks are going to become more and more pervasive, more and more prevalent, especially as [advanced battle management system] and [Joint All-Domain Command and Control] get implemented out into*

a wider context. The amount of data is going to explode beyond anybody's expectations. ... I'm talking about the sheer collection of that data and what that enables our adversaries to do and to think about."[10] According to a recent Microsoft survey of machine learning developers, 25 of the 28 groups interviewed did not know how to secure their machine learning systems, making their security awareness and posture effectively zero.[11]

At the end of the day, the brave new world of emerging technologies will only be as secure as the highly insecure underlying substrate on which they are being built, in which they are being embedded, and through which they are expected to operate transparently, accurately, and resiliently. Until we collectively transform that substrate in order to secure it, though, our future with emerging technologies will be one of exquisite and inexplicable further failures and ceding technological dominance ground to adversaries able to manipulate key national systems. This equally self-evident future path is apparently also not obvious.

"NMPY" – Not My Problem Yet

Why emerging technology developers would put cybersecurity aside during their project creation is understandable. First, most AI developers have been trained as computer scientists or mathematicians or they have grown up in the IT capital goods communities.[12] Little to none of their training specifically focused on the inherent security complexity of the computer languages they know, the programs they run, or the products they make. Indeed, security for computer products in general is relegated to minimal efforts inserted late in the development process with after-market updates left to be someone else's problem.

Second, it is hard enough to get a new machine learning (ML) model to work at all and to gather all the data needed just to train it, let alone slow down its development to worry about what exploit might attack the model and/or its data once it is deployed. One analogy – although a weak one – might be that of designing a ground-breaking new kind of car engine. No producer wants to ask their engineers working on the engine to slow down development in order to first figure out what stresses the new marvel will put on the braking system or even how exactly the braking system will work in the final design. That problem is left to others to solve once the new engine actually works and provides the additional thrust, efficiency, or mileage that then undermines the current brake design. Similarly, the incentive is enormous to just get the ML model to do what was envisioned. For developers, worrying about someone or something later messing that up has long been considered a step too far.

Third, this lesson is tough to learn because of the inherent opacity of the product, even more difficult to disaggregate than the code of a classic computer program. The latter is largely readable by a human, but the model's operating neural networks – modern AI's machine learning workhorses – are not. Machine learning is so heavily dependent on complex structures of opaque computerized calculations. Leaving aside adversaries, the AI/ML models can perform poorly for a wide range of obscure reasons constituting the normal accidents of complex systems.[13] Intentional or accidental corruption of one or more of its three main

development components – the tools, the models, and the data used to train the models – is exceptionally difficult to detect.[14]

Fourth, the inherent optimism of the computer science culture pushes against expecting the worst. Of course, it is recognized that at some point someone in the real world will connect the new AI product, new robot, or new autonomous system to old(er) networks, even if only for updates. Yet. this recognition does not invoke security concerns; indeed, AI and ML are usually characterized as separate from cybersecurity in casual discussion. Generally, the terms are connected with "and" as though they are two different topics, independent of each other. As Lt. Gen. Mary O'Brien, Air Force deputy chief of staff for intelligence, surveillance, reconnaissance, and cyber effects operations, recently noted: *"[t]there's an assumption that once we have the AI, we develop the algorithm, we've had the training data, you know, it's giving us whatever it is we want it to do, that there's no risk, that there's no threat."*[15]

While the Air Force is commonly assumed to be the most advanced US military service in terms of artificial intelligence,[16] Lt. Gen. O'Brien also noted that the service has no unit dedicated to defending the ML algorithms once they are deployed for use and called an officer suggesting the creation of such an AI red team a "maverick."[17] While this term was used with respect, If an officer need, to be a maverick to make the obvious connection that these newer systems will be intrinsically vulnerable, then the lesson of cybersecurity – so poorly and slowly learned in the parent field of computer science – has even less penetrated the more opaque world of AI/ML development.

