Governing Cyberspace during a Crisis in Trust

An essay series on the economic potential — and vulnerability — of transformative technologies and cyber security
## Contents

- **Governing Cyberspace during a Crisis in Trust** ........................................ 4  
  Aaron Shull

- **Tackling Cyber-enabled Crime Will Require Public-Private Leadership** ........ 9  
  Neil Desai

- **Election Cyber Security Challenges for Canada** ...................................... 16  
  Elizabeth F. Judge and Michael Pal

- **State and Surveillance** ............................................................................. 21  
  David Lyon

- **Trust and Data** ....................................................................................... 26  
  Paul Vallée

- **Beware Fake News** .................................................................................. 31  
  Eric Jardine

- **The Need for a National Digital Identity Infrastructure** ......................... 36  
  Andre Boysen

- **The Emerging Internet of Things** .......................................................... 41  
  Christopher S. Yoo

- **The Cyber Security Battlefield** ............................................................... 45  
  Robert Fay and Wallace Trenholm

- **TLD Operator Perspective on the Changing Cyber Security Landscape** .... 49  
  Byron Holland

- **Strategic Stability, Cyber Operations and International Security** ............ 55  
  David Mussington

- **The Quantum Threat to Cyber Security** ................................................ 60  
  Michele Mosca and Bill Munson

- **Mitigating Cyber Risk across the Financial Sector** ............................... 64  
  Christian Leuprecht

- **The Danger of Critical Infrastructure Interdependency** ......................... 69  
  Tyson Macaulay

- **Programmable Trust** .............................................................................. 74  
  Michael Mason and Matthew Spoke

- **Patching Our Digital Future Is Unsustainable and Dangerous** ............... 80  
  Melissa Hathaway

- **Canada and Cyber Governance** ............................................................... 91  
  Stephanie Carvin
It is almost impossible to read the news without coming across a lead story cataloguing the latest cyber breach or misuse of data. Intellectual property is being stolen from companies at an alarming rate. Foreign actors are meddling in elections through fake social media accounts, along with other more nefarious means — including the surreptitious access of internal campaign emails. Criminals use the dark recesses of the internet to sell drugs, guns and even people. And terrorist groups use digital media to recruit and inspire prospective adherents the world over.

This is not even the worst of it. Whether one views Edward Snowden as a hero or a traitor, the fact is that the information he revealed regarding the extent of governmental surveillance and the close relationship between traditionally distinct public and private entities has damaged systemic trust in a profound way. On top of this more robust surveillance state, countries are creating advanced cyber weapons capable of devastating real-world effects. At the same time, more and more critical infrastructure is being digitally enabled and is also, therefore, capable of being digitally disabled by bad actors.

The number of companies and governments that have fallen prey to digital hacking are almost too numerous to count — Ashley Madison, the Bank of Montreal, CIBC, eBay, Equifax and JP Morgan Chase offer ready examples. The volume of these events lays bare the paradox of the digital economy and cyber security. On the one hand, technology has led to convenience, efficiency and wealth creation — and so, companies push to connect everything that can be connected. On the other hand, this great push to digitize society has meant building inherent vulnerability into the core of the economic model. This is all taking place atop a deeply fragmented and underdeveloped system of global rules.

Given this paradox, the purpose behind this essay series on security in cyberspace is threefold. First, it brings together an interdisciplinary team, including the private sector, academics and leading experts to provide creative ideas and fresh thinking in these emergent areas surrounding data governance, cyber security and new technology. Second, it aims to advance a public policy debate that recognizes that while cyber security threats are increasing in both number and sophistication, there is economic potential for Canadian firms to capitalize...
on a growing market. Third, it argues for the advancement of a more stable international institutional order. The international rules-based system in cyberspace is still in its infancy, and innovative thinking is needed to make sure that Canada can play a leadership role in crafting the governance architecture.

The National Cyber Security Strategy released in June 2018 marks an important step forward for Canada in the cyber domain — but there is much left to do. The strategy advances Canadian interests in a number of ways. For example, it recognizes that “cyber security is the companion to innovation and the protector of prosperity” and cyber security is now an essential element to a functioning innovation economy (Public Safety Canada 2018).

The Government of Canada’s efforts in this area are set out in three themes:

- security and resilience (to enhance cyber security capabilities to better protect Canadians and defend critical government and private sector systems);
- cyber innovation (to position Canada as a global leader in cyber security); and
- leadership and collaboration (to have the federal government act as a leadership point in Canada and work to shape the international cyber security environment in Canada’s favour).

Given this national and international context, this essay series takes a broad view of cyber security and addresses a range of topics, from the governance of emerging technology, including artificial intelligence and quantum computers, to the dark Web and cyber weapons. The unifying question underlying this effort is: how can we build a system of governance that enhances global systemic trust and creates opportunities for “middle” power countries such as Canada to advance their strategic and economic interests through enhanced governance arrangements? Cyberspace presents both threats and opportunities — at the same time — and the collective challenge is to advance policy that can best maximize the opportunities while mitigating the threats in a constantly changing global environment.

While at first glance the themes set out in the National Cyber Security Strategy seem categorically discrete, if implemented properly in a way that also accounts for...
credible effort for Canada to try and position itself as a global leader in the field.

On the theme of cyber innovation, the unfortunate fact is that the cyber security industry is growing, based on necessity. This creates both an economic imperative and an opportunity. According to the recent Canadian Survey of Cyber Security and Cybercrime, conducted by Statistics Canada, Canadian businesses are spending approximately $14 billion on cyber security per year. At the same time, various estimates present staggering figures representing the loss to the Canadian economy because of cyber crime and espionage. In this way, the cyber security industry contributes significantly to Canada’s economy, while cyber criminals and foreign adversaries act like a parasite on that value. Moreover, while the growth projections vary, it seems clear that the industry and the economic opportunities created for Canadian companies will continue to grow along with the threat. It will be important to advance domestic policy that allows the best Canadian firms to grow at home, but also to reach international markets.

It will be equally imperative that Canada push to advance global rules or norms in cyber space that foster greater stability. This is particularly important because a lack of clear rules or norms contributes — at least in some way — to a permissive environment where adversarial actors take advantage of the ambiguity in the rules to launch offensive cyber operations against both Canadian business and governmental actors.

On the theme of global rules, there are a number of important observations reflected in Canada’s defence policy. First, that the “most sophisticated cyber threats come from the intelligence and military services of foreign states” (National Defence 2017, 56). Second, that it is the technologically advanced governments and private businesses that are the most vulnerable to state-sponsored cyber espionage and other forms of cyber aggression. Third, that the threat from these forms of state-sponsored cyber aggression will likely continue for the foreseeable future. Finally, that addressing “the threat is complicated by the difficulties involved in identifying the source of cyber attacks with certainty and the jurisdictional challenges caused by the possible remoteness of cyber attacks” (ibid.).
This has led to a deeply contested operational environment in cyber space. According to Strong, Secure, Engaged: Canada’s Defence Policy:

State and non-state actors are increasingly pursuing their agendas using hybrid methods in the “grey zone” that exists just below the threshold of armed conflict. Hybrid methods involve the coordinated application of diplomatic, informational, cyber, military and economic instruments to achieve strategic or operational objectives. They often rely on the deliberate spread of misinformation to sow confusion and discord in the international community, create ambiguity and maintain deniability. The use of hybrid methods increases the potential for misperception and miscalculation. Hybrid methods are frequently used to undermine the credibility and legitimacy of a national government or international alliance. By staying in the fog of the grey zone, states can influence events in their favour without triggering outright armed conflict. The use of hybrid methods presents challenges in terms of detection, attribution and response for Canada and its allies, including the understanding and application of NATO’s Article 5. (ibid., 53)

The ability of foreign adversaries to operate in this grey zone is possible because there is no universally understood set of norms that apply in cyber space. Rather, actors stretch to interpret existing legal rules, such as the United Nations Charter and international humanitarian law, as applicable to technologies and actions that could not have been contemplated at the time that the law was written.

This trend is worrying, although it is not surprising because it is simply a digital continuation of geopolitical rivalry. David Vigneault, director of the Canadian Security Intelligence Service, recently remarked that “economic espionage represents a long-term threat to Canada’s economy and to our prosperity” (Vigneault 2018). He based this assessment on “a trend of state-sponsored espionage in fields that are crucial to Canada’s ability to build and sustain a prosperous, knowledge-based economy [including] areas such as A.I., quantum technology, 5G, biopharma, and clean tech” (ibid.). Owing to the highly sophisticated nature of these efforts, the reality is that adversarial nations are targeting “the foundation of Canada’s future economic growth” (ibid.).

This is all leading to a crisis in trust at the individual level, with individuals feeling vulnerable online. In 2018, CIGI conducted a Global Survey on Internet Security and Trust in partnership with Ipsos, a global market research and a consulting firm. The firm surveyed internet users in 25 countries: Australia, Brazil, Canada, China, Egypt, France, Germany, Great Britain, Hong Kong (China), India, Indonesia, Italy, Japan, Kenya, Mexico, Nigeria, Pakistan, Poland, Russia, South Africa, the Republic of Korea, Sweden, Tunisia, Turkey and the United States. The results were telling: over half of internet users surveyed around the world are more concerned about their online privacy than they were a year ago, reflecting growing concern around the world about online privacy (CIGI and Ipsos 2018).

Even more dangerous is the breakdown in trust between states. Perhaps there is no clearer Canadian example than recent events surrounding the China-based company Huawei, and the discussions related to banning the company from Canada’s 5G networks. 5G is the term used to represent the fifth-generation cellular mobile communications, succeeding LTE (long-term evolution standard). It will be the backbone of the digital economy and lay the foundation for mobile computing, smart sensors, advanced automation and the Internet of Things — including connected cars. In fact, the issue has become so politically charged that two members of the US Senate Select Committee on Intelligence, Republican Senator Marco Rubio and Democrat Senator Mark Warner, wrote to Canadian Prime Minister Justin Trudeau urging him to ban Huawei from Canada’s mobile network. The two senators wrote: “We are concerned about the impact that any decision to include Huawei in Canada’s 5G networks will have on both Canadian national security and ‘Five Eyes’ joint intelligence cooperation among the United States, United Kingdom, Australia, New Zealand, and Canada” (Warner and Rubio 2018).
Clearly there is a deep geopolitical rivalry at play that will come to define the distribution of both wealth and power in the future. This has led to the Paris Call for Trust and Security in Cyberspace, an effort that Canada — along with 65 other states — supports. This broad call for trust is an effort to get states to agree to a set of international rules for cyberspace. However, it falls well short of a detailed treaty. Rather, it is a very high-level, non-binding document that, at its highest, is a call for states to "promote the widespread acceptance and implementation of international norms of responsible behavior as well as confidence-building measures in cyberspace" (France Diplomatie 2018, 3).

Given that existing international arrangements have created an environment in which trust is being routinely undermined and where there is heightened potential for miscalculation, it seems clear that more needs to be done. More robust global rules could create enhanced stability and foster increased trust. Pursuing this agenda is a national imperative, because nothing short of the future of the Canadian economy hangs in the balance.

**Works Cited**


**Endnotes**

1 See www150.statcan.gc.ca/n1/daily-quotidien/181015/dq181015a-eng.htm.

2 See www150.statcan.gc.ca/n1/daily-quotidien/181015/dq181015a-eng.htm.
Picture an ordinary town. On one side of the community there are flourishing businesses and beautiful homes, a lush park, and a school. But they are cordoned off as a gated community with their own private security.

The other side of the wall is completely different. Levels of crime and despair are high. Drugs and weapons are widely available. The sex trade is rampant, including trafficked children. And violence stemming from hatred is a constant fear. Police presence beyond the occasional drive-by is non-existent.

This may sound like an opening scene from a futuristic dystopian film. However, if we look to areas of the internet today, the analogue analogy isn’t far-fetched relative to our current digital reality.

As a result of the rapid pace of technological innovation and the decisions embedded in the software that make up social networks, online marketplaces and other parts of the internet and digital devices, existing laws are often no longer enforceable and the principles that form the bedrock of the rule of law are eroding. This, coupled with the limited legal and technological tools available to law enforcement in the digital age, places ordinary citizens — in particular vulnerable populations such as children, seniors and individuals suffering from mental health challenges — at risk.

Traditional crimes such as burglary and motor vehicle theft are generally down in the liberal-democratic world. In Canada, Criminal Code violations are down significantly. In 2006, there were 2.4 million offences compared to 1.9 million in 2016.1 Breaking and entering dropped from 250,000 occurrences to 160,000; motor vehicle theft decreased from 159,000 to 79,000 over that period.2

These statistics would lead us to believe we’re safer. But looking at the crime statistics through another lens paints a much different picture. Crimes such as the sexual exploitation of children, human trafficking, fraud, and terrorism and mass casualty incidents are all up. For example, Statistics Canada reported in 2016 that child sexual exploitation increased for eight years in a row at a rate of 233 percent (Harris 2017).

Neil Desai is an executive with Magnet Forensics, a Canadian technology company that develops digital forensics software for more than 4,000 police, national security and other public and private agencies with investigative authorities in 93 countries. He also serves as a fellow of the Munk School of Global Affairs and Public Policy at the University of Toronto and faculty at Singularity University.
These crimes are by no means new. What ties their contemporary growth together is the ease by which they can be committed with relative impunity in the digital age. This isn’t to say there aren’t laws against the sexual exploitation of children, fraud or other crimes enabled by digital connectivity. Nor is it to say police agencies don’t have personnel and tools to investigate digitally enabled crimes. In fact, most agencies of scale in advanced industrialized countries have what are known as digital forensics labs to investigate cyber-enabled crimes by collecting digital devices with critical evidence and examining them. This is costly and deeply technical, and is reserved for major crimes such as homicides and organized crime. The general trend is that these investigations are becoming even more complex due to emerging technological trends and the lack of either a technical or public policy response. This has been further exacerbated by a growing chasm between police agencies and emerging technology companies on their respective societal roles.

Technologies Enabling Crime in the Digital Age

Cloud Computing

Cloud computing is transforming the internet and its application — in large part for the better. Simply put, cloud computing off-loads the cumbersome and expensive digital infrastructure required to run applications, store data and enable advanced analytics to

Comparison between Traditional Criminal Techniques and Cybercrime

<table>
<thead>
<tr>
<th>Traditional criminal techniques</th>
<th>Cybercrime</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Burglary</strong>: Breaking into a building with the intent to steal.</td>
<td><strong>Hacking</strong>: Computer or network intrusion providing unauthorized access.</td>
</tr>
<tr>
<td><strong>Deceptive callers</strong>: Criminals who telephone their victims and ask for their financial and/or personal identity information.</td>
<td><strong>Phishing</strong>: A high-tech scam that frequently uses unsolicited messages to deceive people into disclosing their financial and/or personal identity information.</td>
</tr>
<tr>
<td><strong>Extortion</strong>: Illegal use of force or one’s official position or powers to obtain property, funds or patronage.</td>
<td><strong>Internet extortion</strong>: Hacking into and controlling various industry databases (or the threat of), promising to release control back to the company if funds are received or some other demand satisfied.</td>
</tr>
<tr>
<td><strong>Fraud</strong>: Deceit, trickery, sharp practice, or breach of confidence, perpetrated for profit or to gain some unfair or dishonest advantage.</td>
<td><strong>Internet fraud</strong>: A broad category of fraud schemes that use one or more components of the Internet to defraud prospective victims, conduct fraudulent transactions, or transmit fraudulent transactions to financial institutions or other parties.</td>
</tr>
<tr>
<td><strong>Identity theft</strong>: Impersonating or presenting oneself as another in order to gain access, information, or reward.</td>
<td><strong>Identity theft</strong>: The wrongful obtaining and using of another person’s identifying information in some way that involves fraud or deception, typically for economic gain.</td>
</tr>
<tr>
<td><strong>Child exploitation</strong>: Criminal victimization of minors for indecent purposes such as pornography and sexual abuse.</td>
<td><strong>Child exploitation</strong>: Using computers and networks to facilitate the criminal victimization of minors.</td>
</tr>
</tbody>
</table>

third parties, including some of the largest corporations in the world, such as Amazon and Microsoft. This has significantly reduced the operational costs of starting a digitally enabled or wholly online business given the cost and technical advantages.

While the economic advantages of the cloud are clear, the downsides to the criminal justice system are lesser known. Prior to the advent of cloud computing, investigators could petition courts, through warrants, to confiscate digital devices such as computers and smartphones where it was suspected they contain critical evidence regarding a case. Using their technical background and software tools, they could recover critical evidence that resided on the local device, analyze it and report on it to the broader justice sector.

Police agencies still have the ability to petition the court. However, cloud adoption is making the gathering of digital evidence difficult. Many data centres that store cloud-based data are located in foreign jurisdictions, often where the company can get economies of scale, better tax treatment, access to technical talent and favourable data policies. This adds a layer of complexity as law enforcement agencies must work in concert with their national governments and their foreign partners to request such data from private vendors. In the case of most liberal democracies, Mutual Legal Assistance Treaties (MLATs) exist to help streamline these judicially authorized requests for evidence from a vendor in another country.

This approach is not a perfect solution, as the requests must also comply with domestic laws, including those that pertain to data privacy. For example, in surrendering data about a customer, a private vendor could not turn over data about another individual. Often such data isn’t housed in a neat manner, placing the vendor at risk of breaking domestic law if they were to comply with the MLAT process. The process is opaque in cases where a treaty does not exist with a country where critical data for a law enforcement investigation resides.

Even when critical data regarding investigations can be tracked down to a cloud service provider and an MLAT exists between the countries in question, the process is quite time consuming and costly, and places investigations at risk. For these reasons, this process is often reserved for major crimes.

This issue reached a tipping point in the United States. In 2018, Congress passed legislation titled the Clarifying Lawful Overseas Use of Data (CLOUD) Act. Principally, the act asserts that US technology companies must provide data about US citizens on any server they own or operate when requested by a legitimate court order. The CLOUD Act was informed heavily by a 2013 Federal Bureau of Investigation (FBI)-organized crime investigation, which moved slowly through the courts, over access to data about a suspect who had been protected by critical evidence on a Microsoft server in Ireland. This case made its way into the queue at the Supreme Court before Congress acted. Microsoft welcomed a legislative response to this gap in public policy given the risk posed to the reputation of its emerging cloud and cloud-enabled business lines (Cheng 2018).

**Encryption**

While cloud computing is transforming technology in general, for law enforcement there is still often critical evidence on smartphones, computers and now some Internet of Things devices, including text messages, pictures and videos, as well as behavioural data points such as geolocation information or time stamps. Much of the efforts of digital forensics labs in police agencies are focused on this area.

However, in recent years, there has been a technological transformation to device security, specifically around smartphones, that has hindered law enforcement investigations. All major smartphone manufacturers now include encryption security on the physical device as a default setting and link it to a passcode or other form of unlocking, such as fingerprint or facial recognition software. If too many attempts to gain access occur, the device is wiped of all data on the physical drive. This has led to fewer stolen smartphones, but has had significant implications for other law enforcement investigations.

A 2016 terrorism-related case in San Bernardino, California, is the most prominent instance where smartphone encryption has rendered digital forensics futile. The FBI stated it could not access critical evidence on the suspect’s device and petitioned the court to compel the device manufacturer, Apple, to assist in unlocking the device. Apple stated that it did not have the technical capability to
comply with such an order at that time and would not re-engineer its encryption to allow for that ability in the future.3

The issue goes well beyond a single terrorism case. For example, in 2016, the New York County District Attorney (DA), Cyrus Vance, stated there are more than 400 iPhones related to serious crimes that the DA office could not access even with a court order. Other jurisdictions report relatively similar outcomes (Tung 2016).

Vance also called on the US government to introduce legislation that would address this issue. Laws proposed included ones compelling device manufacturers to maintain encryption keys or suspects to unlock their devices should a court order exist. However, a constitutional challenge could arise from the latter, under the US Constitution’s Fifth Amendment, which states that no American shall be compelled in a criminal case to be a witness against themselves. Such legislation has been debated in the United States and the United Kingdom, but no device encryption laws have been enacted to date.

In the absence of enabling public policy, law enforcement has utilized technologies that find vulnerabilities in smartphone encryption to unlock devices and perform digital forensics. While this approach is successful in some cases, it is a costly and imperfect science, given that the device manufacturers regularly update software.

Application encryption is another challenge. Commonly used messaging, social media, marketplace and business productivity “apps” employ encryption technology to provide security for users. This is a lesser challenge for digital forensics professionals, relative to device encryption, as decryption keys are required for the intended audience to consume the data. They can be utilized in law enforcement investigations as well. However, certain applications are architectured to maintain user anonymity throughout the life cycle of data creation and storage. Further, some app developers change the way they store critical data on a regular basis, through their product updates, which can seize the abilities of digital forensics tools and professionals. Criminals, including child sexual predators, terrorists, fraudsters and those involved in organized crime, who understand such technologies can and have leveraged them.

In late 2018, Australia became the first jurisdiction to legislate against application encryption. The law enables Australian police...
agencies to compel companies to create a technical function to give them access to messages related to a specific user via the technology provider without the knowledge of the user (BBC News 2018).

**The Deep and Dark Web**
Sophisticated criminals have also benefited from the “deep” and “dark” Web. The origins of this cavernous, unindexed area of the common internet is often debated. The US Naval Research Laboratory is attributed as the creator of technology behind “The Onion Router,” shortened to Tor, which protects user identity by routing traditional location services and Internet Protocols (IPs) through a number of countries. The US government’s intended use was to provide a mechanism for dissidents in repressive regimes to be able to communicate without tipping off authorities (McCormick 2013).

Today, the deep and dark Web has become a safe haven for criminal enterprise. A 2016 study by King’s College London researchers found that of 2,723 sites they were able to analyze over the deep Web, 1,547 hosted illicit content including drugs, child sexual exploitation images and videos (often referred to as child pornography), weapons and money laundering (Moore and Rid 2016).

While investigating cases on the deep Web can be challenging for law enforcement due to the anonymity it provides users, some technological advancements have been made. Thorn, a US-based, global non-governmental organization, has developed technologies to assist law enforcement agencies in identifying victims of child sexual exploitation and human trafficking through images and videos found on the dark Web and open internet, leading to some successful, high-profile criminal prosecutions. However, the deep and dark Web remains a relatively safe haven for the criminal enterprise due to the anonymity it provides users.

**Cryptocurrencies**
Following the exchange of money in a criminal investigation has long been a tactic for police investigations; however, the advent of cryptocurrencies has hindered this approach. The key distinguishing feature of cryptocurrencies relative to traditional forms of currency is that they are a wholly digital asset class leveraging cryptography to secure transactions, control the creation of additional units and verify transactions. The other differentiator is that they are a decentralized currency, without a central bank or clearing house for transactions, regardless of the size of exchange. The latter feature has posed a great challenge for law enforcement investigators as they must find leads other than lawful tips from the banking sector or regulators to identify currency exchanges.

The most common form of cryptocurrency is Bitcoin. Given its popularity and prevalence in criminal investigations, larger law enforcement and national security agencies have developed some techniques to recover Bitcoin-related evidence from physical devices, known as “Bitcoin wallets.”

While those techniques have worked in some cases, it is an expensive and highly technical investigative area, limited to major crime investigations by some of the world’s largest police agencies. It also doesn’t address the fact that there are now over 1,500 known cryptocurrencies, some of which are layering other forms of security and anonymity of its users, nullifying investigative techniques.

The technical underpinning of cryptocurrencies, known as “blockchain,” is also gaining adoption in other areas of
The US government’s intended use for Tor, which protects user identity by routing traditional location services and IPs through a number of countries, was as a means of anonymous communication for dissidents in repressive regimes; today, the deep and dark Web is a safe haven for criminal enterprise. (Photo: Jarretera / Shutterstock.com)

computing. This will pose an even more complex challenge for digital forensics labs in the future, especially when blockchain is coupled with cloud computing.

The Way Forward: Understanding the Magnitude of the Challenge and Bridging Public and Private

The connectivity provided by the internet and digital devices is accelerating the volume and variety of products and services available to consumers at an unprecedented velocity. This has brought great progress in terms of social connectivity and commercial opportunities. These very technologies are challenging police agencies and governments in their societal role to keep citizens secure and upholding the rule of law. But the magnitude of the impact these technologies are having on the fundamental pillars of liberal-democratic societies, specifically in the criminal justice sector, is less known.

A key part of the response to this challenge is better data. Crime statistics in most democracies are woefully inadequate. They often don’t capture how digital technologies enable crime or categorize crimes based on pre-digital-revolution categories. This data is integral to police agencies as their funding is often tied to such crime statistics.

One area that has seen some innovation on gathering statistics is child sexual exploitation online. The Canadian Centre for Child Protection (CP3), a non-governmental organization, which operates Cybertip.ca, has developed an online portal to capture citizen reporting of such incidents. The portal receives more than 4,000 reports of online child sexual abuse images and videos each month (CP3 2017). Understanding the specific nature of this challenge with data has allowed the Government of Canada and provinces to invest strategically. For example, with additional resources, CP3 has developed an automated technology to notify hosts of illicit content to take it down within the confines of existing laws. While such data has not led to legislative change in Canada, greater public awareness could lead to political action as it permeates through Canadian society.

That being said, legislating for or against specific technologies may not necessarily have the impacts required to wholly reverse the current trend in digitally enabled crime. Technological innovation happens at a much more rapid pace than legislative change in most democracies. Further, most internet-enabled technologies can be utilized across jurisdictions with ease, which often renders domestic laws futile and international cooperation remains a slow and costly process. Legislation in this area should remain based on principles as opposed to prescriptive in terms of technological specifications, such as the removal of encryption from smartphones.

The greatest innovation required in addressing the growth in digitally enabled crimes is the relationship and feedback loops between the public and private sector — specifically law enforcement agencies and technology companies.

Technology companies have latched on to recent public opinion that law enforcement and national security agencies can’t be trusted to have access to private citizens’ digital
communications. They, in concert with privacy advocates, have banked on this sentiment to justify their development, which has rendered court orders futile and placed the rule of law in a precarious position. Conversely, law enforcement agencies have rarely highlighted the day-to-day challenges they face when it comes to cyber-enabled crimes and the growing complexity they pose. Without that level of transparency, it is difficult for citizens to make informed decisions as consumers and citizens.

This is not a sustainable approach. It’s important to consider that we’re still early in the digital age. As more citizens are affected or learn of cyber-enabled crimes, public opinion will inevitably shift. Both government and technology companies will have to rethink their positions.

Fundamentally, liberal-democratic societies cede reasonable amounts of individual civil liberties in exchange for societal security. There isn’t a static balance. It’s highly dependent on the current state of affairs. Without meaningful and regular dialogue between all vested interests, the current chasm will only grow. These parties will have to work with lawmakers around the world to reshape relevant legislation, in a principle-based fashion.

Large technology companies also have a leadership role to bear. They too must take steps to ensure their platforms are not enabling crime and despair. Balancing their intended technological outcome with preserving the rule of law and the protection of vulnerable populations must be at the centre of their development strategies. Collectively, this can be achieved if all parties agree with the maxim that technological innovation, at its best, improves people’s lives while preserving societies’ fundamental values.

