Understanding National Security Threats Enabled by Artificial Intelligence: Implications for CSIS

Prepared for
Canadian Security Intelligence Service

Prepared by
Linda Xu
Edi Qereshniku
Hisham Hazari
Mackenzie Edwards

University of British Columbia School of Public Policy and Global Affairs

About this Report

Project Overview

The Global Policy Project is an intensive capstone project for second-year students of the Master of Public Policy and Global Affairs program (MPPGA) at the University of British Columbia (UBC). This report was produced for the Canadian Security Intelligence Service (CSIS). This project was guided by CSIS' Academic Outreach and Stakeholder Engagement branch in partnership with the Strategic Policy directorate. This project was supervised by Professors Andrea Reimer and Chris Tenove of the UBC School of Public Policy and Global Affairs.

Project description

CSIS' core mandate is to investigate threats to the security of Canada (*CSIS Act*). These threats were first defined in 1984. Since then, the world has transformed significantly, and threats have evolved over time.

The development of emerging and disruptive technologies like Artificial Intelligence (AI) can have great benefits for Canada's prosperity. However, they can also be leveraged by hostile actors against Canada's national interests.

This project will seek to identify how threat-related activities enabled by AI technologies could be used to threaten Canada's national security, how AI-enabled threat activities may challenge CSIS' existing mandate and operations, as well as how AI-enabled threats and CSIS' responses to them may have an impact on individual rights and freedoms.

Project scoping

The scope of this research project has been narrowed in three ways. First, the analysis is limited to threat-related activities enabled by AI technologies. In other words, the threat is not AI itself, but rather the threat activities enabled and exacerbated by AI. This is because pursuant to s. 12 of the *CSIS Act*, "the Service shall collect, by investigation or otherwise, to the extent that it is strictly necessary, and analyse and retain information and intelligence respecting activities that may on reasonable grounds be suspected of constituting threats to the security of Canada and, in relation thereto, shall report to and advise the Government of Canada" (*CSIS Act*, 1984). Additionally, pursuant to s. 12.1 of the *CSIS Act*: "if there are reasonable grounds to believe that a particular activity constitutes a threat to the security of Canada, the Service may take measures, within or outside Canada, to reduce the threat" (*CSIS Act*, 1984). Second, the report will not include military applications of AI because military operations fall under the jurisdiction of the Department of National Defense and not CSIS. Third, the focus of the report is on AI-enabled threats to Canada's national security rather than AI-enabled opportunities for CSIS to capture. The reason is that AI technologies have significant implications on Canada's national security, especially when leveraged by adversarial actors.

Authors

The project team is comprised of four students who are currently completing the MPPGA program at UBC. Below there is more information about each member:

Linda Xu holds a BA in Law from Sciences Po Paris and a BCom in Finance and International Business from UBC's Sauder School of Business. Linda brings experience working in strategy roles.

Edi Qereshniku holds a BA in Economics and a BCom in Finance from the University of Calgary. Edi brings over ten years of corporate banking experience where he provided financing and strategic solutions to corporate clients in various industries.

Mackenzie Edwards holds a BA in Politics, Philosophy, and Economics from UBC. Mackenzie currently works as a policy analyst for the Public Health Agency of Canada.

Hisham Hazari holds a BA in Global Affairs from Jindal Global University. Hisham brings experience working as a security specialist/analyst in the global/corporate security industry. Hisham currently works as a program development officer for Public Safety Canada.

Client Description

The Canadian Security Intelligence Service (CSIS) is a vital agency within Canada's national security infrastructure. Its core mandate is to investigate activities suspected of constituting threats to the security of Canada, inform and advise the Government of Canada of these threats, and take measures to reduce these threats.

The CSIS Academic Outreach and Stakeholder Engagement program supports initiatives from the Government of Canada and the public through hosting workshops, presentations, round-table discussions, commissioning open-source research, and participating in inter-governmental committees. These activities provide valuable insights to inform evidence-based decision making and policy development in government. Additionally, the program leads in coordinating a government-wide approach to academic outreach. The program is frequently consulted on the development and implementation of similar programs within Canada and internationally.

Executive Summary

The core mandate of CSIS is to investigate threats to the security of Canada. CSIS derives its duties and functions from the 1984 *CSIS Act*, in addition to observing laws like the *Privacy Act* and the *Canadian Charter of Rights and Freedoms*. Since 1984, technological advancements have transformed the world significantly. One of the most rapidly advancing and disruptive technologies is Artificial Intelligence (AI). While AI enables opportunities for economic prosperity and societal benefits, it also enables and exacerbates three key national security threats.

AI enables more sophisticated and impactful cyberattacks, increases the generation and spread of disinformation, and precipitates issues with open-source intelligence (OSINT). It can make cyberattacks more effective through faster identification of software vulnerabilities and more precise targeting of phishing campaigns. Utilizing bots and deepfakes, AI can create more believable fake content that will make it harder to distinguish between what is real and what is not. Declining costs for data storage and processing coupled with AI advancements renders OSINT activities more accessible. This increases both the number of non-state actors participating in intelligence gathering and the potential for errors to occur.

CSIS's intelligence cycle includes detecting, defining, analyzing, and investigating threats to Canada's national security. AI-enabled threats are complex and continuously changing which presents several challenges to CSIS's operations. AI-enabled threats make it more difficult for CSIS to interpret which of the four threat definitions in the CSIS Act are impacted, to operate under current data collection restrictions, and to effectively use the judicial oversight system which currently relies on one type of warrant.

Whether AI is used for offensive or defensive purposes, it interacts with Canadian rights and freedoms such as privacy, equity & equality, freedom of expression, and security of person. The proliferation of AI for threat activities or defensive purposes increases the need for large amounts of data which in turn increases risks for biases, as well as the increased need for transparency and accountability.

Given the challenges posed by AI-enabled threats to CSIS's mandate and operations, this report has identified 5 recommendations for CSIS to consider:

1. Increase information sharing by enhancing the IT infrastructure.
2. Upgrade Canada's cybersecurity strategy by taking a multi-agency approach.
3. Advocate for an amendment to the *CSIS Act* to include different types of judicial orders.
4. Revisit funding requirements for hiring, training, and in-house AI expertise.
5. Build collaborative public-private relationships with organizations developing AI technologies to keep pace with technological advancements.

Introduction

The *CSIS Act* provides overall direction to CSIS in conducting its day-to-day operations. The Act was drafted in 1984 to respond to predominantly physical threats, rather than cyber threats. However, the threat landscape has evolved tremendously since the creation of the Act. Canada's national security has been threatened by actors operating with modern technology. Amongst modern technologies, Artificial Intelligence (AI) is one of the most complex, powerful, and rapidly evolving technologies (Employees, CSIS).

Problem Statement

Given the development and deployment of AI in the modern threat environment, the following challenge has been identified:

The implications of AI technologies within a national security context coupled with CSIS' approach to threats within the framework of the *CSIS Act* may not allow CSIS to effectively address the dimensions of national security risks posed by threat activities enabled and exacerbated by AI technologies.

Research Questions

This report explores the following 3 key questions to address the identified challenge:

1. How could AI-enabled technologies be used to threaten Canada's national security?
2. Within the framework of the *CSIS Act*, can CSIS effectively define, investigate, and protect Canada against AI-enabled threats given the emergence and proliferation of AI technologies?
3. How do threat-related activities enabled by AI technologies interact with individual rights and freedoms? What legislation and best practices currently exist to address this interaction?

Methodology

To explore the 3 research questions, our team conducted a comprehensive literature review and 11 semi-structured interviews to investigate these themes.

The literature review was comprised of 62 academic articles, 40 government documents, and 51 grey literature publications. Key pieces of literature include publications from Brookings Institution, Center for Security and Emerging Technologies, Oxford's Future of Humanity Institute, Foreign Affairs magazine, several academic journals, and various government publications from CSIS, Public Safety Canada and the Canadian Center for Cyber Security of the Communications Security Establishment.

The literature highlighted the pace of AI advancement, its broad-ranging applications, and issues around its governance. The literature also revealed various ways in which AI could enable and exacerbate existing threat activities. Furthermore, the literature explored potential challenges that AI-enabled threats could pose to existing legislative authorities under the *CSIS Act* to respond operationally to the speed and scope of AI-enabled threats. Additionally, the literature explored the interaction between AI-enabled threats and activities as well as the rights and freedoms of Canadians; it provided the academic, ethical, and practical backbone on which the analysis of these dynamics was conducted.

The interviews provided crucial insight into AI's foreseeable impact on Canada's national security landscape. They also contributed valuable insight into AI's potential implications on the way in which CSIS currently approaches threat activities. Interviewees included individuals with expertise in AI and Canada's national security from government, academia, and the private sector: CSIS, Justice Canada, Carleton University, Privy Council, Ernest & Young, and the University of British Columbia. Please see Appendix 1 for the full interview list.

Background

Artificial Intelligence 101

There is no unanimously accepted definition of AI. In this report, AI refers to "machines which respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment, and intention" (West & Allen, 2018). In other words, AI systems can perform tasks and make decisions which typically require human intelligence.

Our definition of AI includes both Artificial Narrow Intelligence (ANI) and Artificial General Intelligence (AGI) – although the latter does not currently exist but is theorized to be possible in the near future.

Artificial Narrow Intelligence	Artificial Narrow Intelligence refers to AI systems which are limited to performing specific tasks they are programmed to carry out (Joshi, 2021). Although these tasks are performed autonomously with human-like capabilities, they are limited to a narrow range of tasks (Joshi, 2021). Currently, all AI technologies, even the most advanced AI developments, are classified under ANI (Joshi, 2019).
Artificial General Intelligence	Artificial General Intelligence refers to AI systems which can understand and learn any intellectual task a human can perform (Joshi, 2021). They can autonomously develop multi-functional competencies across various domains (Joshi, 2021). Fifty percent of AI experts predict that AGI will be realized before 2060 (Simon, 2022).

This report will focus on two sub-fields of AI, including machine learning (ML) and natural language processing (NLP).

Machine Learning	ML is a sub-field of AI which focuses on the "use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy" (IBM, 2023). ML generally requires humans to structure and label the data with which the algorithm is trained (IBM, 2023). A subset of machine learning is deep learning (DL). DL analyzes raw, unstructured data to inform and improve its algorithm (IBM, 2023). A great example of ML being applied is Netflix. ML enables Netflix's recommendation engine, which is responsible for customizing each viewer's homepage, to recommend unique content to each viewer (Steck et al, 2022). A pertinent example of DL being applied is Chat-GPT. Chat-GPT is an AI-powered chatbot developed by OpenAI which "uses deep learning techniques to generate human-like responses to text inputs in a conversational manner" (Browne, 2023).

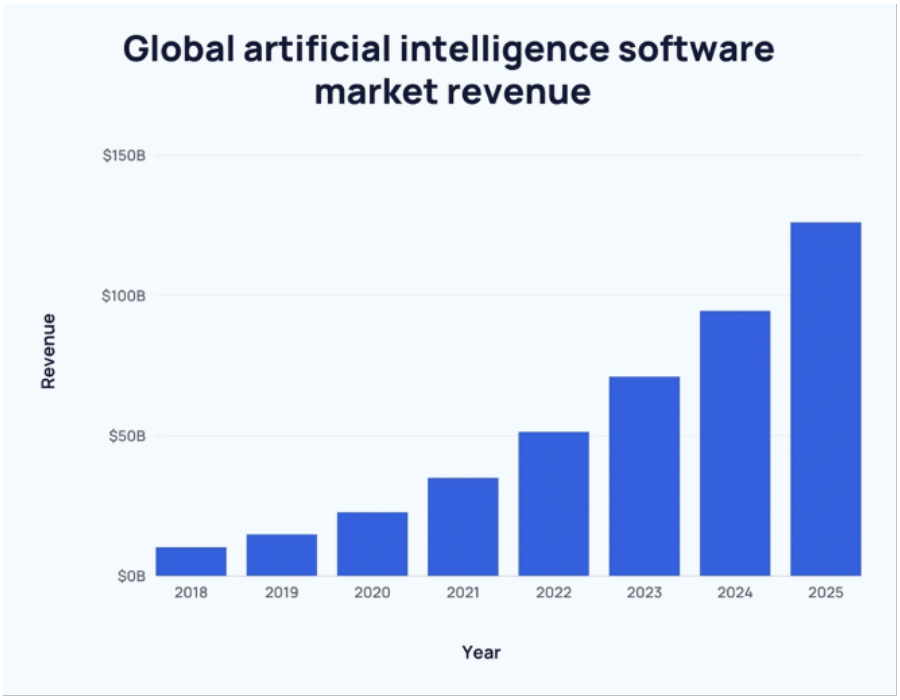
Natural Language Processing	NLP processing is a sub-field of AI which focuses on enabling computers to process and understand written and spoken words in the same way humans can, including intent and sentiment (IBM, 2023). For example, NLP enables voice assistants like Siri and Alexa to understand, respond to, and perform tasks based on a human's voice command. NLP also enables customer service chatbots to have text or speech conversations with the user and to understand, assess, and respond to their queries.

AI is frequently encountered in everyday life. One encounters AI when opening their phone using facial recognition, receiving targeted advertisements online, scrolling through Amazon's recommended-for-you section, using a search engine like Google, and more (Marr, 2019). AI is already integrated and deployed in various sectors including finance, national security, health care, criminal justice, transportation, and smart cities (West & Allen, 2018). Notably, in the finance sector, loan decisions use AI software to analyze a "variety of finely parsed data" about the applicant (West & Allen, 2018). This results in more optimal decision-making compared to purely basing the decision on an applicant's credit score and background check (West & Allen, 2018). Moreover, in the transportation sector, autonomous vehicles use high-performance computing, advanced algorithms, and DL systems to analyze information and adapt to new scenarios (West & Allen, 2018).

Past, Present, and Future of AI Progress

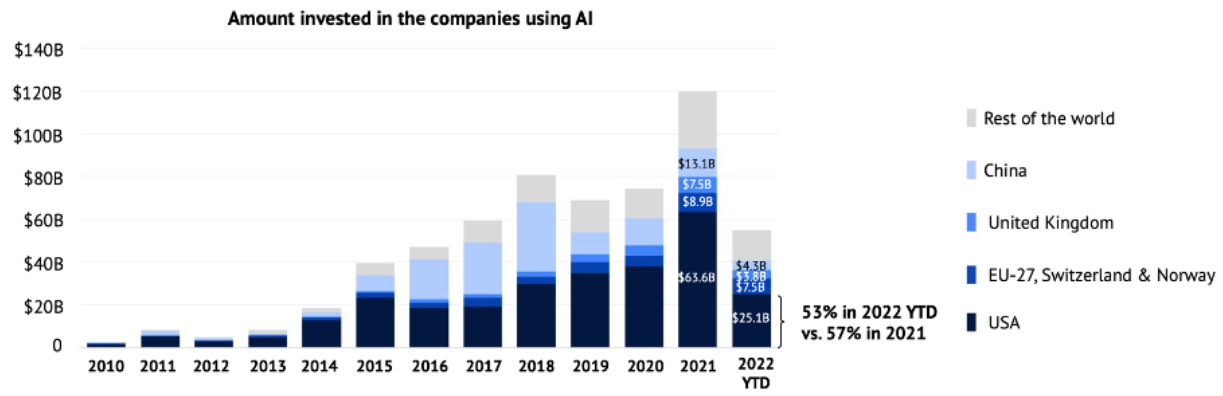
One proxy for the AI sector's growth is consumption of AI software globally. As demonstrated in Figure 1, global AI software market revenue has risen significantly since 2018 and is forecasted to continue to rise through 2025 (Omida, 2023).

Figure 1: Global AI software market revenues from 2018 to 2025 (Omida, 2023)



Since AI development occurs predominately in the private sector, another proxy for the AI sector's growth is the amount of private investment directed towards companies using AI. As illustrated in Figure 2, the amount invested in companies using AI has grown at a rapid pace over the past decade (Benaich & Hogarth, 2022). Figure 2 also highlights the concentration of investment in the United States (US) and China. Notable AI players include large technology companies such as America's Microsoft, Google, IBM, Amazon, and China's Baidu, Tencent, and Alibaba. Investments are also made in smaller-size AI firms such as OpenAI. OpenAI is the creator of ChatGPT, an AI-powered chatbot which garnered over 100 million users within 2 months of launching (Hu, 2023).

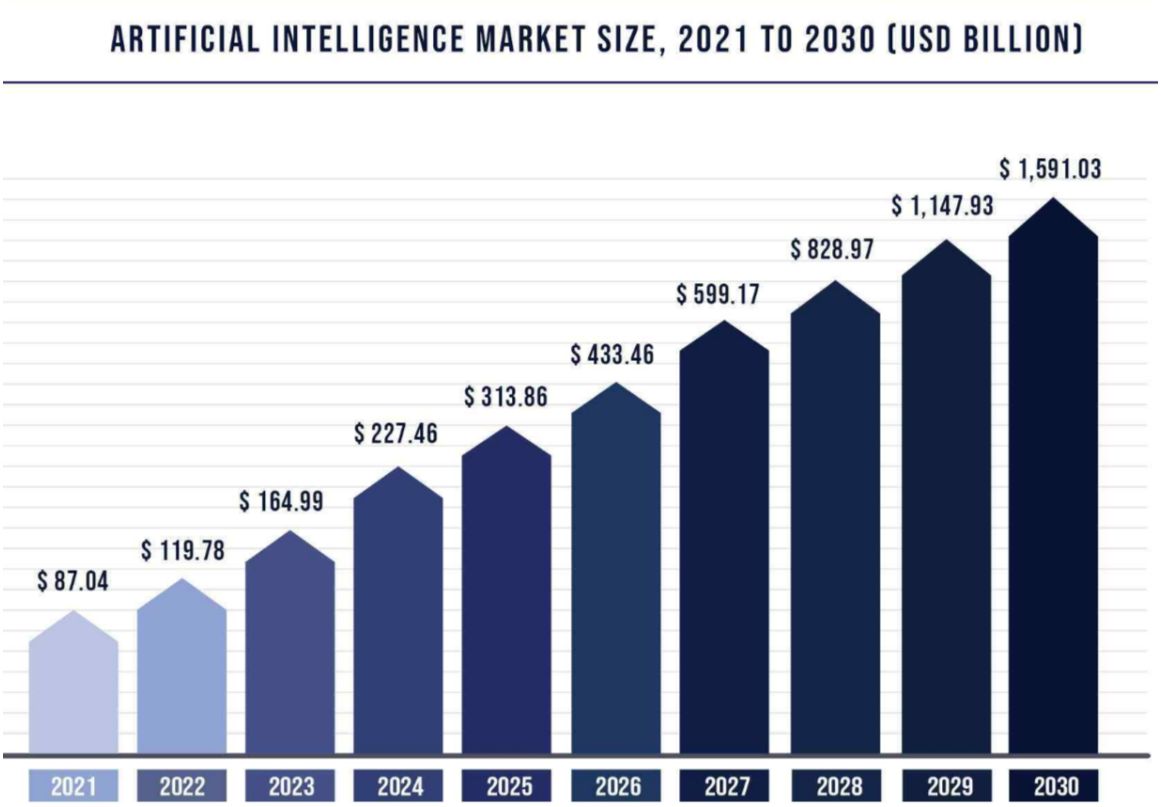
Figure 2: Amount invested in companies using AI from 2010 to 2022 (Benaich & Hogarth, 2022)



Although AI technologies are primarily being adopted by private sector companies, it is expected that governments will eventually catch up on adopting AI technologies (Partner, EY). Whilst governments have largely not adopted AI technologies, they do recognize the importance of having a strong AI sector in their national economy. This is because AI is a key component to enhancing national competitiveness and protecting national security (Knight, 2019). Although government investment in AI is a mere drop in the bucket compared to private investment, it is still important to acknowledge the growth over the past two decades. Between 2001 to 2019, investment in AI-related research and development (R&D) funding from government agencies in Australia, Canada, France, Spain, Japan, the Netherlands, the US, and the European Commission increased by approximately ten-fold in aggregate (Galindo-Rueda & Cairns, 2021).