The Peer Adversary Knows the Lesson, at Least as an Opportunity for Attacks

The time left to learn the lesson about cybered defense of AI/ML shortens daily as westernized democracies fall behind in having a skilled and well-educated workforce able to both develop and find ways to defend AI/ML systems. Aside from the sheer scale of the 5:1 advantage in the overall population compared to the US, China's strategic and well-funded dedication to technology dominance in the last half-century is paying off. Not only has the US' main peer adversary developed a commanding lead over the US in several emerging technologies, but it is also outpacing the US in the number of graduates in STEM subjects. China outspends and outgraduates the US in STEM graduates and Ph.D.'s.[18] There are more Chinese computer science graduates in a year than the US has graduated in all of its STEM programs.[19] Concerning AI, in particular, the US' top twenty accredited universities' Centers of Academic Excellence in Cyber Operations (CAE-CO) offer fewer AI/ML classes for their students than the equivalent eleven schools in China.[20]

Furthermore, China uses its centralized authority to mandate AI education in its high school curricula and requires AI companies to partner with schools and universities to help train students. Whereas the US federal approach is piecemeal, hampered by the Constitutional allocation of education sovereignty to individual states and even districts. Since 2018, the Chinese government approved at least 345 universities to offer an AI major – now

the country's most popular – and at least 34 universities have launched their own AI institutes.[21] The US, by comparison, is experimenting with AI education curricula and industry partnership initiatives, but still in a piecemeal way that varies by state and places a heavier emphasis on computer science education.

Other emerging technologies such as quantum or advanced bio-engineering fare little better. At least for the developers and potential clients of those new and emerging technologies, the awareness exists that the new technology can be corrupted. Quantum computing, for example, is widely viewed as both a benefit to cybersecurity due to the difficulty of breaking its encryption or obtaining unauthorized decryption but also as a potential disaster for the currently dominant asymmetric, public-key encryption that could, in the near future, be made obsolete by quantum decryption.[22] The main adversary for consolidated democracies, China, has noted the connection between cyberspace and quantum or AI/ML or space. It is making massive strides in keeping with its scale and strategic coherence advantage, funding, mandating, and facilitating quantum research and production, along with its computer science, artificial intelligence, and other emerging technology initiatives. The chief adversary has learned the lesson, still just being admired by the democratic defenders.

If the Apple Does Not Fall Far from the Tree – Learn to Transform the Tree

The consequences of not learning how to defend both the cyberspace substrate and its offspring, such as AI, can be dire in the future. While most emerging technology initiatives across consolidated democracies are waiting on lagging market forces to notice, care about, catch up with, and then act to defend cyber/AI systems and any other new technologies, a horde of AI-familiar students exit Chinese high schools and then college to transform their national socio-technical-economic system (STES) and their future prosperity. In the near-medium future, the US and its allies will face not only a tsunami of overwhelming cyber-attacks, which they will have to defend but will also experience vastly greater losses as their advanced and emerging technologies are stolen, corrupted, disrupted, or remotely controlled. The exponential growth of adversaries' skilled attackers, processes, and advanced technologies offers opportunities for more accidents, errors, and malicious interventions.

China also suffers from the same NMPY attitude among its developers of emerging technology. Still, China's leadership has already moved from problem admiration to action to protect its substrate and, by extension, its own development and use of emerging technologies. Certainly, the leadership has learned the lesson about the connection between poor underlying cybersecurity and the future prosperity and defense of the national STES. Its own expertise in cyber-attacks on other nations' cyberspace and emerging technologies indicates what could be done to China itself should the democratic states engage in the same cyber campaigns and mirror China's poor state behaviors in hacking. Fearing those reprisals or exploitation in return, the Chinese regime has created a National Cybersecurity Center to address the underlying insecurity of its cyberspace substrate with generous funding and government top-down authorities.[23]

One of its goals is to displace foreign technology, a process that could mean replicating shoddy existing technology with mirrored equivalents built in China and perhaps coded in native Mandarin. However, it is likely also to mean that, given the surge in the national development of computer science and related talent, this pursuit will generate at scale a bevy of discoveries and innovations. If so, Chinese computer science will transform its own national substrate into something more secure for the benefit of China alone, further ensuring the gap between its technological vigor and the digital security of the consolidated democracies.

