SOVEREIGN DATA

A Primer for Discussion.

Canada is about to embark on a significant effort that will rely on data holdings as evidence. Ensuring that these data remains available, trustworthy, and appropriately accessed is a key step to protecting Canada's sovereignty.

Contents

Limitations OF Examples

Conceptual in nature, this work seeks to provide an input into the discussion on Sovereign Data. It is conceptual in nature and not intended to provide any legal or commercial guidance.

Purpose

This document argues that Canada needs to clearly identify and appropriately control certain kinds of data as a key element in its efforts to open and protect its sovereignty, including in the North.

Background

Understanding this discussion begins with grasping how decisions are made and debated. Several structures can be involved; two of which are examined here. The first is when the entity has already formed an opinion and is seeking evidence that supports their position. The second involves collecting data, contextualising it (generating information), and building understanding and drawing scientific conclusions or inferring patterns (intelligence).

Security Attributes of Sovereign Data

We can approach this using the security attributes of confidentiality, integrity, and availability.

- **Confidentiality**[1]: There will be certain kinds of data that will have value because they are limited to a restricted, authorized community). The value of this security attribute is preserved when the asset (data, information, or intelligence) is limited exclusively to the community approved to access it. Concurrently, part of its value lies in the statement that it has not been made available to those outside of that community in any usable form.

- **Integrity**[2]: There will be certain kinds of data/information/intelligence that need to be trustworthy. The value is that it presents a description that is reliable, credible, and upon which individuals can base decisions. To accomplish this, we need to ensure that what is being protected has not been added to, had elements deleted from, or been modified (either deliberately or through untrustworthy processes), or that it remains within its context.

- **Availability**[3]: Those kinds of data/information/intelligence that must be present in usable form. This speaks to being available on demand, or in at least a timely manner.

[1] As per the NIST Computer Security Resource Center definition at https://csrc.nist.gov/glossary/term/confidentiality.

[2] As per the NIST Computer Security Resource Center definition at https://csrc.nist.gov/glossary/term/integrity

[3] As per the NIST Computer Security Resource Center definition at https://csrc.nist.gov/glossary/term/availability

These attributes do not work in isolation. They often overlap. For example, intelligence must be timely but must also be trustworthy.

Non-Repudiation

The concept of non-repudiation is significant. As we consider the possible influences on the different security attributes, understanding what forces or factors have influenced them becomes critical. These may be routine events that are transparent. Adversaries, however, may seek to hide their transactions. Non-repudiation means that all parties involved in the system cannot dispute the authenticity of the data or the processes handling them.[4]

This concept becomes important when examining the data repositories i.e.. can the data held in these repositories be trusted? Without the concept of non-repudiation, it becomes difficult to say that no unauthorized (including hostile) actors have accessed or corrupted the data.

Non-repudiation might be simplified into five major elements:

- **Proof of origin**: This not only positively identifies the individual but also provides evidence should the sender attempt to deny sending something.
- **Proof of integrity**: ensuring that the message or data has not been tampered with.
- **Proof of delivery**: providing the sender with the evidence necessary to show that the message was received.
- **Verification**: helping confirm the identity of the sender or entities involved.
- **Timestamping**: to assist in building timelines and helping prevent issues like backdating.

Goals and Objectives

The goal is to have sovereign control over data necessary to preserve Canada's autonomy and key decisions.

This consists of three major elements.

1. Protecting the asset, data. Data, in this context, refers to the "building blocks" that are assembled (to provide context) to become information, which, in turn, evolves into intelligence through processing.

[4] The alignment here is in line with the use of non-repudiation in FIPS 204 and similar standards. NIST CSRC at https://csrc.nist.gov/glossary/term/non_repudiation

2. Having final authority and effective control. When the nation does not have the final authority over this asset, then "sovereignty" has not been achieved.

3. Control. This consists in setting limitations on appropriate and inappropriate access (or use) and enforcing them effectively.

Canada will never be completely sovereign. It belongs to alliances that also exercise control over their membership (e.g., NATO). Nations participate in these alliances either by choice or out of perceived need. In this context, this concept of sovereignty considers the potential need to share in such partnerships, without ceding certain levels of control over the data.