Looking forward, AI's market size is projected to grow to more than USD 1.5 trillion in 2030, nearly 10x its current market size in 2023 (Precedence Research, 2023). By 2030, artificial narrow intelligence will be adopted across all realms of life and work (Agrawal, 2021). By 2060, there are likely to be significant advancements towards realizing artificial general intelligence (Simon, 2022).

Figure 3: AI market size from 2021 to 2030 (Precedence Research, 2023)



Canada's Position in the Global AI Context

In 2017, Canada was the first country to adopt a national AI strategy (Pascoe et al, 2017). The Pan-Canadian AI Strategy was launched by the Canadian Institute for Advanced Research (Pascoe et al, 2017). The first phase of the strategy was launched in 2017 with CAD 125 million of funding (Pascoe et al, 2017). It aimed "to build a strong Canadian talent pipeline and ecosystem, including the establishment of centers of research, innovation and training at the national AI institutes" (Government of Canada, 2022). The second phase of the strategy was launched in 2022 with CAD 443 million of funding (Government of Canada, 2022). It aims "to bridge world-class talent and cutting-edge research capacity with commercialization and adoption" (Government of Canada, 2022).

Despite Canadian efforts to be a global leader in AI, it is the United States and China that currently lead the way. These countries have invested significantly more public funding to advance their national AI innovation and commercialization efforts relative to Canada (Silcoff & O'Kane, 2023).

Research Findings

The following section of the report will present the findings of each research question and provide analyses of the findings in relation to the problem statement. This section aims to facilitate an understanding of the development and deployment of AI in the context of national security, its implications for Canada's national security interests, the extent to which CSIS may be able effectively address the threat activities enabled by AI, and the impacts these activities may have on individual rights and freedoms.

AI-enabled Threats to Canada's National Security

For the purposes of this report, national security relates to "any action or event that could materially impact the health, safety, security, or economic well-being of Canadians, or the effective functioning of Canada's governments" (Fasken, 2021).

AI's development and deployment have significant implications for national security (Sayler, 2020). For example, AI could have disruptive impacts on Canada's national interests if used by Canada's adversaries to facilitate intelligence collection (CSIS, 2021). AI development and deployment also have significant implications for intelligence agencies. Specifically, AI could drastically change the intelligence life cycle (Employee, CSIS). Given these implications, there is a need to think critically about the impacts of AI on CSIS' current operations (Employees, CSIS).

CSIS already faces a very broad threat horizon in the online environment (Employees, CSIS). The innovation of AI creates an even broader surface of attack for threat actors (Partner, EY). To encapsulate this broad threat horizon, CSIS' most recent public report identified 12 national security threats to Canada. Of these 12 threats, this project's research, through literature review and subject-matter expert interviews, identified cyberthreats and election security to be particularly vulnerable to threat activities enabled by AI technologies.

Within the realm of cyberthreats and election security, this report identifies three threats that are exacerbated when enabled by AI:

1. cyberattacks,
2. spread of disinformation, and
3. misuse of open-source intelligence (OSINT).

Although these three threats have existed for decades without AI, the application of AI can increase their speed, scope, and sophistication. The subject matter experts from the interviews validated that cyberattacks, disinformation, and issues with OSINT are indeed the three key categories of threats exacerbated by AI (Employees, CSIS). Focusing on cyberattacks, disinformation, and issues with OSINT allows this research

question to investigate in greater depth how AI exacerbates these threats and what the implications are for Canada's national security.

Outside of the scope of this report, there exist two additional threats identified in CSIS' most recent public report that are exacerbated by AI: extremism and terrorism. If further research opportunities arise, these two threats would be worthwhile to investigate.

Threat #1 – Cyberattacks

The application of AI-enabled technologies will increase the number, scale, and diversity of cyberattacks which can compromise public and private computer systems and critical infrastructure.

Cyberattacks are attempts to access computer systems without permission with the goal of stealing, exposing, altering, or destroying information (IBM, 2022). The most common types of cyberattacks include the following:

Type	Description
Malware	Any program or code that is created with the intent of doing harm to a computer, network, or server. This is the most common type of cyberattack because it includes many subsets such as: ransomware, spyware, bots etc.
Denial-of-Service	An attack that floods a network with false requests to disrupt business operations. Whereas DoS attacks are launched from just one system, they can also be Distributed Denial-of-Service (DDoS) attacks that are launched from multiple systems.
Phishing	An attack that uses email, SMS, phone, social media, or social engineering techniques to entice a target to share sensitive information or to download a malicious file that will install viruses on their computer or phone.
Spoofing	A technique where cybercriminals disguise themselves as a known or trusted source and in doing so can engage with the target with the goal of stealing information, extorting money or installing malware or other harmful software.
Identity-Based Attacks	When a valid user's credentials have been compromised and a threat actor is pretending to be that user. These types of attacks are hard to detect because it is often difficult to differentiate between the user's typical behavior and that of the threat actor.

Source: Baker, 2023.

Canada remains a target for malicious cyber-enabled espionage, sabotage, foreign influence, and terrorism-related activities where threat actors seek to compromise government and private sector computer systems by manipulating their users (CSIS, 2021). This presents significant threats to Canada's national security.

As one of the largest users of the internet in the world, Canadians are consistently using online devices for multiple purposes, such as financial transactions, communication, and work. In 2022, approximately 95% of Canadian adults reported using the internet, with data consumption increasing by over 3x the volume since 2015 (CWTA, 2023). As Canadians spend more time on these devices and are continuously connected to the internet, opportunities for cyber threats grow. In 2022, the Canadian Center for Cyber Security (Cyber Center) published its annual report outlining the trends in Canada with respect to cyber threats and highlighted the following five narratives:

1. ransomware and its impact on an organization's ability to function is one of the most persistent attacks in Canada,
2. critical infrastructure is increasingly at risk from cyber threat activity including state-sponsored actors,
3. the state-sponsored cyber programs of China, Russia, Iran, and North Korea pose the greatest strategic cyber threats to Canada,
4. ML technologies are making fake content easier to manufacture and harder to detect and Canada's exposure to disinformation is expected to grow, and
5. disruptive technologies such as ML bring new opportunities and new threats (CCCS, 2022).

AI capabilities and applications can be used for both defensive and offensive purposes with varying degrees of impact to national security. There are also various non-state threat actors in this space which include politically motivated groups that seek to both intimidate and recruit, hacktivists that seek to create political, social, and cultural change and the most common, cybercriminal who are financially motivated (Kreps, 2021). Because cyberattacks have been part of the threat environment for some time, it is important to further analyze which type of cyberattacks are particularly threatening to national security when AI-enabled. Within the threats posed by AI-enabled cyberattacks, the following analysis will focus on software vulnerability detection, spear phishing and data poisoning.

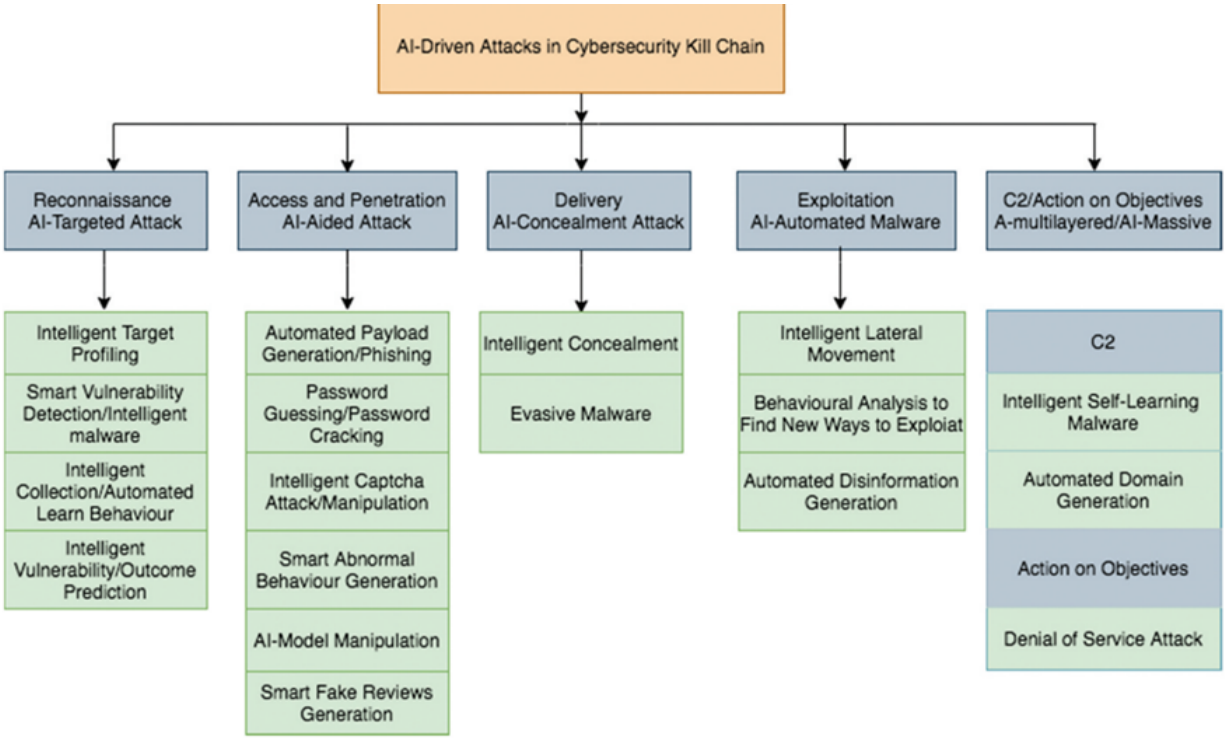
Software Vulnerability Detection

Effective cyberattacks depend on finding vulnerabilities in organizations' computer systems. These are primarily identified by using software that searches for weaknesses in computers, networks, and communications within systems (Kreps, 2021). These vulnerabilities often exist in outdated systems that have not been patched, and AI can help identify these issues, reduce costs by automating the process and increase accuracy with more convincing and targeted capabilities (Kreps, 2021). Because of AI's ability to process data and information much faster, threat actors can better exploit the

window of opportunity between when a threat vulnerability is identified and the system is patched.

The sequence of steps taken by threat actors to achieve their goals is known as the “kill chain” (Buchanan, 2020). As such, a key consideration for national security becomes understanding how AI-enabled technologies can reshape and supercharge this kill chain (Buchanan, 2022). The figure below illustrates the ways in which AI-enabled attacks would modify and affect the “cybersecurity kill chain”.

Figure 4: AI-driven modifications to the cybersecurity kill chain (Guembe et al, 2022 modified from Kaloudi and Li, 2020)

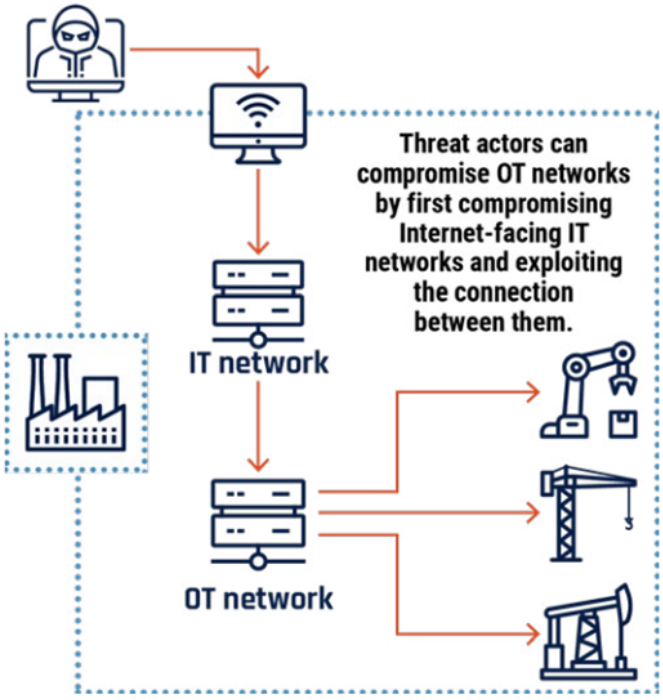


Current research has identified that in the above kill chain, AI-enabled technologies are most used in the access and penetration portion of the attack (Guembe et al, 2022). For example, when measuring the success rate for the password guessing/password cracking technique, it was found that AI-driven methods were able to outperform the traditional algorithms (Guembe et al, 2022). Another key difference between traditional cyberattacks and AI-enabled ones is the way in which decisions are made during the attack. Traditional cyberattacks are based on “if-then” logic which means that it asks whether it has found the target, and if the answer is “yes” the malicious program will execute, but if the answer is “no” it will end (Guembe et al, 2022). Enabling these attacks with AI means that threat actors can use a much more complicated decision logic that goes beyond simply yes or no to help decide whether to attack or not, while simultaneously making it extremely hard for the system’s defense to recognize the malicious code (Guembe et al, 2022). This presents threats because not all organizations are equally protected or have the same resources to do so (Partner, EY).

Canada's Cyber Center (2022) expects that even after patches are developed for systems, threat actors will almost certainly continue to scan the internet for opportunities of finding unpatched systems.

Understanding AI-enabled threats is even more important when considering critical infrastructure because cyberattacks in this space can have a physical impact on the world. Public Safety Canada (2011) identified the following sectors as critical infrastructure: energy and utilities, finance, food, health, government, safety, water, transportation, information and communication technology, and manufacturing. As the Operational Technology (OT) that underpins the industrial processes of these sectors is exposed to the internet, the threat surface increases and becomes more opportunistic for cyberattacks (Associate Professor, Carleton University, & CCCS, 2022).

Figure 5: Cyberattacks on Critical Infrastructure (CCCS, 2022)



While the security of IT hardware and software has improved recently, the security of the internet of things (IoT) and OT has not kept pace, and attacks (specifically malware) have moved into large-scale operations targeting infrastructure, utilities, and corporate networks (Microsoft, 2022). Critical infrastructure has both dependencies and interdependencies and becomes increasingly interconnected because when one is impacted many more follow (Associate Professor, Carleton University). A recent example of the potential impact on critical infrastructure is the 2021 attack on the Colonial Pipeline Company, which was breached by one leaked password in an older system (Morrison, 2021). This forced the company to take some systems offline and disable the pipeline, which provides nearly half of the fuel supplied to the east coast of

the US (Morrison, 2021). Although this attack was not specifically AI-enabled, it highlights how with the increased use of AI in detecting software vulnerabilities (especially the first two steps of the kill chain), the tradeoff between scale and efficiency of attacks can improve and allow threat actors to focus on bigger and more impactful targets (Brundage et al, 2018).

Spear Phishing

A phishing attack is an attempt to extract information or initiate action from a target by making them believe that the communication or request comes from a trusted source rather than the attacker (Brundage et al, 2018). Phishing attacks often succeed by their volume: casting these fraudulent messages using a wide net without necessarily worrying about who or what they catch in the net. Spear phishing is similar, but it is much more targeted and focused on targets who often have higher levels of resources and in turn higher possible payout for threat actors. Spear phishing involves more work, as threat actors gather information about the target, craft high-quality and personalized messages, and send the message with a malicious link that aims to steal information (Onelogin, 2022).

This type of attack requires a significant amount of skilled labour because the goal is to be as realistic as possible, which means not all threat actors can dedicate the resources and time required (Brundage et al, 2018). In 2021, the Canadian Anti-Fraud Center received an increased number of reports of spear phishing with losses reaching approximately CAD 54 million, which is an increase from $30 million in 2020 (CAFC, 2022). Spear phishing is considered effective, as at least 30% of the attacks are deemed successful with an estimated rate of return that is 40x higher than regular phishing attacks (Onelogin, 2022).

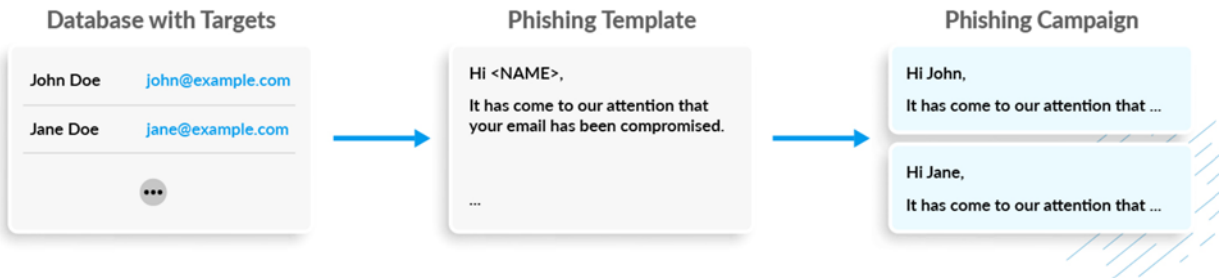
Because phishing and spear phishing are familiar attacks, the introduction of AI-enabled technologies to aid them is expected to expand the set of actors who can carry out the attacks, the rate at which the attacks occur, and increase the set of possible targets (Brundage et al, 2018). The tradeoff between cost and return for threat actors is improved when using AI because: 1) it can quickly gather useful personal information, 2) predict the best targets to approach and for which asset, and 3) use NLP that is crafted in the same manner as the sender without eliciting suspicion (Kreps, 2021, Onelogin, 2022).