Works Cited


Endnotes

1 See www150.statcan.gc.ca/n1/pub/12-581-x/2018000/crr-eng.htm.

2 Ibid.

3 See www.npr.org/series/469827708/the-apple-fbi-debate-over-encryption.
Election Cyber Security Challenges for Canada

Elizabeth F. Judge and Michael Pal

The October 2019 federal election promises to be the first one in Canadian history where “election cyber security” will play a prominent role. Election cyber security can be understood as preventing digital interference with the main actors, institutions and processes of elections. A variety of different threats to Canadian elections, from hacking of political parties to misinformation spread on social media platforms to abuses of voter privacy to foreign interference, are all real risks in 2019. This essay outlines the major election cyber security issues facing Canada by focusing on three key actors — namely, political parties, election administrators and voters. It then analyzes the implications for cyber security of the changes imposed on federal election law by the Elections Modernization Act.

Political Parties

Political parties in Canada are now sophisticated digital operations. They often use the techniques of big data analytics in which sophisticated algorithms generate inferences about voters based on massive amounts of personal information, largely collected from online activities. While still relying on traditional practices such as knocking on doors, parties increasingly operate digitally and integrate voter data into their activities. All federal parties have voter databases that contain sensitive personal information about voters. This information is collected from a variety of sources and is used for fundraising, “get out the vote” efforts and policy development, among other activities. In addition to television and radio advertising, parties now advertise extensively online. This advertising is common on social media platforms and, in particular, on Facebook. Social media advertising allows parties to microtarget messages at particular subsets of voters. Voters may be segmented by postal code, employment or education, or by choices about car models, food, shopping or entertainment, based on the theory that these correlate with political preferences.

The shift of parties into the digital space has expanded the cyber risks that they face. The Communications Security Establishment (CSE) of Canada issued a report in 2017 highlighting the risk of foreign interference to Canadian elections, especially in light of the now well-proven instances of malicious
activities in other democracies in recent years (CSE 2017). The CSE identified political parties as a weak point in election cyber security in Canada. Parties are private actors, with relatively scarce resources given their importance, and are often staffed by volunteers, especially at the riding level.

Political parties are increasingly vulnerable to hacking and present tempting targets for foreign actors. Digital interference with one of Canada’s main political parties would have widespread effects on the trust of Canadians in the electoral process and politics more generally. The hack of the Democratic National Committee in the United States around the time of the 2016 presidential election had negative consequences for American democracy. The passage of a Canadian version of the Magnitsky Act, an US statute that permits the US government to penalize foreign governments for human rights abuses, also potentially raises the likelihood of foreign interference by those state or non-state actors that may see sanctions imposed.2

Party leaders are at risk of impersonation online, if a hostile domestic or foreign entity seizes control of their Facebook page or Twitter account. The stakes involved in malicious impersonation of a party leader or a candidate are very high. Imagine, for example, the potential chaos that could ensue if a foreign entity seized the prime minister’s Twitter account.

This risk is quickly evolving as technology changes. The potential harm caused by impersonation is increasing due to “deep fake” technology in which audio and video recordings are manipulated to create extremely authentic-looking videos of political figures doing or saying damaging things. With the advent of deep fakes, it will be much harder for voters to discern the credibility of a news item or social media post. Voters may inadvertently credit false videos or may doubt the truth of videos that are in fact real, with negative repercussions for democratic debate.

**Election Administration**

Cyber security is also a key concern for Canada’s election administrators. Elections Canada is the non-partisan, independent body that administers federal elections, including managing polling stations, compiling results in ridings, conducting voter registration and so on. The Canadian Centre for Cyber Security’s *National Cyber Threat Assessment 2018* identified public institutions as being at risk of digital interference given the data that they hold and their important roles (Canadian Centre for Cyber Security 2018). The risk of interference in election administration has led the United States to declare electoral institutions to be “critical infrastructure” (U.S. Election Assistance Commission 2018).

Canada’s maintenance of a traditional paper ballot system for federal elections instead of moving to e-voting has fortunately avoided many of the cyber-hacking risks posed by electronic voting machines and internet voting. It is now clear from the experiences of jurisdictions that switched to online voting that these systems cannot yet be secured to the degree of certainty needed for citizens to have trust in the outcome. Although online voting does occur in Canada, notably in some municipalities in Ontario, these races are less likely to generate the attention of hostile foreign powers. The incentives for interference in federal elections are much higher.

Even if Canada’s adherence to the paper ballot has reduced the cyber risk, other forms of digital interference are still a concern. Election administrators have databases that they use for voter registration, which are rich targets for hacking. The internal operations of election administrators such as Elections Canada could be subject to interference to disrupt their activities and to derail elections. Some administrators oversee networked polling stations, which creates some risk.

Like political parties, election administrators are also at risk of impersonation. The “robocalls” scandal in the 2011 election involved fraudulent automated telephone calls, including some purportedly from Elections Canada, which directed voters to the wrong polling station or gave them the incorrect election date. This incident highlights how online misinformation could be spread in future elections by a tactic of impersonating Elections Canada. Social media posts, Twitter feeds, banner advertisements or phishing emails purporting to be from Elections Canada all have the potential to be used by malicious actors to suppress voter turnout.

Elizabeth F. Judge is professor of law and a member of the Centre for Law, Technology and Society at the Faculty of Law at the University of Ottawa, where she specializes in intersections of law, technology and policy.

Michael Pal is an associate professor in the Faculty of Law at the University of Ottawa, where he is the director of the Public Law Group. He specializes in the comparative law of democracy and comparative constitutional law. He has advised governments at all levels on election and constitutional law.

---

1. Judge and Pal 2017
2. Eaves and Lippert 2018

---

Digital interference with one of Canada’s main political parties would have widespread effects on the trust of Canadians in the electoral process and politics more generally.
by sowing confusion. While the automated telephone calls wrought serious damage, the reach of such misinformation through these online mechanisms is potentially much wider.

The risks of misinformation and impersonation of election administrators have been augmented by the increasing use of messaging apps that are end-to-end encrypted, such as WhatsApp. While there are important social benefits for having messaging that is beyond the scrutiny of government, especially in authoritarian regimes, it also means that misleading or false election messages are hard to trace and correct in democracies. Such messages could include content that directs voters to the wrong polling station or gives them the wrong election date for the purposes of voter suppression. For example, in the 2018 presidential election in Brazil, WhatsApp played a crucial role in political advertising and the spread of political information, but also became a mechanism to spread false information and innuendo.³

Voters

No analysis of cyber security threats is complete without considering the impact on voters. The mass collection of information about voters by parties and campaign consultants lays the foundation for major risks to voter privacy. Voter privacy is particularly relevant with regard to political parties and social media platforms.

First, the privacy laws that apply to private and public sector actors do not apply to political parties, which creates huge potential for the misuse of sensitive personal information about voters.⁴ There is no compelling public policy rationale for why political parties should be exempt from robust privacy rules, as nearly every other significant public or private sector organization in Canadian society is subject to them. Voters should know that political parties are abiding by fair information principles, modified to account for other federal election laws, such as the mandatory disclosure of contributors. Fair information principles include accountability, consent and limits on the collection, use and disclosure of personal information. Currently, voters have no way of knowing whether the information that they have knowingly given to parties, or that the parties have collected from social media or private sources, is protected against third-party disclosures or has adequate safeguards, such as encryption. This information is at risk of cyber interference. In response to this problem, the House of Commons Standing Committee on Access to Information, Privacy and Ethics recommended in 2018 that political parties should be included as entities under the existing private sector privacy legislation (House of Commons Standing Committee on Access to Information, Privacy and Ethics 2018).

In our view, the privacy rules that apply to political parties should be tailored to the particular role and function of parties. For example, a “do not call list” that prevented political parties from contacting voters would be disastrous for democratic engagement and could undermine, rather than protect, democratic discourse. Democracy requires contact between parties and voters.

Second, voters’ privacy can also potentially be breached by social media platforms. Their business model is predicated on giving away services in exchange for personal data. The major platforms have been critiqued extensively for the manner in which they collect and disseminate data. This transmission of information about voters held by platforms to app developers, advertisers or other entities,
especially if used for electoral purposes, raises serious risks to voter privacy. In the most notorious example implicating voter data from social media sites, the Cambridge Analytica scandal involved alleged improper third-party uses of Facebook data by campaign consultants, although Facebook disputes the extent.

The Elections Modernization Act

The Elections Modernization Act of 2018 makes a host of changes to federal election law, including some important measures to improve cyber security.

First, social media platforms with a minimum number of users are required to keep a repository of all political advertisements run on their websites. This move offsets, to some extent, the influence of microtargeting. Microtargeted advertisements are only seen by the viewers to whom they are directed, and rules on disclosing the source are easier to evade online. There is, therefore, less public scrutiny of the content and the source than there would be with a traditional advertisement on television or radio. A mandatory repository of advertisements imposes transparency and facilitates public scrutiny of advertisements. This new legislative requirement will not prevent foreign-placed advertisements or domestic ones that otherwise breach campaign finance laws, but it increases oversight as the advertisements will be made available to the public, media and politicians to examine.

Second, the act also creates a host of offences that are directed at digital threats, including interfering with a computer. Social media platforms will not be permitted to take foreign advertisements communicated for the purpose of influencing an elector. The statute also creates new offences of impersonating a politician or Elections Canada. These offences are promising attempts to update the Elections Act to account for digital democracy and existing cyber threats. Yet, they collectively face some challenges, in particular around deterrence and enforcement. It is unlikely that new offences will deter foreign actors funded by a hostile government from hacking into the database of a political party or from placing misleading content on Facebook. Even if the wrongdoers can be identified, if they reside outside of Canada in hostile countries it is unlikely that they would ever be held accountable. It is also unclear whether the provision on impersonation will cover deep fakes.

Finally, the legislation will require political parties to have privacy policies that address specific issues, but does not go so far as to grant voters an enforceable right to their personal information and does not give them a cause of action to combat privacy infringements. This tepid approach to regulating political parties and privacy is a significant missed opportunity, not only for privacy but for cyber security as well. Laws imposing stringent privacy protections would have the salutary indirect effect of requiring parties to strengthen their cyber security protections and would limit the collection of the massive amounts of personal data that underwrite data-driven electoral threats.

Messaging apps that are end-to-end encrypted, such as WhatsApp, increase the risks of misinformation and impersonation of election administrators. (Photo: Tero Vesalainen / Shutterstock.com)
Conclusion

Elections around the globe have been subject to digital interference, both domestic and foreign. Canadian elections are at risk of cyber attacks, and the lead-up to the federal vote in October 2019 has seen government, political parties, election administrators and national security actors try to address the major threats. Election cyber security has come from the margins to the centre of the conversation about democracy and will no doubt continue to grow in importance after 2019. Although Canada has made improvements to election-related cyber security with the enactment of the Elections Modernization Act, much more work remains to preserve the integrity of Canada’s elections from digital threats.

Works Cited


Endnotes

1 Elections Modernization Act, SC 2018, c 31 (Royal Assent 12 December 2018).


3 See, for example, García Martínez (2018).

4 Political parties are not covered by the federal privacy statute pertaining to the private sector because it applies only to “commercial activities.” Personal Information Protection and Electronic Documents Act, SC 2000, c 3 (PIPEDA). Political parties are excluded from the federal privacy statute pertaining to the public sector because they are not within the definition of “government institutions.” Privacy Act, RSC 1985, c P-21. The Canada Elections Act does not significantly oversee the practices of political parties with regard to the collection, use, storage and analysis of data, although it does provide them with an entitlement to basic information about voters.

5 Elections Modernization Act, supra note 1 at s 208.1.

6 Ibid at s 323.

7 Ibid at s 282.4(5).

8 Ibid at s 323.

9 Ibid at s 254(1).
The state has always been engaged in security-related activities, but these have changed over time and especially, in an accelerating way, in the twenty-first century. Such activities include surveillance, understood here as any personal data acquisition and analysis for management, influence or entitlement. Today, state activities cannot be considered without noting the role of data flows between private corporations and government agencies, and of the part played by new technologies themselves that are often permitted a leading role, especially as artificial intelligence (AI) is promoted.

The upshot is that public trust is threatened as governments become preoccupied with issues that do not strike citizens as being central to their own security, and as data breaches and undemocratic practices proliferate. New methods of data analytics demand new approaches to how data is framed, analyzed and used. A duty of care regarding these matters is vital and includes attention to the sources of data and their curation, the algorithms used for analysis and the uses that are permitted for those data. Both internal and external assessment and review should happen periodically and be overhauled as needed, for appropriate data governance to be achieved, under the larger goals of data justice, the common good and human flourishing.

Back Story
In the long history of surveillance, the state has always been the key player. Some notion of security has been a central rationale. Externally, surveillance relates to geopolitical and military purposes or commercial advantage. Internally, surveillance might be pursued for the pacification and administration of the population. This includes the collection and use of data for everything from electoral rolls to health care and welfare provision.

Since the mid-twentieth century, surveillance carried out by state agencies has expanded enormously, both for geopolitical reasons — such as the Cold War and, later, anti-terrorism activities — and because new technologies were developed to enable such expansion. The very technologies, invented and refined for military use, have become the backbone, not only of state surveillance but also of industrial enterprise and everyday commercial and personal activities.

The internet, invented as a Cold War communication network, became public and commercialized in the 1990s, when it became a source of information. Web 2.0 followed, in which not only interactions were facilitated, but where users could provide their own content. Soon it began to morph into the Internet of Things (IoT), which means that surveillance is embedded in all kinds of objects, from buildings and cars to vacuum cleaners and fridges. Thus, data is “skimmed off” from mundane practices including driving, shopping and sending messages (Jeffreys-Jones 2017).

Since the late twentieth century, with the rise of neoliberal policies, the relationship between state agencies and commercial corporations has become deeper and more complex. This is vital for understanding surveillance, not only because corporations supply the know-how and equipment for monitoring and tracking, but also because, today, the data desired for use in policing and intelligence, and in many other
tasks, originates in ordinary online exchanges, searches and interaction, as well as in phone calls. This means routine forms of data exchange, allowing for information to flow between public and private realms, along with many public-private partnerships that have been developing since the 1980s and 1990s, are now normalized and commonplace (Ball and Snider 2013).

Rapid Developments

In the early twenty-first century, the events of 9/11 (that is, September 11, 2001) represented a crucial shift. The rapid securitization of many aspects of government and everyday life in the name of anti-terrorism is now seen as normal. Much of this development depended on the intensified deployment of information technologies from companies that at the end of the twentieth century had feverishly been seeking new markets. Biometrics, for instance, which had been languishing as an idea without an application, suddenly appeared to offer vital and reliable support for identifying suspects (Lyon 2008).

The extent of this began to be clear early on, but the disclosures by Edward Snowden in June 2013 demonstrated beyond doubt that the global security-surveillance network was in high gear. Government agencies were making extensive use of personal telephone and internet data, and it was increasingly hard to distinguish between internal and external surveillance. Consumers and citizens were outraged to know that, somehow, government agencies had access to their personal data. The crucial category was metadata, the details of where and when communications or transactions occurred, between whom and so on. Trust was further eroded by official denials that the metadata involved was consequential, even though it comprised the same data that private detectives would seek; that is, of a very personal nature, just not necessarily of the older date-of-birth or street address type (Lyon 2015; Szoldra 2016).

As a site of user-generated content, the internet was hugely augmented by the growth of social media and then of platforms in general after the success of Facebook, beginning in 2004, and the so-called sharing economy of Airbnb, Uber and the like, a few years later. During these years, first in the corporate sector and then in government, ways to harvest, analyze and monetize this so-called data exhaust on a massive scale were found. The take-up rate accelerated as new data analytics — big data — was developed to utilize this trove.

The apparent possibilities for reviving older dreams of the online world as a harbinger of a new phase of democratic participation served to mask the regulation-resistant, competitive character of these mushrooming corporations. In fact, while ordinary, everyday lives became increasingly transparent to these data-greedy behemoths, their own activities became less and less transparent, something that our research team examined recently, directly relating to Canada (Bennett et al. 2014). Thus, the very bedrock of democratic involvement — trust, based on an informed citizenry — was being eroded from within.

By 2013, the Snowden disclosures indicated how the shift toward big data practices was happening with the National Security Agency (NSA) in the United States, but the shift was occurring simultaneously in Canada and elsewhere in the so-called Five Eyes countries. In Canada, the Communications Security Establishment (CSE) adopted a new analytic method from about 2012, which was described in "scientific revolution" terms. The switch was made from suspicion-led to data-driven approaches, heavily dependent on computing power and algorithmic analytics. Communications were to be monitored and analyzed to discover patterns producing actionable intelligence (Thompson and Lyon, forthcoming 2019).

Although agencies such as the NSA and CSE develop their own methods, they frequently work in tandem with commercial providers and university research hubs to create new surveillance tools within a network of agencies that is far more than the sum of its parts. In Canada, the Tutte Institute for Mathematics and Computing, working with the CSE, is a case in point. Reciprocal relationships are deepened within networks, both for providing expertise and software and for executing surveillance tasks. At the CSE, an education phase gave way to an exploration phase, and from there to engagement, which leans more heavily on machine learning. They work together with private sector commercial organizations as well as universities, developing algorithms for knowledge discovery and data mining.
Big Data Surveillance

Clearly, the phrase “state and surveillance” does not do justice to recent developments in security-surveillance networks following 9/11 and the rise of platforms that generate burgeoning data resources. New relationships mean that once-distinct public and private entities now shade into each other. Government works closely with businesses and research groups, and there is also a sense in which the technological systems themselves participate, especially as AI and machine learning become more significant. Such developments challenge conventional modes of scientific and technological practice, and, of course, the time-honoured approaches to policing and security.

Today, huge amounts of data are sucked into systems that store, combine and analyze them, to create patterns and reveal trends that can be used for security, alongside other uses such as health, marketing, governance and many other areas.1

This is a worldwide trend, seen in global IT companies — now often referred to as surveillance capitalism (Zuboff 2019) — and also in the programs, activities and public documents of the CSE, the Canadian Security Intelligence Service and the Royal Canadian Mounted Police (RCMP). In terms of method, this major shift from causation to correlation raises many questions, for instance about privacy and data protection regimes that sometimes seem to be sidelined.

Some key features include the use of open source intelligence and social media, the geographies of security-surveillance (cables, clouds and data centres), and the implications for international relations of physical communication conduits that are accessible to intelligence agencies (Clement 2018). It also raises questions about how international professional security groups, in public-private partnerships, influence policy and profits. Each of the key features mentioned above relates to further issues, such as how big data practices exploit loopholes in current privacy laws, how security is mobilized as a permanent rationale for increased surveillance and how new channels of power and influence disproportionately disadvantage certain population groups (Dwork and Mulligan 2013; Raley 2013). The latter is clear in “predictive policing,” a parallel field to national security in which scrutiny of those already under suspicion is intensified, and the influence of race, class and neighbourhood are magnified through big data practices (Brayne 2017). This, too, has deleterious effects on public trust.

Underlying and infusing all these, however, is the question of what sorts of knowledge are sought. Time-honoured practices and patterns of research and investigation, in which causes and explanations are sought, give way to inductive analytical methods. Data, collected from disparate sources, are put together in new configurations in order to infer patterns. The result often leans toward correlations.

Since September 11, 2001, the rapid securitization of many aspects of government and everyday life have come to be seen as normal. The intensified deployment of information technologies came primarily from companies that, at the end of the twentieth century, were seeking new markets and applications for ideas such as biometrics.

(Photo: Hayk Shalunts / Shutterstock.com)
that have a much more uneven history in the quest for reliable knowledge. How far can such new methods be trusted, especially when they carry such heavy freight of responsibility for people’s choices, life chances and even human life itself?

Confronting New Questions

The large question to be addressed has to do with data governance. This is closely connected with questions of trust and, thus, also ethics, in both relations with the state and with corporations, in all their early twenty-first-century complexity. Trust has been deeply damaged in both corporate and governmental domains, due to data breaches, surveillance overreach, unfair outcomes in policing and security, and disturbingly protective secrecy. Data governance should not be seen in only a technical or legal sense; data justice in data governance would align this with human flourishing and the common good.

Canada and similarly aligned countries cannot expect to advance their strategic and economic interests, let alone foster human flourishing, without rebuilding trust. This, in turn, relates to the focus of security concerns. If Canadian citizens suspect that the actual focus of security seems to refer to governmental, economic or technological activities and systems alone, then trust is once again threatened. However, if those interests are seen to be under an umbrella of human security (Zedner 2009), where personal, communal and environmental protection are the focus rather than states or national security, this will help to recover trust. These considerations underpin the specific comments that follow.

Turning to specific questions of the digital, and to data in particular, how these are handled is of utmost significance. As the methods of addressing security challenges are shifting fundamentally, so the questions for regulating and overseeing security-surveillance must also change. What was once thought of primarily as a question of data collection is now primarily one of analysis and use of data (Broeders et al. 2017). Along with this is a discernible shift toward data governance in terms of broad ethical frameworks, rather than of privacy alone (Bennett and Raab 2018).

As far as analysis is concerned, duties of care are required both in data collection and curation, and in the use of algorithms that are central to any analysis. Both internal audits and external reviews should be guided by

Data breaches, surveillance overreach, unfair outcomes in policing and security, and disturbingly protective secrecy have damaged citizens’ trust in both corporate and governmental domains. Safeguards against erroneous and malicious use of data, and, importantly, transparency about government use of private surveillance companies to, for example, monitor dissent at pipeline sites, are required. (Photo: arindambanerjee / Shutterstock.com)
the duty of care. If analysis involves profiling and/or automated decision making, or even decision support, then tight regulation is called for. Democratically organized oversight functions are vitally needed at each level.

In Canada, these matters resurface periodically in relation to the regulation of our own security and policing services. Bill C-51 (the Anti-terrorism Act, 2015), for instance, was very controversial for several years due to its permitting certain kinds of access to data without adequate accountability or oversight and its scant regard for civil liberties. Bill C-59 (an Act respecting national security matters) addresses some of these concerns in a manner that is at least somewhat more satisfactory (Forcese 2018), but constant vigilance is required if trust is to be rebuilt to serve the common good and human flourishing. There is an unfortunate history of overreach and obsessive secrecy within the departments charged with security matters, and these do nothing to enhance trust. New modes of transparency and public responsibility are needed throughout.

Conclusion

The “state and surveillance” is a far more complex equation than it may at first appear. Developments in political economy — neoliberal public–private partnerships, for example — and in new data-enabled practices of analytics, machine learning and AI all complicate relationships (Pasquale 2016). This makes it hard to know what exactly transpires within the agencies — security and policing — that are early adopters of new technological and analytic styles of operation. While genuine benefits may well emerge from CSE’s new analytic method or from predictive policing, current trends indicate there is a significant trust deficit and a sense of unfairness, in both procedures and outcomes.

The kinds of operation inspired by corporate practices, such as rating and ranking in credit schemes, by technological activities treated as if data were raw or algorithms were neutral, and that rely on inductive methods that produce correlations rather than explanations, demand radical rethinking. Practices that intensify categorical suspicion, for example, are patently unfair. Thus, requirements for data justice (Hintz, Dencik and Wahl-Jorgensen 2019), as well as for greater transparency, accountability and oversight, need to be part of programs to ensure appropriate data governance. This, too, is only a means to other, societally more significant aims — those of seeking to deepen trust and thus human flourishing and the common good.

Author’s Note

Thanks to Colin Bennett, University of Victoria, who kindly commented on a draft of this article.

Works Cited


Endnote

1 “Big Data Surveillance” is the theme of the current multidisciplinary and international project at the Surveillance Studies Centre at Queen’s University: www.sscqueens.org/projects/big-data-surveillance.
Trust and Data

How Changes to the Privacy Landscape Can Bolster Innovation in Canada

Paul Vallée

Imagine that a team of researchers comes up with a brilliant new idea for predicting the likelihood that a patient will die from an opioid overdose. These researchers identify a set of indicators that, when combined, make it more likely that someone’s drug use will turn deadly — factors such as their health records, whether they have previously been admitted to the emergency room for an overdose, or whether they associate with known drug dealers. By working with data scientists, the researchers develop an algorithm that can identify the highest-risk patients for treatment and intervention. This technology could potentially save hundreds of lives, and the information needed to feed the algorithm already exists in various databases — such as medical files, hospital records, police records or even social media accounts. But this information is also highly sensitive, and while combining it in innovative ways could lead to a public health breakthrough, privacy laws would impede the researchers from cross-referencing these separate databases.

Is there a way for these researchers to securely access the various data sets in a way that serves the public interest while also protecting patients’ rights to privacy? How could they do this while maintaining the trust of the agencies and individuals involved? And what consequences would they face if they failed to uphold that trust?

This is a hypothetical scenario, but this kind of technology already exists, so these questions have real implications for Canada. Massive amounts of data about all aspects of our lives are being collected and stored, and they have the potential both to benefit society and to create innovative new businesses that contribute to our economy. But in order to make the most of this potential, Canada must build on its existing policy and legal framework to make it easier for organizations to prove they can be trusted to keep this sensitive information private and secure, and for the public to evaluate their trustworthiness and hold them to account.

The Global Data Race

Data is a lot like capital — it flows in and out of a country. If we cannot find a way to attract data investments, Canada risks becoming a client state with regard to data. But if Canada becomes the best environment to maximize the economic and social power of data, more data will flow in. To do that, it needs to prove it is better than its global rivals, such as the United States, the European Union and China, in three areas: trust, investment opportunities and data integration.

Trust is the most important factor because it underpins the other two. When a country establishes a climate of trust in its data environment, individuals both at home and abroad have faith that their personal...
information will not be compromised by security breaches or unscrupulous data practices. China is widely seen as having low trust due to its widespread digital surveillance practices. With the United States’ history of avoiding comprehensive regulation, the US approach has traditionally been trusted by its citizens, who are wary of government overreach, but that has started to change with the revelations about social media company practices, such as the Cambridge Analytica scandal. The European Union, by contrast, has its own history behind its comprehensive privacy framework. This framework was recently strengthened by the General Data Protection Regulation (GDPR), which creates even more stringent standards for how companies doing business in the European Union manage users’ data and gain their consent to do so.

The GDPR also introduces massive penalties for companies that fail to comply with its standards, and that has the potential to harm the second area of competition: investment opportunities. To attract investment, a country must make businesses feel confident that the investments they make there will not bankrupt them. The threat of losing up to four percent of annual global turnover in the event of a GDPR violation may create a chill among companies thinking of setting up data-related businesses in the European Union, or simply doing business there (as the GDPR has extraterritorial application). There are also serious concerns that well-established businesses — corporate giants with significant legal and human resources — will be in a much better position to navigate the new system, to the detriment of small start-ups that do not have the same resources. This is where the US system, with its lighter regulatory environment, may look like a better option for companies wanting to set up data-related businesses. China, on the other hand, is considered to have a strong environment for domestic investment, but less so for international investment.

The final aspect of building an attractive data environment is data integration, or the ability to draw from existing data sources, possibly in combination, to create innovative opportunities. China has no shortage of options for achieving this, but it comes at the cost of genuine openness and transparency. By contrast, privacy regulations in the European Union and Canada, for example, are built on principles of openness, transparency and accountability with respect to the personal information companies collect. These principles require consent — that is, individuals must be notified at the time of collection about what personal information is collected about them, how it will be used and the legitimate grounds that the company has for that collection. Under EU laws, using data that individuals have already disclosed, but for purposes other than those for which it was collected, is not permitted, unless the company can demonstrate that the new purpose is compatible with the original purpose.

Canada must build on its existing policy and legal framework to make it easier for organizations to prove they can be trusted to keep this sensitive information private and secure.
Linking personally identifiable information to other sensitive information about users — for example, their health data, locations they have visited or even their consumer purchasing history — poses additional privacy risks. And the possibility that such sensitive information could fall into the wrong hands as a result of a data breach or cyberattack would surely be cause for concern.

Improving Canada’s standing in all three areas is achievable, but will require additional policies that create a more secure, predictable environment for data management. And at the heart of that is changing how we think about and measure trust.