Problem Space Considerations

The Basis of the Need

Why do we require sovereign data? What compels us to believe that Canada actually needs to have its own unique data holdings under its ultimate control?

The roots of this need lie in two concepts: *competition* and *manipulation*.

Competition

Shifts in how nations approach trade have also altered how they view competition. We have seen increased movement backward toward what may be described as "zero-sum" thinking, where one party wins at the expense of others. In brief, certain nations have begun to reinvigorate their need for dominance on the international stage.

Competing in this space requires leadership to have clear situational awareness and as much reliable, credible evidence as possible upon which to decisions can be based . It should be clear that these conditions will never be fully achieved. Situational awareness will have gaps and may have errors. Similarly, the leadership will never have all the evidence necessary to make decisions. The goal, therefore, is to have sufficient reliable, credible data upon which to base decisions.

This challenge directly affects a process similar to the OODA (Observe, Orient, Decide, Act) loop at the strategic level.[5] The gathering of data to form information and intelligence becomes part of the observation process. We must also consider, however, that data presence is only one attribute. We will orient our thinking around the information derived from that data, which underscores the need for the data to be reliable and credible. This may be communicated in terms of accuracy, and being free of external biases.

When this data is accurate (including reliability and credibility), then it can be acted upon. Decisions can be made, and the nation can proceed through the OODA loop more expeditiously.

[5] The context in which the OODA Loop is being used relates to the ability to maintain the initiative during competition (or battle). This can be described more fully at https://www.canada.ca/en/department-national-defence/maple-leaf/defence/2023/11/ooda-loop-halfbeat.html.

This is important when considering maintaining the initiative in discussions, and the movement of technology to market. Where we lose the non-repudiation aspects and the data's reliability and credibility are called into question, the need to circle back and reestablish that credibility can be measured in terms of both direct impacts and losses/delays of realizing opportunities.

Manipulation

We can also argue that we have seen a blurring of the divide between conflict and competition. The presence and penetration of social media and similar technologies have increased the power of information operations to identify, recruit, manipulate, manage, or even dissuade groups that would otherwise be difficult to form. Instead of having to locate susceptible persons physically, nation-states can push ideas onto social media that can attract and hook participants into specific ways of thinking and indoctrinate them through principles like repetition and the complete discouragement of critical thinking, argument, or discussion.

The manipulation challenge impacts most aspects of a nation's activities. It affects the voter bases, public opinion writ large, and can attempt to sway decision-makers by presenting one view as more acceptable than others.. One might theorize that social media ,with its truncated messages and need to cause sensational impacts, stands directly opposed to the necessary approaches of critical thinking and sound judgment. It focuses on building an emotional response that people relate to, and adopt as their own. It recruits those individuals (or their accounts) to serve as rebroadcasting points in echo chambers.

Sensitivity Labels

Canada currently manages several different kinds of sensitive data. These range from the significant values of personal identification data that falls under privacy acts, to the concept of classified information that describes information that falls in the national interest.

The concept of national interest is not understood consistently. For the purposes of this document, the concept of national interest used resides in the Treasury Board of Canada's *Security Organization and Administration Standard*.[6] This focuses on :

- The Canadian economy.
- Defence, Intelligence and certain Investigatory powers.
- Federal-Provincial Affairs.

Many of these efforts have keyed on preventing unauthorized disclosure of sensitive information and the injuries that result from it.

[6] Refer to Appendix C of the Security Organization and Administration standard as found at https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=12333

Single Datum versus Aggregates and Shifting Value[7]

Those who work in the Information Security domain remain focused on two aspects of sensitivity. The first involves the sensitivity of a single instance: the datum or a singular piece of information. These may have a level of sensitivity in their own right.

These groups also remain focused on the shifts in sensitivity that can occur when data or information is aggregated. Aggregation is the process of gathering, collecting, and collating singular entities and bringing them together. This change in sensitivity occurs because the collected data (or information further along the process) can be used to generate context, identify patterns and trends, and so on. That leveraging of the assembled information can then lead to intelligence.