This means that AI can be used for the full cycle of spear phishing and therefore materially reduce the time and resources required by many threat actors (Guembe et al, 2022). For example, researchers at ZeroFox, a cybersecurity firm, demonstrated that a fully automated spear phishing system could create very effective tweets that match the preference of users and lead to clicks on links that are malicious (Brundage et al, 2018). Due to these attacks becoming more and more human like as AI improves and learns from increased amounts of data, it is expected that there will be a significant increase in

network penetrations, personal data theft and intelligent computer viruses from spear fishing attacks. (Brundage et al, 2018, CCCS, 2022).

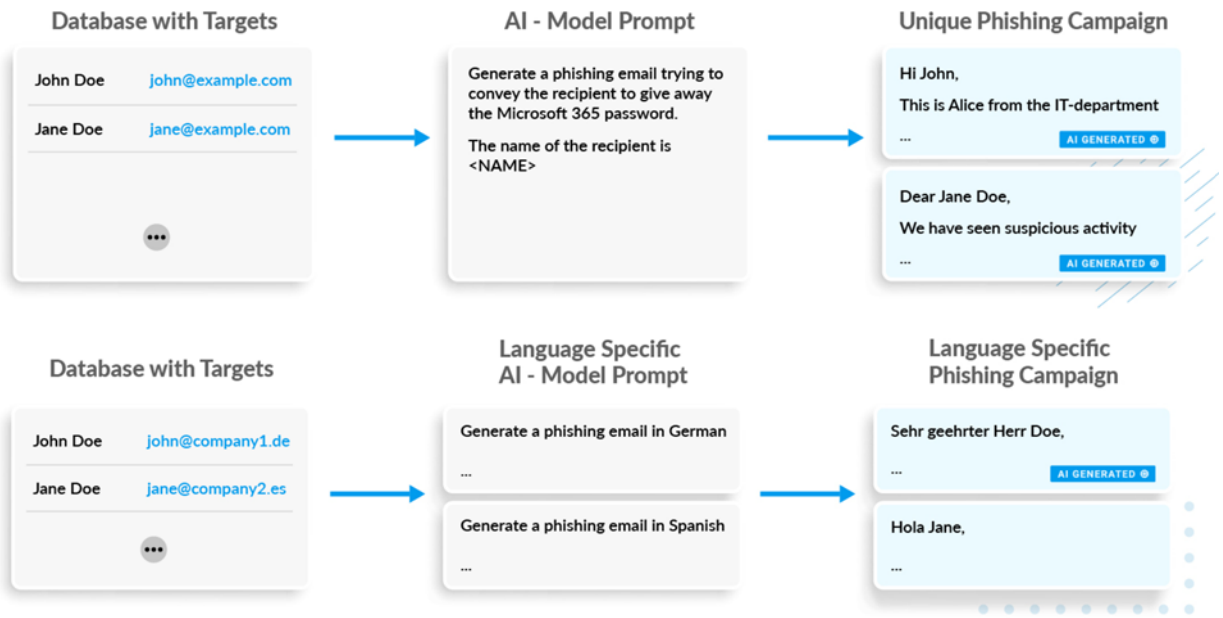
Figures 6 to 8 below, provide an example that highlights the evolution of a traditional phishing attack into an AI-enabled one.

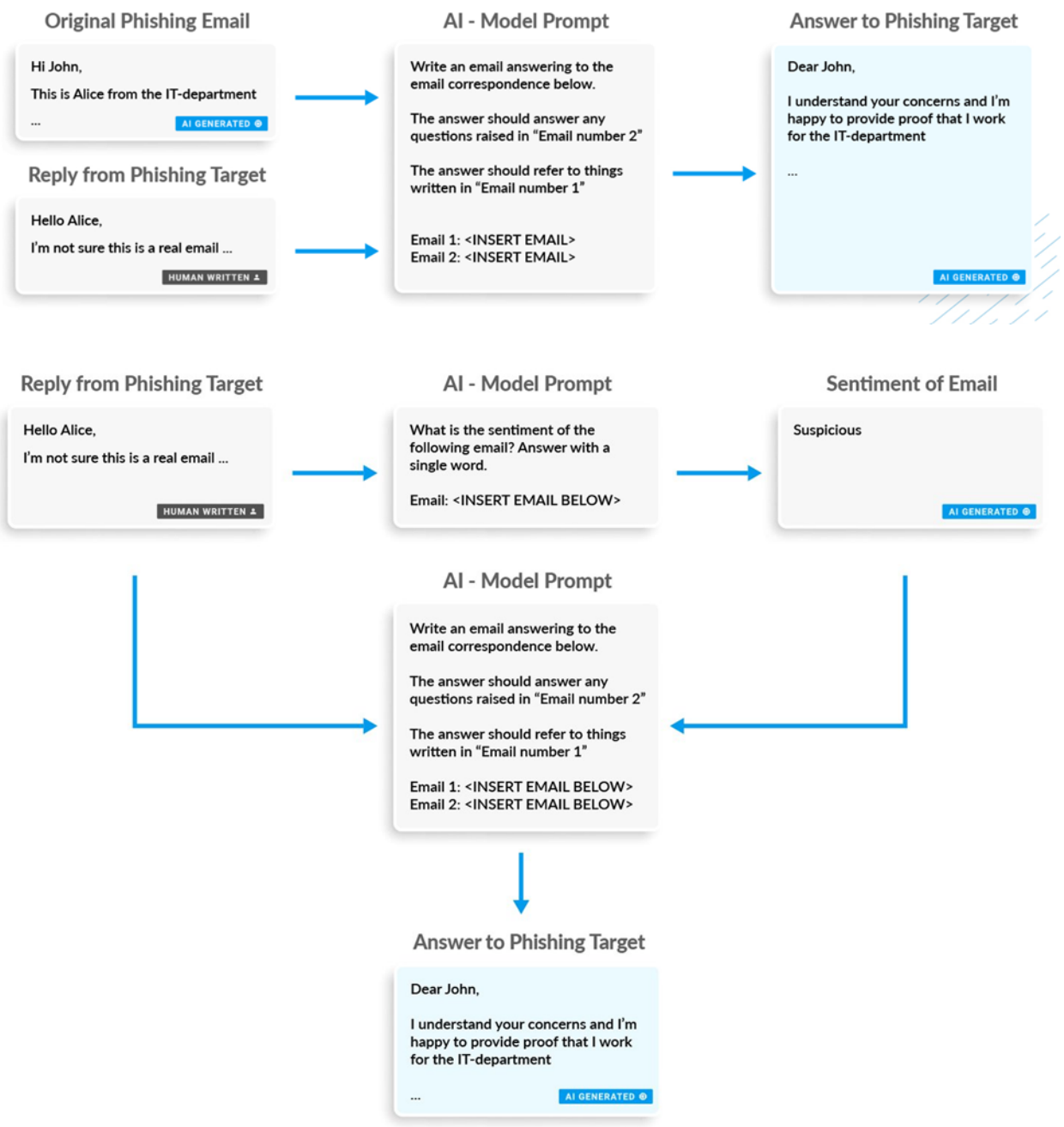
Figure 6: Traditional Phishing Attack (Xorlab, 2023)



Moving from the traditional phishing attack above, to one enabled by AI, there is an improvement in effectiveness. AI using ML and NLP models, improves phishing attacks by differentiating initial messages, using various languages, recognizing sentiment in target's replies, and adjusting the response in order to remain undetected.

Figure 7: AI-enabled Phishing Attack (Xorlab, 2023)

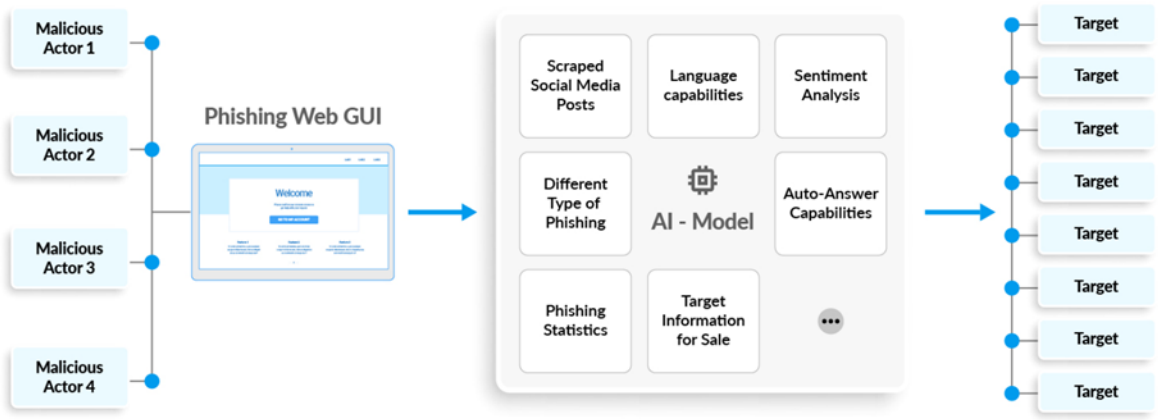




(Xorlab, 2023)

As there are more attempts to stop the use of AI for bad intentions, threat actors could start to host their own AI models that can be integrated into a graphical user interface (GUI) system that makes it easy to use. If used by existing phishing kits, this type of service could significantly escalate cyber threats. The figure below illustrates how it could potentially work:

Figure 8: Illustrative AI-enabled Phishing interface (Xorlab, 2023)



Data Poisoning

Data poisoning involves tampering with the training data that is used by AI with the aim of producing undesirable outcomes (Thrope, 2021). Vast amounts of data are needed to train and improve the decision-making of the AI model. This presents opportunities for threat actors to conduct data poisoning that can impact both the cyber and physical world.

There are two ways in which these attacks can occur. One, is a less sophisticated approach which involves injecting as much bad data as possible; and the second is more precise, and leaves most of the database untouched, except for an undetectable back door that lets attackers control it (Thrope, 2021).

As more Canadian organizations start to rely on these often-unsupervised algorithms, significant damage can be caused from data poisoning before anyone realizes it (Thrope 2021). In recent years, several data poisoning attacks have highlighted this potential threat. In one case, threat actors attempted to poison the Gmail spam filter by sending millions of emails with the intention of confusing the classifier algorithm and modifying its spam classification, which in turn allowed them to send malware without the spam algorithm noticing (Menon, 2023). In another, Microsoft's Twitter chatbot was being trained to engage in Twitter discussions when threat actors poisoned the training dataset, resulting in the bot turning hostile in its communications (Menon, 2023). These types of attacks can cause considerable damage by threat actors because poor quality information will produce subpar results regardless of how advanced the model is (Thrope, 2021).

Threat #2 - Disinformation

Disinformation is "false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit" (European Commission, 2018). Its key characteristics according to the widely adopted "ABC framework" on

disinformation include 1) manipulative actors, 2) deceptive behaviors, and 3) harmful content (Francois, 2019). Disinformation can be conveyed through images, videos, or text and predominantly occurs on social media platforms. These platforms are ideal targets for disinformation because they are privately owned, lack government regulation and oversight, host echo chambers, and run on an engagement-maximizing business model (Bremmer & Kupchan, 2023). Disinformation is expected to continue to proliferate in Canada due to declining trust in traditional sources of information and an increasing blurring of opinion and fact (Rand Corporation, 2022).

Disinformation campaigns have been used to threaten national security in numerous cases, for example:

1. Russia used disinformation to target the 2016 American presidential election. It was able to undermine confidence in the election process, exacerbate social and political divisions among the electorate, and increase the adoption of conspiracy theories (Posard et al, 2020).
2. In the lead up to Brexit, pro-Leave groups used disinformation to sway public opinion in favour of the Leave campaign (Woolley, 2020).
3. In the wake of the 2020 American presidential election, pro-Trump groups used disinformation to question the legitimacy of the election results, leading to the January 6, 2021, U.S. Capitol attack (Rash, 2021).
4. Anti-vaccine movements used disinformation to question the efficacy of vaccines during the COVID-19 pandemic (Wasike, 2022).
5. Russia is using disinformation in the current war in Ukraine to “defend [their] actions, seed doubt about news from the ground, and push misleading or false narratives to undercut support for Ukraine” (Bergengruen, 2023).

Looking forward, threat actors can more effectively spread disinformation to threaten national security by leveraging AI-powered bots and AI-generated deepfakes. These AI-powered bots and AI-generated deepfakes increase the speed, scope, frequency, and sophistication of disinformation, thereby exacerbating its impacts. Therefore, both domestic and foreign, state and non-state actors are empowered to “manipulate public opinion formation, degrade public trust in media and institutions, discredit political leadership, deepen societal divides as well as to influence citizens' voting decisions” (Kertsova, 2018). AI-powered bots and AI-generated deepfakes “pose a clear, present, and evolving threat to national security” (Rand Corporation, 2022).

Bots

A bot is a “software application that operates over a network and is programmed to do a specific, repetitive, predefined work task that a human would typically do” (IBM, 2023). Bots are able to automate, and therefore speed up, simple tasks that have a documented and defined sequence of steps. These tasks would otherwise have to be completed manually. This report will focus on the abilities of bots to both generate and

disseminate computational propaganda disguised as authentic accounts on social media platforms to spread disinformation.

AI enables bots to create more sophisticated content, at faster speeds and greater scales. AI also enables bots to engage in more sophisticated modes of distribution, including targeting. There are four main ways in which AI-powered bots can generate and spread disinformation more effectively:

1. AI-powered bots can generate more disinformation in a shorter time. For example, GPT–4–powered bots can generate content in seconds – a pace that cannot be matched by humans (Heaven, 2020). This allows disinformation campaigns to scale exponentially which significantly increases the volume of disinformation content online. This enables rapid disinformation attacks which can create an immediate, disruptive effect (Villasenor, 2020). As such, AI-powered bots represent a threat to Canada's national security because they enable threat actors to generate disinformation in high volumes, bombarding the social media platforms with noise (Linvill & Warren, 2021). Coping with such a high volume of disinformation presents a serious challenge to intelligence agencies (Employees, CSIS).
2. AI-powered bots can appear more like authentic accounts due to more realistic profiles and posts that disguise themselves in three ways. First, AI technologies enable bots to vary the content and wording of each post. This avoids the sort of replication detected by software designed to identify fake accounts. Second, AI-powered bots do not make the type of linguistic errors that alerts detection software for fake accounts (Renée DiResta, 2020). Third, language models like GPT-3 and GPT-4 enable these bots to closely imitate human language and the ways in which humans communicate. This makes it significantly harder for social media users and moderators to identify what is AI-powered and bot-generated disinformation from what is authentic information (Bremmer & Kupchan, 2023).
3. AI-powered bots are able to generate personalized disinformation. This increases the likelihood of users engaging with its disinformation content for two reasons. First, the use of both ML and NLP enables the bots to algorithmically generate content unique to the target individual. ML enables malicious actors to comprehensively analyze, at unprecedented speed, the profiles and activities of social media users to identify their unique characteristics, tendencies and vulnerabilities (Rosenbach & Mansted, 2019). Second, bots can be programmed with deep learning to read the emotions of the humans they interact with (Joshi, 2020). This enables the AI-powered bots to interact with the target individual in a more personalized manner.
4. AI-powered bots can disseminate disinformation in a more targeted and selective manner. ML increases "the potency of disinformation operations by enhancing the effectiveness of behavioral data tracking, audience segmentation, message targeting/testing, and systemic campaign management" (Ghosh & Scott, 2018). This enables disinformation operators to identify the exact target audience for a

particular message (Ghosh & Scott, 2018). For example, ML algorithms can identify which disinformation narratives successfully trigger responses in which geolocations (Employees, CSIS). The data gathered from this feedback loop improves the dataset with which the ML algorithm is trained on, thus improving the disinformation campaign (Employees, CSIS).

Deepfakes

Deepfakes are AI-generated, "digitally manipulated audio or visual material that is highly realistic" of events which never happened or words that were never spoken (Kertysova, 2018). These synthetically modified audio and visual materials are developed using a class of ML frameworks known as generative adversarial networks (GANs). Deepfakes include deepfake videos, deepfake images, and voice cloning.

Deepfakes are contrasted with cheap fakes which are audio and visual materials that are manually altered with simple, non-AI editing tools to mislead an audience. Cheap fakes are "rendered through Photoshop, lookalikes, re-contextualizing footage, speeding, or slowing" (Paris & Donovan, 2019). For example, one of US House of Representatives Nancy Pelosi's speeches was slowed down to make it seem as if she was slurring her words (Ho, 2019).

Deepfake videos are "altered through some form of machine learning to hybridize or generate human bodies and faces" (Paris & Donovan, 2019). Examples of deepfake videos include a 2018 deepfake of President Barack Obama using profanity and a 2022 deepfake of President Volodymyr Zelensky "appearing to tell his soldiers to lay down their arms and surrender the fight against Russia" (Mak & Temple-Raston, 2020, Allyn, 2022). Neither of these videos happened in real life. Other deepfakes are of AI-generated avatars, rather than real human beings (Satariano & Mozur, 2023).

The AI software used to create deepfakes is now easily accessible and free on GitHub, a Microsoft-owned code repository (Patterson, 2019). Due to the lack of financial and operational barriers, cybersecurity experts worry that deepfakes will be deployed by a range of actors from states to political parties to individual activists with the purpose of smearing their opposition target (Patterson, 2019).

There are three main ways in which deepfakes, created by ML, can spread disinformation more effectively:

1. Deepfakes are more compelling and persuasive than traditional forms of disinformation. In fact, researchers have found in a study of 7,000 participants that deepfake videos are more convincing than fake textual evidence of the same fake event (Wittenberg et al, 2021).
2. Deepfakes are difficult to differentiate from real, authentic content. This will become even more problematic in the future; it is predicted that AI technologies

will become so advanced by 2030 that deepfake audio and visual materials will become indistinguishable from real material (Bayer et al, 2019).