Measures of Trust

Some people may think that the purpose of regulation is to prohibit certain dangerous behaviours in order to protect the public interest — for instance, nuclear operators are not allowed to dump nuclear waste into the water supply or use the technology at their disposal to build weapons. But another way of looking at regulation is as a mechanism to permit certain dangerous behaviours that have a public benefit under certain conditions. Nuclear power can be dangerous, but if it is managed correctly, it can also benefit society as an energy source.

So, rather than outlawing nuclear power, the government imposes certain standards that nuclear facilities must meet in order to operate.

Data activities are essentially the same. Yes, there are inherent risks to collecting and managing individuals’ personal information, including risks of privacy overreach or having the data exposed in a security breach. But there are also potential benefits to combining various data sources as discussed above in order to develop new and innovative uses of this data in the public interest. A change in approach would put more emphasis on the ways that individuals or entities could be verified as trustworthy to carry out these kinds of activities. There are several policy options we could enact individually or in combination to achieve this.

Standards and Certification

Activities that can benefit the public, but also potentially cause harm — such as operating a nuclear plant, practising medicine or even driving a motor vehicle — are generally accompanied by clear standards that set out what requirements a person or organization must meet to carry out these activities. You can tell by looking at someone’s driver’s licence whether they have been deemed capable of safely operating a tractor-trailer or if they are a new driver who must be accompanied by an adult when they get behind the wheel. And, if they fail to meet those standards — or any other rules of the road — they risk losing their licence.

There are always risks to sharing your personal information, but a system of
certification could signal that a person or organization has met the standards of care necessary to be a trustworthy custodian of your data. Canadian law, such as the Privacy Act and the Personal Information Protection and Electronic Documents Act (PIPEDA), already mandates certain privacy standards for private sector organizations. But introducing a tiered rating system, similar to driver’s licences, could be an efficient way of letting customers know what measures a company would use to manage and protect their data in accordance with PIPEDA, without having to read a complicated terms of service agreement. For example, to receive a top rating of A, a company would have to guarantee an audit or supervision capability for the data, meaning it would have a record of exactly which employees have viewed the data and for what purpose. For a lower rating of B or C, the data would be protected by password authorization, but not have an audit trail. If a company chose not to complete the certification process, its lack of a rating in itself would send a strong signal to customers. By making clear exactly how a company would manage and protect personal data, customers would be more likely to trust them with their data.

Introducing new certification systems that are compliant with the GDPR, PIPEDA or other privacy regulations could also pave the way for different organizations to share their data for innovative new programs or services, such as the opioid example from the introduction. Organizations that meet a certain standard of trust could be permitted to undertake data integration activities using multiple data sets, and consent could be addressed and incorporated to make this work. And because of the challenges and risks associated with these kinds of initiatives, there would still need to be some kind of framework that regulates which data integration programs would be permitted.

If it works, the Canadian certification system could become a global benchmark for data management. Companies that secure a Canadian A rating could use it to market themselves worldwide as secure and responsible guardians of data.

Trust Standards for Individuals

It is not always natural to put your trust in a company — in real life, it is often easier to put your trust in individual people, not entities. For example, within the Government of Canada, different employees have different levels of security clearance that permit them to access different kinds of information. Just because the Government of Canada has access to your data does not mean that everybody who works for the government has access to it. You might not trust the groundskeeper at a national park with your personal financial information, for instance, as much as you would trust an accountant with the Canada Revenue Agency.

A set of trust standards centred on individuals could work like the Nexus program for trusted travellers, allowing someone who has been screened and passed tests for reliability to have greater access to certain kinds of data. This could be another way of permitting the integration of data from multiple sources, if the person handling the data has been verified as someone who can be trusted not to misuse it. Individuals could also be approved to access multiple spheres of trust, so that if two or more organizations wanted to collaborate on a data project, rather than making the relevant data available to both groups, they could instead share it with a select number of employees who have been deemed trustworthy by both organizations involved.

Other Innovations

Companies and organizations are getting better about asking users for consent before they collect or store their personal information, but a side effect of this is that many people click on consent forms without giving them any thought. Does anyone really remember which information they have agreed to share with which organizations? If there were a data breach affecting a digital service you used, would you know what information about yourself was vulnerable?

While some people are diligent and thorough when reading various consent messaging, research suggests the vast majority click through without reading much or any of it. In other cases, users may have given their consent long ago to share something with a company that seemed insignificant at the time, but their concerns have changed over the years, or the company later shared it with another party without their knowledge. The GDPR is strict about requiring companies to include a privacy notice specifying the third parties with whom they share personal information.
A national data consent registry that included these third parties could be a one-stop reference for people to keep track of all the permissions they have agreed to. When users click on the “I agree” button for a company or organization, the company would then have to record what information the users have agreed to share on a searchable registry. This could easily be implemented on a blockchain to ensure that service users have a record of all the times they have consented to share their data. By tracking which types of information different organizations collect, a consent registry would also allow anyone who wished to build data integration programs to understand which organizations they would need to approach to obtain data sets.

Another innovative policy option could be to set up data trusts for certain kinds of information that could be put to a socially beneficial use. Data trusts, like fiduciary trusts, would maintain data sets and manage the conditions under which the data could be used and shared. The idea is to make it easier for new start-ups, or anyone who has an idea for using data in the public interest, to access the data they need to make their ideas work, as long as they meet the necessary standards for safeguarding the data.

**Conclusion**

Canada has the opportunity not just to compete with other global markets on data innovation, but to strengthen its leadership role in setting standards for data privacy and security. The benefits could be huge in terms of drawing investment to Canada, building a strong ecosystem for homegrown data companies and developing innovative new ways to use data for the public good. But this can only happen with trust: trust from Canadians that organizations will not misuse or expose their personal information, and trust from companies that want to use data in innovative ways that the investments they make in privacy and security will pay off. There are several ways Canada can foster this environment of trust, but fresh thinking and enhanced outreach will be required to get buy-in from the public and the corporate world.
Beware Fake News

How Influence Operations Challenge Liberal Democratic Governments

Eric Jardine

The 2016 US presidential election was a tumultuous time. In the weeks and months leading up to that Tuesday, November 8, social media sites, such as Twitter and Facebook, were flooded with “fake news” (Howard et al. 2017). Investigations following the election of Donald Trump as the forty-fifth president of the United States revealed that extensive foreign influence had played a role during the campaign, its efforts aimed largely at affecting the course of the election. Most fingers pointed directly to the Russian Federation and the regime of President Vladimir Putin as the most likely culprits (National Intelligence Council 2017).

This was not by any means the first use of social media in influence operations. A few years earlier, for example, the Islamic State terrorist organization (ISIS) used extensive Twitter campaigns to spread propaganda, encourage radicalization and recruit foreign soldiers for its war in Iraq and Syria (Klausen 2015).

Influence operations, whether launched by governments or non-state actors, existed long before social media, but what is new about contemporary influence operations is their scale, severity and impact, all of which are likely to grow more pronounced as digital platforms extend their reach via the internet and become ever more central to our social, economic and political lives. Such efforts represent a clear cyber security challenge. Yet, democracies, which depend on the open and free sharing of information, are particularly susceptible to the poison of influence operations that spread fake news, disinformation and propaganda. The whole edifice of democratic governance is based

Eric Jardine is an assistant professor of political science at Virginia Tech and a fellow at CIGI. Eric’s research focuses on the uses and abuses of the dark Web, measuring trends in cyber security, how people adapt to changing risk perceptions when using new security technologies, and the inherent politics surrounding both anonymity-granting technologies and encryption. He is co-author of the book Look Who’s Watching: Surveillance, Treachery and Trust Online (CIGI Press 2017).
on the assumption of an informed citizenry
with a common sense of facts, shared public
narratives and a solid trust in the information
provided by institutions. This entire assemblage
is threatened by carefully crafted influence
operations and will only grow worse as new
“deep fake” technologies come into play.

The Scope of the Problem

By one false account, in 2016, Democratic
Party nominee Hillary Clinton and her chief
of staff, John Podesta, were operating a child
sex ring out of a pizza parlour’s basement in
Washington, DC. What started as a malicious
internet rumour quickly morphed into a social
media trend. The hashtag #pizzagate went viral
as thousands of accounts tweeted “evidence”
both for and against the story. Many of these
tweets originated outside of the United
States, with disproportionately large clusters
coming from the Czech Republic, Cyprus
and Vietnam. Shortly after the election, this
fictitious online tale made a sinister cross-over
into the physical world, as one of the story’s
followers, Edgar Welch, drove to Washington
with an assault rifle. He entered the pizzeria,
demanding to see the basement (the building
does not have one) and fired off three shots.
What began as online disinformation
had taken a terrible turn (Fisher, Cox and
Hermann 2016).

The pizzagate story is just one illustration
of an increasingly prevalent problem of online
influence operations by foreign governments
and non-state actors. While a healthy
ecosystem involves the free flow of information
and interpretation of facts, large swaths of
online influence operations to date, particularly
as they are directed toward the West, can
be colloquially called “fake news,” meaning
content that is “intentionally and verifiably
false, and [that] could mislead readers”
(Allcott and Gentzkow 2017, 213). Beyond
subverting the facts, fake news also plays
another role. It is crafted to resonate with its
readers. Such resonance does not arise purely
through information. Resonance can also be
based on sentiment or a reader’s sense of its
truth, creating what could be called a folkloric
element (Frank 2015).

If fake news was only about spreading
incorrect information, then those who believed
such stories would have to be either ignorant
or undiscerning about news in general, or
willfully ingesting false content. Viewing fake
news as a genre of folklore, as Russell Frank
has proposed (ibid.), raises a third possibility,
that fake news is appealing because it delivers
a moral narrative or confirms sentiments that
people already hold. From this perspective,
the ISIS social media propaganda about the
corruption of the West (Klausen 2015) or the
fake news stories about the health of Hillary
Clinton during the 2016 election (Milligan
2016) share a common foundation: they
propagate “alternative” information and
present a moral narrative that people holding similar
views can latch on to.

Influence operations using messages
combining these informational and political-
parable-like qualities can be launched by state
actors, non-state actors or some combination
of both. Efforts at influencing information
environments have a long history, but today,
the potential scale of influence operations is
decisively affected by new digital platforms
with vast numbers of users. Facebook alone
has roughly 2.25 billion users. Twitter has
336 million. Mobile messaging applications
that allow users to share threads and stories
likewise capture huge proportions of the
internet-using population, with 100 million Telegram users, 1.5 billion WhatsApp users and 1.0 billion Viber users, not to mention the numerous smaller messaging applications that exist online.

The scaling effect of social media gives a simple boost to terrorist organizations that seek to radicalize individuals or recruit foreign fighters. For example, ISIS ran a highly advanced online influence operation. On Twitter, this process spanned geography, with carefully selected fighters in Syria and Iraq tweeting photos that were then vetted and shared by third parties and individuals linked to ISIS but living in the West (Klausen 2015). Through this simple gatekeeper methodology, ISIS was able to put forward a coordinated influence campaign, designed to showcase a skewed image of the glories of war and life under ISIS.

On other digital platforms, such as YouTube, ISIS used the huge user population (1.8 billion users) and hours of consumed videos (up to one billion hours daily) to spread propaganda videos to glorify its terrorist agenda (Gillespie 2018). As Tarleton Gillespie put it, “ISIS has proven particularly skilled at using social media in this way, circulating glossy recruitment magazines and videos documenting the beheading of political prisoners and journalists” (ibid., 55). The goal was to reach those individuals who might be swayed by ISIS’s messages and encouraged to undertake homegrown operations or become a foreign fighter.

The increased scale of information operations also plays out via new socio-technical algorithmic assemblages. Algorithmic bots, specially designed programs that use computer processing power to spread content via fake user accounts, have helped to generate and pollute the online information ecosystem. Such bots are particularly active during political events. The 2016 US election swung partially due to changed, and somewhat unexpected, shifts in voter preferences in Michigan. Within this key battleground state, as research from the Computational Propaganda program at Oxford indicates, non-professional news (fake news) was shared more frequently via social media than professional, mainstream news (Howard et al. 2017). More troubling still, news produced by reputable media outlets (The New York Times, for example) hit its lowest point as a proportion of content the day before the election (ibid.). These trends were exacerbated by bots’ activity.

The growing sophistication of artificial intelligence (AI) and machine-learning algorithms also points to a potential new qualitative change in influence operations. Generally, people tend to trust the written word somewhat less than they do audio and, in particular, video media. A news story might say that Hillary Clinton is ill, but the story would appear more believable if Clinton were to say so herself — or at least if she were to seem to say so. AI can now be leveraged to generate so-called “deep fake” videos, which actually involve faked video of a person saying fake news (Giles 2019). Deep fakes are hard to spot and will greatly increase the qualitative impact of fake news and foreign influence operations.

With enhanced scale, increasing automation and the capacity for pernicious deep fakes, influence operations by foreign governments and non-state actors have gained a new edge. Operations that would have been manageable in a predigital age are now a very real challenge to liberal democratic regimes.

The Challenge

Democracy is fundamentally based on trust — trust of each other, trust in institutions and trust in the credibility of information. Influence operations, in particular those run by foreign governments or malicious non-state actors, can pollute an information environment, eroding trust and muddying the waters of public debate.

The discourse surrounding the 2016 US presidential election is a case in point. Debate during the campaign was marked by a high level of rancour. Since the election, survey respondents have indicated that they feel civility and trust in major institutions within the United States have declined as the opposing ideological camps have hardened their positions. For example, one survey found that fewer than 30 percent of people trusted media institutions and, more broadly, fully 70 percent of respondents thought that there was less civility (Santhanam 2017).

Fake news and other influence operations are made more powerful by “filter bubbles” (Pariser 2012). The term describes the result of the
algorithmic machinations that lead people into relatively contained online information ecosystems of their own making. Once within such a bubble, people tend to get more of what they like, based on their earlier online choices, whether those are funny YouTube videos of cats or ideologically infused podcasts and posts. The troubling part is that the commercial aim of platform filters — namely, to give people what they want to encourage consumption of content — tends to play out badly in the political space. They lead people to hear their own message rather than others’ points of view, in an echo chamber reinforced by algorithms. While democracy requires the free exchange of information and ideas, filter bubbles tend to isolate users. In a filtered environment, information does not circulate widely and freely.

**Solutions and Ways Forward**

Malicious influence operations are a growing problem, exacerbated by social media platforms that enable the scale-up of misinformation, disinformation, propaganda and information disruption operations and new algorithmic technologies that might potentially cause us to even distrust our own eyes.

Modest, but meaningful, changes are possible and necessary. Broadly, countering the problem means addressing three aspects: exposure, receptivity and counter narrative.

Exposure is at the core of the problem of fake news and other forms of influence operations. A person might be psychologically ripe for radicalization, but, without exposure to ISIS’s message, may never tip over the edge. Likewise, an electorate’s exposure to fake news during an election cycle may affect political discourse and even electoral outcomes. Simply put, reducing exposure to influence operations reduces their effects.

In liberal democracies, where freedom of expression is enshrined as a fundamental right, governments often cannot directly censor the information being shared online. Furthermore, the primary infrastructure for disseminating information during an influence operation (such as social media platforms) is owned and operated by private companies. So, although governments are limited in their ability to constrain exposure, the companies that own the platforms are not. Facebook, Twitter and YouTube (run by Alphabet) can all directly control what sort of information flows across their networks.

While platforms historically avoided explicit content moderation, and to some extent still do, arguing that they are not publishers, consumers have begun to express a desire for some moderation of more extreme and polarizing content, such as white supremacist content or fake news stories. The platforms are able to oblige (Gillespie 2018). These systems can moderate, and so control, exposure to information through two complementary
methods. First, platforms now leverage their vast user bases, encouraging users to flag and report content that is potentially objectionable. The platforms then evaluate the flagged content. If it is found to be in violation of a platform’s terms of service or community guidelines, it can be removed and the account that posted it can be banned (ibid.). Besides these human-driven methods, many firms are using automated detection systems to flag and pull down content. With more data, these approaches will improve further still. Through both measures, the platforms are working to limit the worst effects of malicious influence operations by reducing exposure to such content as ISIS beheading videos, “conspiracy videos” and hate-infused tweets.

Another method for countering influence operations is to build up people’s online “immunity” so that they have less receptivity to misleading, false and polarizing information. Broad-based educational initiatives that aim to increase user awareness of fake content might be helpful, if hugely costly. Inoculating key points (people) within a network is likely more effective and cheaper (Christakis and Fowler 2011). Targeted engagement with individuals at the centre of networks (high network centrality scores, in social network analysis terms) could help promote immunity of the herd and reduce receptivity to fake content (Halloran et al. 2002).

Finally, governments and traditional media institutions can work to create their own narratives of events that can counter the influence operations of others. The effectiveness of such counter narratives is conditional upon the trust that users place in their sources, so initiating these efforts swiftly to stem the tide of disruptive influence operations aimed at diminishing user trust is key. Their effectiveness is also likely a function of how well traditional producers adapt to changing media. The current social networking ecosystem is driven by clickable content. Sending out boring titles into this sort of maelstrom will likely fall flat.

If done right, meeting the messages of foreign influence operations with a counter narrative can have a positive effect on the perceptions of internet users. The public’s willingness to believe climate change denial stories, for example, is reduced if exposure to that disinformation is quickly paired with countering narratives that highlight the flaws in anti-climate change science and point to the climate change consensus that exists within the scientific community (Cook, Lewandowsky and Ecker 2017). In short, while refuting disinformation is an ongoing struggle rather than a quick win, governments — helped by platforms — can counter one influence operation with another. Doing so can help preserve trust, while also retaining the free flow of information that is at the core of liberal democratic governance.

Conclusion
Influence operations targeting liberal democratic regimes are deeply troubling. They disrupt the twin bedrocks of effective democratic governance: the free flow of information and trust. These campaigns can be undertaken by malicious foreign governments who aim to sow chaos, or by non-state actors, such as ISIS, who seek to radicalize disaffected individuals in the West. Countering these operations is both necessary and possible. Such efforts require the engagement of not only governments but also the platforms. Working together, these actors can preserve liberal democratic governance by minimizing exposure to fake news and other influence operations, promoting user immunity and promulgating counter narratives to misinformation.

Works Cited


Halloran, M. Elizabeth, Ira M. Longini, Danielle S. Servan-Schreiber, Steven E.TextColor.secondary_language,TextColor.complementary_color]
The Need for a National Digital Identity Infrastructure

Andre Boysen
Cyber security for health-care data has never been more important nor more vexing than it is today. Across the Group of Twenty countries, health-care spending consumes upward of 50 percent of government revenue, and its share continues to climb. In many places around the world, online access to health-care services is being held back, due to the highly sensitive nature of the data and our collective inability to provide viable protection for online service delivery. As a result, many things that could be done online with greater efficiency, such as seeing our health-care records or getting a new bank account, are instead delivered in person, at a much higher cost. Only the most basic and low-risk services are online today, and even those are beset by the huge overhead cost of data breaches and password resets. This is a global issue that plays out in communities everywhere.

Consider some of the other dynamics at play that contribute to the challenge. Some patients access health care every day; others access services every few years. Some patients are very internet savvy, while others don’t want anything to do with online services that require a perpetual mindset of vigilance and active suspicion, as well as evergreen technical acumen.

The topology of health care is one of the most diffuse sectors of the economy, with no organizing force between government, hospitals, doctors, labs, researchers, patients, medical device makers and health-care foundations and registries (such as Canadian Blood Services and the Canadian Cancer Society). In fact, considering the problem as one of topology rather than one of security might provide some good insight as to the path forward.

It is clear that pushing more and more “point solution” security controls out in response to breaches has not solved the problem — it is chasing the symptoms rather than addressing the issue. In fact, this continuous change in access and control mechanisms has increased the attack surface. 2FA, or two-factor authentication, which uses a device to generate one-time codes, for example, has evolved to real-time intercepts of passcodes by criminals. So, too, have spoofing templates evolved to overcome on-device biometrics.

The current strategies have failed. Pulling back from online service delivery is not possible either — economics and patient safety require innovation. A different approach is needed — one where the security model is strong but hidden from end-users. Such a model could provide simplified access for patients while better mitigating the cyber threats. Hiding the security model improves the patient experience while maintaining proper controls for access.

Are Digital Health Cards the Answer?

Every Canadian has a health-care card, which enables them to access medical care when needed. Importantly, this same card can be used to convey their sharing wishes as regards organ and tissue donation; by registering as an organ and tissue donor, citizens also have the opportunity to save another person’s life.

However, as sophisticated as the health-care system is today, there is a clear opportunity being missed. One of the issues currently plaguing the economy is the privacy and security surrounding digital identity. While digital identity and health care may seem unconnected, digital identity has the potential to change the health-care landscape and the way medical data is shared in Canada.

Canadians are able to give the gift of life through organ donation by registering their consent, but they have no way to share their health-care data — a massive resource that goes untapped, held captive because of its highly personal and sensitive nature. With the right tools and controls, giving patients control of their data will allow them to share it with researchers. We can also transform health-care delivery at the same time.

If a health card can indicate an individual’s willingness to donate their organs, why can’t it also allow individuals to access and donate health-care data? With adequate privacy and security measures in place, sharing that data is something that can be done every day. Data is a gift that keeps on giving.

Data Is the Fuel That Drives Health-care Innovation

Data is the fuel that drives health-care innovation through medical trials — the source for new drugs, devices and therapies and vital to improving health-care outcomes in
Canada. Connecting patients to doctors, and researchers to drug companies, is complex and expensive. Throughout the process, it’s critical to manage consent, ensure privacy and protect access to patient health-care data.

Various studies allow for health-care data to be captured and gathered. From there, the data cannot be used without explicit knowledge and consent from the patient, yet no digital infrastructure is in place to ask individuals to consent to the use of their data. As a result, medical trials toil on, gathering data one study at a time.

According to Toronto University Health Network’s Dr. Joe Cafazzo, setting up one medical trial for a new drug with 200 to 300 patients costs over $1 million, and requires enrolling patients, doctors, nurses and researchers into an online portal to gather and share data over the life of the study. To secure the data, administrators distribute paper forms, gather signatures, confirm participant consent and issue passwords for everyone involved. All of the information-sharing infrastructure is set up, then taken down for each trial to manage protocols around data sensitivity, privacy and consent. While the protocols are necessary, there is no process in place for patients to opt in to participate in subsequent studies or to make data available only as a control sample.

This boils down to one issue that is impacting digital identity today: passwords. Regardless of whether passwords are long, changed several times per minute or composed of random characters, they are not secure enough to keep health-care data private. The data safety issue is preventing the health-care industry from delivering crucial services online — despite provinces spending 40 to 50 percent of budgets on health care today. If health-care data is leaked, there is no way to “refund” privacy, and the consequences may be that an individual is uninsurable or unemployable.

The duty of care with health-care data is high, but finding a solution to enable data sharing needs to be found.

### Health-care Cost Implications of Password Misuse

Searching for a solution for password issues in health care, Dr. Aviv Gladman — chief medical information officer and emergency physician at Mackenzie Health and a trained electrical engineer — conducted a study to analyze how much password friction was costing the health-care sector. Dr. Gladman concluded that three percent of all health-care spending was on inefficient authentication due to doctors and nurses mistyping passwords, resetting passwords or losing password fobs (Gladman 2015).

In 2017, the Canadian Institute for Health Information estimated that total health spending in Canada was expected to reach $242 billion (Canadian Institute for Health Information 2017). According to Dr. Gladman’s findings, that means that upward of $7.3 billion is being spent on password frustrations that are slowing down health-care delivery. This could be easily streamlined, resulting in additional savings and improved patient experience, if individuals could book appointments online or review lab results from a mobile device.

The wave of internet-connected devices that experience a similar password problem is getting worse. Many devices that Canadians are purchasing (stereos, cars, activity trackers, TVs, fridges and so on) are connected to the internet and run with connected apps with passwords. For example, some cars now come with an app that allows the owner to lock the doors, sound the horn or locate it from their phone. There have been cases where the vehicle location feature is still active and shows a past owner where the car is located even after it has been sold to someone else. Better controls are needed for health-care devices.

By contrast, Apple has a strong digital identity scheme for its devices. This gives consumers...
comfort in sharing, for example, the heartbeat and movement data the Apple Watch produces throughout the day. This data is shared with Apple, but there is no way for patients to add this data to their health-care record or share it with their doctor.

Digital identity is bigger than health care — it’s needed right across the economy. But there are crooks — they are out to cause harm, and they are good at it. In 2017 alone, there were 7.8 billion identity records stolen, according to a recent report (Risk Based Security 2018). Today, passwords are the only barrier to accessing sensitive systems and data; however, a good digital identity system will move us beyond this limitation so that consumers can do more online. Health care needs digital identity, and the rest of the economy does, too.

The World before the Electrical Grid

The state of digital identity today can be compared to that of electricity in 1869. Prior to the introduction of the standardized electrical grid in 1870, only the biggest factories had their own electrical generators, which were used to power light bulbs so that factories could run two shifts and increase productivity and output. After the electrical grid was introduced, there were massive efforts to convince businesses to join. There were two groups of businesses that said yes to joining the grid and two groups that said no.

Of those that said yes, the first group consisted of smaller businesses that did not yet have a generator due to their complexity and cost. They could join the grid at a reasonable cost, enabling them to run two shifts and compete with bigger players. The second group consisted of businesses that had a generator but disliked the distraction it represented from the core business of making products. These businesses used a generator because they needed light, but it was not core to the business, and joining the grid allowed them to focus on making products.

Among those that said no, the first group included businesses that were interested in the electrical grid but had recently invested in a new generator. They saw the appeal but took a wait-and-see approach while using the new generator they had already invested in. The second group of businesses believed their generator was core to the business and were worried about relying on a third party for a resource that was key to production.

What is interesting is that in the end, everyone joined the grid. From the standpoint of economics and simplicity, the offering was so compelling that businesses eventually found running their own generators every day to be too inefficient and taxing on the business. At that point, the number of use cases for electricity grew very quickly. Electricity was no longer solely about powering the light bulb — it expanded to many different uses that ultimately transformed the economy.

What Is the Parallel with Accessing Services Online?

Digital identity is a lot like the electrical grid. Every online service delivery organization on the internet is running its own digital identity generator. Facebook, Amazon, Netflix, Google, governments, schools, hospitals, financial institutions and telecommunications providers are all running their own fiefdoms of identity services. Today, the first digital identity grids are starting to emerge, meaning that service delivery organizations no longer have to run their own digital identity generators. Organizations can get out of managing the risky password services that they own and manage.

SecureKey, a leading identity and authentication provider, makes trusted access to online services easier and more private for Canadians, with better integrity and lower costs for business. SecureKey is in the process of developing the digital identity grid in Canada to solve the problems associated with today’s online service delivery organizations. The current system is too difficult for consumers to use, and the costs are unsustainable. Businesses, governments, educational institutions and health-care organizations around the world are regularly experiencing data breaches, because no single organization can afford the massive investments required to make digital identity safe, convenient and private. It takes a village to make digital identity work.

In 2012, Canadian financial institutions partnered and launched the first version of this digital identity grid — a service allowing Canadians to reach Government of Canada services online.
Good digital identity is something you can hold in your hands, simple to use and accepted everywhere, much like a credit card or mobile phone.

The Digital Identity Grid

What is good digital identity? Good digital identity is something you can hold in your hands, simple to use and accepted everywhere, much like a credit card or mobile phone. It is trustworthy and cost-effective for businesses and will provide Canadians with more choice, control and convenience. Through the emerging Canadian model, consumers will be able to combine their financial institution account with their mobile phone and government-issued ID to create a digital identity that is still physical (with the SIM card in their mobile device) and simple to use and can be used everywhere. Digital health cards will be added to this mix.