Leaving this lesson unlearned is a major national security threat for the US and its allies if these numbers hold true in the future. "[B]y the year 2025 more money will be spent on artificial intelligence software and services than on infrastructure-as-a-service and platforms-as-a-service."[24] The race to incorporate artificial intelligence in the form of machine learning and reinforcement learning is showing a pattern of "adopt-before-securing" just like the race to computerize that engulfed the advanced economies in the early 2000s. That earlier wildly enthusiastic develop-and-promote era cemented technological flaws and insecurities deeply within the cyberspace substrate – the same underlayment that threatens these new emerging technologies. The difference now is that a major authoritarian nation is rushing in with the intent to dominate the technology globally. Even if it is rushed and somewhat slipshod, China's progress in replacing foreign technology with indigenous means that the fragmented democracies will be left with emerging technologies more vulnerable to compromise. At the same time, they will find it harder to defend against the more technologically literate, large state. Over time, they will also fall further behind in their ability to reciprocate attacks by disrupting the adversary's increasingly different and likely more defensible cyber substrate.

Action is needed now, and it must be collective action across the consolidated democracies to increase their scale and the pace of transformation of the underlying cyberspace substrate and the emerging technologies themselves.[25] Unfortunately, there is no automatic win for the westernized "likeminded" nations in the future if this lesson remains unlearned. Whatever is done by either authoritarian adversaries or consolidated democracies to transform their underlying vulnerable cyberspace into a more securable substrate will increase the chances of defensible operations for their technological advances emerging now and in the future. The only question is whether the consolidated democracies or the huge rising authoritarian peer power learns or benefits from this lesson sooner. The quicker student wins the better future, and the future of democracy as a desirable model for a prosperous, advanced, and secure society rests on the outcome.