3. Deepfake images of GAN-generated avatars can be used to improve the credibility and believability of fake accounts spreading disinformation on social media platforms. These fake images are preferred to stolen images of real people as they are untraceable (Goldstein & Grossman, 2021). Facebook has identified numerous state-sponsored accounts which used GAN-generated profile photos to disguise themselves (Nimmo et al, 2019).

Applying Deepfakes and AI-Powered Bots to an Election Context

Deepfakes and AI-powered bots can be used to spread disinformation to manipulate elections by domestic and foreign state and non-state actors.

AI-powered bots	AI-powered bots could shape public opinion and control the narrative of political candidates on social media platforms. For example, AI-powered bots could be used by foreign adversaries or domestic actors to create and share computational propaganda to sway public opinion in favor of candidate A or against candidate B.
Deepfakes	Deepfakes could be used to discredit political candidates by swaying the public opinion of them (Bayer et al, European Parliament 2019). A deepfake video could be released that shows a frontrunner candidate engaging in an illegal act, making a controversial statement that is contradictory to their political position, or carrying out socially unacceptable behavior. A deepfake video could also be made to suggest election malfeasance.

Although “artificial intelligence played little role in computation propaganda campaigns to date” in the context of elections, there are “signals that AI-enabled computational propaganda and disinformation are beginning to be used” (Woolley, 2020). Once AI-powered bots and deepfakes are used, the resulting consequences of “elevating fringe candidates, peddling conspiracy theories and fake news, stoking polarization, and exacerbating extremism and even violence” will pose a serious threat to Canada’s national security (Bremmer & Kupchan, 2023). Increased disinformation will contribute to distorting democratic discourse, eroding trust in institutions, jeopardizing public safety, damaging reputations, and undermining journalism (Chesney & Citron, 2018). As disinformation becomes more widespread through AI-powered bots and deepfakes, even content that is real will be looked at through a skeptical lens. How will we know what information to trust when fake information becomes indistinguishable from real information? What are the consequences on national security defined as “action[s] or event[s] that could materially impact the health, safety, security, or economic well-being of Canadians, or the effective functioning of Canada’s governments” (Fasken, 2021)?

Threat #3 – Issues with Open-Source Intelligence

The increasing development and use of AI-enabled technologies has created new opportunities for actors outside the traditional intelligence agencies to provide insight through open-source intelligence (OSINT). OSINT is achieved by acquiring, analyzing, using, and sharing publicly available information, which has grown exponentially due to social media platforms, smartphones and the IoT (Porteous, 2022). OSINT has been improved through cloud solutions that have driven down the cost of data storage, and by AI-enabled technologies such as ML which has improved the ability to analyze and process large amounts of data (Porteous, 2022).

Due to the laws protecting the privacy of Canadians and the limitations for the Canadian intelligence community on collecting and using OSINT, threats arise when other actors become more active in this space. Although many experts have started to advocate for the increased use of AI-powered OSINT in the intelligence community, in Canada there is currently very little intelligence sourced this way (Employees, CSIS). The main national security threats that arise from OSINT in this context include an increase in the number of non-state actors participating in intelligence gathering and an increase in errors due to a lack of established standards around OSINT collection and analysis.

Increased Number of Non-State Actors in Intelligence

In previous periods, intelligence communities and law enforcement had special access to data and information not available to most other actors, but new technologies are enabling non-state actors and individuals today to participate in collecting and analyzing intelligence (Zegart, 2021).

Non-state organizations such as Bellingcat, a collective of researchers, investigators, and citizen journalists, have built a reputation as OSINT wizards (Harding, 2022, The Economist, 2021). They have been able to use OSINT to discover illegal shipping of chemical weapons precursors, identify Russia's involvement in the downing of Malaysian Airlines Flight 17 and identify the Russian officers that were suspected of poisoning two people (Harding, 2022). Because of the vast amount of data being produced every day, and the ability to introduce AI and ML to its analysis, the original understanding and significance of OSINT by intelligence communities is being challenged. Bellingcat is part of a growing ecosystem of non-state actors that have various goals and motivations and include hobbyists, journalists, activists, and conspiracy peddlers, which can create confusion with respect to the motivation behind intelligence information (Zegart, 2021).

Increased Possibility of Errors

This increased level of non-formal and non-standardized participation in the intelligence gathering space is prone to errors. This means that OSINT can both impact policy making and drain time and resources from the intelligence community which will seek to clarify and correct it. By contrast, when considering operations within CSIS, there is a

requirement to explain how information is turned into intelligence and for that analysis to be replicated under review which implies a much higher level of scrutiny on information and analysis (Employees, CSIS). Zegart (2021) provides a valuable example in the illustrated excerpt of how errors can occur in the OSINT world with negative impacts to national security.

Critical aspects of OSINT, when combined with AI-enabled technologies, include faster collection and processing of information with intelligence becoming public without much consideration for standards of disclosure. Further, errors by non-state actors in intelligence can be a challenge for national security, particularly during times of crisis. Rushing to a conclusion can be problematic for solving problems, and due to OSINT's lack of secrecy, decision makers are backed into corners with limited flexibility for de-escalation, negotiation, and compromise (Zegart, 2023).

Impressive but Wrong

"In 2008, a former Pentagon strategist named Phillip Karber was teaching a class at Georgetown University when he decided to guide his students on an open-source intelligence investigation to uncover the purpose of a massive underground tunnel system in China. The existence of the tunnels had been known for years, but their use remained uncertain. Karber's student sleuths produced a 363-page report that concluded that the tunnels were secretly hiding 3,000 nuclear weapons—which would have meant that China possessed a nuclear arsenal around ten times as large as what most experts and U.S. intelligence agencies believed, according to declassified estimates. Experts judged that the report was flat wrong and found the analysis to be riddled with egregious errors. Among them, it relied heavily on an anonymous 1995 post to an Internet forum. Nevertheless, the report was featured in a Washington Post article, was circulated among top Pentagon officials, and led to a congressional hearing in the US. It was all a wild-goose chase that consumed the most valuable resource in Washington: time."

Source: Detailed excerpt taken from Zegart, 2021

Key Takeaways Regarding AI-Enabled Threats

The three threat areas that this report describes are currently part of the threat environment, even when AI is not utilized. Therefore, it is important to highlight the unique characteristics of AI and how it can exacerbate these threats. When these AI-enabled threats are analyzed collectively, three important characteristics of AI stand out:

1. The ability to increase the speed, scope and sophistication of threat activities;
2. The importance and heavy use of all types of data to both improve threat activities and the effectiveness of AI technologies; and

3. Due to the speed, scope, and heavy use of data, there is an increased need to deploy AI-enabled defenses in order to cope with AI-enabled threats.

First, due to the ability to process vast amounts of data and information at much faster speeds than humans and current technology, the main characteristic of AI-enabled threats is that attacks by threat actors can be conducted much faster and often with more accuracy than before. Due to this increased speed of threat activity, barriers to entry for threat actors are reduced, as less labour and resources are required to assemble information and plan more sophisticated attacks. As a result of the lower barriers to entry and higher accuracy, AI-enabled threats can also increase the scope of the attacks, both in number and scale which has implications for all the national security threats defined in the *CSIS Act*.

Second, the main way in which AI capabilities are improved, particularly for ML, includes the collection and use of large amounts of data. Data that in the past may have not seemed important for national security purposes can now be analyzed in aggregate through AI-enabled technologies to provide insights and trends that would not be available to human analysts. This includes the ability to utilize data that previously could not be used to identify someone, into data that can. Therefore, the centrality of data is an important aspect of AI-enabled threats that presents relatively new challenges for CSIS. These challenges arise because of the various restrictions under the *CSIS Act, Privacy Act*, and the *Charter* that create unique barriers to addressing threats that are enabled by AI. These legislations were drafted, however, at a time where the digital revolution and AI could not have been envisioned.

Third, AI-enabled threats are more difficult to address by existing technology and defenses. Many experts have outlined that within the current trajectory of AI development and the way in which threat actors are starting to utilize it, it will become more necessary to use AI for defensive purposes against AI-enabled threats. In other words, it is expected that there will be an escalation within the current trajectory where AI will be needed to combat AI. This presents challenges for CSIS as AI activity both offensively and defensively interacts with the rights and freedoms of Canadians.

The following sections will utilize the three identified threats of cyberattacks, disinformation and OSINT combined with the AI-specific threat characteristics outlined above to analyze CSIS' ability to address national security concerns and the ways in which the rights and freedoms of Canadian's are impacted.

The *CSIS Act* and operational challenges

Part of the intelligence cycle is the ability for CSIS to detect, analyze, define, and investigate threats of all types. Canada's national security, much like that of its Five Eyes partners (intelligence partnership alliance of Australia, Canada, New Zealand, the UK, and the US), has been threatened by actors operating with modern complex technology including AI. These threats have not only evolved but also multiplied alongside existing traditional threats that have existed since the creation of the *CSIS Act*. This drastic shift in the last decade has opened the potential to greatly affect the speed at which investigative agencies need to respond to modern technological threats, more specifically those related to AI. This is primarily due to the nature of AI, which is both complex and rapidly evolving and can convolute the interpretation of threats and the scale at which they might affect the security of Canada (Employees, CSIS). Three main characteristics of AI discussed earlier directly influence the response and investigative direction that CSIS would have to take while dealing with AI enabled threats.

This section will explore the challenges that CSIS may face while dealing with AI-enabled threats within three key areas of focus: 1) interpretation and assessment of AI-enabled threats, 2) the role that dataset rules play in the collection of data, and 3) the role of the judicial oversight and warrant system. This analysis will be followed by a scenario-based case study to better illustrate impacts.

Interpretation and Assessment of AI-Enabled Threats

AI-enabled threats bring complexities that require a better understanding of both the *CSIS Act* and the rapid pace at which technological realities are evolving. The evolution of technology and variety of threat actors is increasing rapidly which requires government agencies to constantly modify and adopt revised approaches to threat detection.

Further, critical infrastructure across the government and private sector has increasingly incorporated AI-reliant algorithms, which remain susceptible to sabotage by both domestic and foreign adversaries (Alhajjar, 2022). The speed at which investigative agencies can operate is largely determined by the scope of threats, their intelligence cycle, and the legal operational framework. The speed, scope, and clandestine nature of many AI-enabled threat activities pose serious problems to CSIS' existing approaches to identifying threats.

The scope of operations for CSIS is well established through existing operational guidelines on how to interpret, investigate, react, and respond to threats prior to the application of threat reduction measures (TRMs). The driving component for interpreting and analyzing threats starts with identifying which of the four definitions of threats to the

security of Canada outlined in s. 2 of the *CSIS Act* is being impacted (*CSIS Act,* drafted in *1984*):

a. “espionage or sabotage that is against Canada or is detrimental to the interests of Canada or activities directed toward or in support of such espionage or sabotage,
b. foreign influenced activities within or relating to Canada that are detrimental to the interests of Canada and are clandestine or deceptive or involve a threat to any person,
c. activities within or relating to Canada directed toward or in support of the threat or use of acts of serious violence against persons or property for the purpose of achieving a political, religious or ideological objective within Canada or a foreign state, and
d. activities directed toward undermining by covert unlawful acts or directed toward or intended ultimately to lead to the destruction or overthrow by violence of, the constitutionally established system of government in Canada,

but does not include lawful advocacy, protest or dissent, unless carried on with any of the activities referred to in paragraphs (a) to (d)”.

These definitions are broad enough to capture many activities and can seem to overlap depending on the activity. Further, adversarial activities could fall under two main categories – foreign and domestic. More important for this report, there is no specific mention or indication of cyberthreats, leaving CSIS operations with a broad enough definition of what constitutes a threat to national security, but at the same time open for contestation from both internal government stakeholders and the public (Employees, CSIS). When placing certain threats within the realm of AI, CSIS is presented with challenges for interpreting a threat and how it impacts national security because of the speed, scope, and access to data that AI-enabled threats possess.

The challenge for CSIS when dealing with AI-enabled threats therefore is whether it has the mandate and operational capacity to identify and act on them. Can it investigate such threats as quickly as the threat activities themselves develop? Can it determine who the actual adversaries are, where they are located and if they fall under foreign or domestic categories? For example, an uptick in disinformation activity by AI-bots on a social media platform may act as a trigger point or catalyst for disruptive events to occur within Canada. In this case, at which point does the activity become a threat and under which of the four definitions does it fall? Additionally, adding a layer of complexity, when can foreign influence be determined while working within an AI-enabled digital space?

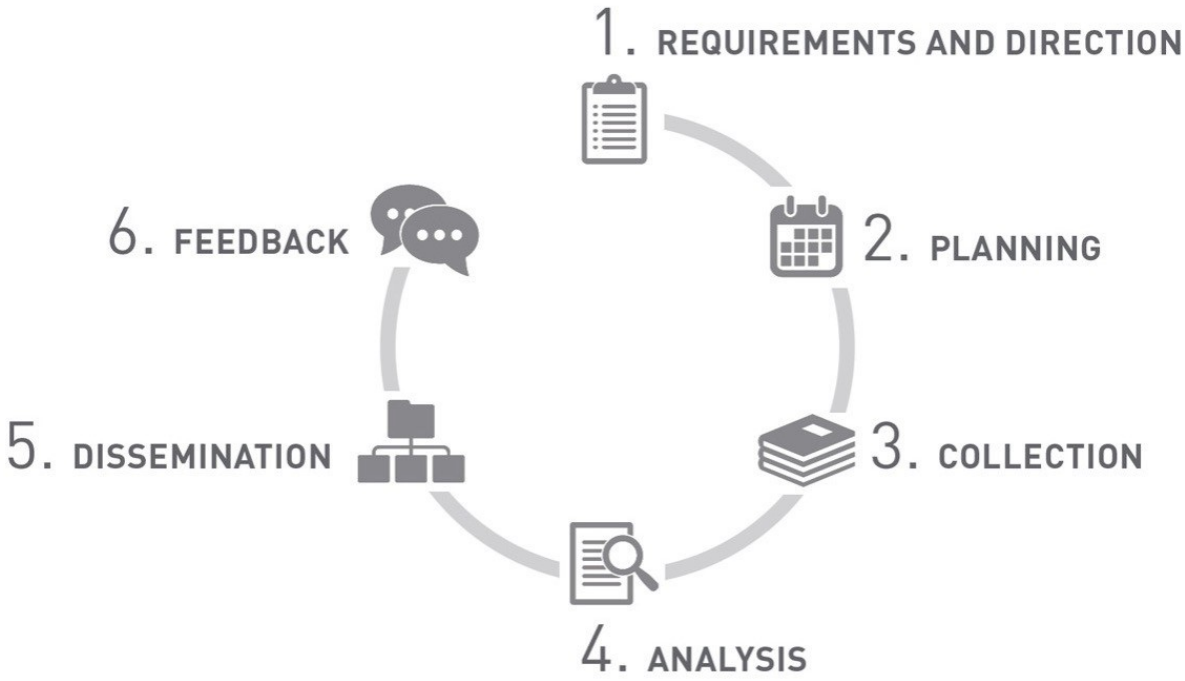
According to the Security Intelligence Review Committee (SIRC) (now National Security and Intelligence Review Agency), there are requirements that perceived threat activity assessments must cover four specific categories in two separate groups. For clarity they are as follows:

Group A	Group B
• Foreign influenced; • within or relating to Canada; • **clandestine *or* deceptive;** and • detrimental to the interests of Canada; or	• Foreign influenced; • within or relating to Canada; • detrimental to the interests of Canada; and • **involve a threat to any person.**

Each of these terminologies are relatively open-ended which is an aspect that makes the *CSIS Act* one of the more unique in Canadian legislation. SIRC has also noted that the terminology, "detrimental to the interests of Canada" could prove to be problematic. Canadian interests are broad and subjective, which allow decision makers a level of discretion and CSIS with a higher degree of responsibility to interpret threats enabled by AI. Additionally, other terms used within Group A and Group B such as "clandestine or deceptive" can also pose challenges to what this constitutes during the assessment of this criteria for CSIS. How would "clandestine or deceptive" differ from "espionage" in (a) and (b) (SIRC, 2010)?

It is important to note that with greater ambiguity in interpreting and defining threats to the national security of Canada, there are greater complexities for CSIS to place (and justify) threats enabled by AI under those definitions. Because this is a critical part of determining the scope and investigative mandate for CSIS, it further impacts the full intelligence cycle.

Figure 9: CSIS' Intelligence cycle (CSIS, 2023)



Datasets and Collection of Data

An essential component of CSIS' intelligence cycle is collection (of data) which directly influences the analytical outputs in the next stage (analysis). As discussed earlier in the report, data is of critical importance to AI for both its improvement and processing. Data also has unique qualities which make it difficult to determine ownership and jurisdiction. Datasets "collected" under ss. 11.01 to 11.25 of the *CSIS Act* can be broadly classified into three categories – Canadian, foreign, and publicly available (*CSIS Act,* 1984*, National Security Act,* 2017).

For CSIS, one of the key elements of investigating AI-enabled threats is the "collection" of data which is based upon "necessity" for the sake of investigation only. This is related to authorities outlined in section 12 of the *CSIS Act* which allows for collection of information and intelligence based on reasonable grounds, directly relating to activities that may pose a threat to the national security of Canada. In this regard, the nature, particularly its origin is not addressed in this section, with a greater emphasis on appropriate interpretation of thresholds requirements (that is left upon CSIS to assess).

Alternatively, limitations also exist when it comes to accessing data related to Canadian citizens and permanent residents and those which are publicly available. The application of these datasets is outlined in Section. 11.01-11.25 of the *CSIS Act* and applies to "all datasets that contain personal information as per section 3 of the *Privacy Act* and that do not directly and immediately relate to activities that represent a threat to the security of Canada" (*CSIS Act,* 1984). It is also important to note that this section also allows for the collection of datasets not directly related to activities that threaten national security. Furthermore, these also 'assist' CSIS' operations in narrowing down investigative leads.

Apart from pre-determined (by the Minister) classes of datasets that CSIS is allowed to collect which must be approved by the Intelligence Commissioner, it also has further dataset collection restrictions under subsection 11.05 (1) and (2) of the *CSIS Act*. They are as follows:

11.05 (1): "Subject to subsection (2), the Service may collect a dataset if it is satisfied that the dataset is relevant to the performance of its duties and functions under sections 12 to 16" (*CSIS Act*, 1984).