For example, consider a grid of temperature sensors monitoring Arctic waters. The grid can provide information covering a large area of the ocean over a period of time, allowing scientists to identify conditions or features of importance. This picture, however, is composed of information gathered from multiple sensors. Each one of these sensors would be unable to paint the whole picture, but each is integral in helping generate a reliable and credible picture at the higher level. Even at the sensor level, the data it sends over time consists of single readings tied to a specific location. Those data may be less valuable than the assembled picture, but they still need to be reliable and credible to ensure the overall picture is trustworthy.

The difference in value is not necessarily linear in nature and may vary from information set to information set.

Shifts in Spans of Control

When transferring data to a third party, the ability to ensure it is handled decreases. When the custodian has direct control over the data, the control decisions are internal and subject to that organization's risk management structures and practices.

This assurance is reduced even with contractual (or similar) documentation that flows down to subcontractors . The past year has seen the rise of self-interest. Companies, even nations, have to offer a greater degree of assurance that data is being handled appropriately and that they will not put their own self-interest above the data custodian's needs. Further, nation-states are invoking laws that require data holders to share that data, even if it resides outside their borders.

In today's climate, once the data has been transferred to a third party, that specific data cannot be assured in terms of its treatment or its further disclosure.

Legislative and Regulatory Shifts

Canada is currently looking to pass Bill C-8, *An Act Respecting Cyber Security, which amends the Telecommunications Act and makes consequential amendments to other Act*s. This Bill is presently

[7] For a description of the relationship between data, information and intelligence, refer to the US Cybersecurity magazine's article by AJ Nash (Sprint 2017) as found at https://www.uscybersecurity.net/csmag/the-differences-between-data-information-and-intelligence/

labelled C-8 and applies to the 45th session of parliament. The bill was formerly Bill C-26 (same name but for the 44th session).

Bill C-8 will introduce the *Critical Cyber Systems Protection Act* as Part 2 of the overall act. That will introduce the concept of critical cyber systems that are defined as a "cyber system that, if its confidentiality, integrity, or availability were compromised, could affect the continuity or security of a vital service or vital system. While data in the national interest would become classified by definition, the concept of nation building projects (often tied to critical projects) may align more closely with this concept.

Key Elements to be Considered in the Framework

"Sovereign Data" can be characterized by the following:

- While it may align with the national interest, it more closely aligns with the concept of a critical service. This critical service aligns with the concepts of Critical Infrastructure Assurance (and Protection), rather than the national interest in the context of data or information becoming classified.

- While sensitive data in the national interest is well-defined in terms of classified with structures (including sensitivity levels and caveats like "Canadian Eyes Only," the sovereign data's value lies primarily in its availability and trustworthiness. This mirrors or approximates the same consideration seen for operational data for critical infrastructure.

- The threats that must be protected against go beyond the Confidentiality security attribute. The need for the data to be present in a trustworthy format invokes access control but also invokes the need for a constant monitoring of state and the handling of that data.

- Addressing this security need does not mean that the data cannot be transmitted, but means that there must be a trustworthy source (repository) of that data that can offer the assurances. That data (such as duplicates of the data) that are communicated to third parties would be seen as less trustworthy should they be returned.

- There would need to be an approach to how such data is stored, processed, and communicated to preserve its value. This raises the question of a sovereign processing capability that consists of both the processing capability but also controls to prevent threats to the availability or integrity of the data from being inserted through the supply chains tied to those services.

The framework for determining if data should be treated as sovereign data would involve the sequential set of questions:

- Does the data relate to something of national importance? Does it relate to the national interest directly (at which point the concept of classified may be invoked) or to a national priority that may support the concepts of defence, intelligence, the management of the economy, or federal-provincial affairs? *It may also be recommended that this be extended to the relationship between the Crown and First Nations as a means of establishing a base level of data on certain topics that both parties agree is trustworthy.*

- Are the core attributes that require preservation or assurance availability and integrity, but not necessarily confidentiality? The focus of these security attributes does not lie in preventing unauthorized disclosure but of ensuring the availability of data with adequate assurance that it exists in a trustworthy state.

Structure of Operations / Use Cases

Use Case 1: National Project Data

In this use case, data is collected that provides key inputs into national projects. This data may include scientific data collected as part of environmental assessments, survey data, or similar data that would be the foundation for decisions to proceed and, if proceeding, for what form future actions would take.