NOTES

1. A seminal Berkeley professor of complexity, technology, and society, Dr. Todd R. LaPorte, commented once that often the most significant role of social scientists is to make the self-evident obvious to those who were not paying close attention. Personal discussion.
2. Barack Obama, White House Press Conference: Remarks by the President on Securing our Nation's Cyber Infrastructure, US Government (Washington DC, 2009), http://www.whitehouse.gov/the_press_office/Remarks-by-the-President-on-Securing-Our-Nations-Cyber-Infrastructure/.
3. Francesca Spidalieri, "One Leader at a Time: The Failure to Educate Future Leaders for an Age of Persistent Cyber Threat," Pell Center for International Relations and Public Policies, March 2013, https://salve.edu/sites/default/files/filesfield/documents/pell_center_one_leader_time_13.pdf.
4. Håvard Myrbakken and Ricardo Colomo-Palacios, "DevSecOps: A multivocal literature review," paper presented at the International Conference on Software Process Improvement and Capability Determination, https://www.researchgate.net/profile/Havard-Myrbakken/publication/319633880_DevSecOps_A_Multivocal_Literature_Review/links/59da4c47a6fdcc2aad12a134/DevSecOps-A-Multivocal-Literature-Review.pdf 2017.
5. Runfeng Mao et al., "Preliminary findings about devsecops from grey literature," paper presented at the 2020 IEEE 20th International Conference on Software Quality, Reliability and Security (QRS), https://qrs20.techconf.org/QRS2020_FULL/pdfs/QRS2020-4LGdOos7NAbR8M2s6S6ezE/891300a450/891300a450.pdf 2020.
6. Melanie Mitchell, Artificial intelligence: A guide for thinking humans (New York: Picador, 2019), https://melaniemitchell.me/aibook/.
7. Pavel Laskov and Richard Lippmann, "Machine learning in adversarial environments," Machine Learning 81, no. 2 (2010), https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.375.4564&rep=rep1&type=pdf.
8. For an example, see the OpenAI blog found at https://openai.com/blog/adversarial-example-research/.
9. Patrick Tucker, "US Needs to Defend Its Artificial Intelligence Better, Says Pentagon No. 2 – AI safety is often overlooked in the private sector, but Deputy Secretary Kathleen Hicks wants the Defense Department to lead a cultural change," Defense One, June 22, 2021, https://www.defenseone.com/technology/2021/06/us-needs-defend-its-artificial-intelligence-better-says-pentagon-no-2/174876/.
10. Patrick Tucker, "Vulnerabilities May Slow Air Force's Adoption of Artificial Intelligence – More data on the battlefield means a wider attack surface, something the Defense Department has yet to prepare for, experts say," Defense One, September 23, 2021, https://www.defenseone.com/threats/2021/09/vulnerabilities-may-slow-air-forces-adoption-artificial-intelligence/185592/.
11. Dr. Hyrum Anderson at Microsoft, Enigma 2021 talk, https://www.youtube.com/watch?v=RWx_DitGF7A.
12. Term originally used by John Mallery, CEO of Work Factors Group in multiple presentations 2010-2020, Cambridge.
13. Zachary Arnold and Helen Toner, AI Accidents: An Emerging Threat – What Could Happen and What to Do, Georgetown University Center for Security and Emerging Technology CSET (July 2021), https://cset.georgetown.edu/publication/ai-accidents-an-emerging-threat/.
14. Andrew Lohn, Poison in the Well: Securing the Shared Resources of Machine Learning, Georgetown University Center for Security and Emerging Technology (CSET), June 2021, https://cset.georgetown.edu/publication/poison-in-the-well/.
15. Billy Mitchell, "As Air Force adopts AI, it must also defend it, intelligence chief says," FedScoop, September 22, 2021, https://www.fedscoop.com/air-force-artificial-intelligence-algorithm-defend-intelligence-chief-mary-obrien/.
16. Author personal conversations within the US AI defense community.
17. Mitchell, "As Air Force adopts AI, it must also defend it, intelligence chief says."
18. Remco Zwetsloot et al., China Is Fast Outpacing U.S. STEM PhD Growth, Georgetown University CSET, August 2021, https://cset.georgetown.edu/publication/china-is-fast-outpacing-u-s-stem-phd-growth/.
19. Steven Overly and Melissa Heikkila, "China wants to dominate AI. The U.S. and Europe need each other to tame it," Politico, March 2, 2021, https://www.politico.com/news/2021/03/02/china-us-europe-ai-regulation-472120.
20. Dahlia Peterson, Kayla Goode, and Diana Gehlhaus, AI Education in China and the United States — A Comparative Assessment, Georgetown University Center for Security and Emerging Technology,September 2021, https://cset.georgetown.edu/publication/ai-education-in-china-and-the-united-states/.
21. Peterson, Goode, and Gehlhaus, AI Education in China and the United States – A Comparative Assessment.

NOTES

22. Neha Sharma and Ramkumar Ketti Ramachandran, "Emerging Trends of Quantum Computing: The Emerging Trends of Quantum Computing Towards Data Security and Key Management," Archives of Computational Methods in Engineering, April 28, 2021.
23. Dakota Cary, China's National Cybersecurity Center – A Base for Military-Civil Fusion in the Cyber Domain, CSET Georgetown University Center for Security and Emerging Technology CSET,July 2021, https://cset.georgetown.edu/publication/chinas-national-cybersecurity-center/.
24. Simon Sharwood, "AI to be bigger than IaaS and PaaS combined by 2025 – $500bn a year to be spent on electro-brain and supporting tech vs $400bn on cloud infrastructure," The Register, August 6, 2021, https://www.theregister.com/2021/08/06/idc_ai_paas_iaas_predictions/.
25. Chris C. Demchak, "Cyber Competition to Cybered Conflict," eds, Emily Goldman, Jacqueline Schneider, and Michael Warner, Ten Years In: Implementing New Strategic Approaches to Cyberspace (Naval War College Press, The Newport Papers, 2020, https://digital-commons.usnwc.edu/usnwc-newport-papers/45/.