11.05 (2): "The Service may collect a dataset only if it reasonably believes that the dataset

a. is a publicly available dataset;
b. belongs to an approved class; or

c. predominantly relates to non-Canadians who are outside Canada" *(CSIS Act,* 1984*).*

In its current structure, CSIS is restricted in its ability to collect, analyze, and retain data. An increase in the number of internet users and many of its applications has led to an overall increase in the amount of data and datasets that are publicly available.

To generate threat-related assessments and inferences using this data, subsection 11.05 (1) & (2) in particular (c) add ambiguity for CSIS when defining what may or may not be "outside of Canada". This coupled with the nature of modern datasets generated by AI (or manipulated) with additional layers of complexities such as the format, language and encryption with which they are obtained.

Furthermore, the use of the word 'predominantly' in this section is vague when it comes to dataset collection, refinement, and analysis. In this context, 'predominantly' refers to datasets that contain information which is primarily foreign in nature and as such does not contain personal information that may belong to Canadians. Therefore, refinement and sanitization of datasets such as segregation of data, to conform with subsection 2 (a) (b) or (c) may cause operational delays including the use of certain datasets that can help with investigative leads. An example would be the use of OSINT data where publicly available information is being leveraged by threat actors to launch AI-enabled threats. Additionally, in terms of establishing 'predominance' to the extent which Canadian data is blended within acquired datasets, section 11.07 (1) obligates CSIS to evaluate the nature of data. The criteria for that evaluation as per that section are as follows:

11.07 (1) – "If the Service collects a dataset under subsection 11.05(1), a designated employee shall, as soon as feasible but no later than the 90th day after the day on which the dataset was collected, evaluate the dataset and confirm if it

a. was publicly available at the time of collection;
b. predominantly relates to individuals within Canada or Canadians; or
c. predominantly relates to individuals who are not Canadians and who are outside Canada or corporations that were not incorporated or continued under the laws of Canada and who are outside Canada" (*CSIS Act*, 1984).

These additional conditions within the *CSIS Act* do not match with how data presents itself, thereby impacting effective exploitation of composite datasets, particularly when AI has the capacity to further augment those datasets. Furthermore, applying s. 11.1 (c), therefore, would entail the modification of acquired datasets which would create two different classes of datasets with separate operative procedures (with regards to retainment, exploitation and warrant application).

Given the increase in the number of users on various social media platforms, including threat actors, personal information is being publicly shared, which connects with the higher availability of sensitive information online. As per CSIS' Data Acquisition Program (DAP) and framework of data collection, the service acquires bulk data for two broad purposes – referential and non-referential. Referential refers to data that is openly available and does not constitute 'collection' within the *CSIS Act.* In contrast non-referential data is collected within the authority of the *CSIS Act* according to internally established thresholds of necessity (SIRC, 2015). AI-enabled threats largely operate on open-source platforms where data is often heterogenous and therefore would largely fall under the referential category for CSIS, limiting their capacity to analytically exploit that data. While CSIS cannot 'collect' and utilize this type of data (referential) without relevant authorization/warrant, non-state adversaries can and are increasingly using AI to operationalize the information.

Actionable inputs collected from various sources require reasonable grounds for collection, analysis, and retention, particularly with threats related to AI which impact various domains such as digital, physical, and political security (Brundage et al, 2018).

Similarly, the Communications Security Establishment (CSE) obtains information on Canadians or those residing in Canada during its collection of foreign signals intelligence (SIGINT) and through Canadian Identifying Information (CII) which includes personal information such as names, emails, and IP addresses (NSIRA, 2021). Section 4 of the *Privacy Act* (1985) for CSIS requires that this type of personal information is removed before analysis, where this information does not directly relay to one of CSIS' operating programs or activities, puts considerable obligations on CSIS to refine data collected to launch investigations. Therefore, data manipulation techniques such as data poisoning by threat actors may result in additional complexities for CSIS (with obligations and responsibilities outlined in the *CSIS Act* and *Privacy Act*) in examining data and creating deviations in threat-modelling outcomes (Marshall et al., 2019). Further, a variety of consultative processes such as post-examination assessment of political, operational, foreign relations and legal risks of proposed TRMs may position CSIS at a disadvantage while dealing with AI-enabled threats.

Exploring possible CSIS responses to national security threats, TRMs can be classified into three categories: 1) messaging, 2) leveraging, and 3) interference (NSIRA, 2020). The establishment of reasonableness and proportionality of TRM application has been a cause for concern, as highlighted by the National Security Intelligence Review Agency (NSIRA) in its 2020 TRM review report. Considering the broad concerns that were highlighted in the report and thinking about AI-enabled threats, the requirement for CSIS to establish reasonableness and rational link in the selection of individuals suspected of posing threats may overlap with aspects of privacy, rights and freedoms outlined in the Charter.

With the advancement of AI-enabled threats and the nature and importance of data, authorities within CSIS require a clear embedded framework within the *CSIS Act* to effectively address AI-enabled threats. The complex authorities outlined in this section with regards to dataset (bulk) classification, retention and the compliance requirements present themselves as inordinate barriers. These are legislated in silo and therefore augment operational responses (adding statutory limitations) to AI-enabled threats. The interpretation of these sections in addition to internal operating procedures and principles would require additional justifications (in terms of application of TRMs based on exigencies and intrusiveness).

Warrants and Judicial Oversight

A core part of balancing CSIS's mandate against fundamental rights and freedoms guaranteed in the *Canadian Charter of Rights and Freedoms* (*Charter*) falls under judicial oversight through the issuance of investigative warrants. However, the ability to respond to AI-enabled threats remains restricted for CSIS due to the existing regime of the warrants system.

As a civilian agency, the *CSIS Act* allows for only one type of warrant which covers all types of threats that CSIS is mandated to investigate and address (*CSIS Act,* 1984). By way of comparison, the RCMP can request various types of warrants specified in *part 15* of the *Criminal Code of Canada* (*Criminal Code of Canada,* 1892). There remains a major difference in the powers granted to each of these agencies to carry out investigations of AI-enabled threats with overlapping mandates and different warrants system. The warrants system for CSIS currently acts as a one-glove-fits-all approach. But as AI develops, the speed of AI-enabled threats continues to increase, which in turn reduces the time available to respond (Allen & Chan, 2017).

According to a CSIS employee familiar with the *CSIS Act*, the current form of legislation was drafted to accommodate all types of threats, however, lacks concrete reliability to narrow down modern-day threats related to AI and other technologies. Therefore, certain sections in the *CSIS Act*, specifically those relating to operational responses through the application of warrants – *part 2 section 21 subsections 2-3-3.1,4 and 4.1*, while necessary, may pose challenges to CSIS's operations.

Another limitation with respect to application of warrants are the obligations outlined in *part 2 section 21.1 (1.1) – 2 – d, e, f* of the *CSIS Act* which state that applications for retention of data requires multiple conditions including clarification of classes of person the warrant is to be directed towards, and a general description of the place where the proposed warrant is to be executed, if it can be provided. With the nature of AI threats, the fulfilment of these conditions increases the layers of obligations upon CSIS in the backdrop of rapid development of both existing threats and newer ones (Employee, CSIS). As such, the obligation to comply with the *CSIS Act* compounds for CSIS while clarifying definitions of threats "within" or "relating to Canada". AI-enabled threats,

especially those containing aspects of collaboration between domestic and foreign threat actors have the potential to amplify challenges for CSIS across all three of the aspects discussed above, placing the threat under a definition, collecting data, and applying for warrants. Especially when considering amendments to legislation such as the *Privacy Act,* and the proposed *Artificial Intelligence and Data Act.*

The *Privacy Act* for example, was first passed into Canadian federal law in 1985 and was done so to extend the reach of laws that were already legislated to protect the privacy of Canadians, but also allow Canadians to have access to personal information about themselves (*Privacy Act*, 1985). Due to AI's heavy reliance on data, the *Privacy Act* interacts with how CSIS investigates AI-enabled threats. Within these investigative activities, the use of personal information plays an important role, and thus the term "identifiable information" is critical.

Within the *Privacy Act*, there are numerous definitions and examples of what "identifiable information" is. This term is defined as information that can be used to identify an individual, and that is recorded in any form including, but not limited to the following (*Privacy Act,* 1985):

1. *An individual's race, national or ethnic origin, colour, religion, age, or marital status.*
2. *Education, medical, criminal or employment history, or information regarding financial transactions.*
3. *Any identifying number, symbol or other assigned to the individual.*
4. *A person's address, fingerprints, or blood type.*
5. *Personal opinions or views of the individual except where they are about another individual or about a proposal for a grant, an award, or a prize to be made to another individual by a government institution or a part of a government institution specified in the regulations.*
6. *Correspondence sent to a government institution by the individual that is implicitly or explicitly of a private or confidential nature and replies to such correspondence that would reveal the contents of the original correspondence.*
7. *The views or opinions of another individual about the individual.*
8. *The views or opinions of another individual about a proposal for a grant, an award, or a prize to be made to the individual by an institution or a part of an institution referred to in paragraph, but excluding the name of the other individual where it appears with the views or opinions of the other individual.*
9. *The name of the individual where it appears with other personal information relating to the individual or where the disclosure of the name itself would reveal information about the individual.*

Much like the *CSIS Act*, the *Privacy Act* faces similar issues of applicability in the face of new technology. Digital transformation has completely transformed the understanding that the average Canadian has about their personal information, and how this

information is considered data. The way that this personal information flows is fast, elaborate, and universal, and as such what is included in any definition of “identifiable information” has widened drastically (Department of Justice Canada, 2019). “Identifiable information” is not the same as what is considered “basic subscriber information”, which is information about a person that is publicly available, like finding an individual's address in a phone book. This information is currently any government organization's main source of information on someone and needs to be distinguished from the “identifiable information” mentioned above (Government of Canada, 2016).

The widened threat landscape and properties with which AI-enabled threats function are likely to continue to challenge CSIS and its operations. Coupled with the various legislations surrounding privacy and data regulation and the need to identify who those threat actors are, will intrinsically make it difficult for CSIS to detect and respond to AI-enabled threats. To better understand how the above analysis can unfold, the hypothetical scenario below will briefly examine various aspects of AI-enabled threats and how CSIS would respond.

Navigating the Complexities: A Hypothetical Scenario of AI-Enabled Threats to Canadian Democracy

Context

Foreign interference in elections is not new. This has been observed in elections in several countries including the US Presidential Election in 2016 where Russia used disinformation to polarize voters (Rand Corporation, 2022). Given the increase in foreign interference in Canada's elections since 2021, authorities in Canada such as, but not limited to CSIS have been tasked with overseeing elections and closely monitoring online activities related to it. The task, however, is more challenging than it appears as it requires CSIS to also consider if general disinformation and polarization can translate to ideologically motivated extremist events.

Hypothetical Scenario

Federal elections in Canada are slated to take place in October 2025. According to the rules and regulations set out by Elections Canada, voters both within and outside of Canada, will have the option to cast ballots electronically online through a specialized portal (e-voting). A large part of canvassing by candidates across all major political parties will also occur on platforms online (predominantly on social media platforms). Past incidents in other countries such as Nigeria and in the case of the Brexit referendum in Britain have proved that online media platforms can be leveraged using AI-enabled techniques to create, shift and drive a particular narrative that could potentially be detrimental to the core values, beliefs, and democratic system upon which the country is built.

Threat

A group of foreign threat actors plan to utilize AI-enabled technologies to polarize and propagate extremist narratives and undermine Canadians' trust in their democratic institutions. Planning on influencing elections in Canada both directly and indirectly, these threat actors will rely on various AI-enabled technologies.

1. Manipulate information available online by constantly building or stitching elements of extremist narratives into existing information (Brundage et al, 2018).
2. Approaching influencers/personalities with a large social media audience (AI-enabled analysis to determine the targets).
3. Denial of information (DoS) attacks where bots are used to generate misinformation narratives (Brundage et al, 2018).
4. Automated hyper personalized disinformation campaigns (Brundage et al, 2018).

Threat actors also are keen to use other AI-enabled threats such as data-poisoning which can be used to influence opinion polling. AI-enabled disinformation campaigns through targeted deployment of AI-bots across public and private platforms stands to essentially amplify misleading campaigns/narratives.

These can be used during elections on social media platforms such as Twitter and Facebook where re-tweets and the number of shares by these automated bots can amplify the popularity of that misleading narrative. Additional components at risk during this scenario are databases of voters (registry), digital voting infrastructure to facilitate online voting and associated IT systems. All these remain vulnerable to AI-enabled threats discussed earlier in the report.

Analysis

For CSIS it is important to understand the motivation of the threat actors in this case, which could be both organized and sponsored by adversarial states. The ability for threat actors to more effectively remain anonymous, and the higher use of social media, make AI-enabled threats more effective. Furthermore, the internet and social media platforms such as Facebook, Twitter, and TikTok, have made the availability and access to important information much easier than traditional media. This has enabled threat actors situated in foreign locations to carry out activities that are clandestine and deceptive in nature.

The first step in the scenario above is for threat actors to outline the vulnerabilities within the socio-political fabric of the country during elections, narrowing down on polarizing topics and building disinformation campaigns. Specifically tailored messages can also be targeted towards a specific community known as “community-targeted spam” (Brundage et al, 2018). As such, AI can be used both to avoid detection and help incorporate ML and NLP to develop wider and more effective disinformation campaigns during elections.

Because these activities can fall under freedom of speech, operational responses to this type of threat would challenge CSIS’ ability to effectively frame this scenario within the context of threats to national security. Furthermore, the high-profile nature of elections and large media focus means that consideration must be given to public option and the political dynamics of any intervention. Therefore, deployment of TRMs specifically for interference (assuming it’s literal translation as a measure in itself) would invite backlash and resistance across all stakeholders in the election process. The nature of DoS attacks and AI-bots involves a large number of social media profiles that are used to push a narrative. For CSIS to investigate this, it becomes challenging because amongst the bots there could be real Canadian users and their personal information. These investigations would require justification though the warrant application before the court. But as discussed in the previous section, this would require CSIS to also differentiate between foreign and domestic data. Further addressing this hypothetical scenario includes issues with voting online, which increases the scope for CSIS to investigate, amplifying the requirement for “predominantly” domestic datasets and in turn increasing the risk of breaching the *Charter*.

As such, scenarios such as these and many more are poised to challenge CSIS operations with the onset of newer AI-enabled threats. Navigating these complexities under the current framework of the CSIS Act can become increasingly complex for CSIS, particularly when intersecting civil liberties are involved.

AI-Enabled Threats and Canadian Rights and Freedoms

As outlined in the previous section, there are various federal government legislations and agreements which include principles that are important to consider when analyzing the AI-enabled threats highlighted in this report: cyberattacks, disinformation, and OSINT. Due to the speed, scope and sophistication of AI-enabled threats, there is also an expectation that AI will be required for defensive purposes to address these AI-enabled threats. Within this context of offensive (threat actors) and defensive (CSIS) uses of AI, this section analyzes the ways in which it impacts the rights and freedoms of Canadians. Below is a high-level summary that outlines which rights, freedoms, and principles (including the various legislations) are impacted by the three AI-enabled threats.

Summary – How AI-enabled Threat-Related Activities Interact with Individual Rights and Freedoms

Type of Threat	Explanation of Threat	Rights, freedoms, and principles effected
Cyberattacks	Threat actors attempt to access computer systems to steal, expose, alter or destroy information. Cyberattacks by foreign or domestic actors can contribute to espionage, sabotage, foreign influence, and terrorism activities.	Privacy, as defined in: • The *Privacy Act* • *Universal Declaration of Human Rights (UDHR)* • *Guiding Principles on the Responsible Use of AI* • *Montreal Declaration.* • *Artificial Intelligence and Data (AIDA)* Security of Person, as defined in: • *Canadian Charter of Rights and Freedoms (Charter)* • *UDHR* • *Montreal Declaration* Transparency, as defined in: • Paradox with national security • *Guiding Principles on the Responsible Use of AI*
Disinformation Threats	Disinformation threats such as deepfakes will likely become a foreign interference tactic used to	Security of Person, as defined in: • The *Charter* • *UDHR* • *Montreal Declaration*

	spread disinformation during election periods. Deepfakes are “digitally manipulated audio or visual material that is highly realistic and virtually indistinguishable from real material” powered by deep learning, a subset of ML.	• *AIDA* Search and Seizure, as defined in: • The *Charter* Good Data, as defined in: • AI Ethics • *Guiding Principles on the Responsible Use of AI* Freedom of Expression, as defined in: • The *Charter* • *UDHR* Equity and Equality: • The *Charter* • *Montreal Declaration*
OSINT	The accessibility and availability of publicly available information have increased exponentially through the use of social media platforms, smartphones, and the internet of things. As a result, an increasing number of actors worldwide are acquiring, analyzing, using, and sharing such information. OSINT presents national security threats due to the involvement of numerous actors with varying motives and loyalties, its informal nature, which increases the potential for errors, and its public nature, which reduces the government's ability to compromise with adversaries.	Good Data, as defined in: • AI Ethics • *Guiding Principles on the Responsible Use of AI* Transparency, as defined in: • Paradox with national security • Pillars of Good data Privacy, as defined in: • The *Privacy Act* • *UDHR* • *Guiding Principles on the Responsible Use of AI* • *Montreal Declaration* • *AIDA* Freedom of Expression, as defined in: • The *Charter* • *UDHR* Prudence, as defined in: • *Montreal Declaration* Equity and Equality: • The *Charter* • *Montreal Declaration*

The Canadian Charter of Rights and Freedoms

It is impossible to talk about rights and freedoms in the Canadian context without first discussing the *Charter.* As such, it is the obvious place to start when we are attempting to understand how AI-enabled activities interact with the rights and freedoms of Canadians.

The *Charter* adopted in 1982, and its 32 sections, "guarantees the rights and freedoms set out in it subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society." (*Canadian Charter of Rights and Freedoms,* 1982*).* Not all 32 of these sections are relevant to AI-enabled activities. To narrow this selection down, several interviewees with relevant *Charter* experience were asked which sections of the *Charter* they believed were most relevant in relation to CSIS, its mandate, and AI-enabled activities. Those participants including a counselor from the Justice Department and an official from the Privy Council Office, highlighted sections 2, 7, 8, 15, and 32 as the key sections of the *Charter* that contribute to the dynamic between AI-enabled activities, threats, and rights and freedoms.