The new digital identity grid is launching in Canada in 2019. It will provide better business confidence for identity registration, use less data and lower costs. It will give consumers the control and convenience to manage their online life, and it will mean that possession of user data will no longer be enough to allow imposters to masquerade as someone that they are not. It will leverage blockchain technology, allowing for transparent, secure data tracking across devices. And it will support the global principles of privacy and security by design developed by Ann Cavoukian, former information and privacy commissioner for Ontario and current distinguished expert-in-residence at Ryerson University’s Privacy by Design Centre of Excellence, who is providing privacy expertise for organizations working in this area.

Finally, it will meet the criteria of Canada’s identity standards organization, the Digital ID and Authentication Council of Canada (DIACC). DIACC is composed of members from across Canada, from governments, financial institutions, telecommunications companies and more, alongside SecureKey, striving to make digital identity work for Canadians across the economy.

Creating the Digital Circle of Care for Patients

Modern medicine best practice holds that the health-care system empowers the patient by putting them in the centre of their own health-care story; each individual creates their own circle of care. We do not yet have the tools to allow patients to do this.

National digital identity infrastructure is what is required to solve the problem. Here in Canada, we are on the cusp of having a world-leading digital identity scheme. It is designed by Canadians for Canadians. And it is designed to work across the economy, so that businesses and consumers can conduct transactions online with trust, confidence and privacy.

It is not a technology problem (the technology exists); it is not a skills problem (we know how to do it); it is not a money problem (health-care costs would come down significantly). It is a problem of focusing national will.

Sharing health-care data is needed, achievable and worthwhile. Digital identity is required to make this happen. Consumers will be able to see and donate their data, allowing them to become health-care heroes every day.
One of the most dynamic and exciting developments in information and communications technology is the advent of the Internet of Things (IoT). Although networking technologies have become increasingly ubiquitous over the past two decades, until recently they have largely been restricted to connecting traditional end-user devices, such as mainframes, desktop and laptop computers, and, more recently, smartphones and tablets.

Recent years have witnessed the attachment of a much broader range of devices to the network. These have included vehicles, household appliances, medical devices, electric meters and controls, street lights, traffic controls, smart TVs and digital assistants such as Amazon Alexa and Google Home. Industry analysts estimate that there are currently more than eight billion such devices connected to the network and project that this number will expand to more than 25 billion by 2020. The increasing deployment of these devices has enabled new use cases for network technologies. Some experts project that the IoT may generate as much as US$13 trillion in revenue by 2025.

Unlike traditional cyber systems, which connect general-purpose computers, IoT systems often link together highly specialized devices designed for specific purposes with only a limited degree of programmability and customizability. In addition, IoT systems often store and process data in a distributed manner, in contrast to the highly centralized approach of consolidating storage and computing power in large data centres. In addition, IoT systems are sometimes called cyber-physical systems, because unlike purely cyber systems, they also include sensors that collect data from the physical world.

The distributed nature and the presence of physical sensors create both new opportunities and vulnerabilities from the standpoint of security and privacy. To date, however, Christopher S. Yoo is a CIGI senior fellow and the John H. Chestnut Professor of Law, Communication, and Computer Information Science and the founding director of the Center for Technology, Innovation and Competition at the University of Pennsylvania. He was previously a professor of law at Vanderbilt University in Tennessee and has held visiting academic appointments in Australia, China, Germany, Italy, Japan, South Korea and Switzerland. Before entering the academy, he clerked for Justice Anthony M. Kennedy of the Supreme Court of the United States.
the industry, end-users and the academic community have only just begun to appreciate what the burgeoning deployment of this technology might mean and to study how to prepare for the challenges posed by this new technological environment.

The Personal Nature of the Information Collected

One of the IoT’s most distinctive aspects is the increasingly personal nature of the information collected. Connecting vehicles to the network means that others can track those vehicles’ movements and the manner in which they are operated. The use of smart devices in homes can reveal a great deal of information about residents’ habits and the ways that they live their lives. Attaching medical devices to the network can yield an immense amount of sensitive information about people’s health care. Combining multiple sources of data together and running predictive analytics on the resulting data can allow interested parties to infer surprisingly detailed levels of personal information about those using IoT devices. Interestingly, a survey of US consumers indicated that they are the most concerned about the sharing of information that reveals their personal habits (Rainie and Duggan 2016).¹

The Distributed Nature of Data Storage and Processing

Another difference between IoT systems and traditional systems is the frequency with which data is stored and processed locally. The fact that many IoT systems have little tolerance for latency often means that they handle many of the data-related functions in the local device instead of transmitting all data to a central location, such as a data centre.

Storing and processing data on a distributed basis has both advantages and disadvantages. The absence of a single large repository of multiple users’ data eliminates the presence of a large tempting target with a single attack surface that can draw the attention of cyber attackers. At the same time, decentralized storage raises the possibility that some locations will not consistently maintain the appropriate levels of security hygiene. Instead of relying on a single, hardened point protected by a small cadre of highly trained security professionals, distributed storage and processing rely on the diligence of individual users to maintain the integrity of the system.

In addition, the lack of centralized control means that any system architect must take into account the fact that the incentives of different actors connected to the system will necessarily vary. Although decentralized decision making often leads to outcomes that maximize the benefits to the system as a whole, that is not always the case. Under certain circumstances, it may be in the selfish best interest of one actor to submit erroneous data into the system in order to try to obtain greater benefits or to bear fewer costs. Even if every actor were to submit accurate information, individual actors may find it advantageous to deviate from their expected response to that data. As a result, IoT systems need some way to ensure the provenance and accuracy of data and to police whether decentralized decision makers are acting in ways that are consistent with the proper functioning of the overall system.

Sensors as a New Attack Vector

Everyone who has used the internet is well aware of the onslaught of cyber attacks that bombard computers nearly every day. Viruses, worms, trojans, botnets and other forms of malware have become all-too-familiar parts of the online experience, as are persistent efforts to hack through security.

The fact that IoT systems necessarily incorporate sensors that collect data from the physical world subjects them to an entirely
new vector of attack. In addition to the range of traditional online threats, flooding a sensor with electromagnetic radiation can cause it to malfunction. Even worse, a more sophisticated attacker can send carefully calibrated erroneous information to the sensor that can cause the system to take actions that are not warranted by the actual situation. For example, something as simple as spoofing location data can cause a connected car to veer far off course.

The Possible Corruption of IoT Devices

The fact that IoT devices are both partially programmable and connected to the network raises the possibility that bad actors may attempt to commandeer them or cause them to malfunction. The reality is that most IoT systems were not designed with security in mind. Video repositories such as YouTube contain numerous videos showing how sophisticated actors can use laptops to take over the driving functions of cars. The trade press abounds with stories where malicious operatives have subverted smart refrigerators, televisions, baby monitors and digital assistants. Perhaps most problematically, many medical devices have no security built into them at all. Many stories document the ease with which hackers can stop critical devices such as pacemakers and insulin pumps.

One can easily conceive of situations that would go beyond mere interference and extend to even more dire situations. The phenomenon of ransomware suggests that an adverse actor could use these capabilities to engage in extortion or worse.

Potential Responses

The existence of these potential threats underscores the need for the IoT industry and the academic community to develop solutions to these problems. Under a recent US National Science Foundation grant, a number of colleagues and I have designed a variety of strategies to address these problems.2

For example, the redundancy inherent in the distributed nature of the IoT can guard against cyber attacks, including zero-day attacks that have never been seen before. Utilizing an emerging approach known as accountability, IoT systems can assign a number of the other nodes to recheck the calculations of each node periodically. If a majority of the other nodes assigned to rerun the calculation come to a different result, the node being checked is declared to be in fault and isolated from the system.

Another technique known as state estimation can protect against sensor attacks. This approach takes the early experiences with a particular environment to estimate the reasonable range of possible values that a sensor might report. If the system receives data from the sensor that falls outside that range, it can flag that sensor for additional scrutiny or even go so far as to isolate it from the system.

With respect to privacy, a scheme known as differential privacy can prevent particular data from being attributed to any specific person in situations when individual data points are combined and reported as an aggregate value, such as a mean, by adding a predefined range of random noise to each data point. If the number of observations being aggregated is large enough, the central limit theorem of statistical analysis dictates that the randomness of the noise will tend to cancel itself out. This key concept of probability theory means that the data associated with different individuals can be obfuscated without materially degrading the quality of the information being sought. However, the resulting mean is more properly regarded as a distribution than as a true value. So long as the designers know how much variation the problem on which they are working can tolerate, they can calibrate the system in a way that preserves anonymity without compromising system performance.

What is perhaps most striking about each of these potential solutions is that none is perfect. Consider the approach reflected in accountability. If all of the nodes assigned to rerun the calculations of the node being checked are themselves compromised, they will come to the same erroneous answer and thus will fail to identify the fact that the node being checked has been corrupted. These errors can be reduced by assigning more nodes to rerun the calculations or by rerunning the calculations more frequently, but these solutions are costly and still will not completely eliminate the possibility that an attack may escape detection.

Similarly, state estimation only provides a probabilistic indication of integrity. It is
possible that an attack might yield values that fall within the range predicted by state estimation or might be successfully spoofed during the initial calibration phase so that the system believes that erroneous data is actually accurate.

The limits of these solutions underscore the fact that no amount of diligence can completely eliminate the security and privacy risks confronting IoT systems. Indeed, system designers could spend their entire development budgets on improving security, in which case they would have no money left to develop product features, and their system would still not be entirely secure. This means that the proper design of privacy and security of the IoT must be conceived as a trade-off that attempts to strike the proper balance between functionality and security.

The limited nature of security also dictates that the quest for perfect protection represents something of a unicorn hunt. Although designers should attempt to protect their systems as well as possible, the impossibility of perfect protection dictates that they should also plan for the inevitable failures by employing a layered security approach that supplements border protections with mechanisms designed to achieve fast detection and remediation of problems as they occur.

The Role of Law and Governance

The need to optimize multiple concerns also necessarily implies that the solution will not turn solely on the available technical solutions. Instead, the ultimate balance will depend on economic and legal considerations as well. For example, policy makers must decide whether to rely on tort law, which involves ex post compensation for wrongful harms suffered, or regulation, which focuses on ex ante prevention of harms.

With respect to tort law, whether product liability will stop short of holding IoT device manufacturers to a perfection standard may depend on how many other courts follow the lead of many US and Canadian courts and adopt the risk-utility standard. This standard explicitly frames the analysis in terms of the costs and benefits of different designs. Regulation will likely follow the existing sector-specific agency structure, which will assign responsibility for different types of IoT to different agencies. This division of authority risks yielding inconsistent outcomes and relying on IoT expertise spread thinly across multiple agencies.

A central question regarding privacy regulation will turn on whether it will follow the sector-specific approach followed in the United States and some Canadian provinces, or the omnibus privacy regulation embraced by the federal government in Canada and in Europe.

In addition, several standard-setting organizations (SSOs) are vying for leadership in IoT standards. The burgeoning significance of the IoT heightens the importance of the governance structures that determine how these SSOs will make decisions. Perhaps most importantly, the primary goal should not be to remediate problems that have occurred, but rather to create high-powered economic incentives to avoid them in the first place. That means that any legal and regulatory interventions must seek to align incentives with good outcomes and should reflect the likely reactions to any policies.

Policy makers have only begun to consider how law and governance will need to adapt to the emergence of the IoT. Hopefully, the academic and industry research communities will continue to provide answers to these questions as the IoT industry continues to mature.

Work Cited

Endnotes
1 The survey found that smart thermostat information was the type of personal information among six scenarios that consumers found the least acceptable to share.
2 This research is supported by the US National Science Foundation Award No. 1505799.
Artificial intelligence (AI) is truly a revolutionary feat of computer science, set to become a core component of all modern software over the coming years and decades. This presents a threat but also an opportunity. AI will be deployed to augment both defensive and offensive cyber operations. Additionally, new means of cyber attack will be invented to take advantage of the particular weaknesses of AI technology. Finally, the importance of data will be amplified by AI’s appetite for large amounts of training data, redefining how we must think about data protection. Prudent governance at the global level will be essential to ensure that this era-defining technology will bring about broadly shared safety and prosperity.

**AI and Big Data**

In general terms, AI refers to computational tools that are able to substitute for human intelligence in the performance of certain tasks. This technology is currently advancing at a breakneck pace, much like the exponential growth experienced by database technology in the late twentieth century. Databases have grown to become the core infrastructure that drives enterprise-level software. Similarly, most of the new value added from software over the coming decades is expected to be driven, at least in part, by AI.

Within the last decade, databases have evolved significantly in order to handle the new phenomenon dubbed “big data.” This refers to the unprecedented size and global scale of modern data sets, largely gathered from the computer systems that have come to mediate nearly every aspect of daily life. For instance, YouTube receives over 400 hours of video content each minute (Brouwer 2015).

Big data and AI have a special relationship. Recent breakthroughs in AI development stem mostly from “machine learning.” Instead of dictating a static set of directions for an AI to follow, this technique trains AI by using large data sets. For example, AI chatbots can be trained on data sets containing text recordings of human conversation collected from messenger apps to learn how to understand what humans say, and to come...
up with appropriate responses (Pandey 2018). One could say that big data is the raw material that fuels AI algorithms and models.

The main constraint on innovation is no longer the difficulty in recording and storing information, but the finding of useful insights among the sheer abundance of data now being collected. AI can notice patterns in mammoth data sets that are beyond the ability of human perception to detect. In this way, the adoption of AI technology can make even mundane and seemingly trivial data valuable. For instance, researchers have trained computer models to identify an individual's personality traits more accurately than their friends can, based exclusively on what Facebook posts the individual had liked (Wu, Kosinski and Stillwell 2015).

**AI and Cyber Security**

Hardly a day passes without a news story about a high-profile data breach or a cyber attack costing millions of dollars in damages. Cyber losses are difficult to estimate, but the International Monetary Fund places them in the range of US$100–$250 billion annually for the global financial sector (Lagarde 2012). Furthermore, with the ever-growing pervasiveness of computers, mobile devices, servers and smart devices, the aggregate threat exposure grows each day. While the business and policy communities are still struggling to wrap their heads around the cyber realm's newfound importance, the application of AI to cyber security is heralding even greater changes.

One of the essential purposes of AI is to automate tasks that previously would have required human intelligence. Cutting down on the labour resources an organization must employ to complete a project, or the time an individual must devote to routine tasks, enables tremendous gains in efficiency. For instance, chatbots can be used to field customer service questions, and medical assistant AI can be used to diagnose diseases based on patients' symptoms.

In a simplified model of how AI could be applied to cyber defence, log lines of recorded activity from servers and network components can be labelled as “hostile” or “non-hostile,” and an AI system can be trained using this data set to classify future observations into one of those two classes. The system can then act as an automated sentinel, singling out unusual observations from the vast background noise of normal activity.

This kind of automated cyber defence is necessary to deal with the overwhelming level of activity that must now be monitored. We have passed the level of complexity at which defence and identification of hostile actors can be performed without the use of AI. Going forward, only systems that apply AI to the task will be able to deal with the complexity and speed found in the cyber security environment.

Continuously retraining such AI models is essential, since just as AI is used to prevent attacks, hostile actors of all types are also using AI to recognize patterns and identify the weak points of their potential targets. The state of play is a battlefield where each side is continually probing the other and devising new defences or new forms of attack, and this battlefield is changing by the minute.

*Automating tasks that previously would have required human intelligence, such as using chatbots to field customer service questions, is one of the essential purposes of AI, and enables tremendous gains in efficiency for organizations.*

*(Photo: Piotr Swat / Shutterstock.com)*
Perhaps the most effective weapon in a hacker’s arsenal is “spear phishing” — using personal information gathered about an intended target to send them an individually tailored message. An email seemingly written by a friend, or a link related to the target’s hobbies, has a high chance of avoiding suspicion. This method is currently quite labour intensive, requiring the would-be hacker to manually conduct detailed research on each of their intended targets. However, an AI similar to chatbots could be used to automatically construct personalized messages for large numbers of people using data obtained from their browsing history, emails and tweets (Brundage et al. 2018, 18). In this way, a hostile actor could use AI to dramatically scale up their offensive operations.

AI can also be used to automate the search for security flaws in software, such as “zero-day vulnerabilities.” This can be done with either lawful or criminal intent. Software designers could use AI to test for holes in their product’s security, just as criminals search for undiscovered exploits in operating systems.

AI will not only augment existing strategies for offence and defence, but also open new fronts in the battle for cyber security as malicious actors seek ways to exploit the technology’s particular weaknesses (ibid., 17). One novel avenue of attack that hostile actors may use is “data poisoning.” Since AI uses data to learn, hostile actors could tamper with the data set used to train the AI in order to make it do as they please. “Adversarial examples” could provide another new form of attack. Analogous to optical illusions, adversarial examples consist of modifying an AI’s input data in a way that would likely be undetectable to a human, but is calculated to cause the AI to misclassify the input in a certain way. In one widely speculated scenario, a stop sign could be subtly altered to make the AI system controlling an autonomous car misidentify it as a yield sign, with potentially deadly results (Geng and Veerapaneni 2018).

The New Value of Data

AI technology will alter the cyber security environment in yet another way as its hunger for data changes what kind of information constitutes a useful asset, transforming troves of information that would not previously have been of interest into tempting targets for hostile actors. While some cyber attacks aim solely to disrupt, inflict damage or wreak havoc, many intend to capture strategic assets such as intellectual property. Increasingly, aggressors in cyberspace are playing a long-term game, looking to acquire data for purposes yet unknown. The ability of AI systems to make use of even innocuous data is giving rise to the tactic of “data hoovering” — harvesting whatever information one can and storing it for future strategic use, even if that use is not well defined at present.

A recent report from The New York Times illustrates an example of this strategy in action (Sanger et al. 2018). The report notes that the Chinese government has been implicated in the theft of personal data from more than 500 million customers of the Marriott hotel chain. Although commonly the chief concern regarding data breaches is the potential misuse of financial information, in this case the information could be used to track down suspected spies by examining travel habits, or to track and detain individuals to use them as bargaining chips in other matters.

Data and AI connect, unify and unlock both intangible and tangible assets; they shouldn’t be thought of as distinct. Quantity of data is becoming a key factor to success in business, national security and even, as the Cambridge Analytica scandal shows, politics. The Marriott incident shows that relatively ordinary information can now provide a strategic asset in the fields of intelligence and national defence, as AI can wring useful insights out of seemingly disparate sources of information. Therefore, this sort of bulk data will likely become a more common target for actors operating in this domain.

Implications for Policy and Governance

These unfolding developments will force a rethinking of prevailing cyber security strategies. In an increasingly interconnected system, identifying the weakest link becomes more challenging, but also more essential. As sensors, machines and people become interwoven providers of data for valuable AI systems, there will be a proliferation of entry points for cyber attacks. Cyber security requires a comprehensive strategy to minimize weakest links; a piecemeal approach to cyber policy will not work. Since the training
data that feeds the most important and revolutionary AI technologies is global in scope, gathered from across many different countries, it is clear that governance at the national level alone will not suffice.

Global policy makers have begun turning their attention to the ramifications of widespread AI technology, and its effect on cyber security in particular. The Group of Seven (G7) turned its attention to the governance of AI during the 2018 summit in Charlevoix, Quebec, pledging to “promote human-centric AI” through appropriate investments in cyber security, while paying heed to privacy and personal information protection regarding the data that serves as the raw input for machine learning (G7 2018).

The application of AI technology to pre-existing cyber attack strategies such as spear phishing will both augment their effectiveness and — by circumventing labour constraints — expand the number of actors capable of undertaking them. This lends a greater urgency to existing efforts to create effective global governance in cyberspace and international data protection, such as the United Nations’ Group of Government Experts’ attempt to establish accepted norms of conduct.

The very same pieces of technology that enable more threatening types of cyber attack are also driving growth in the civilian economy and enabling more effective cyber defence. While commonly thought of as a threat to privacy, AI also has the potential to help preserve privacy and exert control over proprietary data and its derived assets. Policy makers will have to carefully consider how to regulate the use of these technologies, balancing the need to keep powerful weapons out of the hands of malicious actors without stifling innovation. It will be essential to harmonize such policies across national jurisdictions. Since hostile actors are capable of reaching across borders with stupendous ease, any country that unilaterally restricts the use and development of these technologies within its borders would be putting itself at a competitive disadvantage. Moreover, as AI technology becomes more integrated into the general economy and civilian sphere, existing legal and normative frameworks may need to be adjusted to cover novel forms of attack such as data poisoning and adversarial examples. Up to this point, data theft has been the main concern in cyberspace. Going forward, hostile actors will likely try to gain access to databases not only to obtain their information, but also to alter and manipulate them. The legal definition of what constitutes a cyber attack may need to be amended to cover these novel threats (Brundage et al. 2018, 57).

AI algorithms learn from data to produce a valuable new prediction tool, and the output of AI can be separated from the original training data. Therefore, to truly control the data and its value, any assets that are produced from data must also be controlled. The infrastructure that allows the recording, storage and analysis of big data should be treated as an asset just like it is in any other sector. Furthermore, some sectors, such as finance, have systemic implications, and are even more important to protect due to third-party linkages. Governing institutions will need to continue to improve their security posture in these and many other areas, including identity fraud. Since the AI software used for attack purposes is capable of rapidly evolving, this is an ongoing requirement rather than a one-off investment.
TLD Operator Perspective on the Changing Cyber Security Landscape

Byron Holland

Just five years ago, most industries were concerned with simply implementing the necessary technology infrastructure for going digital. We have now arrived at an inflection point where securing those internet-connected digital assets against inevitable cyber attacks is critically important. The cyber security landscape has changed dramatically, with major distributed denial of service (DDoS) attacks and ransomware schemes routinely making headlines. According to a report by the Center for Strategic and International Studies (CSIS) and McAfee, the global cost of cybercrime was estimated at US$600 billion for 2017 — a significant jump from US$445 billion in 2014 (Lewis 2018). As more people come online and access essential services via the internet, bad actors stand to benefit a great deal from a successful attack. The domain name system (DNS) is one of the most critical components of global internet infrastructure. When a portion of the DNS is compromised or unavailable, users cannot reach the associated resources on the internet. This is because the DNS operates as the address book for the internet and is responsible for translating Internet Protocol (IP) addresses, such as 162.219.54.2 or 2001:500:80:2::12, into human-friendly domain names, such as example.ca, and much more. Taking down elements of the DNS means entire swaths of the internet become unreachable. It is no wonder, then, that the DNS is a perennial target for cyber attackers.

Top-level domain (TLD) registry operators play an important role in ensuring the DNS — and, therefore, the internet — functions on a day-to-day basis. The TLD is the string of characters to the right of the dot in a web address, such as .ca, .com or .org. TLD registry operators are responsible for maintaining a database of all domain names for their TLD. Registry operators store information about each person or company that registers a domain name, as well as the administrative and technical information that makes their domain name reachable on the internet. This includes the IP addresses of the name servers associated with each domain name under management.

A country code TLD, or ccTLD, has two letters and is generally designated for use by a particular country, sovereign state or autonomous territory. The Canadian Internet
Registration Authority (CIRA) manages the stewardship of .ca on behalf of all Canadians. In addition to country codes, there are over one thousand generic top-level domains (gTLD), from the ever-popular .com and .org, to newer generic TLD strings such as .sucks or .guru.

As security threats ramp up, it is critical that businesses vigilantly monitor and protect against security threats on two primary axes: databases and information technology (IT) infrastructure. For a TLD registry operator like CIRA, this means securing the databases associated with the domain name registry and the DNS infrastructure that supports public domain name resolution. For a bank, this translates into protecting databases containing customers’ valuable personal information and any internet-connected infrastructure that supports financial transactions. For a social media platform, this involves protecting users’ personal data and securing the infrastructure that ensures global service availability.

**Data Security**

Equifax and Cambridge Analytica may be household names, but for the worst of reasons. An organization’s reputation is only as good as its data security controls. Regulators are cracking down on poor information security practices, and people are acutely aware of the harm associated with personal information divulged in a data breach. Financial institutions and large corporations remain targets of data theft, but phishing schemes and attacks against small businesses, municipalities and universities are increasing.

A 2018 CIRA survey of 500 individuals with responsibility over IT security decisions found that 66 percent of businesses with 250 to 499 employees experienced a cyber attack in the last 12 months, and an estimated 70 percent of data breaches occur at companies with fewer than 100 employees (CIRA 2018). Whether it is collected for the purposes of providing a service, selling advertising or for analytics more generally, personal information retained by any organization is stored in a database. Regardless of the size or nature of a business, these databases are prime targets for cyber attackers.

**Threat Vectors**

For a TLD operator, the integrity of the registry is critical to its operations. Depending on its operational model, a TLD can provide internet-accessible services to its domain name distribution channel partners (registrars) and/or to individual domain name owners (registrants). In all cases, the integrity of the registry is dependent on the ability of the registrar and registrant to protect their system access credentials. Like other service providers who allow users to access systems online, many TLD operators are enforcing enhanced security mechanisms such as two-factor authentication and IP address whitelisting.

It is not uncommon for a company to notify customers of an undetected breach that occurred months or even years earlier. The primary threat CIRA encounters as a registry is compromised registrars or registrants, where a bad actor infiltrates their systems and then lies in wait before changing the attributes of domain names in order to carry out the attack of choice. This often takes the form of pointing a domain name toward a compromised site instead of the rightful registrant’s intended content. Attackers tend to target high-profile, high-traffic domain names and redirect unsuspecting users toward malicious sites that prompt them to enter personal information or that implant malware into their browser.

Also related to registry security is the prevention of phishing and malware distribution via doppelgänger domains or “typosquatting.” These attack vectors involve new domain registrations that imitate existing, usually high-traffic, domains with similar spellings or easily mistyped permutations. Doppelgänger sites are generally used for distributing malware and executing phishing schemes, usually by imitating financial institutions or government agencies to collect valuable personal information that can be used to steal identities and drain bank accounts.

Typosquatters also harvest personal information via email by registering domains that omit the dot between a company’s host name and their subdomain. When a user sends an email to hello@caexample.com rather than hello@ca.example.com, the contents of the email are shared with the typosquatter, thereby opening a phishing opportunity.

**Prevention and Mitigation**

To prevent attackers from accessing valuable personal information, the domain name industry has responded by implementing...
state-of-the-art controls to protect security management systems and lock down access, so that only authorized registrars have the ability to access systems. To prevent abuse in the registry, many TLD registries have implemented these controls.

At CIRA, for example, we actively monitor new and existing domain registrations for malicious activity. Registrants also have the option of locking their domains to prevent domain abuse by using a registry lock mechanism, where changes to the domain can only be performed by an authorized person using multifactor authentication. CIRA also investigates registrations based on information we receive from partners in the cyber security ecosystem, including the Canadian Centre for Cyber Security (CCCS) and the Canadian Cyber Threat Exchange (CCTX).

Where a registration is determined to be dubious, CIRA may audit the registration via the Registrant Information Validation process. This process is in place to ensure that a given registrant meets Canadian presence requirements as per CIRA policy, which dictates that persons who wish to register a .ca domain name or sub-domain name require a legitimate connection to Canada in at least one of 18 categories. Malicious .ca registrations frequently originate from foreign registrants who do not provide documentation to prove they meet Canadian presence requirements. When a registrant fails to demonstrate compliance with the Canadian presence requirements, the domain is suspended and then cancelled.

