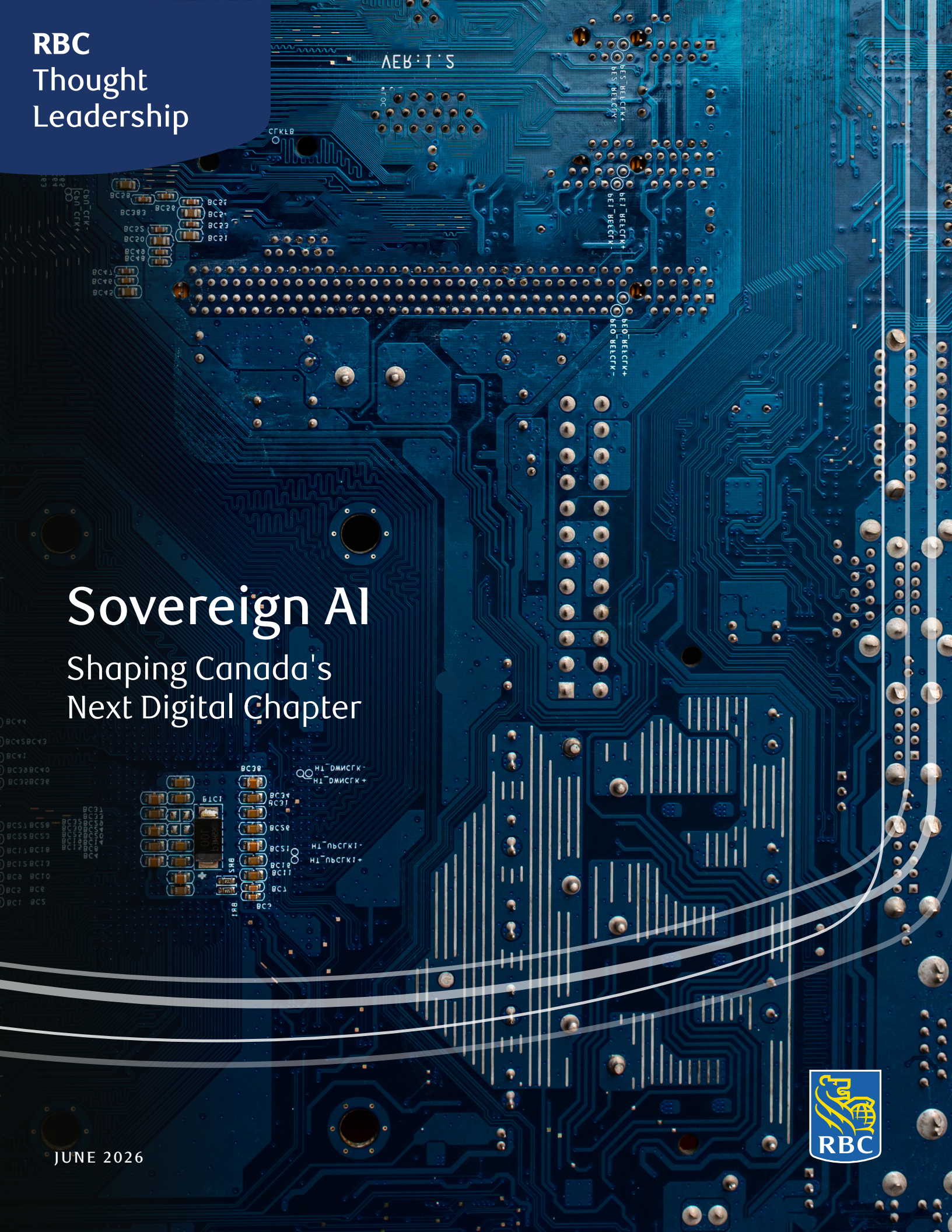
RBC
Thought
Leadership

Sovereign AI

Shaping Canada's Next Digital Chapter



JUNE 2026

American excellence in artificial intelligence has had an unintended side-effect: a hyper concentration in computing software and hardware. Three U.S. tech firms account for around 85% of Canada's cloud infrastructure spending, while another three account for roughly 88% of enterprise foundation model usage. Meanwhile, NVIDIA makes up about 80-90% of the advanced AI chip market.