The *Charter* applies only to actions by government organizations, not companies or individuals. CSIS must balance the imperative of protecting the *Charter and* privacy rights of Canadians with the imperative of protecting the security of Canada and Canadians. The attempt at balancing these imperatives may, however, be impacted by AI-enabled threats. While the legal relevance of AI-enabled activities only applies to actions taken by CSIS, the principles remain relevant to the interaction between AI-enabled threats, and rights and freedoms.

Relevant *Charter* Rights and Freedoms	Why it's relevant to AI-enabled activities	Example
Section 2 – 2b: Freedom of expression, 2c: Freedom of peaceful assembly, and 2d: Freedom of association	AI could be used to undermine anonymity of the crowd and target social movements and marginalized communities.	The increasing trajectory of AI technology suggests that it will eventually need to be used defensively, resulting in a scenario where AI combats AI. As per the example used in the previous section of this report, if CSIS were trying to investigate discourse on social media platforms that may (or may not) be foreign interference, then this could impinge on S. 2 of the *Charter*.
Section 7 – Life, liberty,	While CSIS is not a policing agency or organization, case	S. 7 of the *Charter* could be impacted by AI-enabled activities

and security of person	law interpretation around the definition of security has been moving to widen the scope of what security of a person is (Counselor, Justice Canada). Security can be interpreted to mean health, safety, and personal autonomy, (*R. v. Morgentaler* (No. 2), [1988] 1 S.C.R. 30; *Carter v. Canada*, 2015 SCC 5). Given that AI-enabled activities could be used by both CSIS and threat actors to undermine this security, S. 7 becomes relevant.	by CSIS if attempts to investigate national security threats lead to using AI to assess the possibility of recidivism in previously charged criminals (Morgan et al, 2023).
Section 8 – Search and Seizure	S. 8 of the *Charter* states that Canadians have the right to be "secure against unreasonable search or seizure", which limits the options available to CSIS to obtain evidence of wrongdoing. This links with privacy rights, and it is applicable to AI-enabled activities given the reliance of these activities on data, specifically "identifiable information". Case law has also allowed for S. 8 to be applied to electronic information because of R. v. Spencer, 2014 SCC 43, thus opening it up to the realms of AI and its algorithms.	As AI-enabled deepfakes start to become more prevalent, it will be harder for CSIS to know the difference between real and deep faked video. Failure to comprehend the distinction between authentic and deep faked videos may result in the unintentional utilization of deep fakes as evidence during investigations. This could result in the violation of S.8 in the event that a search warrant is issued based on false information. This may worsen the already discussed problem of delays in warrant approval and may encourage ineffective measures that increase the risk of successful AI-enabled threats.
Section 9 – Right not to be "arbitrarily detained or imprisoned"	S. 9 of the *Charter* faces similar dynamics with AI-enabled activities as S. 8 in that it comes down to how information is gathered, and the data or	Law enforcement agencies utilizing AI systems which involve technologies such as facial-recognition or profiling, are required to demonstrate sufficient justification for suspecting an

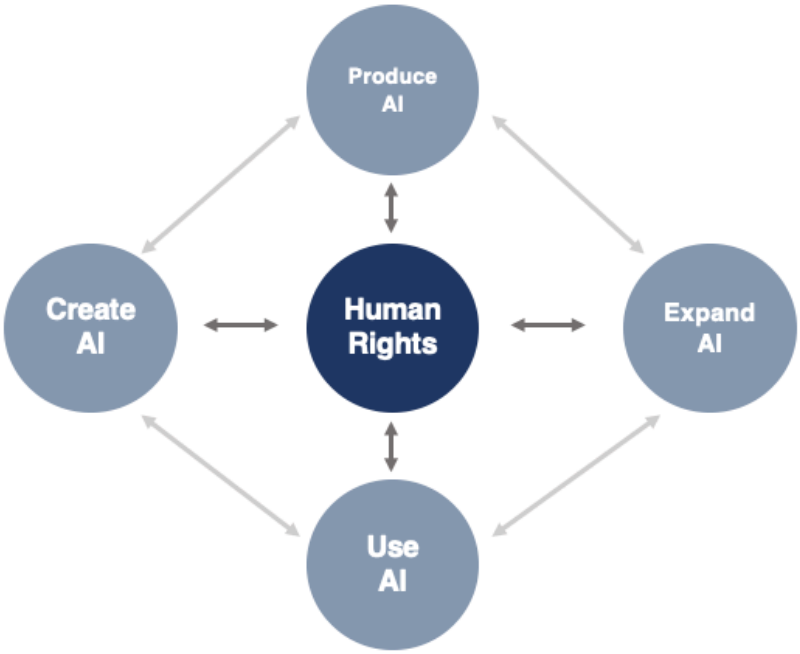
	algorithms that were used in the investigation.	individual's involvement in a crime. In addition, utilization of biased data that results in unintentional racial profiling may lead to unconstitutional detention. This is because the Supreme Court of Canada has ruled that the utilization of racial profiling during the detention of an individual may be considered when determining whether such detention is arbitrary, as demonstrated in the case of R. v. Le, 2019 SCC 34 (Morgan et al, 2023).
Section 15 – Equality before and under law and equal protection and benefit of law	Algorithmic bias is a pervasive issue that affects various sectors, including healthcare, criminal justice, and finance. (Obermeyer, 2021) If this sort of data were to be used within an investigation, it could leave CSIS without the ability to use the information because doing so would violate the C*harter*.	When AI is used to provide insights and trends based on information that human analysts would not have been able to develop into identifiable information, there is the possibility that false or biased data is used to make decisions. If this happens it could lead to biases within AI-enabled activities by CSIS, thus violating S. 15 of the *Charter.*
Section 32 – Application of Charter	Data that is used for, or accessed by AI is often stored on servers outside of Canada, even if the target of an investigation is within Canada. As such, current interpretations of S. 32 affect the ability to apply charter rights to extraterritorial activities of Canadian government actions.	R. v. McGregor, 2023 SCC 4 provides an example of a supreme court decision in which the court had to rule on whether a Canadians S. 32 rights are applicable when Canadian government organizations are acting abroad. Given that data can often be based offshore, even if it's Canadian data, this discussion becomes more relevant as more AI-enabled activities and threats based on the use of data become mainstream (Counselor, Justice Canada).

Universal Declaration of Human Rights

Three articles from the *UDHR* will be discussed. Two of the three *UDHR* Articles were chosen for this report based on a rapid review by Mpinga et al. (2022) of 157 academic articles that studied AI and human rights in some fashion. This was done in order to understand if there is an emerging academic discipline focusing on that dynamic. Within this review, the authors found that Articles 3 and 12 of the *UDHR* were the most addressed across these 157 articles, with the right to life appearing in 21% of the articles, the right to security/safety appearing in 19%, and the protection of privacy appearing in 14% of the articles. The authors also proposed a conceptual framework for the mutual and dynamic linkages between AI and human rights. Mpinga et al. (2022) argue that human rights take a central stage while AI is an evolving reality around them. This reality depends on where in the life cycle AI is, either creation, production, commercialization, or utilization. In this framework, each of these stages is seen as having an impact on human rights and are also feeding back on each other.

In the context of this report, the utilization of AI is most important, given the focus on AI-enabled activities. Figure 10 shows that the utilization of AI impacts both the creation and expansion of AI, as well as being impacted by them, with a similar relationship with human rights. This means that CSIS' own use of AI and the activities it enables requires recognition that it not only impacts human rights but also plays a role in the expansion and creation of AI technologies. This can widen the possibility of threats from actors willing to use AI-enabled activities to pursue their goals. These dynamics increase the threat surface by allowing more access to information and providing more opportunities to threat actors.

Figure 10: Framework of AI and Human Rights (Mpinga et al, 2022)



The accessibility of internal government networks to threat actors has significantly increased with the evolution of technology. In the past, physical access to buildings was necessary to siphon digital information from vulnerable organizations. With the introduction of the internet and Wi-Fi, the attack surface increased as networks became more widely accessible beyond the physical realm. As government organizations,

including CSIS, adopt AI technology for internal and investigative purposes, the threat surface is likely to widen further (Partner, EY). While AI technologies, and in turn, threats will continue to grow regardless of CSIS' actions, the use of AI by CSIS could potentially lead to the creation and expansion of other AI technologies by threat actors, resulting in a wider attack surface and possible compromise of human rights.

Guiding Principles on the Responsible Use of AI

As AI has become more pervasive throughout society, governments have attempted to create best practices by developing principles and guidelines that allow them to navigate this new landscape without the requirement of formal legislation. The Government of Canada's 2018 *Guiding Principles on the Responsible Use of AI* is a great example of this. While originally developed in relation to all AI use by federal government organizations, four of the five principles can be used to identify and prioritize how CSIS can both use and deal with AI-enabled activities in the national security setting.

Relevant *Guiding Principles on the Responsible Use of AI*	Why it's relevant to AI-enabled activities	Example
Understand and measure the impact of using AI by developing and sharing tools and approaches.	CSIS must evaluate the impact of AI activities, considering the potential risks to identifiable Canadian information and increased attack surface. Sharing tools and approaches may increase the risk of cyberattacks, requiring a balance with security measures. However, sharing tools and approaches can aid in evaluating their impact, improving understanding of CSIS systems' effectiveness and potential discrimination, and reducing the risk of AI-enabled threats infringing on rights and freedoms.	CSIS sharing tools and approaches, even in an effort to increase transparency, could potentially provide threat actors engaged in AI-enabled activities with access points. Such vulnerabilities in software contribute to increasing the speed and scope of AI-enabled threats that compromise CSIS's security.
Be transparent about how and	Transparency is a vital principle in national security	In the case that an AI vs. AI scenario, CSIS would be

when we are using AI, starting with a clear user need and public benefit.	and AI fields, specifically regarding the data collection and usage in AI-related activities of CSIS. Even with the public's limited tech knowledge and understanding, demonstrating precise user needs and public benefit in national security can be challenging. Moreover, CSIS's requirement to keep much of its internal operations' secret further compounds the difficulty. (Privy Council Office, 2019).	required to show a clear user need and public benefit for participating in this AI-enabled activity. Doing so could widen the attack surface for threat actors that use cyberattacks such as spear phishing, or data poisoning that could lead to infringements on rights and freedoms.
Provide meaningful explanations about AI decision making, while also offering opportunities to review results and challenge these decisions.	The accountability principle pertains to the responsibility of AI and the decision-making process for its utilization. Although challenging to implement in the context of national security, it can be achieved through existing privacy legislation (Counselor, Justice Canada).	Effective AI-enabled defenses against AI-enabled threats require CSIS to implement review mechanisms capable of fast decision-making given the speed, scope, and sophistication of these threats. A new warrant that permits quick action when investigating AI-enabled threats can provide more opportunities for meaningful explanations, enabling better utilization of AI by CSIS, and reducing the likelihood of AI-enabled activities and threats negatively impacting rights and freedoms.
Be as open as we can by sharing source code, training data, and other relevant information, all while protecting	While this principle focuses on national security and defense, the dynamics of data sharing suggest that it will likely become increasingly relevant as CSIS expands its use of AI technology.	Sharing source code increases the risk of data leaks, which could be exploited by threat actors to compromise the security of individuals' rights and freedoms. In this context,

personal information, system integration, and national security and defense.		CSIS must exercise precautionary decision-making principles to protect the rights of Canadians from AI-enabled threats.

Montreal Declaration for a Responsible Development of Artificial Intelligence

The *Montreal Declaration* was established in 2018, following a stakeholder engagement process that included 15 workshops over three months. More than 500 experts, citizens, and stakeholders with diverse backgrounds contributed to the development of ten principles that should guide the use of AI systems. These principles are aimed at promoting and preserving the interests of people and groups and include the following (Université de Montréal, 2018):

1. Well-Being
2. Respect for Autonomy
3. Protection of Privacy and Intimacy
4. Solidarity
5. Democratic Participation
6. Equity
7. Diversity Inclusion
8. Prudence
9. Responsibility
10. Sustainable Development

The *Montreal Declaration* principles are applicable to the development and deployment of AI in various fields, including national security. As the deployment of AI leads to the creation and expansion of AI-enabled activities and threats, the principles are relevant to the national security scope. More specifically for the purposes of this report the significant principles are 3, 5, 6, and 8.

Relevant Principles	What is it?	Why it's relevant to AI-enabled activities or threats	Example
Principle 3: Protection of Privacy and Intimacy Principle	Privacy and intimacy must be protected from AI's intrusion and data	Privacy is one of the most relevant values when it comes to the use of AI by any party, including CSIS and other government organizations.	While threat actors using cyberattacks such as ransomware or spear phishing currently are mostly interested in the use

	acquisition and archiving systems (DAAS).	Given AI's requirement to use significant amounts of data, the kind of data that can be amalgamated to identify a person has expanded.	of data collection to raise funds, steal intellectual property, and against political dissidents. These activities enhanced by AI in the hands of threat actors could cause infringements on privacy rights aligned with this principle.
Principle 5: Democratic Participation Principle	AI's must meet intelligibility, justifiability, and accessibility criteria, and must be subjected to democratic scrutiny, debate, and control.	There are two issues surrounding the use of algorithms and AI software by CSIS: Who has ownership? Who is liable when things go wrong? The decision-making algorithms used by public authorities, including CSIS, should be accessible, except in cases where it could pose a risk. This principle aligns with CSIS's use of AI in national security, emphasizing the importance of transparency and accountability. However, a subprinciple of principle 5 also recognizes that this democratic scrutiny may not always be possible in the context of national security.	The issue of liability would be crucial if CSIS engages in AI-enabled activity based on poisoned or tampered data from a threat actor, resulting in an error. The opaque nature of AI and the need to combat AI-enabled threats could require regulatory bodies to oversee and debate the use of certain data. These bodies, established through democratic processes, would play a critical role in controlling which data is permissible for use.
Principle 6: Equity Principle	The development and use of AI's must contribute to	A just and equitable society requires even treatment of all those within it. AI-enabled threats like	AI-enabled disinformation, created using biased algorithms, can target marginalized

	the creation of a just and equitable society**.**	disinformation can pose issues with this principle. Disinformation can specifically target marginalized communities, while investigative decisions based on bias or incorrect data, or even data based on disinformation, can lead to inequitable decisions (Nava-Schellinger, 2021).	communities and exacerbate tensions. The "Freedom Convoy" protest demonstrated the potential harm of the spread of disinformation. If AI were involved in this spread in a future protest, it could lead to violent conflict that undermines the goal of a just and equitable society.
Principle 8: Prudence Principle	Every person involved in AI development must exercise caution by anticipating, as far as possible, the adverse consequences of the use of AI and by taking the appropriate measures to avoid them.	This relates to the human rights conceptual framework discussed previously in the report. The use of AI, and thus the participation in AI-enabled activities, creates feedback for the creation and expansion of AI. Therefore, when CSIS participates in AI-enabled activities, those activities could lead to threat actors expanding their AI-enabled threats capabilities. As such, CSIS needs to have prudency when dealing with AI-enabled activities or threats.	This principle is similar to the first federal government *guiding principle for the use of AI* which highlights the need for the understanding and measuring of impacts of AI.

Canada's First Attempt at Governing AI: Bill C-27 and the *Artificial Intelligence and Data Act (AIDA)*

Bill C-27, also known as the *Digital Charter Implementation Act*, introduced in Parliament on June 16th, 2022, includes Part 3, the Artificial Intelligence and Data Act (*AIDA).* The legislation aims to establish consistent requirements across Canada for the design, development, and use of AI and prohibit actions that could cause significant harm to Canadians or undermine their rights and freedoms (Landry et al, 2022).

AIDA may not be legally applicable to CSIS, however. First *it* is still in its second reading in the House of Commons and may not be enacted. Second, in its current incarnation it only applies to the private sector, not the public sector. Nevertheless, some of the definitions within *AIDA* are valuable for analyzing CSIS' approach of AI-enabled threats and activities.

One key term is "biased output", defined as:

"...content that is generated, or a decision, recommendation or prediction that is made, by an artificial intelligence system and that adversely differentiates, directly or indirectly and without justification, in relation to an individual on one or more of the prohibited grounds of discrimination set out in section 3 of the Canadian Human Rights Act, or on a combination of such prohibited grounds." (Bill C-27, 2022).

This definition interacts with possible AI-enabled activities that CSIS could conduct and provides legislative and practical depth to the idea of bad input data equals bad and possibly illegal output (Greiman, 2021). It also applies to OSINT, as its informal nature can lead to errors in data that could then cause adverse effects to Canadians, undermining their rights and freedoms.

Another relevant definition within *AIDA* is 'harm' defined as: "(a) physical or psychological harm to an individual; (b) damage to an individual's property; or (c) economic loss to an individual" (Bill C-27, 2022). This definition highlights further possibilities of how AI-enabled activities and threats can interfere with rights and freedoms. Physical or psychological harm is obvious, but damage to an individual's property less so, given the question of how much a person's identifiable information or data is considered their property. If using a person's data inappropriately can be considered damage to property caused by an AI-enabled threat or activity. Thus, a form of harming the person, then this definition could open a whole new form of thinking regarding data privacy, depending on how it is to be used and interpreted when inevitably discussed in the legal system.

*AIDA a*lso has a relationship with the *Charter.* For example, *AIDA* helps protect S. 7 of the *Charter,* and the security of a person, as Part 2 of *AIDA* would prohibit knowingly or recklessly making an AI available for use if the system is or does cause serious harm. Another example is that of S. 8 of the *Charter, which covers* search and seizure. Under *AIDA*, the Minister responsible may compel the production of certain information from persons subject to the *Act* to verify their compliance with the *Act*. This person may also need to provide auditors with records (Government of Canada, D. of J., 2022). This starts to provide a possible starting point for the creation of a structure of transparency for CSIS and its use of AI.

AI Ethics: What is Good Data?

The effectiveness of AI-enabled activities depends on the quality of input data. If there is bias in the input data, it will be reflected in the output of the system or activity. Human decisions in the design of an AI system can also have significant consequences on rights and freedoms. For example, prioritizing certain variables of data in a national security investigation using AI-enabled activity can introduce biases that shape the AI's decisions. Therefore, the outcome of AI-enabled activities, even when based on factually correct data, can harm human rights if data is analyzed or prioritized in a biased manner (Greiman, 2021).

There are four "pillars" of good data: community, rights, useability, and politics (Daly et al, 2021). When we consider these pillars in the context of CSIS, national security, and AI-enabled activities or threats, three of the four pillars are most significant.