*Trends*
Both the CCCS and the CCTX were established in the last two years in order to facilitate knowledge sharing of known threats and respond to the growing threat landscape. CIRA and its registrar partners have become increasingly security savvy, implementing stringent security protocols in order to prevent bad actors from infiltrating the registry. As a result, we experience very few registry compromises, but this doesn’t mean attempts are decreasing. Bad actors consistently probe the registry. Like any system that is connected to the internet, we experience steady probing and attacks against our databases.

As in any industry, domain name holders want to know that their data is being adequately protected when they entrust a company with their personal information. A key element of the value proposition of any given TLD is trust in the reputation and ability of the registry operator to police its name space and enforce its policies for registrars and registrants. Given the nature of .ca’s Canadian presence requirements and the processes we have implemented to protect .ca’s reputation, malicious registrations are less common in CIRA’s registry than in many other TLDs. According to Spamhaus, the most abused TLD is currently .loan, with 30,399 of its 33,328 visible domains under management linked to malicious spam or malware distribution. Unfortunately, TLDs that do not protect their namespaces against misuse are enabling cyber attackers to engage in malicious activity such as spam attacks and infrastructure abuse.

**Infrastructure Security**
DDoS attacks represent the single biggest threat to internet-connected infrastructure, including the infrastructure that supports the DNS. The objective of a denial of service attack is to exhaust the computational or bandwidth resources of the target website or digital service by overwhelming the infrastructure that supports it. With a DDoS attack, the attack traffic originates from a distributed network of compromised systems recruited to simultaneously overwhelm the target with internet traffic. DDoS attacks typically require thousands of devices working in concert. This is known as a botnet, and users are often unaware their devices are participating in a botnet attack due to malware installed on their machines.
The burgeoning Internet of Things (IoT) presents a major threat to internet infrastructure and is particularly problematic for network operators running critical infrastructure. Any internet-connected device with a processor can be recruited into a DDoS attack. Attackers target vulnerabilities in the unsecured IoT devices that have flooded the consumer market, including everything from internet-connected routers and cameras to toasters and doorbells.

**Threat Vectors**

A TLD’s DNS infrastructure is comprised of a network of public DNS servers, located in a number of strategic geographic locations. Many public and private networks follow a similar architecture. These servers are frequently the targets of DDoS attacks.

There are two broad types of DDoS attacks: brute force and amplification. The 2016 Mirai attack against Dyn’s managed DNS infrastructure (Dyn 2016) is an example of a brute force attack. This attack leveraged hundreds of thousands of compromised IoT devices to send traffic directly to Dyn’s DNS servers and represented the largest DDoS attack to that point in the history of the internet (Woolf 2016).

Amplification-based DDoS attacks are particularly effective against DNS infrastructure. Amplification attacks involve three elements: spoofing, reflection and then amplification. An attacker’s goal is to saturate a given server, thereby taking it offline and preventing legitimate queries from getting through. An attacker achieves this by imitating, or “spoofing,” the IP address of the target machine, then manufacturing and directing a high number of User Datagram Protocol-based queries at open public DNS, Simple Network Management Protocol and Network Time Protocol servers. Operating public DNS servers makes DNS registry operators a prime target for large-scale amplification attacks.

Thousands of open servers perceive the attack queries as originating from the target server, then reflect the attack from the source toward the attacker’s intended target. In the case of attacks that harness DNS infrastructure, the problem is compounded because a small 64-byte DNS query can be crafted to generate a large answer with thousands of bytes in response, thereby swamping the target with a high volume of junk traffic.

A further type of amplification attack involves querying thousands of open Memcached servers, which are typically used to improve the
A further threat associated with the DNS is DNS hijacking and Border Gateway Protocol (BGP) hijacking. This occurs when attackers wrongly and purposely announce ownership of internet resources (for example, nameservers or IP addresses) that they do not control, in effect impersonating the true managers of the resources. In a 2018 paper, researchers at the US Naval War College and Tel Aviv University described unusual patterns in BGP announcements involving China Telecom misdirecting traffic through China before delivering it to the rightful destinations in North America and Europe (Demchak and Shavitt 2018). Such attacks require extensive planning, and some are more pervasive than others.

**Prevention and Mitigation**

In the wake of the Mirai attack against Dyn, many stakeholders in the network operator community ramped up efforts to protect critical internet infrastructure from IoT attacks. From a TLD operator’s perspective, the current best practice to mitigate against large-scale DDoS attacks is to utilize multiple globally distributed DNS anycast providers. Local anycast nodes are not globally accessible and only serve local network peers and internet service providers (ISPs), making the nodes significantly less vulnerable to globally distributed DDoS attacks. For ccTLDs, it is a best practice to implement local anycast as close as possible to the country’s users. In Canada, CIRA operates .ca local anycast nodes connected to local internet exchange points in cities across the country.
In terms of DNS and BGP hijacking, the internet operator community has responded by developing new protocols and standards such as Domain Name System Security Extensions and DNS-based Authentication of Named Entities, which allow the rightful resource managers to sign and authenticate resources and detect traffic redirection. However, these protocols have not yet been widely adopted.

Trends

Over the last five years, attacks against the DNS have increased in frequency and intensity. This rise is directly correlated with the proliferation of botnets that are made possible by low-cost IoT consumer devices. Many of these consumer devices are visible on the public internet and have notoriously low security settings, making them easily hacked with default passwords. The aggregated bandwidth of millions of compromised “zombie” devices in a botnet has proven disastrous for the targets of attacks. Even sophisticated targets cannot sustain the brute force of upward of a terabit per second of traffic.

While the industry has responded to the problem by implementing changes to network architecture and introducing rate limits on the number of consecutive queries that a server answers, bad actors are innovating, too. Attackers are simulating traffic in ways that make sham queries appear very realistic, making it difficult to differentiate between legitimate and fraudulent traffic. Attacks also tend to generate a flurry of legitimate retry traffic as DNS servers refresh their caches in response to being prevented from completing a legitimate query. This retry activity serves to further swamp the target.

The nature of attackers has also changed. While the mischievous teenagers and organized hacktivist groups of the early Web do continue to inflict some damage, organized crime rings and state-sponsored actors are capable of much larger attacks. These sophisticated actors engage in a range of malicious activity, including major DDoS campaigns, cyber espionage and election tampering. In the wake of cyber threat activity against the election processes in the United States and Europe, the newly created Cybersecurity and Infrastructure Security Agency (CISA) has warned against the threat of state-controlled actors’ attempts to influence the democratic process in Canada (CSE 2018).

Hyperlocalization of infrastructure is a new trend developing in response to ever-growing DDoS attacks. For a TLD operator, the goal of hyperlocalizing infrastructure is to situate root and TLD zone files on or as close as possible to an ISP’s recursive DNS servers. Network operators are also closely monitoring new encryption protocols coming out of the Internet Engineering Task Force such as DNS over Transmission Control Protocol and DNS over Hyper Text Transfer Protocol Secure. The implementation of these new protocols is still very new, and the associated potential threats remain unknown at this time.

Conclusion

Prevention of data theft remains a high priority in the technology industry, but threats to infrastructure posed by IoT-enabled botnets are growing in frequency and severity. There is no silver bullet for mitigating the threats that cyber attacks pose to data security or critical internet infrastructure. The network operator industry and wider internet community have responded to new threats with a multi-layered approach, including everything from threat monitoring and knowledge-sharing to redesigning networks and developing entirely new protocols.

Cyber security requires not just a single solution, but an array of approaches that reinforce one another. As a TLD operator, CIRA constantly monitors its systems in order to detect attacks and mitigate risk. We’re also acutely aware that the world of cyber security does not stand still. New attack vectors and seemingly “black swan” events are constantly cropping up, requiring ongoing vigilance and adjustments to the changing landscape.

Works Cited


Endnotes

Strategic Stability, Cyber Operations and International Security

David Mussington
It is easy to be taken aback by how quickly digital information and communication technology (ICT) has become indispensable to government, the economy and everyday life. Vital infrastructure such as electrical grids, hospitals, media and transportation networks have become ICT reliant. The weapons and defensive systems of most advanced economies have followed suit. But the same flowering of ICT infrastructure that has produced wondrous gains in efficiency carries with it an inherent vulnerability, presenting a novel avenue of attack through cyberspace by which hostile actors can strike. Governments have been slow to rouse themselves to this threat; a recent report from the US Government Accountability Office (2018) admonished the Department of Defense for its lax standards, asserting that many US weapons systems could be disabled through simplistic cyber attacks. This pervasive vulnerability to threats from cyberspace has worrying implications for national security and international stability.

The technical and political difficulties of accurately attributing cyber attacks offer hostile actors the ability to avoid punishment, creating an “offence-dominant” environment. Shared supply chains and reliance on a small number of ICT platforms ensure that government infrastructure and security systems possess the same technical vulnerabilities as the private sector, many of which are well known or easily discoverable. Antiquated global governance surrounding the use of force has allowed malicious actors to perpetrate mischief while staying just below the threshold that would provoke a response. In combination, these factors present a challenge to the maintenance of global stability, which both national governments and international organizations are struggling to cope with.

The rapid rate of technological change inevitably outpaces government and society’s ability to comprehend that change. This is true at both the national and the international level. National governments and international organizations are now struggling to understand the vulnerabilities posed by the world’s unprecedented reliance on digital infrastructure, and the destabilizing effect this may have on the current international order.

Well-established concepts within international security, such as the effectiveness of deterrence strategies, have been cast into doubt. Nonetheless, a few broad implications of the new importance of the cyber domain can be observed from within a general climate of uncertainty.

### Attribution Uncertainty

Strategic stability at the global level relies on the concept of deterrence — preventing aggression by threatening harsh punishment, or by imposing costs that exceed any benefits from attack. The anonymity granted to actors in cyberspace makes it tough to identify the culprit of a given attack with a high degree of certainty (the origin of a piece of malware is much less obvious than the origin of a missile strike), undermining the effectiveness of deterrence strategies and emboldening attackers (Solomon 2011).

While there has been some progress in improving the technical aspect of cyber attack attribution, political difficulties remain. After all, for a deterrence strategy to work, a state must retaliate once an attack is identified, and allies committed to collective defence must come to their aid. Despite traditional rhetoric, such assistance is never automatic, and the added problem of convincingly attributing cyber attacks adds another layer of uncertainty to the political calculus. Honouring commitments to allies can be costly, and states will be reluctant to bear this burden if there remain any doubts about the identity of the attacker. In this way, the cyber-attribution problem can undermine the cohesiveness of alliances and, by extension, international stability.

Another factor that plays into attribution difficulties is the growing technological capability of the private sector. This has empowered a plethora of actors, such as cyber security firms, to perform their own cyber attribution and contest the attribution claims of state governments (Romanosky 2017). Claims made by states must now survive inspection by subject matter experts in the private sector (many of whom have experience in the defence and intelligence communities), who question all factual disclosures and explanations. The waters of attribution are further muddied by politicians and members of the media who are often quick to denounce suspected culprits despite lacking technical evidence. When the French media outlet TV5Monde was infiltrated by hackers claiming to be affiliated with the Islamic State group, certain politicians and members of the media...
were quick to run with this story, although the French prosecutor’s office later found the evidence pointed toward a Russian espionage group (Soesanto 2017). This has had the effect of eroding national governments’ authority over such matters, aggravating uncertainty.

**Offence Dominance**

An assertion frequently made about cyberspace is that attacking is relatively easy, and protection and defence much more difficult, compared to conventional theatres of war (Kello 2013). Attacks and espionage in cyberspace can generally be perpetrated at lower cost compared to traditional methods. The 10 million daily intrusion attempts at the Pentagon speak volumes about the affordability of offensive cyber operations (Fung 2013). This allows traditionally weaker actors to pose a threat to the United States and its allies in ways not previously possible.

Furthermore, our ever-increasing reliance upon ICT infrastructure in defence systems and the civilian economy has dramatically multiplied the number of vulnerable points that must be defended. While the size of a state’s physical territory, defended by its conventional forces, usually stays the same over time, the number of “entry points” in cyberspace that it must defend is constantly growing (Singer and Friedman 2014). The arrival of cloud services and the Internet of Things will only add to this difficulty. Compounding this problem is the failure of civilians to adopt safer digital habits. Reliance on a small number of technology platforms ensures that common exploitable vulnerabilities are widely shared, while public disclosures of compromised systems spread knowledge of these common exploits to potential hostile parties.

Tools of cyberwar are largely non-physical and therefore easier to conceal than conventional forces, making it difficult for actors to assess each other’s capabilities. Offensive military cyber doctrines in the United States, Russia, China and elsewhere show that states are imitating neighbours and competitors when they develop their own cyber capabilities. However, these doctrines are not widely understood, feeding mistrust and the perceived need to gain a “first mover advantage” (ibid.). This in turn heightens the danger of escalation and reduces stability. Under the circumstances, a stable and persistent advantage in cyberspace seems unattainable.

**Intelligence Tools as Offensive Cyber Weapons**

Many observers fail to fully appreciate how much current cyber operations owe to innovation by intelligence agencies charged with obtaining information about the political, economic and defence postures of potential competitors (and allies). The malware and signals intelligence capacity of these agencies grants the ability to maintain an accurate awareness of changes in the cyber environment, as well as the power to reshape it. States feel that access to foreign information systems and critical infrastructures is necessary for them to be aware of incoming attacks and to retaliate against them. However, the ability to degrade an opponent’s conventional military capabilities through cyber-enabled espionage may actually weaken deterrence in other domains such as air, maritime, land and outer space. The timely and coordinated deployment of these conventional forces has become dependent on ICT infrastructure. The possibility that these communication and early warning systems may have been covertly infiltrated erodes actors’ confidence in their defensive abilities, increasing mistrust and the potential for conflict. The effect this may have on the behaviour of nuclear armed states is especially worrying (Cimbala 2016).

**Pervasive Infrastructure Cyber Vulnerability**

The private sector remains at the forefront of ICT development. Therefore, governments must rely on the same “commercial off-the-shelf
Restoring clarity to the “action-reaction” dynamic is necessary both to dissuade hostile actors by guaranteeing reprisal for certain offences, and to solidify an understanding among allies as to when they must come to one another’s assistance.

Current trends appear to be pushing the ICT supply chain toward greater homogeneity. This is partially a consequence of laws and regulations, but also of industry convergence around common standards driven by commercial incentives. Best-practice guidelines issued by governments generally call for the maintenance of secure system configurations, but this will not solve the problem of vulnerable legacy technology, or the “undirected” nature of technical change driven by commercial competition. This presents another weakness that governments have been slow to acknowledge.

Antiquated Legal Regime and “Grey Zone” Conflict

International law surrounding the use of force is now more contested, with disputes over whether it can properly address the threat posed by cyber attacks in a world rife with vulnerable ICT infrastructure. Existing norms and laws (for instance, as articulated in article 2(4) of the United Nations Charter) were created at a time when the use of force took the form of obvious, more easily attributable discrete events, such as the movement of troops or a missile strike. In contrast, a widely accepted understanding of what would constitute “use of force” in cyberspace has yet to be found (Tsagourias 2017). Many types of hostile action conducted in cyberspace, such as theft of a corporation’s intellectual property or spreading misinformation to influence foreign elections, do not cause direct harm to people in the same way as conventional weapons, leaving doubts as to whether they constitute a “use of force,” and, therefore, whether the victim may invoke their right to self-defence.

This ambiguity, combined with the difficulties of cyber attribution, has been voraciously exploited by an assortment of state, non-state and suspected proxy actors, as part of a strategy sometimes referred to as “hybrid warfare” (Cantwell 2017). The result has been a near constant drizzle of activity in cyberspace calculated to fall into a “grey zone”—undoubtedly hostile, but falling below the threshold of intensity that would provoke retaliation. Russia’s aggressive activities over the past decade provide a prime example of this tactic. These are widely believed to include attempts to influence elections in the United States and Western Europe, and denial of service attacks on government service websites (ibid.). Yet, despite these provocations, Russia’s adversaries — actual and potential — appear hesitant to respond decisively.

Coping with Instability

As a response to grey zone conflict and offence dominance in cyberspace, many national governments, such as the United States, Germany and Canada, have concluded that a static defence is no longer adequate and have been adjusting to allow pre-emptive cyber operations intended to disrupt hostile actors before they can act (Herpig 2018; Nakashima 2018; Grigsby 2017). Organizations at the international level have mirrored this trend. In 2017, the North Atlantic Treaty Organization (NATO) adjusted its policy away from ambiguity on cyber effects to a more responsive stance, establishing a Cyber Operations Centre to integrate the cyber capabilities of its members into military operations (Ricks and Ali 2017). While this may be necessary to cope with the attribution problem and grey zone hostilities, whether or not this will re-enable effective deterrence or cause further destabilization through tit-for-tat escalation remains unclear.

Due to many of its members being on the receiving end of grey zone cyber attacks, NATO has been a leading light in trying to resolve the current uncertainty plaguing international governance of cyber conflict. It has attempted, through efforts such as the establishment of the Cooperative Cyber Defence Centre of Excellence and publication of the Tallinn Manual, to arrive at a clear interpretation of which acts in cyberspace are permissible or not under current international law (Arts 2018). The alliance relies on all members following through on their commitment to collective defence as stipulated under article 5 of the alliance’s treaty. This makes the attribution challenge in cyberwarfare especially problematic, as it can give members a plausible reason to demur on this potentially costly commitment. This is forcing NATO to consider what kind of activity in cyberspace would be serious enough to invoke the collective
defence clause. While NATO has affirmed that article 5 could be triggered by a significant cyber attack, as of yet it has not determined a precise threshold (ibid.).

Implications and Policy Consequences

Global strategic stability is undermined by the failure of states to take seriously the erosion of defence capabilities caused by growing reliance on ICT technologies in critical infrastructures and weapon systems. At present, COTS and the ICT supply chain that services critical infrastructure present a particularly vulnerable point of entry for malicious actors. Existing governance and oversight mechanisms concerning the deployment of ICT will prove too lenient for the developing threat environment. Enhanced communication and tighter cooperation between government and the private sector will prove crucial to bolstering defences in this area. More arrangements like the Information Sharing and Analysis Centers, which facilitate intelligence sharing on cyber threats between the public and private sector, would be of great benefit (Lord and Mussioning 2017).

Superior coordination and information sharing are also required at the international level. In the absence of an offence-dominant environment, efforts must be taken to assure the uncertainties felt by various actors as to each other’s capabilities and intentions. The technical and political difficulties in attributing cyber attacks, combined with their affordability, will continue to encourage attackers. Those defending against cyber attacks must therefore take a firmer, less equivocal stance than they have so far displayed. Absent an international consensus on what constitutes use of force in cyberspace, the United States and fellow NATO members must collectively decide upon a clear code of conduct for responding to grey zone activities, in order to banish ambiguity and the risk of miscalculation. A red line should be drawn around the most pernicious types of cyber hostilities now being perpetrated, such as attempts to sway foreign elections, the violation of which should trigger a measured yet firm response. Restoring clarity to the “action-reaction” dynamic is necessary both to dissuade hostile actors by guaranteeing reprimand for certain offences, and to solidify an understanding among allies as to when they must come to another’s assistance. In the long term, the United States and its allies should promote more effective international governance by pushing to have these red lines enshrined as international norms in fora such as the United Nations. There is an urgency to this effort — failure to do so will only entrench the idea that the constant grey zone hostilities we are now witnessing have become a tolerable part of international behaviour.

Works Cited


The Quantum Threat to Cyber Security

Michele Mosca and Bill Munson

Canada’s cyber security strategy, *National Cyber Security Strategy: Canada’s Vision for Security and Prosperity in the Digital Age* (June 2018), stresses the need to prepare for increasingly sophisticated threats to the cyber systems that our critical infrastructure and democratic institutions rely on. The strategy commits the government — in this context, its cyber security efforts — to “focus on emerging areas of Canadian excellence, such as quantum computing” (Public Safety Canada 2018, 24).

Many people have heard of quantum computing, know that it’s coming and are aware that it will bring an almost unimaginable speed-up in the ability of computers to perform many kinds of calculations. This will allow wonderful advances in, for example, our ability to discover new materials and design new life-saving drugs. Unfortunately, powerful quantum computers will also enable the hacking of today’s “unbreakable” encryption in minutes.

As things stand, the encryption that underpins the security of society’s critical infrastructure is at serious risk of being undermined by quantum computers within the next eight to 15 years. This is the “quantum threat” — that Canada’s national security and economic prosperity will be jeopardized as government and other critical infrastructure systems become vulnerable to hostile actions because our cryptography is no longer strong enough to protect us. Even now, bad actors are able to copy and store encrypted data until a quantum computer is available to decrypt it.

This essay outlines how achieving a quantum-safe Canada is a natural cornerstone of a national strategy to protect Canadians and the economy from cyber attacks while also reaping the economic benefits of those efforts.

The Quantum Threat to Cyber Security

There is growing recognition of the need for society to prepare for increasingly sophisticated threats to the cyber systems that our critical infrastructure and democratic institutions rely on. Doing so will require substantial investments in cyber security tools, services and skills, including those necessary to address the quantum threat.

At the same time, cyber security is not only a means of protection but also an important source of innovation that will help ensure competitiveness. There are calls for governments to focus efforts on supporting emerging areas of local, regional or national excellence. In Canada, these areas clearly include quantum computing.

Addressing the Quantum Threat

Canada must respond proactively to the quantum threat and implement the elements that will enable an orderly and timely transition to cryptography that is designed to resist quantum attacks (i.e., “quantum-safe” cryptography). Otherwise, our security and economic prosperity will be jeopardized as government and other critical infrastructure systems become vulnerable to hostile actions because of weak cryptography.
The most common forms of cryptography — those used in widely deployed “public-key infrastructure” (PKI) — happen to be based on mathematical problems that are the most vulnerable to ready solution by a full quantum computer. This is a source of great concern, as PKI applications have universal importance by providing assurances such as key agreement (so that only the intended parties have access to a specific communication or transaction) and authentication (so that each party to a transaction knows that the other parties are who they say they are and that messages are legitimate). Without such assurances, there will be no trust and few transactions online, whether they involve humans or the devices that make up the Internet of Things.

The challenge is that a replacement suite of mature, tested quantum-safe cryptographic algorithms are not yet available. Nor are the tools based on them. Nor are the cyber security experts with quantum-safe skills who will use the tools to diagnose and fix each system separately. Without a strong impetus to focus efforts on a long-term campaign to meet the quantum threat, Canada will lose ground as vulnerabilities are exploited and the potential for global leadership is undermined.

Quantum-safe Solutions

An effective response to the quantum threat will necessarily involve a range of stakeholders working together to identify opportunities to translate cutting-edge research into innovative quantum-safe products. An infusion of targeted financial support for infrastructure and personnel is needed to accelerate work on the discovery, testing and deployment of quantum-safe solutions in two areas: post-quantum cryptography and quantum key distribution.

Post-Quantum Cryptography
Quantum readiness demands that new quantum-safe algorithms and cryptographic tools be discovered and developed to replace those now in place. In 2016, the US National Institute for Standards and Technology (NIST) began a multi-year project to identify a standardized suite of viable quantum-resistant cryptographic systems by 2024. The announcement of NIST standards for post-quantum cryptography is expected to result in a retooling of the information and communications technology infrastructure worldwide.

Canadian researchers are active in the NIST effort and have contributed a number of the systems now under consideration. It will be to Canada’s long-term economic advantage if its researchers participate centrally at every stage of the NIST process and beyond, so their efforts should be encouraged and supported. Canada’s researchers and technologists are also at the forefront in developing software and services for post-quantum cryptography, including open-source software, commercial software and professional services. In response to advances in quantum computing, researchers will need to continue their work as successive generations of increasingly efficient and effective quantum-safe cryptography are deployed.

Quantum Key Distribution
The goal in quantum key distribution (QKD) initiatives is a scalable, tamper-proof tool for...
The important key-agreement mechanisms that protect digital transactions. The properties of quantum physics enable two parties to exchange signals that cannot be viewed, copied or tampered with by any third party without being detected immediately. This fundamental ability to detect an eavesdropper can be leveraged to achieve key agreement through untrusted communication channels. Since QKD does not rely on assumptions about the computational difficulty of mathematical problems, the keys cannot be mathematically cryptanalyzed (i.e., broken). This eliminates the risk of an unexpected mathematical advance leading to the systemic compromise of critical infrastructures, or the decryption of past messages that were protected with quantum-vulnerable keys. Research and development related to practical QKD requires substantial investment in essential physical components — such as satellites and ground stations — as well as software, related applications and skilled personnel.

There is a clear need for QKD to be integrated into a real-world network in three to five years. This would enable the testing of QKD with a national satellite-based network linking individual collaboration centres. Preliminary work is already under way at universities across Canada. Not only are some of the critical physical elements in place, but leading researchers have also already coalesced and can mobilize quickly.

These researchers will continue innovating to make QKD more effective and less expensive. Fully reaping the benefits for Canada and Canadians requires additional targeted financial investments to accelerate this work and integrate it into a broader effort to address the impending threat. This would likely first entail the completion of several collaboration centres on separate networks in cities across Canada, the most likely being:

- Calgary (near energy sector, to be enhanced);
- Waterloo/Toronto (near financial sector and government, to be developed);
- Ottawa (near government, to be completed); and
- Montreal (for example, tied to aerospace or the artificial intelligence sector, to be developed).

The separate networks would subsequently be integrated into a single functioning Canadian QKD network, which may eventually be linked into a global QKD network.

### Expanding the Quantum-safe Skills Base

The National Cyber Security Strategy recognizes the need to expand Canada’s capacity to undertake the requisite research and commercialization activities. Significant steps must be taken to strengthen Canada’s skills base, without which the desired facets of cyber security — protection and economic development — cannot be achieved.

Programs and courses offering professional training will need to be established if Canada is to have the necessary cadre of cyber security experts with superior quantum-safe skills. These experts would perform tasks such as cyber risk assessment and systems integration to ensure that the appropriate quantum-safe solutions have been properly installed and integrated into complex legacy systems.

Development of a large pool of systems integrators and cyber security professionals with strong quantum-safe skills will take several years. A number of Canadian post-secondary institutions have indicated interest in augmenting their cyber security programs.
with courses focusing on the migration to post-quantum cryptography. Ideally, they will collaborate on a standard quantum-safe module for incorporation into existing cyber security programs.

In addition, possibilities around outreach to industry should be explored. There is likely to be an appetite for training courses to familiarize technical staff with quantum-safe technologies and how best to work with external quantum-safe experts. There will also be a need for certification schemes to allow the quality of the training and the expertise of the trainees to be evaluated on an ongoing basis.

While education is a provincial responsibility, there is a need for the federal government to play strategic and funding roles to ensure that the provinces and territories, and the agencies and regulatory bodies they control, move with a sense of urgency.

Using Government Policy Levers

Governments have access to numerous policy powers that may be useful in encouraging and even ensuring that digitally enabled infrastructure — such as smart roads, smart bridges and smart cities — is designed, built and installed to be quantum-safe. These levers include approval, planning, procurement and funding powers, none of which need to be costly.

A simple example would be a federal policy that any proposal for federal support for an infrastructure project must be accompanied by a cyber security strategy. This would necessarily include a quantum-safe strategy for infrastructure expected to be in service for decades.

Taking Advantage of Opportunities for Canadian Leadership

As noted above, the National Cyber Security Strategy stresses the need to prepare for increasingly sophisticated threats to Canada’s cyber systems. At the same time, it points out that cyber security is not just a means of protection but also an important source of innovation that will help ensure Canada’s competitiveness. Both sides of the coin are in play when it comes to the quantum threat.