Canadian firms and governments do not play a meaningful role in the global AI supply chain. But there is both a commercial and national demand to exert more control over our digital infrastructure to ensure deep AI capability as a country. The issue of "sovereign AI" has emerged as a critical issue as the world is swept up by the new technology. At the same time, choosing to build sovereign infrastructure and sovereign AI systems, or creating sovereign requirements for existing cloud infrastructure could involve a cost premium and could reduce technological competitiveness. Canadian businesses pursuing sovereign AI initiatives should consider which workloads to keep on existing cloud infrastructure, and which may require new architecture.

Canada is not alone in its pursuit of sovereign AI. More than 70% of global executives, investors and government consider sovereign AI as an "existential concern" or strategic imperative" to their goals, according to McKinsey & Co., which projects global sovereign AI to become a US$600-billion market by 2030.

For several years, the Canadian government has pursued new legislation to strengthen privacy protections and modernize its digital regulation. At the same time, a more assertive posture from Washington heading into the forthcoming Canada-U.S.-Mexico Agreement (CUSMA) review has cast some of Canada's digital regulatory efforts as potential trade irritants. The result is a live debate—on both sides of the border—over how Canadian businesses can continue to access the latest AI innovations while maintaining robust safeguards and protections.

Canadian firms have found themselves making infrastructure, data, and vendor decisions in an environment that has materially changed in the last 18 months. The AI stack—cloud, compute, foundation models, and the data those systems run on—has simultaneously become a top subject of trade negotiations, regulatory design, procurement strategy, and operational risk management. The choices being made now at the negotiating table and in Canadian boardrooms will set the conditions for the next decade of the digital economy.

For a start, the U.S. is now more aggressively pushing for its interests and tech sector dominance, with an immediate challenge being CUSMA's digital trade chapter, with a July 1, 2026, being the milestone for the six-year joint review.

Unlike several other trade agreements, on data localization specifically, CUSMA goes further than comparable agreements: unlike the Comprehensive and Progressive Agreement for Trans-Pacific Partnership, its rule against requiring domestic computing facilities (Article 19.12) contains no "legitimate public policy objective" exception, narrowing the space Canada would otherwise have to mandate local data storage for public services or citizen data.

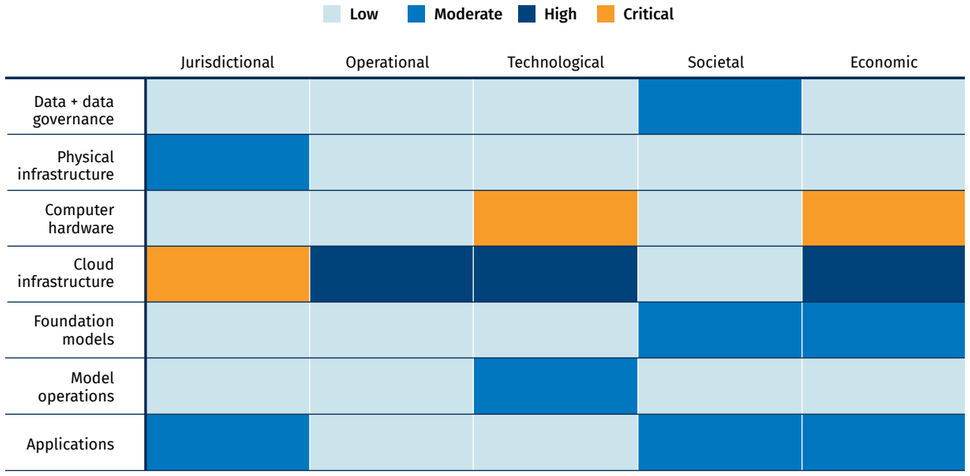
Four provisions matter most for AI and data sovereignty: 1) limits on less-favourable treatment of foreign digital products, 2) restrictions on blocking cross-border data flows for business conduct, 3) prohibitions on requiring domestic computing facilities as a condition of operating, and 4) limits on source-code disclosure requirements. There is still room to move, but in narrower channels: the federal government's national security exception, the federal procurement exclusion, and prudential carve-outs in financial services to preserve system stability. The federal government is also working on over a dozen targeted exceptions for defence.

Through a series of executive orders and policy shifts, the U.S. has hardened its position on digital trade. Washington is increasingly treating digital and AI regulation, including rules adopted outside its borders, as a trade irritant and barrier to market access. In March 2026 the United States Trade Representative (USTR) named Canada's sovereign computing initiative, digital rules, and proposed regulations as trade barriers in its National Trade Estimate report. U.S. Trade Representative Jamieson Greer has noted that all options are on the table regarding the future of CUSMA. A U.S. executive order in 2025 creating a federal AI Litigation Task Force, while a White House recommendation calling for a "minimally burdensome national standard," point in the same direction: AI regulation is being framed as commercial friction. In a CUSMA submission, the American industry-led Computer & Communications Industry Association (CCIA) labelled Canada's Online News Act and Online Streaming Act as discriminatory, suggesting major headwinds ahead for the Canadian government and businesses to protect their interests.