In this case, the data types would be identified, and it would be the responsibility to create a trustworthy and immutable source of data (i.e., once added, it cannot be changed) that could be shared amongst the different decision-making parties.

In this context, the declaration of "Sovereign Data" would mean that an external party could be given a certified copy of the data but would not be able to access the data source with manage, edit, or delete privileges. It could also involve a trusted data custodian making a copy of the data to be sent, ensuring that no external party can affect it.

Sample Structure of Operations

In this use case, the nation-building project requires that studies are performed that will provide data that supports decisions made to proceed, to not proceed, or with respect to what adjustments may be needed. This satisfies the first criteria in terms of general sensitivity.

The data is intended to be shared across different communities but needs to be available in a trustworthy state. In this context, the information may need to be shared across communities that are not security cleared as part of the efforts to ensure full and meaningful consultation. Given this balance across the security attributes, the assurance is with respect to the data's availability and trustworthiness.

The data is also seen in the context of being applicable to the general conditions of data necessary to support a critical service under Bill C-8. The compromise of this data is more in terms of its

improper handling, corruption, or loss of availability than the kinds of conditions found in Section 4 (1) of the *Foreign Interference and Security of Information Act*.

To preserve the assurance that the data is reliable and trustworthy, the onus is on the project to create an immutable copy of the data that becomes the official source. This immutable source becomes what might be described as a "read-only" copy that is preserved against deletion to prevent its loss or corruption from

Limitations Assumed to Apply

As with any project, some limitations could occur at the point between the sensor and the creation of the data. This would need to be addressed to ensure that the protective controls accurately reflect the data that should have been received from the sensor.

This also assumes that any system that would interact with Sovereign Data would necessarily be free of elements that would undermine the assurance that it could not be leveraged to attack the availability and/or integrity of the data. This would require an in-depth understanding of any service that directly affects the data, or services that preserve its availability, including external services that operate in the background.

Ownership and Accountability

The ownership of the data would necessarily fall under the federal government in trust. The federal government would become accountable for the preservation of the availability and integrity of the data.

This ownership, however, would need to fall under the oversight of the major stakeholders of a particular project to ensure that the federal government did not itself seek to cause the loss of availability or trustworthiness for its own purposes. This speaks to a need for transparency in the day-to-day operations of how such data is preserved in a trustworthy state and in terms of any changes to the infrastructure or supporting services that seeks to preserve it.

About the Author

Allan McDougall has focused on critical infrastructure protection and assurance over the past 30 years across military, public service, and the private sector. These have included within the Department of Fisheries and Oceans/Canadian Coast Guard, Transport Canada (as the Senior Inspector for Ports), and the Canada Border Services Agency. Within the private sector, he has worked to support transportation networks and maritime operations both domestically and in higher-risk environments. He is a founding member of the International Association of Maritime Security Professionals and one of Canada's original trainers under the IMO Train the Trainer program. He has worked across the lifecycle from design activities within the CSC Program at Irving Shipbuilding.

Allan has co-authored four books on Critical Infrastructure Protection and one book on Transportation Systems Security, which focus on establishing and managing resilient networks. These works have been used as graduate texts at several universities.

Allan holds an MA in Security Management from the American Military University (focusing on autonomous shipping) and a BMASc from the Royal Military College of Canada. Additionally, he holds a BA from the University of Western Ontario. He has several security-related certifications, including the Certified Protection Professional (ASIS), Physical Security Professional (PSP), Professional in Critical Infrastructure Protection (PCIP), Certified Master Anti-Terrorism Specialist (CMAS), and Computer and Information Systems Security Professional (CISSP).

Allan is a director at the National Center of Excellence and Innovation in Maritime Security and a Senior Security Program Manager at ADGA, located on the Eastern Shore of Nova Scotia.

About the National Center of Excellence and Innovation

Founded in 2024, the National Center of Excellence and Innovation is a multi-disciplinary focal point for communities coming together to address complex maritime security challenges. It draws together academics, practitioners, and others who have related experience ranging from law enforcement to the impacts of severe weather and changing ocean conditions. You can visit the website at https://marseccoe.com.