Pillar	What is it?	Why it's relevant to AI-enabled activities or threats
Community	Data collection, analysis, and utilization should prioritize the data and technological sovereignty of data subjects and communities, rather than being determined solely by those in power. AI should be constructed by communities to assist their participation in data-related decision-making and governance (Daly et al, 2021).	Given AI's ability to develop identifiable information with a wider speed and scope than a human analyst, the need for data collection to be orchestrated and mediated in a community setting is important for AI-enabled activities and threats. This pillar suggests that data being used as the basis for AI-enabled activities conducted by CSIS and its partners need to have updated regulation and considerations on a consistent basis and done so in collaboration with a diverse range of communities in Canada.
Rights	Data collection, analysis, and use should prioritize the sovereignty of data subjects and communities, rather than being determined solely by those in power. AI technology should be designed to facilitate community participation in decision-making and governance related to	This pillar illustrates that good data can be used to not only protect these rights and freedoms from AI-enabled threats, but also enhance and strengthen these freedoms.

	data, constructed in collaboration with communities (Daly et al, 2021).	
Useability	The concept of Good Data requires that it is consensual, transparent, and fit for purpose. Additionally, measures of fairness and other values attributed to data should extend beyond technical explanations and challenge broader societal unfairness. Good Data is dependent on context and, with reasonable exceptions, should be open, published, revisable, and form useful social capital where appropriate (Daly et al, 2021).	The term transparent has already been established as an important principle across most of the legislation and guidelines attempting to prevent AI-enabled activities and threats from interfering with rights and freedoms. Also, this fit for purpose concept is consistent with analysis about the need for good input to AI in order to create better investigative outputs.

The Apparent Paradox of Transparency and National Security

Transparency is often held as one of the critical principles of liberal democracies. It is also a key component of strong privacy practices. It holds institutions accountable for their actions, promotes individual agency, and can be used to create a sense of trust between these individuals and institutions (Government of Canada, D. of J., 2022). However, regarding national security, transparency is rarely a high priority. Traditionally, national security has been used to limit access to certain information. This action is logical, given that actions of transparency by the government tend to lead various actors to focus their attention on these available pieces of information (Meijer, 2013). Disclosing information regarding national security operations, vulnerabilities, and intelligence can expose the nation to increased security risks through exploitation by malicious entities. Transparency has been placed alongside all of the principles, rights and freedoms discussed in this section. As such, CSIS must implement a distinct set of actions to address public concerns over transparency in government organizations.

Parliamentary, agency, and advisory group mechanisms can help increase transparency in AI-enabled activities and threats. In 2019, CSIS launched the Stakeholder Engagement Program to engage with non-traditional sectors such as the Canadian industry, civil society, and provincial and municipal officials to them to threats and mitigate risks. The National Security Transparency Advisory Group, also launched in 2019, advises the federal government on increasing transparency across national security and intelligence departments, promoting transparency and democratic accountability, and increasing public awareness and access to related information. There are various accountability organizations in place, including but not limited to the Standing Senate Committee on National Security and Defense, the Standing Committee on Public Safety and National Security, the Auditor General of Canada, the Privacy

Commissioner, and the National Security and Intelligence Review Agency (CSIS, 2020). External bodies can assist in ensuring that CSIS adheres to AI ethics, principles, and regulations, but addressing this issue will require filling the knowledge gap related to AI-enabled activities and their processes. By doing so, CSIS' accountability and transparency models will be enhanced, enabling the organization to effectively communicate information to the public about its actions and how it is held accountable. Ultimately, this could increase transparency.

When considering CSIS' relationship with transparency, many of the interviewees for this report noted that CSIS and other national security organizations already have the regulations, accountabilities, capabilities, and authorities that provide solid transparency in a national security space (Counselor, Justice Canada & Official, Privy Council Office). However, it is essential to note that transparency in and of itself does not inherently create openness. Although these regulations seem sufficient, the problem is the lack of strictness and public display of transparency measures. Recent commentary by Wark (2023) has brought attention to the National Security Transparency Commitment's underwhelming outcome. Despite the potential of new review entities, their primary audience is viewed as the government rather than the public.

Additionally, the government needs to show more regard for the recommendations of the National Security and Intelligence Committee of Parliamentarians, especially evident in the 2020 report on foreign interference. Additionally, the higher use of the internet and various public engagements with how governments handle information has increased the demand for more transparency. Because of the public's limited understanding of national security, technology and government actions undertaken by CSIS are often thought of as threatening and requiring transparency. This means that even if absolute transparency is not the problem, how the public perceives the lack of transparency can become one. Increasing dialogue and interactions with the public and broader society and providing education on these topics can help increase the perceived transparency of CSIS (Counselor, Justice Canada & Employee, CSIS).

Key Takeaways:

1. **The rights, freedoms, and principles of AI discussed in this section should be at the forefront of frameworks and guidelines relating to AI use and governance.** AI is an ever-evolving reality, and each stage of the AI lifecycle has mutual and dynamic links with human rights. AI-enabled activities used by CSIS, and AI-enabled threats used by threat actors can potentially lead to a broader attack surface and compromise human rights. Government organizations must recognize the increased threat surface of AI and prioritize protecting human rights in AI development and use. Increasing dialogue and education on national security, technology, and government policy can help achieve this goal.

2. **The quality of input data is critical to the effectiveness of AI-enabled activities. Any bias in the input data will be reflected in the system's output or activity.** Human decisions in the design of an AI can have significant consequences on people's rights and freedoms. Therefore, the outcome of AI-enabled activities, even when based on factually correct data, can infringe upon human rights if data is analyzed or prioritized in a biased manner. Government organizations such as CSIS must be mindful of these potential biases and take steps to mitigate them while dealing with AI-enabled threats.

3. **Transparency and accountability are crucial in understanding possible AI use by CSIS.** The rights and freedoms of Canadian citizens are enshrined in legislation, while principles such as transparency and accountability are often left to bureaucratic structures to handle. The public is now interested in knowing how government organizations make decisions, but AI adds another layer of complexity by creating multiple layers of secrecy. CSIS has launched initiatives like the Stakeholder Engagement Program and the National Security Transparency Advisory Group to address public concerns and increase transparency. However, the public still demands more transparency, and increasing dialogue and education on national security, technology, and government policy can help achieve that. The need for accountability and transparency in AI use will become even more critical, especially as we face AI vs AI situations.

Conclusion and Recommendations

AI-enabled threats such as cyberattacks, disinformation, and OSINT pose unprecedented challenges to intelligence agencies, including CSIS. AI increases the speed and scope of threat activities, relies heavily on large datasets, and necessitates AI to defend against AI. These are three critical characteristics of AI-enabled threats that make them more difficult to address using existing technology and defenses. Moreover, these challenges are exacerbated by the various restrictions under the CSIS Act, Privacy Act, and the Charter that create unique barriers to addressing AI-enabled threats.

CSIS may face challenges in dealing with AI-enabled threats in three key areas: interpretation and assessment of AI-enabled threats, the issues dataset rules play in data collection, and challenges concerning judicial oversight and the warrants system. In addition, the current legislation may pose challenges to CSIS' operations when addressing AI-enabled threat activities. For example, certain sections of the CSIS Act, specifically those relating to operational responses through the application of warrants, may not be fast enough to address modern-day threats related to AI and other AI technologies.

This report highlights the need for prioritizing the protection of human rights in the development and use of AI technology. AI-enabled investigative activities used by government agencies and threat actors pose a risk to human rights such as privacy, equity, equality, and security of person. Frameworks and guidelines relating to AI use and governance should prioritize the protection of human rights. It is important for government organizations to recognize the increased threat surface of AI and take necessary measures to protect human rights in AI development and use. Increasing dialogue and education is also important to ensure that AI technology is developed and used in a way that respects and upholds the rights and freedoms of Canadians.

Based on these conclusions, and with consideration of this report in its entirety, the following recommendations are offered as a way for CSIS to begin addressing the implications of AI-enabled threats on Canadian national security:

1. **CSIS should enhance information technology architecture within the government to streamline sharing of critical information related to AI-enabled threats across departments.** There are various government stakeholders involved in national security, which makes access to internal information more important for CSIS. With the exponential growth of AI technologies, there is an increased need for streamlined sharing of critical information across departments to effectively tackle AI-enabled threats to national security. CSIS can enhance Information Technology Architecture within the government by leveraging technological solutions to enable the sharing of

critical information related to AI-enabled threats across departments. This can be done through the development of a robust information-sharing framework, which ensures that all departments can access critical information in a secure and transparent manner. Continued cooperation with other domestic agencies will be necessary to address technological constraints and facilitate data sharing. This approach will increase transparency and build trust between agencies, while also safeguarding the rights and freedoms of Canadians.

2. **CSIS should upgrade Canada's cyber security strategies to include AI-enabled threats and create a multi-agency long-term strategy with law enforcement agencies to detect and mitigate these threats.** The problem is the increasing speed, scope, and sophistication of AI-enabled threats, which pose a significant risk to national security and the privacy of Canadians. CSIS can collaborate with law enforcement agencies to develop a long-term strategy based on Canadians' rights and freedoms, alongside ethical and guiding principles such as transparency, equity, and prudence. This strategy can include regular assessments of cyber risks and proactive measures to mitigate them. Upgrading cyber security strategies to include AI-enabled threats and creating a long-term strategy with law enforcement agencies will enable CSIS to effectively detect and mitigate cyber threats. Collaboration with other Canadian organizations will increase the speed, scope, and sophistication of the response to these threats, ensuring the protection of Canadians' rights and freedoms. The strategy's guiding principles will ensure that it is implemented in a transparent and equitable manner.

3. **CSIS should advocate for an amendment to the *CSIS Act* to include a different type of judicial authorization that can enable its intelligence-gathering operations, including being able to investigate more expediently AI-enabled threats.** The exponential growth of AI technologies presents a problem for CSIS' ability to react to AI-enabled threats, which requires a multi-layered approach that is time-consuming and resource-intensive. A special warrant for AI-enabled threats would enable more efficient intelligence gathering by providing a legal basis for expedited procedures. Strict regulation and oversight would also be necessary to prevent the infringement of rights and freedoms, such as S. 8 of the *Charter*. This would also enable CSIS to better fulfill its obligations to deal with AI-enabled threats, by providing a legal basis for expedited procedures to keep pace with the speed and scope of these threats, while ensuring protection of rights and freedoms through regulation and oversight.

4. **CSIS should revisit their funding requirements to support the hiring, training, and in-house development of AI expertise.** The problem CSIS faces includes shortages in people with AI and technical expertise, which private entities are readily recruiting and retaining with more attractive compensation

packages. AI-enabled threats can develop insights, techniques, and tactics faster than human analysts, and in-house AI expertise is necessary for timely and effective threat reduction measures. This approach will enable CSIS to utilize AI optimally, recognize potential privacy violations, and safeguard against other possible infringement to rights. By ensuring that CSIS has a sufficient number of AI experts, the organization will have the necessary knowledge to monitor and investigate AI-enabled threats, allowing them to stay ahead of potential threats in the future.

5. **CSIS should build and nurture relationships with private sector organizations involved in the development of AI to increase information sharing and remain on par with technological advancements within the AI landscape.** The private sector's majority ownership in the development of AI places government agencies like CSIS in a reactive position, making it difficult to operate effectively. CSIS should initiate dialogue with private sector organizations involved in AI development and explore opportunities for collaboration. CSIS should look to establish a framework for sharing information on emerging AI threats and mitigation techniques, create joint initiatives to address such threats, and establish channels for continued knowledge transfer. By sharing information and working collaboratively, CSIS will be able to keep pace with the technological advancements within the AI landscape, build transparency and trust, and facilitate continued knowledge transfer in subject fields extending from AI-enabled threats.

Glossary of Terms

Term	Definition	Term	Definition
AI	Artificial Intelligence	**Five Eyes Partners**	An intelligence partnership alliance between Australia, Canada, New Zealand, United Kingdom and the USA.
AGI	Artificial General Intelligence	**GAN**	Generative Adversarial Networks
AIDA	Artificial Intelligence and Data Act	**GUI**	Graphical User Interface
ANI	Artificial Narrow Intelligence	**IT Systems**	Information Technology Systems
CII	Canadian Identifying Information	**IoT**	Internet of Things
CSE	Communications Security Establishment	**Kill Chain**	Sequence of steps taken by threat actors to achieve their goal
CSIS	Canadian Security and Intelligence Service	**ML**	Machine Learning
DAP	Data Acquisition Program	**NLP**	Natural Language Processing
DAAS	Data Acquisition and Archiving Systems	**NSIRA**	National Security Intelligence Review Agency
DL	Deep Learning	**OSINT**	Open
DoS	Denial of Service	**OT**	Operational Technology

Term	Definition	Term	Definition
RCMP	Royal Canadian Mounted Police	**The *Charter***	Canadian Charter of Rights and Freedoms
R&D	Research and Development	**TRM**	Threat Reduction Measures
SIGNIT	Signals Intelligence	**UDHR**	Universal Declaration of Human Rights
SIRC	Security Intelligence Review Committee		

Appendix 1

Position	Department/Organization
CSIS Employee	Canadian Security and Intelligence Service
CSIS Employee	Canadian Security and Intelligence Service
CSIS Employee	Canadian Security and Intelligence Service
CSIS Employee	Canadian Security and Intelligence Service
CSIS Employee	Canadian Security and Intelligence Service
CSIS Employee	Canadian Security and Intelligence Service
Official	Privy Council Office
Professor	University of British Columbia
Associate Professor	Carleton University
Counselor	Justice Canada
Partner	EY

References

Agrawal, M. (2021). *The possibilities of AI in 2030: Transformation across dimensions*. Forbes. Retrieved 2023, from https://www.forbes.com/sites/forbesbusinesscouncil/2021/08/23/the-possibilities-of-ai-in-2030-transformation-across-dimensions/?sh=646ef9056b67

Alhajjar, E. (2022). *Adversarial machine learning poses a new threat to national security*. AFCEA International. Retrieved 2023, from https://www.afcea.org/signal-media/cyber-edge/adversarial-machine-learning-poses-new-threat-national-security

Allen, G., & Chan, T. (2017). *Artificial Intelligence and national security*. Belfer Center for Science and International Affairs. Retrieved 2023, from https://www.belfercenter.org/publication/artificial-intelligence-and-national-security

Allyn, B. (2022). *Deepfake video of Zelenskyy could be 'tip of the iceberg' in Info War, experts warn*. NPR. Retrieved 2023, from https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia

Altman, M., Wood, A., O'Brien, D. R., Vadhan, S., & Gasser, U. (2015). Towards a Modern Approach to Privacy-Aware Government Data Releases. *Berkeley Technology Law Journal*, *30*(3), 1967–2072. https://www.jstor.org/stable/26377584

Attard-Frost, B. (2022). *Once a leader, Canada's AI strategy is now a fragmented laggard*. Faculty of Information (iSchool). Retrieved 2023, from https://ischool.utoronto.ca/news/once-a-leader-canadas-artificial-intelligence-strategy-is-now-a-fragmented-laggard/

Baker, K. (2023). *10 most common types of cyberattacks*. Crowdstrike. Retrieved 2023, from https://www.crowdstrike.com/cybersecurity-101/cyberattacks/most-common-types-of-cyberattacks/

Bayer, J., Bitiukova, N., Szakacs, J., Alemanno, A., & Uszkiewicz, E. (2019). *Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States*. Policy Department for Citizens' Rights and Constitutional Affairs. Retrieved 2023, from https://www.europarl.europa.eu/RegData/etudes/STUD/2019/608864/IPOL_STU(2019)608864_EN.pdf

Benaich, N., & Hogarth, I. (2022). *State of ai report 2022*. State of AI Report 2022. Retrieved 2023, from https://www.stateof.ai/

Bergengruen, V. (2023). *Inside Russia's year of Ukraine propaganda*. Time. Retrieved 2023, from https://time.com/6257372/russia-ukraine-war-disinformation/

Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts, 1st Session, 44th Parliament, 2022.

Bremmer, I., & Kupchan, C. (2023). *Top Risks 2023*. Eurasia Group. Retrieved 2023, from https://www.eurasiagroup.net/files/upload/EurasiaGroup_TopRisks2023.pdf

Browne, R. (2023). *All you need to know about chatgpt, the A.I. chatbot that's got the world talking and Tech Giants clashing*. CNBC. Retrieved 2023, from https://www.cnbc.com/2023/02/08/what-is-chatgpt-viral-ai-chatbot-at-heart-of-microsoft-google-fight.html

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., hÉigeartaigh Sean O, Beard, S., Belfield, H., Farquhar, S., … Amodei, D. (2018). (publication). T*he Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.* University of Oxford;;University of Cambridge;Center for a New American Security;Electronic Frontier Foundation;OpenAI. Retrieved 2023, from ttps://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/MaliciousUseofAI.pdf?ver=1553030594217.

Buchanan, B. (2020). *A National Security Research Agenda for Cybersecurity and Artificial Intelligence*. Center for Security and Emerging Technology. Retrieved 2023, from https://cset.georgetown.edu/wp-content/uploads/CSET-A-National-Security-Research-Agenda-for-Cybersecurity-and-Artificial-Intelligence.pdf

Canada Wireless Telecommunications Association. (2023). *Canadians among global leaders in internet usage and smartphone ownership, Pew research center study shows.* Cision Canada. Retrieved 2023 from https://www.newswire.ca/news-releases/canadians-among-global-leaders-in-internet-usage-and-smartphone-ownership-pew-research-center-study-shows-833834298.html

Canadian Anti-Fraud Center. (2022). *Annual Report 2021.* Retrieved 2023, from https://publications.gc.ca/collections/collection_2022/grc-rcmp/PS61-46-2021-eng.pdf

Canadian Center for Cyber Security. (2022). *National Cyber Threat. Assessment.* Communications Security Establishment (CSE). https://www.cyber.gc.ca/sites/default/files/ncta-2023-24-web.pdf

Canadian Charter of Rights and Freedoms, Part 1 of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), 1982, c 11. Canadian Security and Intelligence Service Act, 1984 – Part 2, Judicial Control https://laws-lois.justice.gc.ca/eng/acts/C-23/page-6.html#h-76429

Carter v. Canada, 2015 SCC 5

Chesney, R., & Citron, D. K. (2018). Deep Fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(1753). https://doi.org/10.2139/ssrn.3213954

Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., Kak, A., Mathur, V., Mcelroy, E., Nill Sánchez, A., Raji, D., Rankin, J., Richardson, R., Schultz, J., West, S., & Whittaker, M. (2019). AI Now 2019 Report. Retrieved 2023, from https://ainowinstitute.org/AI_Now_2019_Report.pdf

CSIS. (2020). *CSIS 2019 Public Report.* CSIS. Received 2023, from https://www.canada.ca/en/security-intelligence-service/corporate/publications/2019-public-report.html

CSIS. (2022). *2021 Public Report*. CSIS. Retrieved 2023, from https://www.canada.ca/en/security-intelligence-service/corporate/publications/csis-2021-public-report.html.