Ottawa moved on June 4, 2026, releasing AI for All, its national AI strategy, organized around trust, opportunity, and sovereignty. Two of its six pillars target exactly these dependencies—a "sovereign AI foundation" in compute, data, and talent, and scaling Canadian champions—under a "build-partner-buy" approach. The headline commitments are concrete: infrastructure "operated under Canadian control and Canadian law," sovereign-compute partnerships of 850 megawatts by 2030 scaling toward 2.3 gigawatts, a public supercomputer by 2031, and $700 million in affordable compute for smaller firms. Depending on execution, this may help address some of the gaps.

For Canadian leaders, sovereignty in the AI era should not mean isolation. It means freedom from coercion: the ability to choose which AI models to use, whose hardware runs AI inference, which jurisdiction governs data, and which providers can be substituted under pressure. Notably, Canada's domestic stack for infrastructure is developing faster than the prevailing narrative suggests, including Cohere's CoreWeave build out, Bell AI Fabric, TELUS AI data centres, the ThinkOn–Hypertec–Aptum–eStruxture consortium, and a new Canadian supercomputer are all building domestic AI and data infrastructure. Furthermore, Canada has many successful, global enterprise technology companies across AI models, financial services, healthcare technology, knowledge management and data storage, and beyond.

Canadian AI Sovereignty Risk Heat Map



In *Sovereign by Design* (Mullin & Khan, 2026) the above heat map rates "the severity of risk to Canadian AI sovereignty" at each layer/dimension, where sovereignty means freedom from coercion, and the ability to structure dependencies so they can't be used as leverage. The four tiers map to how concentrated the dependency is and whether Canada has a viable substitute:

Reading the heat map. Each cell rates the severity of sovereignty risk where a layer of the AI stack (rows) meets a dimension of digital sovereignty (columns). Severity reflects the degree of foreign dependency, supplier concentration, and substitutability—and how much leverage a disruption would hand to an outside actor.

- **Low—**Minimal exposure. Canada has domestic capacity, diversified supply, or ready substitutes; the dependency creates little room for foreign leverage.
- **Moderate—**A real but manageable dependency. Concentration or foreign exposure exists, but allied alternatives, partial domestic capacity, or workarounds limit the leverage it confers.
- **High—**A significant vulnerability. Heavy reliance on a small number of foreign providers or jurisdictions with few near-term substitutes; a disruption or coercive act would impose serious costs and be slow to route around.
- **Critical—**A severe chokepoint in Canada's supply chain. Dependence on a single (or a few) foreign-owned sources with few viable domestic or allied substitutes. Loss of access could halt or severely degrade AI capability and cascade through the stack.

Key considerations for firms

Four questions sit on the table for any Canadian firm of meaningful scale considering measures to increase its AI sovereignty. The stakes vary by sector—financial services, healthcare, and defence and critical infrastructure face the sharpest versions—but the underlying architectural choices are increasingly shared.

1. Plan for three regulatory scenarios, not one

The CUSMA review could produce continuity on digital trade (status quo), modest tightening of restrictions, or material renegotiation on digital trade. Firms could build their AI and digital strategies with each of these futures in mind. Domestic regulatory uncertainty compounds the trade variable. Canada's patchwork of AI and data laws could get clearer—or more tangled—as the federal government reforms the Privacy Act, PIPEDA, and other digital rules all likely at the same time. The Connected Care for Canadians Act (now Bill S-5, reintroduced in the Senate in February 2026) intersects health data interoperability with AI training in ways the legislation leaves largely undefined; and Bill C-8 is poised to extend supervisory expectations to telecommunications and adjacent critical infrastructure operators. The April 2026 Canadian Financial Sector Resiliency Group (CFRG) convening on Mythos signalled that frontier AI is now treated as a financial stability and cybersecurity concern, not simply a technology one. This is a shift that companies in energy, telecom, transportation, and water should expect on their desks next, likely ahead of a finalized supervisory framework for dealing with super powerful AI.