Working in our favour is the fact that Canada is in the vanguard globally in both cryptography and quantum information science, and strong in cyber security applications and services. There is a significant history of collaboration among these realms, so Canada should be able to get its house in order ahead of other countries and then export its quantum-safe products and expertise abroad. Taking advantage of this opportunity would enhance both Canada’s national security and its economic prospects.

Implementation of the key elements discussed above will enable Canada to take advantage of the opportunities for innovation, prosperity and competitiveness that are inherent in moving quickly to address the quantum threat. A number of complementary actions should also be taken in support of the core elements:

- Name an advisory committee of top scientists in cryptography and cyber security to provide expert advice on research priorities and parameters for projects and proposals.
- Identify the technical expertise needed to monitor relevant international standards development work and participate as necessary.
- Identify the program management expertise needed to advance innovation and commercialization activities, the market research exercises needed to quantify the national and global requirements for quantum-safe expertise, and the necessary export-development initiatives related to quantum-safe technology, expertise and training.

Without a strong impetus to focus efforts on a long-term campaign to meet the quantum threat, Canada will lose ground as vulnerabilities are exploited and the potential for global leadership is undermined. We cannot afford to be a follower, facing massive security vulnerabilities and prohibitive upgrading costs simply because we delayed taking action. At the same time, we should not be blind to the economic benefits of vibrant cyber security and quantum-safe industries, or to the danger that we will lose our current edge if we delay action.

A version of this essay was first delivered to the Standing Committee on Public Safety and National Security on February 22, 2019.

Canadais in the vanguard globally in both cryptography and quantum information science, and strong in cyber security applications and services.

Work Cited


Endnote

1 PKI is a system that binds “public keys” to various entities. These public keys are numbers (representing, large integers, or points satisfying the mathematical equations of an “elliptic curve,” for example) that are publicly available. For each public key, the respective entity retains a “private key” that is kept secret and should be infeasible to compute by someone who only possesses the public key. Two entities, each knowing only its own private key and the other’s public key, can exchange non-confidential random numbers (which depend on their private keys) and derive a shared secret key. One entity can also use its private key to digitally sign a message such that any third party in possession of the entity’s public key can validate the origin and integrity of the message.

The Quantum Threat to Cyber Security
Mitigating Cyber Risk across the Financial Sector

Christian Leuprecht

Most critical functions of twenty-first-century society have become inextricably dependent on digital infrastructure, in particular the financial industry, whose business model relies on consumer confidence in the overall financial system. The internet is now the primary mechanism for financial transfers between banks and other institutions; most customers rely on online banking to manage their accounts and for the majority of point of sale payments. In fact, Canada ranks among the most cashless societies in the world (ForexBonuses 2017). The more reliant on digital technology the financial system becomes, the more interconnected it is and the more vulnerable it is to cyber exploitation. Consumers notoriously prefer convenience over security, and financial institutions encourage consumers to use online technology as a way of harnessing efficiencies and reducing operating costs. Malicious actors are not targeting the industry for mere financial gain: because the financial industry is systemically significant, adversaries are actively looking to exploit vulnerabilities that could be used to bring it down, thereby undermining confidence in the financial system and causing social chaos and turmoil to threaten the democratic way of life. The financial industry’s dense interconnectivities, broad digital footprint with consumers and extensive reliance on technological infrastructure expose it to a disproportionately large attack surface. Governance at both the national and the international level has not kept up.

The Threat Landscape

Canada’s financial sector is an appealing target for profit-motivated cybercriminals: it is subject to millions of infiltration attempts each day, compounded by cyber-enabled crime such as credit card fraud. The financial industry experiences greater losses from cybercrime than any other sector, reportedly experiencing attacks three times as often as other industries (Raytheon Company 2015, 3). A recent report from the International Monetary Fund (IMF) estimated that banks’ average annual potential losses from cybercrime could amount to nine percent of their net income, equivalent to US$97 billion (Bouveret 2018, 21).

Cybercriminals attempt to steal credentials and obtain information such as the passwords and personal information of bank staff and customers, allowing them to access accounts.
and place fraudulent payment orders. Phishing is a low-risk, low-cost instrument for even the least-skilled cybercriminals. Distributed denial of service attacks can disable financial services, preventing customers from accessing accounts and payments from being processed. The reams of sensitive customer data held by financial institutions contain a motherlode of high-value personal information. The consequences of large-scale data breaches, such as the 2017 theft of the financial records held by Equifax of more than 140 million people, undermine the mutual trust and confidence on which the financial system relies (Fleishman 2018). Although difficult to quantify, the cost of this shaken faith means the true burden of cyber heists extends beyond mere monetary losses.

Hackers working at the behest of states are now a serious cyber threat to the financial sector. Backed by the resources of state governments, they have the ability to cause significant disruption to the financial system. North Korea maintains dedicated teams focused on cyber operations against financial institutions. The attempted theft of more than one billion dollars through the Society for Worldwide Interbank Financial Telecommunication (SWIFT) network, including brazen attempts on central banks, has been attributed to teams such as the “Lazarus Group,” whose infamous “WannaCry” ransomware attacks have resulted in damages estimated at up to US$4 billion (Symantec 2016; Berr 2017; FireEye 2018).

Previous cyber operations against the financial sector were mainly carried out by financially motivated criminals. Their schemes aimed for quick profit before escaping, emphasizing speed and seeking to cause minimal collateral damage so as not to draw the attention of law enforcement. The growing cyber capabilities of state and non-state actors, however, are primarily driven by geopolitical goals (Leuprecht, Szeman and Skillicorn 2019). That scenario raises the prospect of genuine cyberattacks — defined as meeting the threshold of the use of force under international law — on the financial sector to wreak havoc and provoke instability as an end in itself (Healey et al. 2018).

Developed states have a mutual stake in upholding a functioning global financial system, but actors such as terrorist organizations and isolated rogue states may feel that they stand to gain from financial instability by holding developed countries ransom. North Korea-backed hackers operate with the aim of generating revenue for the regime in Pyongyang. By contrast, the so-called “DarkSeoul” attacks of 2013 followed joint military exercises between South Korea and the United States, targeting South Korean banks and television networks and paralyzing victims by disabling their computer systems (BBC News 2013). The accompanying bellicosity from Pyongyang (threatening pre-emptive nuclear strikes), and the fact that television networks were targeted alongside banks, indicates that the financial system was targeted as a means to a geopolitical end.

The financial system is fragile, resting upon a foundation of mutual trust and confidence. Modern history has plenty of examples of prolonged economic malaise provoked by a negative shock that caused confidence to evaporate, sending the economy into a downward spiral. It is not difficult to imagine that this sort of shock could be deliberately induced by an adversary or hostile actor.

In 2013, the Twitter feed of the Associated Press was hacked, reporting that an explosion at the White House had injured President Barack Obama. The ruse was quickly exposed, but the momentary shock provoked panic in the financial sector, causing the Standard & Poor’s 500 Index to drop 0.9 percent (equivalent to US$130 billion) (Matthews 2013). These losses were quickly recuperated, but the incident demonstrates that actors in cyberspace can intentionally undermine the stability of the financial system. Simple methods of exploitation could have far-reaching consequences.

**Structural Vulnerabilities of the Financial Sector**

The global scale, complex interconnectivity and systemic significance of the financial industry pose a unique cyber security challenge. Large multinational financial institutions tend to house their data across different countries, rendering them vulnerable to compromise in transit and at rest in jurisdictions with lax security standards. Banks are now often encouraged by host governments to keep customer and transaction data stored within the host country’s borders through measures such as data localization laws, and some institutions...
The financial industry experiences greater losses from cybercrime than any other sector, reportedly experiencing attacks three times as often as other industries.

Generating revenue for the regime in Pyongyang is usually the aim of North Korea-backed hackers; however, the 2013 DarkSeoul attacks demonstrated that these same capabilities can be deployed to achieve geopolitical goals.

(Photo: LMsparencer / Shutterstock.com)

have already made data localization part of their business model. This can be difficult, however, as operations in the financial sector span the globe and it may not be clear where a given customer’s data should be stored or how to control the path taken by the data (Leuprecht, Skillicorn and Cockfield 2019).

Global interconnectivity raises the threat of “contagion” in the wake of a cyber operation. The most recent financial crisis shows how losses can cascade. This is true for losses incurred in the course of doing business and for losses caused by cyber intrusions. The SWIFT interbank communication system reaches banks in almost every country on the planet. Circumventing the national borders of the physical world, the SWIFT network can act as a vector for cyber operations. Banks in developed states with relatively robust security precautions are exposed to hackers in jurisdictions where security regulations and enforcement are less stringent (ibid.). In 2016, cybercriminals (possibly the Lazarus Group) acting through the SWIFT network convinced the Federal Reserve Bank of New York to transfer US$81 million from the Central Bank of Bangladesh’s account to recipient accounts in the Philippines (Corkery and Goldstein 2017). Contagion is also the result of the virulent nature of cybercriminals’ tools. In 2017, the WannaCry ransomware spread to hundreds of thousands of computers in a matter of days (Jones and Bradshaw 2017). The structure of the global financial system means that a single compromised node can have disproportionate consequences for the integrity of the network as a whole.

Notwithstanding the densely interwoven structure of the financial system, essential functions such as trade matching and custody of securities are concentrated in select hubs. These activities are also highly dependent on information and communications technology infrastructure, such as cloud computing services, which have the potential to be infiltrated or disabled by cyberattacks. These “single points of failure” can grind the whole system to a halt (Healey et al. 2018). In many instances, there are no clear alternatives or workarounds that financial actors could use in the event of a crisis.

Moral Hazard

A cyber operation’s likelihood of success can be affected by the security efforts of the targeted institution as well as by the digital hygiene followed by users and customers. The typical end-user of an online chequing account prefers convenience over security. Asking end-users to cover their own losses in the event of a heist seems intuitively unfair. Even if they were to adjust their behaviour by adopting measures such as dual sign-in authentication and not using wireless networks, they would remain vulnerable if their financial institution did not follow suit, and they have little power to force it to do so. As a result, Canadian banks currently bear the costs of consumer losses, as long as the victim was not negligent (Leuprecht, Skillicorn and Cockfield 2019). However, leaving banks to cover end-user losses in this way gives rise to a moral hazard:
since they are assured that they will not be out of pocket in the event of a heist, end-users have little incentive to follow better security protocols. This leaves banks holding the bag, which exposes them to perverse incentives for greater cyber exploitation.

**Policy Approaches**

Faced with persistent and sophisticated actors launching increasingly ambitious and sophisticated attacks on financial institutions, governments must signal a willingness to punish and deter offensive action. If hostile actors are enjoying the backing of states, it is in the interest of Canada and its allies to project power and stability in cyberspace. Governments will need to commit to deterrence through punishment in the case of a debilitating attack against critical infrastructure. Hostile actors need to be put on notice that even attacks that do not necessarily meet the threshold of the use of force under international law or the North Atlantic Treaty Organization's Article 5 may meet with reprisal. Bill C-59, Canada’s new national security bill, proposes to grant Canada’s signals intelligence agency, the Communications Security Establishment, the ability to conduct “active cyber operations” aimed at disrupting and disabling hostile actors. Canada and other friendly governments should develop policies to pursue guilty parties within the boundaries of international law, much of which does not apply in cyberspace, where operations largely fall below the threshold of the use of force. This necessitates enhanced international cooperation to enable extraterritorial investigation and prosecution. Mutual legal assistance treaties facilitate the sharing of information in attribution and prosecution. Greater track two and track 1.5 diplomacy, such as the United Nations’ Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, might eventually encourage more countries to sign on to the Budapest Convention on Cybercrime (Eoyang et al. 2018). Coordinated diplomatic pressure, backed up by a credible threat of sanctions or other punishment, will be needed to ensure compliance by rogue states (ibid).

Due to the interconnectedness between financial institutions and the risk of contagion, improving cyber security in the financial sector will require strengthening its weakest links. The inability of small and medium-sized financial institutions to properly take advantage of the same security measures as the major banks is one such blind spot. Complemented by financial intelligence networks, the Canadian Cyber Threat Exchange provides cyber domain awareness to bolster the defences of the major banks. However, many small and medium-sized enterprises have not had the same access to such intelligence (Leuprecht, Skillicorn and Cockfield 2019). In fact, smaller financial institutions have incurred disproportionately large losses from cyber heists, which bear equally disproportionate existential risks, suggesting that “economies of scale” are at work in cyber security (Bouveret 2018). Although any one such institution may appear systematically inconsequential, the interconnectedness of the financial sector means that small actors actually present a systemic risk. Policies will need to be amended to bolster the defences of the industry’s smaller institutions and enable them to benefit from timely threat intelligence.

Governments could also do more to protect the technological infrastructure upon which the financial industry is dependent. The Canadian government’s role and obligation in rebuilding critical infrastructure if it were disabled by a cyber attack is unclear. Citizens expect government to respond to naturogenic or anthropogenic disasters, and government should anticipate the possibility of having to similarly respond to a catastrophic failure of critical infrastructure in the event of a crisis as a way of mitigating the danger inherent in these “single points of failure.”

The typical end-user of financial services such as an online chequing account prefers convenience over security. (Photo: Shutterstock.com)
As detailed above, the distribution between banks and customers of the costs incurred by successful cyber attacks is problematic. Placing the burden on customers when they have little power to affect their banks’ security efforts may be unfair, but making the banks responsible for covering consumer losses raises the problem of moral hazard. Both the banks and their customers would benefit from a more mature cyber security insurance sector as a way to monetize risky behaviour by firms and individuals and incentivize good behaviour. Due to the novelty of cyber risk, cyber security insurance remains a fledgling industry that needs government attention. It will need detailed data on cyber exploits to properly quantify risk. Yet, banks currently have little incentive to share the frequency with which they are attacked, as that may have a negative impact on a firm’s reputation. Since February of this year, Canada’s prudential regulator, the Office of the Superintendent of Financial Institutions, has required federally regulated banks and insurers to report technology and cyber security incidents, although more robust requirements for the disclosure of breaches of the sort found in the European Union’s General Data Protection Regulation would be even more beneficial (Middleton 2018). These efforts will need to be complemented by coordination at the international level to confront the transnational nature of cyber threats by promoting common standards and information sharing. The Group of Seven (G7) has begun the process of harmonizing cyber security standards for financial institutions, formulating the “G7 Fundamental Elements of Cybersecurity in the Financial Sector” (G7 2016), the Group of Twenty, through the Financial Stability Board (FSB), has likewise started to consider the risk that cyber operations pose to financial stability and has made attempts at developing a common lexicon to ensure consistent classification and reporting of cyber breaches (FSB 2018). Ultimately, the global community has a collective interest in defending the integrity of the international financial system. In an interconnected world, robust common regulatory standards are essential to this effort.

Works Cited


The Danger of Critical Infrastructure Interdependency

Tyson Macaulay

Those of us who live in the Western world have had the tremendous privilege of being able to take functioning critical infrastructure for granted. Clean water, reliable roads, high-quality health care, dependable electricity, telephones and email are all so fundamental to modern existence that it is impossible to picture life without them. The advent of the internet has made critical infrastructure more complex, more interdependent and, therefore, more fragile. We have become complacent in our reliance on critical infrastructure, but recent developments have been a rude awakening. It is now clear that cyberwarfare can have an impact on the physical world through attacks on critical infrastructure. The often dimly understood interdependencies between critical infrastructure sectors pose a grave risk. A blow to one critical infrastructure sector could cause cascading second-order effects on other sectors, leading to a large-scale catastrophe that spirals out of control. This essay will discuss a method to assess the risks to critical infrastructure that result from interdependencies related to data flows and will examine how gaps in security can have impacts in other critical infrastructure sectors.

Cyber Attacks on Critical Infrastructure

Critical infrastructure consists of the systems that have been deemed fundamental to the functioning of a society and an economy, such as energy, transportation, telecoms, the provision of food and water, and vital health services. The disruption or destruction of critical infrastructure would have an immediate and direct impact on the economic activity, day-to-day life and safety of those affected.

The advancement of cyberweapons and hackers’ tool kits now permit malicious actors to attack critical infrastructure in ways that have an immediate and frightening effect on the physical world. In 2015, Ukraine was the subject of a shocking cyber attack that managed to disable a portion of the nation’s electrical grid (Greenberg 2017). The attack, widely believed to have been carried out by Russia, intentionally caused widespread blackouts for hundreds of thousands of people. Although the attack and similar incidents directed at Ukraine in the years that followed were only temporary disruptions, they provide ample evidence of the scale of damage that cyber attacks could inflict on critical infrastructure. Hospitals had to return to using pens and paper during the attack, jeopardizing the delivery of services to those who urgently needed them (Borys 2017).

Digital technology has made the world smaller, and critical infrastructure in Western nations is not safe from this new danger. The US government has denounced Russia Tyson Macaulay is a veteran of the cyber security industry with over 25 years of technical and international management experience. As chief product officer at InfoSec Global, he sets direction and strategy for all products, including full accountability for engineering and services, research and development, marketing, and customer satisfaction. Tyson has been a security researcher and lecturer since the beginning of his career in 1992, with a personal syllabus of four books, dozens of periodical publications and international standards contributions, and two registered patents. He continues to support the development of engineering and security standards through volunteer efforts in the International Standards Organization and Professional Engineers Ontario, and various board positions.
for infiltrating the country’s power grid and gaining remote access to energy sector computer networks (US Department of Homeland Security 2018). In 2017, the WannaCry ransomware epidemic disabled Britain’s National Health Service for several days, leading to the cancellation of 19,000 appointments (Field 2018). In Denmark, the headquarters of Maersk, responsible for around one-fifth of the world’s shipping, was brought to a standstill by the NotPetya malware, causing transportation disruptions at port facilities worldwide (Matthews 2017). Sophisticated actors can insinuate themselves into vital control systems and remain dormant and undetected for long periods of time before the right moment to strike presents itself.

As a consequence of these events, cyber attacks on critical infrastructure have become a pre-eminent concern for national security. This year’s Worldwide Threat Assessment by the US intelligence community emphasized the ability of nation-state adversaries to launch successful cyber attacks against critical infrastructure in the United States (Coats 2019). More starkly, the report of the US National Defense Strategy Commission describes how the United States is unprepared for cyber attacks on its critical infrastructure, and how this may seriously threaten the country’s military supremacy (National Defense Strategy Commission 2018).

What makes the potential consequences of a major cyber attack on critical infrastructure difficult to predict is the interdependence between various sectors. An attack on one sector could have spillover effects on the other sectors that depend on it. The transportation sector depends on the provision of electricity by the energy sector to power trains and traffic control systems, just as the energy sector relies on the timely delivery of fuel and other inputs through the transportation sector. With respect to the threat of remote infiltration, a working group of industry experts and government officials at MIT’s Internet Policy Research Initiative warned that no one currently understands the extent to which electricity generation is coupled with other sectors, and therefore the risk of “catastrophic macroeconomic failure” in the event of a cyber attack is not adequately known (Brenner 2017, 28). Arriving at a proper comprehension of the interdependency between critical infrastructure sectors is vital in order to fully appreciate the inherent risks.

**Data Interdependency Assessment**

The continued functioning of critical infrastructure is highly dependent on communication, which underpins everything from the logistics of order fulfillment to the financial transfers of funds and sharing of intellectual property. As a result, addressing threats to data exchanges between critical infrastructure sectors is important to their protection. Understanding the flows of data is a useful tool for identifying hidden risks. This essay reviews a method for assessing risks driven by data-dependency relationships between different industries designated as critical infrastructure, and examines how these relationships form interdependencies within critical infrastructure sectors.

There are two basic varieties of threats to data flows: threats to availability affect whether the
data can be accessed when it is needed; and threats to confidentiality and integrity concern data being disclosed or changed without the reliant party's knowledge or approval.

Data dependency is a measure of how sensitive a critical infrastructure sector is to the availability, integrity and confidentiality of data flowing between the sectors. More specifically, "dependency" reflects the one-way data-security requirements of one critical infrastructure sector on another. "Interdependency" refers to the bidirectional system of data and information being shared between critical infrastructure sectors (Macaulay 2008).

**Defining Dependency in Critical Infrastructure**

The metrics and analysis presented here are drawn from earlier work, in which the survey and data collection methodology are documented (ibid.). In sum, more than 100 security and communications executives from all critical infrastructure sectors were asked detailed questions about the sensitivity of information they send and receive from all other critical infrastructure sectors, in order to quantify "inbound" versus "outbound" data dependency.

Inbound data dependency is about information and data being delivered to, and consumed by, a critical infrastructure organization. Information and data arrive in the form of voice calls, internet-based business systems and services, and even social media and other employee activities. Inbound dependency, therefore, involves the cyber security properties of information needed by critical infrastructure organizations to continue the production of goods or services. For instance, how long can a water treatment plant continue to operate safely without information from testing laboratories in the health sector? Inbound data dependency is related to the vulnerabilities of a sector that are caused by interdependency.

Outbound data dependency is about information from a given critical infrastructure sector that is sent to other critical infrastructure sectors. Websites are information assets established in part to address outbound data on a self-serve basis. Outbound dependency concerns the security requirements that other, consuming critical infrastructure sectors place on the suppliers of information. To return to the example of the water treatment plant, outbound dependency is concerned with how long the health sector can safely operate without information from the water treatment plant. Outbound dependency concerns the threat that a given critical infrastructure can pose to other critical infrastructure sectors due to interdependency.

**Dependency Matrices**

A dependency matrix is a means of visualizing the cyber risks associated with critical infrastructure interdependency. The dependency matrix reveals the potential vulnerability of a given critical infrastructure to threats from other critical infrastructure sectors. The table below is an example of a dependency matrix. Both inbound and outbound dependencies are presented through this single tool. Together, inbound and outbound dependence equal "interdependence."

<table>
<thead>
<tr>
<th>Critical Infrastructure sector</th>
<th>Energy</th>
<th>Communications &amp; IT</th>
<th>Finance</th>
<th>Health Care</th>
<th>Food</th>
<th>Water</th>
<th>Transportation</th>
<th>Safety</th>
<th>Government</th>
<th>Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>9.37</td>
<td>8.62</td>
<td>5.01</td>
<td>5.42</td>
<td>4.12</td>
<td>4.01</td>
<td>3.33</td>
<td>2.43</td>
<td>2.17</td>
<td>1.96</td>
</tr>
<tr>
<td>Communications &amp; IT</td>
<td>6.96</td>
<td>8.02</td>
<td>6.11</td>
<td>5.11</td>
<td>4.23</td>
<td>7.83</td>
<td>5.01</td>
<td>3.51</td>
<td>3.24</td>
<td>1.42</td>
</tr>
<tr>
<td>Finance</td>
<td>7.13</td>
<td>7.19</td>
<td>8.02</td>
<td>8.23</td>
<td>5.01</td>
<td>6.76</td>
<td>4.02</td>
<td>5.18</td>
<td>7.96</td>
<td>1.75</td>
</tr>
<tr>
<td>Health Care</td>
<td>4.12</td>
<td>2.43</td>
<td>9.99</td>
<td>8.25</td>
<td>1.8</td>
<td>4.43</td>
<td>3.33</td>
<td>5.76</td>
<td>5.06</td>
<td>2.57</td>
</tr>
<tr>
<td>Food</td>
<td>1.47</td>
<td>1.66</td>
<td>1.96</td>
<td>2.76</td>
<td>6.45</td>
<td>1.83</td>
<td>2.48</td>
<td>1.05</td>
<td>2.71</td>
<td>1.99</td>
</tr>
<tr>
<td>Water</td>
<td>4.90</td>
<td>1.84</td>
<td>1.96</td>
<td>3.6</td>
<td>3.76</td>
<td>4.43</td>
<td>3.33</td>
<td>5.78</td>
<td>5.06</td>
<td>2.57</td>
</tr>
<tr>
<td>Transportation</td>
<td>6.82</td>
<td>3.95</td>
<td>4.23</td>
<td>4.95</td>
<td>5.01</td>
<td>2.96</td>
<td>7.49</td>
<td>3.78</td>
<td>4.66</td>
<td>5.84</td>
</tr>
<tr>
<td>Safety</td>
<td>7.85</td>
<td>3.96</td>
<td>3.6</td>
<td>5.71</td>
<td>1.02</td>
<td>4.54</td>
<td>3.25</td>
<td>8.23</td>
<td>5.73</td>
<td>4.96</td>
</tr>
<tr>
<td>Government</td>
<td>5.85</td>
<td>3.03</td>
<td>7</td>
<td>6.12</td>
<td>4.76</td>
<td>5.05</td>
<td>7.61</td>
<td>6.43</td>
<td>8.78</td>
<td>5.96</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>3.87</td>
<td>3.75</td>
<td>4.66</td>
<td>3.01</td>
<td>4.5</td>
<td>3.43</td>
<td>4.53</td>
<td>1.17</td>
<td>3.63</td>
<td>7.15</td>
</tr>
</tbody>
</table>

Source: Author.
The columns for each critical infrastructure sector represent how a sector self-rates (according to interviews conducted with stakeholders) (ibid.) its dependency on information coming from other critical infrastructure sectors — the inbound dependency. Most organizations will intuitively understand their vulnerabilities regarding the information they consume. The rows represent how dependent information-receiving critical infrastructure sectors are on information and data from a given critical infrastructure sector, according to their own assessments — the outbound dependency for each sector. Unlike inbound dependency, most organizations do not have a great deal of insight into how all other critical infrastructure sectors actually need the information and data they produce.

In total, more than 4,000 distinct data dependency metrics were gathered from critical infrastructure stakeholders; dependency was ranked on a scale from 1 to 10. The higher the number in the column, the greater the dependency (vulnerability) on data flowing into a sector; the higher the number in a row, the greater the dependency (threat) of others on data flowing out of the sector. By mapping out critical infrastructure interdependencies in this way, we can begin to understand — and take precautions against — the sort of cascading effects that might follow a major cyber attack on a critical infrastructure sector.

**Energy Sector Data Dependency**

![Energy Sector Data Dependency Diagram]

**Source:** Author.

---

**Sector-specific Dependency Analysis: Energy Sector**

The energy sector is primarily concerned with electric power generation and transmission, as well as oil and gas production and storage. Energy is often considered a “super critical” infrastructure because most other critical infrastructure sectors cannot operate if energy is not functioning. The following tornado diagram (see figure below) is used to illustrate the energy sector’s inbound and outbound data dependencies and the resulting cyber vulnerabilities and threats. The diagram is divided vertically by an axis that is valued at zero. The left side displays the median inbound data dependency values for the energy sector, in descending order from highest to lowest. The right side displays the outbound data dependency values for each of the sectors (that is, how do other critical infrastructure sectors depend on the data from the energy sector?).

Intra-sector data dependencies are typically strongest among critical infrastructure sectors. Energy is the largest consumer of its own information and data as a result of the tight supply chain linkages between different organizations, for instance, in production versus distribution of both electricity and fossil energies.

As shown on the left side of the diagram, the energy sector is a large consumer of data from other sectors and expresses the highest dependency.
overall inbound dependency and, therefore, cyber vulnerability to critical infrastructure interdependency. Energy’s inbound data and cyber requirements are heightened relative to other sectors by the high standards around energy supplies mandated by clients and government. Loss of data flows from within the energy sector itself is a serious vulnerability that might be exploited by a variety of different threats, such as unauthorized access to control systems, or an impact on energy generation when critical data flows required for plant operation are rendered unavailable.

