Atlantic Council
SCOWCROFT CENTER
FOR STRATEGY AND SECURITY

Emergent
Futures
Lab

Scenario Two: AI and Insecurity for All: The Future of Cyber Conflict

Trends

Trend 1: *AI Development*

As the proliferation and integration of narrow-AI applications penetrate our daily lives, they alter security relationships. The degree and manner depend on which sector develops AI-enabled cyber capabilities at scale first: nation-states or private sector.

Trend 2: *AI Fragmentation*

The use of AI-enabled cyber capabilities will create trust and coordination challenges depending on the extent to which the Internet becomes increasingly fragmented along sovereign lines or if it remains defined by an open architecture.

Scenario Overview

Artificial intelligence (AI) is the label given to technologies that learn from their environment to dictate and extend the boundaries of possibility for tasks traditionally done by man. At the high end, AI will empower nations to achieve their goals in new and innovative ways, and to escape many of the shackles of both practicality and law that have historically dictated the development of the mechanisms of state power.

There are two major types of AI: narrow AI, where a specific algorithm is developed to address a task or function, and general AI, where a machine intelligence is built which is indistinguishable from human cognition or sentience. In terms of the actual technologies involved, AI is a basket of techniques and systems designed to enable sophisticated machine interaction with environments. In addition to deep reasoning and learning capacities, AI also includes any technology that allows machines to move independently or to sense their environment. Together, AI technologies enhance the speed of decision making, increase potential independence from human-in-the-loop decision making, and expand the range of complex activities that are possible in short time scales. With cyber capabilities, AI augmentation of existing methods and architectures will introduce new opportunities for strategic operations, but also new challenges.

First, AI will potentially lower the barriers to entry for new, technologically less adept adversaries to develop and employ sophisticated capabilities. Second, there is prospect of an arms race of sorts between defenders and offensive actors to automate and make intelligent this automation ahead of each other. Finally, the spread of AI might facilitate cyber-attacks which regularly take place along multiple vectors (or impact multiple related systems), making it difficult to identify and disrupt the sequence of events supporting an attack.

AI also factors into broader, external geostrategic dynamics, the combination of which will then dictate the dynamics of future cyber conflict and the Internet. Two sets of conditions include (1) whether states or private industry drive AI developments; and (2) whether the Internet is open or fragmented; these two sets mix to create four possible future worlds.

The first future world occurs when the current world order is subverted, leading to an Internet is no longer global, but fragmented among Western and authoritarian camps, and the AI arms race is driven by national political interest. The second future world focuses on defending the global commons, where the world compromises to maintain a global Internet but state interests drive AI development albeit with international coordination on issues like cybercrime.

The third future world revolves around micro-fragmentation, where an open global Internet exists but proprietary pockets of AI research and development, led by the private sector, cause increasing vulnerability of services and IoT, which forces successful companies to take an active role in deterring attacks on their products. Finally, the fourth world exists within a crisis of sovereignty, where the Internet is divided. Private sector AI development conflicts with widespread internet fragmentation, and routing and supporting infrastructure is exposed worldwide as companies choose which Internet fragments to support thereby permitting or denying access of their research to military and intelligence operations.

20XX Vignette #2

Captain Wiggins couldn't tell if it was day or night. She had been shifting schedules on short notice in the operation center, surging to chase shadows for over twenty months. It started with a new tailored spearphishing campaign sold on the dark web that leveraged off-the-shelf AI/ML scripts to identify recurring patterns from social media and optimize tailored solicitations. It was increasingly hard to find the line between nation-states and cyber mercenaries. After the last tech bust, large numbers of data scientists and neural network experts sold their talents to a mix of private firms that occupied a gray space between illicit networks, fronts for nation-states, and legitimate corporate clients. As a result, cyber arsenals expanded with a wide range of malware, often optimized to learn in contact.

The ensuing wave of attacks changed the strategic landscapes. There were just too many cyberattacks from a range of state and non-state actors for national governments to track them all, much less engineer the type of sophisticated countermeasures and defensive layers required to stop malware that learned as it probed the network. Adding to the chaos due to stalled political debates about the cost and vulnerability of building 5G networks, major cities had been left to fill a security vacuum. Larger cities started to pay for "digital lances," a new term for firms that provide high-end defenses for municipal networks. These twenty-first century hacker knights provided mixed results and some even moonlighted conducting attacks for the highest bidder. Captain Wiggins' response team did what they could to identify which of these private firms could be trusted so they could focus limited resources on targeting the more sophisticated threat actors, it was hard to keep up.

Near-Term Implications (6 months to 1 year)

- Possibility of new classes of persistent threat tools that can adapt to countermeasures

- Further blurring of the line between influence operations and cyber operations that creates new approaches to social engineering and spearphishing at scale

- New vulnerability vectors associated with 5G at scale and the Internet of Things

Long-Term Implications (> 5 years)

- Increased probability of an AI arms race
- New opportunity costs; the increased need to harden targets diverts money from offense to defensive tools and tool kits

Scenario Author:

Chris Whyte is an assistant professor of homeland security and emergency preparedness at VCU. His research interests include a range of international security topics related to the use of information technology in war and peace, political communication and cybersecurity doctrine/policy.

Trey Herr is the Director of the Scowcroft Center for Strategy and Security's Cyber Statecraft Initiative at the Atlantic Council.

Will Loomis is a Program Assistant in the Scowcroft Center.

Twitter: @CyberStatecraft