Daly, A., Devitt, S. K., & Mann, M. (2021). AI Ethics Needs Good Data (P. Verdegem, Ed.). University of Westminster Press. http://www.jstor.org/stable/j.ctv26qjjhj.9

Department of Justice Canada. (2019). Privacy Act Modernization: A Discussion Paper. Retrieved 2023, from https://www.justice.gc.ca/eng/csj-sjc/pa-lprp/dp-dd/pdf/dp-1.pdf

DiResta, R. (2022). *The supply of disinformation will soon be infinite*. The Atlantic. Retrieved 2023, from https://www.theatlantic.com/ideas/archive/2020/09/future-propaganda-will-be-computer-generated/616400/

European Commission. (2018). (rep.). *A Multidimensional Approach to Disinformation. Report of the Independent High-Level Group on Fake News and Online Disinformation.* Publications Office of the European Union. Retrieved 2023.

Fasken. (2023). *National security law*. Canadian National Security Law. Retrieved 2023, from https://www.fasken.com/en/knowledge/doing-business-canada/2021/10/23-national-security-law

François, C. (2019). *Actors, Behaviors, Content: A Disinformation ABC: Highlighting Three Vectors of Viral Deception to Guide Industry & Regulatory Responses*. Transatlantic Working Group. Retrieved 2023, from https://www.ivir.nl/publicaties/download/ABC_Framework_2019_Sept_2019.pdf

Galindo-Rueda, F., & Cairns, S. (2021). *A new approach to measuring government investment in AI-related R&D*. OECD.AI. Retrieved 2023, from https://oecd.ai/en/wonk/government-investment-ai-related-r-and-d

Ghosh, D., & Scott, B. (2018). *Digital Deceit: The technologies behind Precision Propaganda on the internet*. New America. Retrieved 2023, from https://www.newamerica.org/pit/policy-papers/digitaldeceit/

Goldstein, J. A., & Grossman, S. (2021). *How disinformation evolved in 2020*. Brookings. Retrieved 2023, from https://www.brookings.edu/techstream/how-disinformation-evolved-in-2020/

Government of Canada, D. of J. (2022). Statement of Potential Charter Impacts. Department of Justice. Retrieved 2023, from https://www.justice.gc.ca/eng/csj-sjc/pl/charter-charte/c27_1.html

Government of Canada. (2016). Our Security, Our Rights. Retrieved 2023, from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ntnl-scrt-grn-ppr-2016/ntnl-scrt-grn-ppr-2016-en.pdf

Government of Canada. (2022). *Government of Canada launches second phase of the Pan-Canadian Artificial Intelligence Strategy*. Canada.ca. Retrieved 2023, from https://www.canada.ca/en/innovation-science-economic-development/news/2022/06/government-of-canada-launches-second-phase-of-the-pan-canadian-artificial-intelligence-strategy.html

Greiman, V. (2021). Human Rights and Artificial Intelligence: A Universal Challenge. Journal of Information Warfare, 20(1), 50–62. https://www.jstor.org/stable/27036518

Guembe B., Azeta A., Misra S., Chukwudi Osamor V., Fernandez-Sanz L., Pospelova V. (2022) *The Emerging Threat of Ai-driven Cyber Attacks: A Review.* Applied Artificial Intelligence, 36(1). https://doi.org/10.1080/08839514.2022.2037254

Heaven, W. D. (2020). *A GPT-3 bot posted comments on Reddit for a week and no one noticed*. MIT Technology Review. Retrieved 2023, from https://www.technologyreview.com/2020/10/08/1009845/a-gpt-3-bot-posted-comments-on-reddit-for-a-week-and-no-one-noticed/

Ho, V. (2019). *Nancy Pelosi condemns Facebook as 'willing enablers of Russian interference'*. The Guardian. Retrieved 2023, from https://www.theguardian.com/us-news/2019/may/30/nancy-pelosi-calls-facebook-willing-enablers-of-russian-interference

Hu, K. (2023). *CHATGPT sets record for fastest-growing user base - analyst note*. Reuters. Retrieved 2023, from https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

IBM. (2022). *What is a cyberattack.* Retrieved 2023 from https://www.ibm.com/topics/cyber-attack

IBM. (2023). *Digital Workers vs. Chatbots vs. bots: What's the difference?* IBM. Retrieved 2023, from https://www.ibm.com/cloud/blog/digital-workers-vs-chatbots-vs-bots-whats-the-difference

IBM. (2023). *What is Artificial Intelligence (AI) ?* IBM. Retrieved 2023, from https://www.ibm.com/topics/artificial-intelligence

IBM. (2023). *What is machine learning?* IBM. Retrieved 2023, from https://www.ibm.com/topics/machine-learning

IBM. (2023). *What is natural language processing?* IBM. Retrieved 2023, from https://www.ibm.com/topics/natural-language-processing

Joshi, N. (2020). *Choosing between rule-based bots and AI Bots*. Forbes. Retrieved 2023, from https://www.forbes.com/sites/cognitiveworld/2020/02/23/choosing-between-rule-based-bots-and-ai-bots/?sh=70a27b20353d

Joshi, N. (2022). *7 types of artificial intelligence*. Forbes. Retrieved 2023, from https://www.forbes.com/sites/cognitiveworld/2019/06/19/7-types-of-artificial-intelligence/?sh=7c6c24cf233e

Kertysova, K. (2018). Artificial Intelligence and Disinformation: How AI Challenges the Way Disinfornation Is Produced, Disseminated, and Can Be Countered. Security and Human Rights, 29, 55-82.

Knight, W. (2021). *Military artificial intelligence can be easily and dangerously fooled*. MIT Technology Review. Retrieved 2023, from https://www.technologyreview.com/2019/10/21/132277/military-artificial-intelligence-can-be-easily-and-dangerously-fooled/

Kreps, S. (2021). *Democratizing harm: Artificial intelligence in the hands of nonstate actors.* Brookings. Retrieved 2023, from https://www.brookings.edu/research/democratizing-harm-artificial-intelligence-in-the-hands-of-non-state-actors/

Landry, K., Henderson, C., & Pinchak, J. (2022). Bill C-27 – Canada's proposed Artificial Intelligence and Data Act. Stewart McKelvey. Retrieved 2023, from https://www.stewartmckelvey.com/thought-leadership/bill-c-27-canadas-proposed-artificial-intelligence-and-data-act/

Linvill, D., & Warren, P. (2021). *Understanding the pro-china propaganda and disinformation tool set in Xinjiang*. Lawfare. Retrieved 2023, from https://www.lawfareblog.com/understanding-pro-china-propaganda-and-disinformation-tool-set-xinjiang

Mak, T., & Temple-Raston, D. (2020). *Where are the deepfakes in this presidential election?* NPR. Retrieved 2023, from https://www.npr.org/2020/10/01/918223033/where-are-the-deepfakes-in-this-presidential-election

Marr, B. (2022). *The 10 best examples of how AI is already used in our everyday life*. Forbes. Retrieved 2023, from https://www.forbes.com/sites/bernardmarr/2019/12/16/the-10-best-examples-of-how-ai-is-already-used-in-our-everyday-life/?sh=3d6062661171

Marshall, A., Parikh, J., Kiciman, E., & Kumar, R. S. S. (2019). *Threat Modeling AI/ML Systems and Dependencies.* Retrieved 2023, from https://learn.microsoft.com/en-us/security/engineering/threat-modeling-aiml

Max. (2023). *Why AI-powered phishing will become a serious security issue for your organization.* Xorlab. Retrieved 2023, from https://www.xorlab.com/en/blog/why-ai-powered-phishing-will-become-a-serious-security-issue-for-your-organization

Meijer, A. (2013). Understanding the Complex Dynamics of Transparency. *Public Administration Review*, *73*(3), 429–439. http://www.jstor.org/stable/42002946

Menon, A. (2023). *Data poisoning and its impact on the AI ecosystem.* TheMathCompany. Retrieved 2023, from https://themathcompany.com/blog/data-poisoning-and-its- impact-on-the-ai-ecosystem

Morgan, C., Rothman, C., Glover, D., Thérien, D., Hantho, E., Keogh, E., Langlois, F., Waschuk, G., Choi, J., Lan, J., Adessky, J., Joizil, K., & Lim, W. (2023). Artificial Intelligence. Global Legal Post. Retrieved 2023, from https://www.globallegalpost.com/lawoverborders/artificial-intelligence-1272919708/canada-1511789807#1

Morrison, S. (2021). *How a major oil pipeline got held for ransom. The largest petroleum pipeline in the country was reportedly breached by a single leaked password.* Vox. Retrieved 2023 from, https://www.vox.com/recode/22428774/ransomeware-pipeline-colonial-darkside-gas-prices

Mpinga, E. K., Bukonda, N. K., Qailouli, S., & Chastonay, P. (2022). Artificial Intelligence and Human Rights: Are There Signs of an Emerging Discipline? A Systematic Review. Journal of Multidisciplinary Healthcare, Volume 15, 235–246. https://doi.org/10.2147/jmdh.s315314

National Security and Intelligence Review Agency,*Threat Reduction Measures review 2020*. https://www.nsira-ossnr.gc.ca/wp-content/uploads/Redacted-TRM-Review-e-Updated.pdf

Nava-Schellinger, V. (2021) How Misinformation Hurts Communities of Color Most. NCOA. Retrieved 2023, from https://www.ncoa.org/article/how-misinformation-hurts-communities-of-color-most

Nimmo, B., Eib, C. S., Tamora, L., Johnson, K., Smith, I., Buziashvili, E., Kann, A., Karan, K., Rosas, E. P. de L., & Rizzuto, M. (2019). (rep.). *#OperationFFS: Fake Face Swarm*. Graphika and Atlantic Council's Digital Forensic Research Lab. Retrieved 2023, from https://public-assets.graphika.com/reports/graphika_report_operation_ffs_fake_face_storm.pdf .

NSIRA (2020). *Review of CSIS threat reduction activities*. National Security and Intelligence Review Agency. Retrieved 2023, from https://www.nsira-ossnr.gc.ca/wp-content/uploads/Redacted-TRM-Review-e-Updated.pdf.

NSIRA (2021). *Review of the Communications Security Establishment's Disclosures of Canadian Identifying Information*. National Security and Intelligence Review Agency. Retrieved 2023, from https://www.nsira-ossnr.gc.ca/wp-content/uploads/2021/06/10397868-001-EN-CII-Review-2018-19-1.pdf.

Obermeyer, E. B., Rebecca Nissan, and Ziad. (2021). To stop algorithmic bias, we first have to define it. Brookings. Retrieved 2023, from https://www.brookings.edu/research/to-stop-algorithmic-bias-we-first-have-to-define-it/

Omdia. (2023). *Artificial Intelligence*. Omdia. Retrieved 2023, from https://omdia.tech.informa.com/topic-pages/artificial-intelligence

Onelogin. (2022). *Watch out for AI-powered spear phishing. Retrieved 2023, from* https://www.onelogin.com/resource-center/infographics/cybersecurity-ai-spear-phishing

Paris, B., & Donovan, J. (2019). *Deepfakes and cheap fakes*. Data & Society. Retrieved 2023, from https://datasociety.net/library/deepfakes-and-cheap-fakes/

Pascoe, J., Murray, S., & Proulx, M. (2017). *UAlberta to play prominent role in pan-Canadian AI strategy*. University of Alberta. Retrieved 2023, from https://www.ualberta.ca/folio/2017/03/ualberta-to-play-prominent-role-in-pan-canadian-ai-strategy.html

Patterson, D. (2019). *From deepfake to "Cheap fake," it's getting harder to tell what's true on your favorite apps and websites*. CBS News. Retrieved 2023, from https://www.cbsnews.com/news/what-are-deepfakes-how-to-tell-if-video-is-fake/

Perteous, H. (2022). *The growing importance of open-source intelligence to national security.* Library of Parliament. Retrieved 2023, from https://hillnotes.ca/2022/02/17/the-growing-importance-of-open-source-intelligence-to-national-security/

Posard, M., Kepe, M., Reininger, H., Marrone, J., Helmus, T., & Reimer, J. (2020). (rep.). *From Consensus to Conflict: Understanding Foreign Measures Targeting U.S. Elections*. Rand Corporation. Retrieved 2023, from https://www.rand.org/pubs/research_reports/RRA704-1.html.

Precedence Research. (2023). *Artificial Intelligence (AI) Market*. Precedence Research. Retrieved 2023, from https://www.precedenceresearch.com/artificial-intelligence-market

Privy Council Office. (2019). Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service. Privy Council Office Retrieved 2023, from https://www.canada.ca/en/privy-council/corporate/clerk/publications/data-strategy.html

Public Safety Canada. (2021). *National Strategy for Critical Infrastructure*. Retrieved 2023, from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/srtg-crtcl-nfrstrctr/index-en.aspx

R. v. Le, 2019 SCC 34.

R. v. McGregor, 2023 SCC 4

R. v. Morgentaler (No. 2), [1988] 1 S.C.R. 30

R. v. Spencer, 2014 SCC 43

Rash, W. (2021). *Disinformation propelled by social media and conspiracy theories led to insurrection*. Forbes. Retrieved 2023, from https://www.forbes.com/sites/waynerash/2021/01/19/disinformation-propelled-by-social-media-and-conspiracy-theories-led-to-insurrection/?sh=9495c0434e05

Satariano, A., & Mozur, P. (2023). *The people onscreen are fake. the disinformation is real.* The New York Times. Retrieved 2023, from https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html

Satter, R. (2019). *Experts: Spy used AI-generated face to connect with targets*. Associated Press. Retrieved 2023, from https://apnews.com/bc2f19097a4c4fffaa00de6770b8a60d

Sayler, K. (2020). (rep.). *Artificial Intelligence and National Security*. Congressional Research Service . Retrieved 2023, from https://sgp.fas.org/crs/natsec/R45178.pdf.

Security and Intelligence Review Committee, *Amending the CSIS Act, http://www.sirc-csars.gc.ca/csiscr/amdmod-eng.html#a140*

Security and Intelligence Review Committee, Review of CSIS' use of data management and exploitation activities (2015). Retrieved April 11, 2023, from http://www.sirc-csars.gc.ca/opbapb/lsrlse/2015/2015-02-eng.pdf.

Silcoff, S., & O'Kane, J. (2023). *Canada has leading AI experts. but does Ottawa have the right plan to support an AI industry?* The Globe and Mail. Retrieved 2023,

from https://www.theglobeandmail.com/business/article-canada-support-ai-industry/

Simon, C. (2022). *Council post: Why artificial general intelligence isn't further along*. Forbes. Retrieved 2023, from https://www.forbes.com/sites/forbestechcouncil/2022/08/29/why-artificial-general-intelligence-isnt-further-along/

Steck, H., Baltrunas, L., Elahi, E., Liang, D., Raimond, Y., & Basilico, J. (2022). *Deep Learning for Recommender Systems: A Netflix Case Study*. Netflix Research. Retrieved 2023, from https://research.netflix.com/publication/%20Deep%20Learning%20for%20Recommender%20Systems%3A%20A%20Netflix%20Case%20Study

Thrope, J. (2021). *What is data poising and why should we be concerned.* International Security Journal. Retrieved 2023 from https://internationalsecurityjournal.com/what-is-data-poisoning/

United Nations. (1948). OHCHR | Universal Declaration of Human Rights. OHCHR. Retrieved 2023, from https://www.ohchr.org/en/human-rights/universal-declaration/translations/english

Université de Montréal. (2018). Declaration for a Responsible Development of Artificial Intelligence. Retrieved 2023, from https://monoskop.org/images/b/b2/Report_Montreal_Declaration_for_a_Responsible_Development_of_Artificial_Intelligence_2018.pdf

Villasenor, J. (2022). *How to deal with ai-enabled disinformation*. Brookings. Retrieved 2023, from https://www.brookings.edu/research/how-to-deal-with-ai-enabled-disinformation/

Waltzmann, R. 2017. *The Weaponization of Information: The Need for Cognitive Security*, Retrieved 2023, https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT473/RAND_CT473.pdf

Wark, W. (2023). *Robust, Reliable National Security Requires Transparency*. CIGI. Retrieved 2023, from https://www.cigionline.org/articles/robust-reliable-national-security-requires-transparency/

Wasike, B. (2022). Memes, memes, everywhere, nor any meme to trust: Examining the credibility and persuasiveness of covid-19-related memes. *Journal of Computer-Mediated Communication*, *27*(2). https://doi.org/10.1093/jcmc/zmab024

West, D. M., & Allen, J. R. (2022). *How artificial intelligence is transforming the world*. Brookings. Retrieved 2023, from https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/

Wittenberg, C., Tappin, B. M., Berinsky, A. J., & Rand, D. G. (2021). The (minimal) persuasive advantage of political video over text. *Proceedings of the National Academy of Sciences*, *118*(47). https://doi.org/10.1073/pnas.2114388118

Wong, S.-L., Liu, Q., & Shepherd, C. (2019). *Old messages, new memes: Beijing's Propaganda Playbook on the Hong Kong protests*. Financial Times. Retrieved 2023, from https://www.ft.com/content/7ed90e60-ce89-11e9-99a4-b5ded7a7fe3f

Woolley, S. (2023). *We're fighting fake news AI bots by using more AI. that's a mistake.* MIT Technology Review. Retrieved 2023, from https://www.technologyreview.com/2020/01/08/130983/were-fighting-fake-news-ai-bots-by-using-more-ai-thats-a-mistake/

Zegrat, A. (2021). *Spies like us. The promise and peril of crowdsourced intelligence.* Foreign Affairs. Retrieved 2023, from https://www.foreignaffairs.com/reviews/review-essay/2021-06-22/spies-us

Zegrat, A. (2023). *How technology is disrupting the intelligence world. A conversation with Amy Zegart.* Foreign Affairs Podcast. Retrieved 2023, from https://www.foreignaffairs.com/podcasts/how-technology-disrupting-intelligenceworld