Information and data from the safety sector (which includes law enforcement and first responders) is the second-highest inbound dependency for energy. The inability to receive security reports would significantly impair the energy sector’s ability to prepare and respond to incidents that might impact operations. This may also suggest that unauthorized or spoofed data purporting to be from the safety sector could represent a potential pathway for cyber attacks.

Every sector except for the energy sector itself consistently rates the information and data from the energy sector (outbound information) as a lower priority than the energy sector rates the reciprocal information (inbound information). Note that this does not indicate that other sectors think of the supply of energy itself as a low priority. Rather, energy supplies are somewhat taken for granted, with backup generators able to cover the brief power outages other sectors are accustomed to. They are likely unprepared to deal with the sort of extended outages a major cyber attack could potentially cause. This creates a vulnerability in other sectors that place a reduced emphasis on the data from the energy sector, as well as an opportunity to compromise data flows at critical times. A low outbound score also represents a threat to the energy sector itself: it may be overestimating other sectors’ responsiveness and collaboration in the event of a crisis.

Conclusions and Policy Considerations
It is obvious that cyber attacks on critical infrastructure are capable of inflicting real-world damage. The frequency and severity of such incidents will likely only increase. Western countries are beginning to formulate strategies and policy responses to meet this challenge. The United States has moved toward a more aggressive posture to defend its critical infrastructure systems against cyber attacks. The most recent Department of Defense Cyber Strategy outlines a “defend forward” policy for addressing cyber threats to US critical infrastructure, possibly including pre-emptive action (Department of Defense 2018). The US government’s most recent cyber strategy details a growing emphasis on offensive cyber operations by certain branches of the US government, and the United States has not ruled out responding to major cyber attacks on critical infrastructure through conventional forces (President of the United States of America 2018). Other governments have been more muted regarding threats of retaliation; however, all are taking seriously the cyber threat toward critical infrastructure.

Governments have an urgent need to achieve a clearer understanding of the often-opaque interdependencies between critical infrastructure sectors, and to take steps to mitigate the chances of cascading chain reactions. Industry experts have suggested that governments could have a role to play in coordinating stress test exercises and simulations between critical infrastructure sectors, which could illuminate present weaknesses and build resilience (Brenner 2017).

Critical infrastructure interdependencies should also be considered at the international level. Canadian and US energy grids are intertwined, such that a cyber attack that disables electricity supply in the United States could have second-order effects for Canadian critical infrastructure. The example of Maersk’s encounter with NotPetya shows how many critical infrastructure functions are provided by multinational corporations, which could lead to worldwide disruptions if they fell victim to a cyber attack. The degree of critical infrastructure interdependence across national borders suggests a need for greater degrees of policy coordination and information sharing on a global level.

Works Cited
History provides evidence that macro-level shifts in how we organize our economies has at times required radical shifts in how we govern our people. Agriculture destroyed the chief, forcing a transition to oligarchy. Industrialization destroyed the king, forcing a transition to democracy. The information revolution, coupled with globalization, is eroding representative government and forcing a transition toward something completely new.

At the centre of this shift is a new technology that has the potential to reinforce our existing political scaffolding with tamper-proof, censorship-resistant, incorruptible “programs.” This technology, which builds on 20 years of research into cryptographic currency, and 40 years of research in cryptography, is the blockchain. Today — at this moment — developers across the globe are using blockchain-based smart-contract platforms to build the person-less institutions of the future.

Why Is This Happening?

The economic underpinnings of any society are the means and mechanisms of production and the distribution and allocation of goods, services and resources.

Every so often something new comes along that profoundly changes our economic foundations, forcing us to adapt how we govern ourselves in order to cope with these changes. With a historical lens it is apparent that these shifts can span decades, and sometimes centuries, but occasionally a change is big enough that it is labelled a “revolution.” The first agricultural revolution (the Neolithic Revolution) started in 10,000 BC and was the result of novel discoveries in the selective breeding of crops and farming, the domestication of animals as a means of labour, population consolidation and, as a result, population growth on an unprecedented scale.

A second agricultural revolution in Britain in the mid-seventeenth century saw transformational economic impacts through improved plough technology, land ownership rights, infrastructure (in the form of roads and canals) and a national tariff-free market system. This shift in the means and mechanisms of production resulted in massive changes in population density. These changes forced society to adopt centralized administrations, hierarchical ideologies, depersonalized systems of knowledge such as writing, divisions of labour, non-portable architecture and art, property ownership and more. These constructs may seem common to us today, but at the time they were intellectual innovations with competing ideologies that took generations to spread and become mainstream. Perhaps the most important of the changes caused by the Neolithic Revolution was how people were governed. There was a shift from tribal leadership to oligarchs. Humans would experiment with many more forms of government up until the 1700s; however, the dominant form of governance at the turn of the eighteenth century was the feudalist model with the king or queen as its focal point.
In the middle of the nineteenth century, a second period of change began when machines replaced hand production, chemical and iron production processes were invented, steam power and factory systems became available and accessible port systems for trade were built. This shift, the Industrial Revolution, brought increased social mobility and a boom not only in production but also in population. Among many new phenomena, including “the crisis of the family,” in which the pre-existing institution of the working family collapsed, was a greater shift in societal power structures.

Industrialist business owners, with their new fortunes, became the recipients in a power transition from the nobility class (land owners) to the business class (industrialist entrepreneurs) as the middle class grew. Feudalism was being eroded by the new capitalists and, as a result, elected “representative” government was put in place to govern the people and keep the industrialist business owners content. Closer lines of accountability were drawn between the government and the people.

More than 150 years later, this style of democratically elected representative government is a staple of the free developed world. However, as the twentieth century came to a close, several distinct innovations have put humanity on a new economic growth trajectory: the personal computer and then the handheld computer (smartphone); the internet as the connective tissue and network that every personal computer connects to; and software applications and their ability to capture and perform our routinized behaviours. And finally, e-commerce, the first truly global, borderless competitive arena for the distribution and allocation of goods, services and resources.

History has shown us again and again that if the economic forces that govern our production and consumption are altered by a set of technological and intellectual innovations to a substantial degree, new social governance methods must be adopted to cope with the change. In the modern age, the information revolution and globalization are radically altering our economic foundation and putting national politics under pressure. If we are not careful during this transitional period, we run the risk of producing one of two equally unattractive outcomes: the collapse of government as we know it or the emergence of techno-totalitarian nation-states.

National Politics Is Not Equipped to Deal with New and Complex Global Problems

Current political systems, and the social evolution that led to their formation, all share common characteristics, resulting in universal challenges seemingly not addressed by any current form of government. Namely, these systems are designed around jurisdiction-based politics, which essentially divide the world into sovereign parcels of land and govern people based primarily on their geographical location.

Historically, this made sense, because commerce occurred and information flowed
within hyper-local geographical areas, long before the world became globalized and the internet was created. On top of this confining legacy feature, our political and bureaucratic systems have, for the most part, become bloated to a point of snail-paced change and evolution, all while the world that they govern, and in which they operate, is changing at unprecedented speed.

In short, the way people are organized and controlled through forms of political governance no longer reflects the boundaries of commerce and communication, and, increasingly, there is evidence that our current frameworks won’t work for a rapidly changing world.

Today, an educated millennial is likely to have travelled to and even lived in multiple countries, have friends distributed around the world and interact heavily within online communities and markets — and the trend is continuing in this direction. On the other end of the spectrum, many of the historical barriers that prevented the inclusion of people living in secluded and poor regions of the world into the global economy have ceased to exist. What remains are artificial boundaries and restrictions to joining the rest of the world in prosperity. Take music as an example of this. In the past, the barrier to making music was production — it involved an expensive studio with equipment and distribution that was monopolized by a few geographically connected record labels. In 2019, you can record music on a phone and distribute it across the globe on YouTube, SoundCloud and Spotify in minutes. Innovation is driving this trend in the private sector; in the public sector, our governing frameworks will also see changes.

As a result of the current boundaries and restrictions, there exists a very rigid and inflexible system that seems difficult to adjust without force, even though there is more and more evidence pointing to its growing ineffectiveness. That said, although we can likely imagine a system of human organization and governance that is designed from a different set of starting assumptions — potentially something that results in a highly connected network of city states and global digital economies — the transition process may seem too insurmountable to be realistic.

In the meantime, efforts should be made to build toward systemic, incremental improvements within our current structures.

Social Governance Is, at Its Core, Powered by Trust

“A leader you can trust” — this is a message you are likely to hear from politicians during an election. People generally choose to vote for a candidate they trust. Our systems function and maintain the confidence of people based
on that trust, which needs to exist within all layers of social governance in order for civilized society to function. Trust in politicians, trust in money, trust in education, trust in the free press, trust in infrastructure — the list goes on.

In its simplest form, social governance consists of the mechanisms employed to organize human behaviour toward commonly accepted positive outcomes. These mechanisms span commerce, education, the environment and social interactions, among other areas. Social institutions were designed to address these domains and have historically been adequate in meeting this requirement because the people they served trusted their competence and importance.

Governments, among other things, collect taxes, provide public goods and spend on economic externalities, redistribute income and provide social security. At its core, this behemoth of a bureaucratic engine is a “trust machine” that functions on the notion that if we collectively pay for its operation, then the institution will allocate funds appropriately. There are 195 of these engines operating with varying levels of success worldwide. Shared global economic externalities such as the ocean, polar ice caps and the biosphere are largely ignored, to our collective detriment. In addition to these challenges, the economic dependence on digital markets and the internet more broadly is increasing around the world; this is a domain that is not easily governed by individual nation-states.

The question is not whether change is needed, rather, what does the solution for a new global governance space look like in a world that is clearly changing? Many of us believe “programmable trust” plays a very important role in this new world. In order to understand it we must understand its inception — bitcoin.

The real innovation in bitcoin was that for the first time, we could engage in commerce on the internet without the need for trusted third parties to process electronic payments. Completely non-reversible transactions were not really possible before bitcoin since financial institutions could not avoid mediating commerce-related disputes. The cost of mediating disputes, building and hosting infrastructure, and setting up institutions to run and manage this infrastructure all result in economic waste and rent-seeking at the centre of our markets. According to Satoshi Nakamoto’s “white paper” on bitcoin, “the cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions” (Nakamoto 2018). The solution was to use cryptographic proof instead of third-party institutional trust.

This novel insight opened the door for unmediated transactions at the high and low end of the economic spectrum. Bitcoin’s design consideration is unique in that further exploration of the concept allows us to look at aspects of societal governance where third-party trust can be replaced by types of decentralized, ownerless, cryptographic platforms.

Although it’s difficult to predict how deep and fundamental some of these changes will be to the structures of society, we could begin by imagining how blockchain systems and programmable trust could play a role within our current systems. Remember, the transition from monarchs ruling over city states to democracies governing nation-states was a major paradigm shift. There are possible scenarios where structural transformations will be required for us to navigate this period in history and we should keep an open mind.

However, without calling for a revolution, here are some key practical areas that could see impact and improvement without significant structural change.

**Taxation**

All government services are paid for by taxes. Without taxes, the government would have no capital to operate. In the United States, the majority of federal tax revenue comes from income tax, while the majority of state revenue comes from a general sales tax for products and services.

For tax collectors, blockchains, smart contracts and cryptographic currencies represent infrastructure that brings us toward a permanent cashless system with reduced tax fraud and increased compliance. Blockchains and smart contracts can record real-time transactions along the value chain, create smart contracts and calculate, withhold and remit taxes automatically to the tax authority. In
Canada, this would mean a shrinking of the HST/GST tax gap that currently accounts for an estimated loss of $5 billion in annual revenues.

For individuals and businesses, tax compliance will become an “automated procedure” that is programmed to execute as a result of economic behaviours. The option to direct your contributions to public goods that you feel passionate about or that directly affect you could perhaps be embedded in tax collection. There is more on this below in the discussion of public goods.

A more reformist perspective could also include questioning the fundamental economics of taxation. Bearing in mind that income tax in Canada and the United States is relatively new (it was introduced in the United States during the Civil War, and in Canada during World War I to pay for wartime national debt), new forms of revenue could be made possible through the use of these systems. One such example is from the Zcash project, which has essentially created a system-wide programmatic inflation to pay for the core development and maintenance of its system.

Income Redistribution

As efforts are made toward retooling and retraining contributing members of society to their changing economic conditions, wealth and income distribution become important topics for public discourse. We are significantly underprepared for a world where jobs are replaced by artificial intelligence and machine labour; however, these are the economic and social realities of the next two decades.

Ideas such as universal basic income (UBI) are gaining traction in the wake of increased uncertainty around the longevity of our careers. Pilot projects are in development in Finland, the Netherlands and Kenya to test study these concepts. A candidate for the American Democratic presidential nomination, Andrew Yang (#YangGang), is making this a cornerstone of his campaign. The advent of cryptocurrencies and smart contracts means the potential to create unconditional basic income systems that operate by creating a blockchain that supports a fungible token...
and employs a decentralized identity system. A number of projects are experimenting with these systems, including Manna, BrightID, Swift Demand, Kuwa and Raha. It is difficult to imagine fair and cost-effective systems of income distribution that do not involve the programmatic allocation of capital.

Many, including Y-Combinator’s president, Sam Altman, who has been personally financing a UBI pilot, see a world in which decentralized finance networks operating on blockchain infrastructure will transcend local government. They will distribute coins or tokens to any provable unique human. If that network became big enough and people believed in it, it would be hard to stop.

**Economic Externalities**

Most notable among our market’s greatest externalities is the environment. The economic means of production has and continues to negatively affect our planet at an accelerated rate. Scientific consensus leads us to believe that we have 12 years to reverse a point of no return on this issue. As individuals and institutions act in their own self-interests, this tragedy of the commons continues; however, mechanism design and cryptocurrency may have an answer.

Similar to the function of carbon credits, enforced mechanisms can be made to incentivize behaviours that promote positive economic externalities and punish negative ones. These mechanisms will use cryptocurrency and smart contracts on a distributed, transparent shared platform. Core to this problem is a deep-rooted challenge that nation-states are not well equipped or incentivized to care about or have an impact on a global problem of this scale.

Incentive schemes will need to focus on water scarcity, natural disasters, climate change, biodiversity loss, air pollution and deterioration in ocean health if we are to succeed in this problem domain.

**Public Goods**

Cities are havens for an extensive set of public goods: sewage treatment, roads, parks, sidewalks, traffic lights, culture and services that are hard to create natural competitive markets around. If public funding for them ceased, they would not exist. Analogous to these public goods on the internet is open-source software, which often acts as the foundational underbelly for the creation of for-profit software companies. Open-source software is typically funded by “donations” (optional tax); however, new advances in the use of smart contract-based blockchains has generated promising new funding mechanisms such as CLR matching.1 Individuals make public goods contributions to projects of value to them. These individual contributions are “matched” or “topped off” by a government, grants program, or private philanthropist — programatically. By making an individual donation, you contribute to a public good. This funding is guaranteed to be met by the matching fund, widening the reach of your donation. The contributions you make become immutable “law.”

Experiments such as this in open-source funding, if successful, will scale to have societal impacts on the allocation of capital in scenarios of broadened scope. Now more than ever, it is time to experiment with new forms of social governance using programmable trust as a technology for good.

With blockchain protocols as natural incubation environments for this experimentation, we need more of our best minds in the social sciences, policy, mechanism design and economics working alongside our brilliant computer scientists and cryptographers, focusing on international issues of coordination and leveraging advances in technology to imagine and create a world that can help us navigate this transition.

Join us.

**Work Cited**


**Endnote**

1 CLR stands for Capital-constrained Liberal Radicalism. For more on CLR, see https://papers.ssrn.com/sol3/papers.cfm/abstract_id=3243656.
Patching Our Digital Future Is Unsustainable and Dangerous

Melissa Hathaway
Innovative technologies of the twentieth century have profoundly transformed society and the economy. The first electronic message was sent nearly 50 years ago on October 29, 1969 (over the Advanced Research Projects Agency Network, or ARPANET, the network that became the basis for the internet), but the internet did not become an engine of commerce until 1985 with the introduction of the .com top-level domain (Hathaway 2012). E-commerce was made easier with the launch of the World Wide Web in 1990 and was further accelerated by affordable computing power embedded with functionality and a wide range of applications in the palm of our hands (mobile phones). Nations and corporations alike have since embraced, adopted and embedded information and communications technology (ICT) into their networked environments and infrastructures, and realized phenomenal business and economic growth through improved services, increased productivity and decreased costs. Today, the digital economy represents about 20 percent of global GDP (Wladawsky-Berger 2017; Huawei and Oxford Economics 2017) and, by 2020, at least 30 billion Internet of Things (IoT) devices will hyper-connect our countries’ infrastructures and businesses and generate US$8 trillion in global revenue (Cleo 2018).

Yet this digital transformation — underpinned by affordable communications and cheap devices — has introduced new risks that cannot be ignored. The decision to embrace and embed poorly coded or engineered, commercial-off-the-shelf technologies into every part of our connected society — from government systems to critical infrastructures and services to businesses and households — is not without consequences. The providers of these technologies — the ICT vendors — are incentivized to be first to market with their products, and the marketplace has simply accepted the vendors’ promise that they will fix or “patch” the flaws in their products later. For example, Microsoft formalized this regular patching process in October 2003 — it has become known as “patch Tuesday.” Other vendors patch on a less frequent basis with little transparency on the known vulnerabilities that they have transferred to our digital products and services. Patch Tuesday is inevitably followed by a “vulnerable Wednesday” — where malicious actors, who are now also aware of those newly disclosed vulnerabilities, can exploit unpatched systems and steal sensitive data, knock businesses offline and, in some cases, destroy the IT systems that power businesses and essential services. Most organizations are not able to promptly update their systems when patches are released, further heightening our collective vulnerability to cyber harm.

The gold standard for implementing a software patch is 30 days (Proviti 2017). Other organizations may take longer to implement a software update in order to complete proper testing of systems to ensure that other business applications or processes are not negatively impacted. Still others may choose not to update their software for fear of breaking legacy applications within older and most likely end-of-life systems. To put this in perspective, Microsoft’s patch Tuesday in April 2019 included 15 software patches to address at least 74 vulnerabilities in its Windows operating systems and supporting software, including two zero-day bugs (Krebs 2019a). The previous patching update, in March 2019, similarly addressed more than five dozen vulnerabilities in Windows operating systems, Internet Explorer, Edge, Office and Sharepoint (Krebs 2019b). This “field it fast, fix it later” ethos has increased our exposure and is leading to real economic losses. For example, cybercrime is growing at 26 percent per year and is estimated to cost the global economy at least US$2.1 trillion in 2019 — or two percent of global GDP (Symantec 2018). Moreover, IoT attacks have increased by 600 percent between 2016 and 2017, in large part because of the ease to exploit connected devices (ibid.).

The flagrant ease with which these vulnerabilities can be exploited is often lost on both the general public and policy makers. For instance, Shodan — a free and publicly available search engine developed to locate digitally connected devices — can be used to easily find unpatched systems (Hill 2013). The tools needed to exploit known vulnerabilities are also inexpensive and easy to wield. Whether you purchase the book Hacking for Dummies, or hire a professional dark-web-market service, the ability to cause harm is no longer solely the purview of nation-states. Distributed denial-of-service attacks can be executed for as little as US$700, while stolen bank credentials can be purchased for the price of a cup of coffee (Barysevich 2017). Unauthorized access to accounts on Instagram,
In the United Kingdom, WannaCry affected at least 81 of the 236 National Health Service trusts — rendering medical equipment inoperable and significantly affecting public health and safety. (Photo: Pegasus Pics / Shutterstock.com)

The Cost of Global Cyber Insecurity

The economic and societal consequences of this widespread vulnerability are becoming increasingly acute. The world bears witness to a growing number of high-profile cyber incidents resulting in risks to public health and safety, global transportation and commerce and key industrial manufacturers. For example, in May 2017, a particularly simple strain of ransomware called WannaCry targeted flaws in Microsoft Windows operating systems, affecting millions of computers in 150 countries across every business sector. This global attack halted manufacturing operations, transportation systems and telecommunications systems. According to the National Audit Office in the United Kingdom, WannaCry affected at least 81 of the 236 National Health Service trusts — rendering medical equipment inoperable and significantly affecting public health and safety (National Audit Office 2017).

Six weeks later, in June 2017, a destructive malicious software called NotPetya swept the world, destroying the capital assets of hundreds of companies in minutes. Business operations halted in many companies, including Maersk (shipping), Merck (pharmaceuticals), Mondelez (confections) and DLA-Piper (legal services). Shipping giant A.P. Moller-Maersk was one of the companies most affected by this attack. It is responsible for the management of 76 port facilities worldwide and roughly 20 percent of the world’s container shipping capacity (Reuters 2017). It was figuratively and literally dead in the water after NotPetya spread across its entire global network. Within minutes, the virus encrypted and wiped the company’s information technology systems globally, including 4,000 servers, 45,000 computers and 2,500 applications across 600 locations in 130 countries. Maersk’s systems were offline for more than 150 hours (Maersk books an average revenue of US$2.9 million per hour) and the company reported first-quarter losses in the order of US$435 million to replace the IT systems that powered its digital business (A.P. Moller-Maersk 2017). Ultimately, it lost 10 percent of its market share to China Ocean Shipping Company. Maersk’s shareholder value depreciated by 30 percent within nine months.
of the incident and depreciated more than 50 percent 18 months post incident. In addition, Denmark’s GDP was also negatively impacted as Maersk contributes at least seven percent of the country’s GDP. The second- and third-order consequences to global shipping and the global economy have not been quantified (Greenberg 2018).

Now, Maersk executives talk about the importance of recovery operations since it took a whole-of-company effort to get the business back online (Palmer 2019). However, Maersk was aware of its digital vulnerabilities and the need for cyber security improvements prior to NotPetya’s release (A.P. Moller-Maersk 2016). Maersk may have weathered the storm better if it had implemented standard security procedures, such as regular updates to its software and operating systems and development of network segmentation.

The economic damages caused by NotPetya and WannaCry can be measured in the hundreds of billions of dollars. Yet, there are fears that global businesses are still unprepared for a global outbreak of another ransomware or destructive attack. In the first quarter of 2019, the new LockerGoga ransomware exploited unpatched Microsoft systems, knocking offline French engineering consultancy Altran Technologies, Japanese optical products manufacturer HOYA Corporation and American chemical companies Hexion and Momentive (Franceschi-Bicchierai 2019), as well as Norwegian Norsk Hydro — one of the world’s largest aluminum manufacturers (Ashford 2019).

As more companies connect and instrument their businesses to the IoT, their exposure to product vulnerabilities and exploitation thereof will also increase — putting their business operations at risk. Software and hardware design vulnerabilities should be addressed in those products’ design and development phases prior to debuting in active, high-stakes industrial operations. Critical infrastructure such as energy grids, manufacturing centres and petrochemical plants are increasingly coming under attack from malware designed to infiltrate industrial control systems (ICS) in order to disable, disrupt or seize control of the hardware. For example, the Triton malware was designed to sabotage critical operational technology in ICS, map the industrial network, and allow attackers to remotely control systems (Sobczak 2019). The first instance of its use was discovered in a Middle Eastern petrochemical facility in 2017. Although Triton was foiled by a flaw in its own design, it could have been used to override the shutdown procedures, which normally prevent disasters such as explosions or leakage of toxic chemicals (Giles 2019; Vijayan 2017; Jackson Higgins 2018). The malware exploited a vulnerability in Schneider Electric’s Triconex safety instrumented system. The system is deployed in 73 countries across numerous sectors including refining, petrochemicals, chemicals and specialty chemicals, power generation and pharmaceuticals (Desruisseaux 2018). As industrial manufacturers embark on their digital transformation, automating their processes and embedding IoTs in their business lines, their risk of digital disruption and asset destruction also increases. The use of sophisticated malicious software to target these systems is on the rise — and is alarming.

**Interstate Behaviour in Cyberspace: Hostility on the Rise**

The danger of interstate cyber hostility is also imminent. According to the 2019 US National Intelligence Strategy, “cyber threats will pose an increasing risk to public health, safety, and prosperity as information technologies are integrated into critical infrastructure, vital
National cyber security strategies, no matter how comprehensive, will fail unless clear lines of accountability are drawn, delineating security obligations among relevant parties.

National networks, and consumer devices” (Office of the Director of National Intelligence 2019). Cyber insecurity is taxing our economy and destabilizing our security. Each vulnerability is only a keystroke away from being exploited with weapons and services that are easily accessible and affordable online. Individuals, organizations and nation-states are increasingly taking advantage of these vulnerabilities to illegally copy intellectual property to advance economic interests; seize personal identifiable information to monetize in the dark market and pilfer universities’ research to advance sovereign interests; steal money or cryptocurrency to skirt the impacts of sanctions; and seed distrust among political parties, leaders and countries. As a testament to the growing anxiety around interstate cyber hostilities, in 2018, the United States and the United Kingdom took the unprecedented step of jointly calling out another state, warning that Russia had been infiltrating energy and transportation infrastructure, nuclear facilities and critically important private sector firms (US Department of Homeland Security 2018).

Numerous multilateral institutions have been promoting the responsible use of technology and advocating for normative or “responsible” behaviour among nations. Ensuring international agreement on what is proper and what is not proper behaviour in cyberspace is a priority for almost every country seeking to create stability and safety in cyberspace (Finnemore and Hollis 2016; Henriksen 2019). The first set of discussions in this regard was proposed by Russia in 1998. The UN Secretary General established a Group of Governmental Experts (GGE) to study the “developments in the field of information and telecommunications in the context of international security.” Since 2004, five GGEs have continued to study the threats posed by the misuse of information and communications technologies (ICTs) in the context of international security and how these threats should be addressed. Three of these groups have agreed on substantive reports with conclusions and recommendations.

In July 2015, member countries of the UN GGE endorsed and adopted a new set of voluntary, non-binding norms of responsible state behaviour in cyberspace. One of the most important norms agreed to by the group stated that “a State should not conduct or knowingly support ICT activity contrary to its obligations under international law that intentionally damages critical infrastructure or otherwise impairs the use and operation of critical infrastructure to provide services to the public” (UN General Assembly 2015, para. 13[f]). However, as demonstrated by the WannaCry incident (attributed to North Korea), the NotPetya destructive attack (attributed to Russia) and other similar attacks against companies and countries, states’ actions often do not match their professed ideals and norms of conduct are routinely ignored (Hathaway 2017, 2). Intentional damage of other nations’ infrastructure is becoming tacitly accepted as the normal state of affairs.

In September 2017, UN Secretary General António Guterres stated that “cyber war is becoming less and less a hidden reality — and more and more able to disrupt relations among States and destroy some of the structures and systems of modern life” (UN Secretary General 2017). He acknowledged that traditional forms of regulations do not apply, signalling a need for strategic thinking, ethical reflection and thoughtful regulation (ibid.). At the December 2018 UN General Assembly plenary meeting, two processes were launched to discuss the issue of security in the ICT environment for the period 2019–2021. Resolution 73/27, proposed by the Russian delegation, established an Open-Ended Working Group, which will be comprised of the entire UN membership (UN General Assembly 2018a; 2018b). It will further the development of norms and principles for responsible state behaviour in cyberspace and will look for meaningful ways to implement them. The group will deliver a final report at the seventy-fifth session of the UN General Assembly in September 2020. Another resolution, proposed by the United States, established a new GGE on “advancing responsible state behavior in cyberspace in the context of international security” (UN General Assembly 2019). This group will continue to study possible cooperative measures to address information security threats.

Other international organizations have also been promoting the responsible use of technology in order to build trust and confidence in the use of ICTs and minimize cyber harm. The 57 member states of the Organization for Security and Cooperation in Europe (OSCE), for example, have adopted 16 confidence-building measures (CBMs) to reduce the risks of conflict stemming from the misuse of ICTs and to increase cooperation among states to protect their critical infrastructures. The OSCE believes that increasing direct communication among states will defuse conflicts and prevent unintentional
escalation. The language in the document is that of a non-legally binding agreement, but it is a step toward advancing international cooperation in cyberspace in order to promote best practices and address vulnerabilities affecting our economy. Other multilateral institutions have adopted these CBMs, including the Organization of American States and the Association of Southeast Asian Nations.

Yet, in parallel to these confidence-building and norm-setting efforts, countries are also developing their own offensive cyber capabilities to deter or possibly respond to cyber attacks. The problem is that they are fighting fire with fire. For example, in 2016, at the North Atlantic Treaty Organization (NATO) Warsaw Summit, the alliance declared cyberspace as the fifth domain of warfare. Since that time, seven members have pledged their offensive cyber weapons to the alliance and stand ready to employ the full force of their arsenal should one member fall victim to a particularly grievous cyber attack. From now on, NATO will integrate the sovereign effects from the nations that are capable and willing to provide them (Freedberg 2018).

The Role of Governance in Reducing Cyber Risk

The digital environment continues to underpin our homes, businesses and countries with products and services that are pre-packaged with exploitable weaknesses. The high-profile cyber security incidents of recent years are symptomatic of the attitude that continues to dominate the development and commercialization of digital technology, in which companies strive to release products as quickly as possible and worry about security flaws after they have already been deployed. Ultimately, the paradigm of “field it fast, fix it later,” which continues to hold sway in the technology industry, must be overcome. If we are to achieve a stronger level of security or at least significantly reduce cyber risk in the digital age, governments will need to step in and hold digital service providers and the manufacturers of ICT technology accountable for ensuring their products maintain adequate cyber safety standards.

As the scale of the threat has become more apparent, governments around the world have turned to developing frameworks for understanding the nature of their digital dependency, cyber security strategies for fending off these threats and policies to establish standards of safe behaviour.

For example, in the United States, the Department of Commerce is launching an initiative to improve transparency around software components. The so-called Software Bill of Materials intends to drive a disclosure process for all software and IoT vendors to share the details on the underlying components, libraries and dependencies of their software with their customers. According to Allan Friedman, director of cyber security for the National Telecommunications and Information Administration, “this transparency can catalyze a more efficient
market for security by allowing vendors to signal quality and giving enterprise customers key knowledge — you can’t defend what you don’t know about” (Friedman quoted in Epper Hoffman 2018). It would also give enterprises more insight into the risks to their digital businesses (i.e., patch Tuesday).

If this initiative does not catalyze industry to take more responsibility for the inherent flaws in their products, the state of California has taken an even more proactive approach. In anticipation of the unfolding IoT vulnerabilities, California passed a connected devices law, which lays out the security features that must be included in all digitally connected devices. The law will go into effect on January 1, 2020. It requires vendors that intend to sell connected devices (i.e., IoT) in California to implement enhanced security measures for all those products. It broadly defines devices as any device that connects directly or indirectly to the internet and has an Internet Protocol or Bluetooth address. These security measures include device attestation, code signing and a security audit for firmware in low-level components.

In Europe, the General Data Protection Regulation (Council of the European Union 2016), which entered into force in May 2018, aims to hold companies accountable for the digital security of personal information. The Network and Information Security Directive stipulates minimum standards of care for the cyber security of critical infrastructure, including energy, transport, banking, finance, health, water and digital infrastructures such as online marketplaces (for example, eBay and Amazon), search engines (for example, Google) and clouds. Companies that suffer a significant breach or service outage must notify the relevant national authority within 48 hours and include the following data points: duration of incident; number of affected parties (for example, customers, vendors, and so on); geographic spread; extent of disruption of service; and impact on economic (calculated in GDP terms) and societal activities.

Similarly, China passed a national cyber security law that went into effect in June 2017. It contains 79 different articles detailing data protection requirements and cross-border data flow guidelines, as well as specific guidelines for “critical information infrastructures” (CII). This includes information services and the law establishes a broad definition of CII as a service that may cause serious damage to national security, the national economy and public interest if destroyed, if functionality is lost or if data is leaked (Creemers, Triolo and Webster 2018).

The common thread between all these policies is that destructive and disruptive cyber activities require urgent attention and action. National cyber security strategies, no matter how comprehensive, will fail unless clear lines of accountability are drawn, delineating security obligations among relevant parties. Presently, the delegation of duties between government and the private sector remains unclear in many areas, such as the protection of critical infrastructure. This ambiguity makes it particularly difficult to hold organizations responsible for lax security standards. Comprehensive, methodical assessments of cyber risk at the national level will be required to correctly identify the greatest areas of vulnerability and address the gaps in current defensive strategies. Policy makers need to ascertain what risks they are willing to bear and what would be considered intolerable. Risk reduction activities also require the allocation of dedicated and appropriate resources, both human and financial, for their implementation. Only with a concerted and coordinated effort across national stakeholders will it be possible to significantly reduce cyber risk and move forward to ensure the future safety and security of a nation.

All governments are operating under resource constraints and will need to engage in sincere, honest reflections in order to set digital security priorities. Many current policy approaches cast a wide net in terms of which systems are deemed critical to national and economic security. However, by focusing their attention too broadly, countries risk devoting insufficient attention and resources to those few indispensable infrastructures, services, companies and assets upon which everything else depends. The fact of the matter is that some are more important than others. The provision of energy and telecommunications, for example, is essential to the economic health and national security at the most fundamental level, as nearly all other systems would cease to function without them. Certain companies, which comprise a large proportion of the total economy of a country, may also warrant special attention.
For instance, A.P. Moller–Maersk contributes a large share of Denmark's GDP, such that when the company fell victim to NotPetya in 2017, the Danish economy suffered significant collateral damage. The United States and Germany have proceeded by identifying companies contributing more than two percent of their national GDP and forging better information-sharing arrangements with them to ensure cyber security concerns are given due consideration in corporate protective measures (Hathaway 2018, 9).

Despite a nearly universal agreement about the importance of shielding critical services and assets from digital harm, governments have thus far had difficulty in accurately assessing where the greatest vulnerabilities lie, and therefore knowing exactly what warrants their immediate attention or is the highest priority. For example, the city of Atlanta — one of the top 100 resilient cities globally — was knocked offline in March 2018 by the SamSam ransomware (Schwartz 2018). Its January 2019 audit showed that the city had known gaps in its security that had not been addressed. Less than six months later, another critical asset in the United States — the port of San Diego — suffered a ransomware attack that used the same variant of malware as the one in Atlanta. SamSam affected IT systems and disrupted public services (Kan 2018). In the Netherlands, despite efforts by the Dutch government to bolster the cyber security of its critical infrastructures and services, officials were caught off guard when the port of Rotterdam (the largest in Europe) fell victim to the NotPetya malware in 2017. Upon further review, Dutch officials discovered that they had not classified ports as critical infrastructure under their infrastructure protection policies (Hathaway 2018). Many critical assets of great importance to economic vitality and national security have been overlooked by current cyber security strategies, necessitating more rigorous countrywide assessments.

**Time to Get Strategic**

As things currently stand, countries at the cutting edge of the technological frontier are moving forward with the development and deployment of IoT and other innovative technologies at a breakneck pace. First-mover advantages are perceived to be so great that most relevant actors have not stopped to consider the potentially destabilizing effects of these technologies for fear of falling behind their economic and geopolitical rivals. Yet, by attaining an advantage in this “technological arms race,” countries are rendering themselves more dependent on technologies that are increasingly complex and opaque — and thus vulnerable — leading to a higher risk of accidents and unanticipated negative effects. As a recent report from the Center for a New American Security put it, “superiority is not synonymous with security” (Danzig 2018, 7). In the long run, it will be those nations that have given pause to consider the possibilities for adversarial use of the technologies in question that will be best placed to reap rewards in terms of wealth and influence.

Increased automation, interconnectedness and reliance on the internet require that we embrace a new form of cooperation, in which vulnerabilities are reported to the owner of the information system, allowing the organization at stake the opportunity to diagnose and remedy the vulnerability in question before detailed vulnerability information is disclosed to third parties or the public. This is called responsible disclosure. Ideally, vulnerabilities are largely prevented through a design process that gives security higher priority. So far, the ICT industry has followed a different path and many vulnerabilities are repaired only after the product has been embedded in an operational environment and supporting business-critical systems (Internet Engineering Task Force 2002).
Each vulnerability is only a keystroke away from being exploited with weapons and services that are easily accessible and affordable online.

The United States maintains a National Vulnerability Database; 78 organizations in 14 countries use the data. Vulnerabilities reported to the Department of Homeland Security’s Cyber and Infrastructure Security Agency by way of the US Computer and Emergency Readiness Team are disclosed to the public within 45 days of the initial reporting, regardless of the existence or availability of patches or workarounds from affected vendors. China has a similar system, but it operates twice as fast as the American process, averaging just 13 days after public disclosure. China proactively scours the web and other sources of information, looking for vulnerability information, whereas the United States waits for reports from vendors to be processed through the Common Vulnerabilities and Exposures database (Waterman 2017).

The trade-offs between prompt disclosure and withholding knowledge of some vulnerabilities for a limited time can have significant consequences. From a government point of view, disclosing a vulnerability can mean that intelligence agencies forego an opportunity to collect crucial intelligence that could thwart a terrorist attack, stop the theft of a nation’s intellectual property or even discover more dangerous vulnerabilities that are being used by hackers or other adversaries to exploit our networks (The White House 2014). But when a corporation decides not to disclose critical unfixed vulnerabilities in its software, should that be considered okay? What about when the database of all known vulnerabilities is illegally copied by malicious actors? Is there an obligation to disclose the loss and begin addressing the risk that the corporation has now transferred to society? In 2013, hackers obtained unauthorized access to a Microsoft database that contained descriptions of critical and unfixed vulnerabilities in its software, including the Windows operating system (Menn 2017). In August 2016, government tools that were largely focused on exploiting these Microsoft vulnerabilities began to be publicly released — presenting a real risk to global corporations and the global economy. Some of these tools (or weapons) were ultimately behind the WannaCry and NotPetya attacks in 2017 (Patel 2017; Hay Newman 2017; Schneier 2017).

In February 2017, Microsoft launched a campaign to deflect attention from its flawed products and put the responsibility for the exploitation of those vulnerabilities back onto nations. It launched its “Digital Geneva Convention” campaign, stating that governments should commit to “protecting civilians from nation-state attacks in times of peace.” The document asserts that “just as the Fourth Geneva Convention recognized that the protection of civilians required the active involvement of the Red Cross… protection against nation-state cyber attacks requires the active assistance of technology companies.” Microsoft affirmed that the tech sector plays a unique role as the internet’s first responders, and the technology companies, therefore, should commit themselves to collective action that will make the internet a safer place, affirming a role as a neutral “digital Switzerland” that assists customers everywhere and retains the world’s trust (Smith 2017). However, it is too bad that the company chose to pursue a convention about normative state behaviour vice fixing its own flawed products. Microsoft has gone on to advocate for a “Cybersecurity Tech Accord,” launched in 2018, that promises to defend and advance the benefits of ICTs to society. It assumes that technology companies are the rightful heirs that create and operate online technologies. Finally, Microsoft’s efforts were highlighted again in the “Paris Call for Trust and Security” that was announced at the opening of the Internet Governance Forum in November 2018. It was supported by governments and private sector organizations around the world. But are we to believe the charlatan who quietly hides their negligence and shifts responsibility to another party? Society needs responsible, ethical and serious corporate leaders who are dedicated to delivering a secure and resilient digital future for all (Tech Accord 2019).

We must become much more strategic in how new digital technologies are created and deployed. Over the last 30 years, we have created a unique and strategic vulnerability to society — an inherently insecure internet supported by poorly engineered products. It is an existential threat to our economy and our sovereign security. To address this immediate threat, an emergency counter-measures board and mitigation process should be initiated that is global and convenes the best talent, regardless of nationality. The industry has fielded us vulnerable products fast — now, we must work together to reduce the risks and heal our digital environment as quickly as society can.
Our governments should require: a new vulnerability disclosure process (and operational requirements); a duty to warn of imminent danger, such as in the case of an emerging attack; and a duty to assist in the case of cyber emergencies (Hathaway and Savage 2012). ICT purveyors of products should be required to implement a new communications and warning system for urgent patches, adding “emergency” to their repertoire of categories (emergency, critical, important, moderate and low).

Consumer protection agencies must also engage. We have been conditioned to marketplace recalls related to food, medicine, automobiles and even children’s toys — IT products are not recalled, even when it is known that they can cause serious harm to society. The consumer protection agencies can drive accountability by eliminating or significantly reducing after-market repairs (patch Tuesday) to a market that drives accountability through product recalls. Vendors should have to deliver well-engineered products and services and present the buyer with a list of the underlying components, libraries and dependencies — a “software bill of materials” — to drive transparency and accountability. This process could also inform the emerging revisions of ISO/IEC 29147:2014, Information technology — Security technology — Vulnerability disclosure.7

Finally, the UN General Assembly has recognized the importance of reducing the ICT threat to society by launching two new fora to deliberate on normative state behaviours and to look for meaningful cooperative measures to address information security threats. These efforts are essential to develop pathways for direct communications among states and to help prevent unintentional escalation in cyberspace.

The world has witnessed an alarming number of harmful ICT practices and internationally wrongful acts through the misuse of ICTs in recent years. There has been a large, perhaps unwarranted, degree of faith in novel technologies. We tend to trust that technology will always work as intended — and only as intended — often failing to give much thought to how the technologies that are created to solve our problems could be turned to nefarious ends. The time has come to recognize this overarching problem and subject technological development to greater scrutiny. The downsides of novel technologies should be contemplated along with the benefits they may bring. Only then will we be able to start eradicating the vulnerabilities from the core of our digital future.

Works Cited


Giles, Martin. 2019. “Triton is the World’s Most Murderous Malware, and it’s Spreading.” MIT Technology Review, March 5. www.technologyreview.com/s/618054/cybersecurity/critical-infrastructure-triton-malware-utm_campaign=the-download-unpaid-engagement&utm_source=ls-email&utm_medium=email&utm_content=70515982&_hsmi=29884963&hmcid=p2ANzgz--8FAh9vccvlQK85SFZP0EvAenjo0rLTS_s7xEmUJD-5a6fsM71GrH8g19DBrbc2Dfe1yeU0C0wefC3r3zOn63fRDJ3Kq4XWtWdfl8W8gqV7Y_hkm=70515982.


Governing Cyberspace during a Crisis in Trust


Menn, Joseph. 2017. “Exclusive: Microsoft responded quietly after detecting NotPetya attack and the NHS.”


Kan, Michael. 2018. “Ransomware Strikes the Port of San Diego.”


Menn, Joseph. 2017. “Exclusive: Microsoft responded quietly after detecting secret database hack.”


Reuter. 2017. “Global Shipping Giant Maersk is Reeling from the Ransomsware Fallout.”


Schwarz, Mathew. 2018. “Atlantis’ Ransomware Cleanup Costs Hit $2.6 Million.”


Subczak, Blake. 2019. “The inside story of the world’s most dangerous malware attack.”

UN General Assembly. 2015. “Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security.”

Endnotes

1. Maersk share price was at a high of around 14,000 Danish kroner just before the NotPetya attack. Six months after the event, its share price had dropped to around 10,000 Danish kroner. One year post incident, the share price dropped further to 8,000 Danish kroner.


4. The seven NATO members that have pledged their offensive cyber weapons to the alliance are Estonia, Denmark, France, Germany, Netherlands, United Kingdom and the United States.


9. This international standard ISO/IEC DIS 29147 revision is currently under development.
What kind of place is cyberspace? It is tempting, and perhaps even romantic, to think of it as a void without laws or governance. Certainly, that was the vision of the first cyberpunks, who promoted a kind of digital anarchy that was free of rules and where humans could prosper away from the eyes and ears of governments, censorship and control (Rid 2016, in particular chapter 5). A second view is that cyberspace is a haven for criminals who clandestinely sell their services and illegal wares on the “dark Web” (Bartlett 2015). A third vision of cyberspace is as a tool of surveillance used by technology and social media companies who monitor, store and sell our data — sometimes in partnership or competition with national security agencies (Lyon 2019).

There may be a kernel of truth in all of these visions, but they overlook the important fact that the internet is not outside of territorial control — it is subject to rules, regulations and, fortunately, the development of norms, including privacy. And yet, as this collection of essays on cyber security and trust shows, creating governance for cyberspace is one of the greatest global challenges in the twenty-first century.

Four notable key themes run throughout the essays in this collection, which form an important background for thinking about finding a path forward to promote responsible policies in this space:

• **Threats at the speed of cyberspace:**
  The theme that unites the essays is that technological change is evolving the cyber-threat landscape at a pace that both the public and the private sectors are having difficulty keeping up with in terms of security and managing privacy. Whether it is the innovations of cyber criminals, the actions of state-backed malicious actors or the “fake news” sent out by armies of bots to further undermine trust, the array of challenges is staggering and will require multi-faceted creative solutions.

• **Regulation:** It is notable that none of the authors in this series believe that regulation is an impediment to a better cyber future. While they differ in their approaches (some favouring more protections and some concerned about the effects of going too far), there seems to be a consensus that a regulatory environment that creates the legal foundation for innovation to take place...
Balancing our digital future with respect for privacy has become a key political issue. An individual uploading a photo on a social media platform or sharing personal data wants to be able to trust that it will be treated appropriately. (Photo: pixinoo / Shutterstock.com)

is vital to securing the future of Canada's digital economy. In some cases, this might mean creating standards by which private companies can prove themselves to be trusted with sensitive data (such as health information) so it can be better used to provide more targeted, potentially life-saving services. However, creating legislation that balances the need for innovation with the need to protect citizens is difficult, and governments have often fallen behind.

- **Public-private partnerships**: Many of the authors argue that addressing rapidly evolving threats and developing and operationalizing solutions will require robust private-public partnerships. As Aaron Shull notes in his introduction, Canada's 2018 National Cyber Security Strategy is as much about fostering innovation responsibly in the private sector as it is about protecting Canadians. No one sector will be able to achieve this alone, but how cooperation should take place is far from clear.

- **Privacy**: In the wake of scandals ranging from Edward Snowden's revelations to Cambridge Analytica, the need to balance our digital future with respect for privacy has become a key political issue. While individual citizens may not think twice about uploading a photo on a social media platform, they want to be able to trust that the information they are providing will be treated appropriately. While some countries, such as China, are moving ahead with a comprehensive surveillance state with little concern for privacy, it is unlikely that the West will be able to secure its digital future without ensuring that the increasing information citizens put online will be respected.

Importantly, these four themes cannot be thought of as independent from one another — they intersect in ways that amplify risk and make finding policy solutions difficult. For example, as Christopher Yoo (2019) notes in his essay on the Internet of Things (IoT), IoT devices collect large amounts of personal information, may store it in a distributed way and were not designed for security. This leaves the devices — and personal data — vulnerable to cyber criminals and malicious state actors who can hijack IoT gadgets for their own purposes. However, few (if any) states have found ways to regulate IoT systems in a practical way.

**Canada’s Cyber Policies and Practices**

Since 2016, Canada has taken a number of steps toward addressing some of the challenges outlined in this essay series that involve new policies, powers, institutions and coordinated international action. First, as discussed above, a new cyber security strategy that links safety to innovation was released in 2018. Notably, it is the first cyber security policy since 2010 and represents an important and much-needed update. In addition, in 2017, the Government of Canada introduced its new defence policy, *Strong, Secure, Engaged*, which prominently features cyber-related issues, including challenges (such as hybrid warfare), recruitment needs and implications for research and development.

Second, Bill C-59 is the most significant reform to Canada’s national security architecture since 1984. Cyber security is at the heart of many of the bill’s reforms, including the need for greater state capacity to defend Canada against threats with enhanced review and privacy protections. Significantly, Bill C-59 grants Canada’s signals intelligence agency, the Communications Security Establishment (CSE), the ability to defend designated critical infrastructure from attack (“defensive cyber”) as well as an offensive capability (“active cyber”). It also grants the Canadian Security Intelligence Service (CSIS) the legal grounds to take in
public data (ingestion) and then to refine and use it (digestion) (Forcese 2018). Data sets comprised primarily of Canadian information will require annual approval of the minister of public safety and an intelligence commissioner — a quasi-judicial position also created under the legislation. There will also be further internal vetting by CSIS, and further retention of the data must be approved by the Federal Court, which is empowered to impose conditions on subsequent use (ibid.).

Third, a number of new domestic institutions have been established to bolster cyber security. In 2016, the Canadian Cyber Threat Exchange (CCTX) became operational with a mandate to improve information sharing on cyber threats faced by the private sector. Importantly, the CCTX is a private sector initiative to improve cyber security across the board so that Canadians are confident in doing business online. In 2018, the federal government created an outward-facing arm of the CSE, the Canadian Centre for Cyber Security (CCCS), to improve communication on cyber issues with small and large businesses and the general public. Notably, the CCCS is the government’s point of contact with the CCTX. The 2018 federal budget also promised the creation of a national cybercrime coordination unit, although it is anticipated that it will not be fully operational until at least 2023 (Solomon 2019).

Finally, Canada has quietly developed a multilateral diplomatic approach that promotes cyber security and defends norms in cyberspace through coordinated action. Notably, many of these steps were outlined in the communiqué following the June 2018 Group of Seven (G7) summit in Charlevoix, Quebec. First, along with its allies, Canada has called out malicious cyber activity by North Korea, Russia, and China on several occasions. Notably, this coordinated diplomatic activity goes beyond the “Five Eyes” alliance (Australia, Canada, New Zealand, the United Kingdom and the United States), and includes Denmark, the Netherlands and Japan. Second, in January 2019, the government created the “Rapid Response Mechanism” that will share information and threat analysis with other G7 countries, as well as identify opportunities for coordinated responses when cyber attacks occur. These diplomatic actions are significant for at least three reasons. First, Canada has been very reluctant to call out states for malicious behaviour. Unlike the United States, which has frequently indicted foreigners and even foreign officials who are believed to have engaged in cyber espionage, Canada has very rarely identified malicious state actors or even spoken out against this kind of behaviour. In this sense, Canada’s statements demonstrate a willingness to “name and shame” in a way it has not done before.

Second, as noted above, this coordinated diplomatic action shows a willingness to work with other liberal democratic countries beyond the Five Eyes. It suggests a broadening of potential partners to ensure global cyber security. Finally, these statements made by a growing number of allied states are helping to contribute to the creation of norms for cyberspace. Calling out malicious activity as counter to the expectations of international behaviour is important for the development of standards and perhaps for laying the foundations of an international order, if not international law and regulations.

**Securing Trust and a Safe Cyber Future**

While these first steps are important and go some way to creating the legal and policy grounds to promote cyber security domestically and internationally, there is more that can — and needs to — be done.

First, while Canada is developing a multilateral approach to cyber diplomacy with an emphasis on “naming and shaming” behaviour that it considers to be malicious or illegal, there are more steps it can take. As noted above, international law and regulations in cyberspace are still in an early stage, but it is important to remember that Canada will have to live with whatever legal norms develop. The Government of Canada will need to make its understanding of international norms and law in cyberspace known so that its views are represented as these standards develop. For example, in May 2018, the United Kingdom’s Attorney General Jeremy Wright gave a speech that outlined the United Kingdom’s views on applying international law to cyberspace. Canada, which has a strong incentive for a rules-based international order — even in the digital realm — should take note and do the same.
Second, clarifying its position on international cyber norms will be helpful for government and policy leaders in thinking about what kind of cyber future they wish to live in. It is very likely that in the future Canada will have to navigate a “splinternet” between three worlds: a state-dominated China, a regulated Europe and a relatively unregulated United States. While it is unlikely that China provides the kind of model that Canada will want to emulate, it will be much harder to navigate between the US and European approaches. In making its decision, Canada will have to balance its economic requirements with the need to protect privacy. This will likely involve an ongoing dialogue with stakeholders in the private and non-governmental sector.

Third, the potential for innovation can only be met if Canadians are willing to trust the digital services presented to them by the public and the private sectors. This requires policies and regulations that protect the privacy of Canadians and the security of the systems that hold their information. Unfortunately, this is an area where Canada is currently failing on two fronts. As noted above, there is little to nothing in the way of standards or regulations for the private sector. Worse, Canada’s federal legislation on information sharing is overly broad and will remain so, even with Bill C-59’s national security overhaul. The amount of information shared between government agencies without the knowledge or consent of Canadians and with virtually no review or oversight is significant. A 2017 investigation by the Office of the Privacy Commissioner found that there were “significant procedural deficiencies” in the way information was being handled and that the current information-sharing regime “will remain a threat to the privacy of individuals” (Office of the Privacy Commissioner of Canada 2017).

As digital technologies make it easier to gather, share and store personal information, this problem is only going to get worse if no steps are taken. The Canadian government and private businesses need to find a way to ensure the protection and safety of information. They should create policies and regulations that allow for agile standards that can evolve with changing technologies. Creating robust review and oversight mechanisms of the entities that provide digital services to Canadians would enhance public confidence that their information is safe and correctly stored.

Finally, government policies that foster not only innovation but also a diversity of companies working in this field are needed. Diversity is important for two reasons. First, a lack of competition means that there is less incentive for technology companies to invest in robust cyber security because they do not have to worry about their reputation. Additionally, a company that is able to dominate a particular area will likely become a target.
of cybercriminals and malicious state actors seeking to find and exploit vulnerabilities (National Cyber Security Centre 2019). In both cases, the lack of competition makes it easier for harmful cyber activity to occur. The government should find ways to ensure that innovation results in diversity as well as economic benefits for Canada.

Creating cyber policy that balances security, privacy, innovation and trust is an imperative for Canada. Our economy and society will not be able to harness the benefits of the next industrial revolution without agile yet robust policies that create room for experimentation but safeguard the rights of citizens. Complicating an already difficult problem is the challenge of trying to accomplish all of this in a time of contested international norms and malicious state behaviour in cyberspace. Nevertheless, as a highly connected society with the technological skills to innovate and malicious state behaviour in cyberspace. Nevertheless, as a highly connected society with the technological skills to innovate and a legal framework that provides guidance on protecting the rights of individuals, Canada can become a leader in this space. The question now becomes, will we be bold enough to act?

Works Cited


Endnotes

1  See Desai (2019).

2  See Jardine (2019).

3  See Lyon (2019).

4  See Vallée (2019).

5  See Boysen (2019).


7  See Department of National Defence (2017).

8  Bill C-59 (“An act respecting national security matters”) has not passed the Senate at time of writing. It is slated to be voted on by the end of May 2019, but there remains the risk that it will not pass in this legislative session.

9  See CSE (2017).

10  See CSE (2018a); Global Affairs Canada (2018).

11  See CSE (2018b).


13  Two rare exceptions prior to 2017 include identifying China in 2014 (following an article on the National Research Council) and Iran in 2015 (for a hack into Canadian government systems). However, these attribution statements were either low-key affairs (China) or the result of media investigations (Iran).

Governing Cyberspace during a Crisis in Trust

While technology has led to convenience, efficiency and wealth creation, the push to digitize society quickly and relentlessly has meant building inherent vulnerability into the core of the global economic model. With this paradox in mind, CIGI convened a group of interdisciplinary experts to delineate the looming gaps in the governance of cyberspace. Pursuing this agenda is a national imperative; nothing short of the future of the Canadian economy hangs in the balance.

Watch videos with series authors at cigionline.org/cyberspace