2. Treat sovereign AI as an opportunity, not just compliance

Canada's strengths in energy, expanding data-centre capacity, and emerging AI champions could form the basis of a sovereign stack that did not exist a few years ago. Firms in adjacent sectors—such as legal, professional services, and insurance—that adopt sovereign infrastructure could build a stronger position for future Canadian regulatory shifts or trade developments. Whether that stack reaches commercial scale will depend on procurement decisions by large anchor buyers. With bank AI adoption rising from about 30% in 2019 to 50% in 2023, and projected to reach 70% by 2026, the choices the Big Six and other major financial firms in the near future could determine whether a Canadian sovereign cloud ecosystem becomes truly viable. Canada's Defence Industrial Strategy and new NATO commitments also create a parallel growth path for dual-use firms serving Canadian, Five Eyes, and allied demand.

3. Get the talent and IP question right

Many technology-focused STEM graduates leave Canada, particularly top university software engineering graduates. But without people who have deep AI capability within government, Canadian companies and institutions will struggle to realize their potential. Beyond source code, the IP, model weights, fine-tuning datasets, and prompt instructions accumulated in production deployments are increasingly proprietary and valuable. Companies could also treat top AI graduates and people with significant AI skills even as strategic assets. The point lands hardest in healthcare, where procurement cycles for AI scribing, triage support, and administrative automation are top use cases.

4. Act collectively

No Canadian firm acting alone can move these questions. Industry associations—the Canadian Bankers Association, the Canadian Marketing Association, the Canadian Council of Innovators, and sector-specific bodies—are natural vehicles for some of the conversations now needed with government. The Business Council of Canada has made CUSMA review a central advocacy priority. Firms and industries that have not yet defined what they want from those conversations should do so now.

> Canada's strengths in energy, expanding data-centre capacity, and emerging AI champions could form the basis of a sovereign stack

Scaling Canadian AI champions

The Canadian entities below are operating, contracted, or rapidly building, and are among the major domestic firms currently available for sovereign AI procurement and partnership.

+ Cohere — frontier model capability

The Toronto-based company was founded in 2019. Cohere is the only Canadian-headquartered company building frontier-class language models with enterprise traction in regulated sectors, and recently reaching a combined ~US$20 billion valuation through its merger with Germany's Aleph Alpha. A federal MoU recognizes Cohere as a strategically important Large Language Model (LLM) provider, with $240 million in committed federal funding, and it is the anchor tenant of a new Cambridge, Ontario, AI compute facility operated by U.S.-based CoreWeave under the Sovereign AI Compute Strategy. Cohere's enterprise positioning includes sovereign deployment options for customer Virtual Private Clouds (VPC) and on-premises environments. The recent merger with Aleph Alpha extends reach into regulated European markets. The CoreWeave operating relationship has prompted reasonable questions in the ecosystem about how much Canadian ownership across the stack is required, achievable, or desirable; on balance, strengthening Cohere's competitive position by available means likely improves Canada's overall AI standing. RBC is a national partner and user of Cohere's North platform.

+ Bell AI Fabric—sovereign cloud and AI compute

Bell's $2 billion+ investment in Canadian AI compute, announced in 2024 and expanded in 2025, is anchored by NVIDIA infrastructure. It's positioned as Canadian-jurisdiction sovereign capacity for enterprise customers, with initial capacity targeted at federal, provincial, financial services, and health customers. The U.K., Germany, and France have adopted the telco-anchored sovereign cloud model, and Bell's scale makes it the largest single domestically controlled compute investment outside the federal program.

+ TELUS Sovereign AI Factory—carrier grade reliability

Announced in 2024, the second major Canadian telco offering domestic AI infrastructure began operations in 2025, in Rimouski, Quebec. The facility targets customers that require Canadian-resident, Canadian-operated AI compute with carrier-grade reliability. Two telco-anchored options mean meaningful procurement competition for Canadian sovereign cloud.

+ Sovereign Cloud Consortium—attaining critical mass

It's a coalition of mid-sized Canadian-owned data centre and cloud operators—ThinkOn, Hypertec, Aptum, eStruxture—offering sovereign cloud services for federal and regulated workloads. ThinkOn describes itself as the only Canadian-owned cloud service provider approved under the Shared Services Canada Framework Agreement for Secure Workloads at Protected B. The consortium could be an answer to the public-sector buyer's problem of needing scale without single-vendor lock-in.

+ Vector, Mila, AMII—research and talent

The three CIFAR-funded Pan-Canadian AI Strategy institutes—Vector (Toronto), Mila (Montreal), AMII (Edmonton)—generated much of the research base that produced Cohere, Element AI (now part of ServiceNow), and a deep bench of senior AI talent.

Other Canadian-owned providers, universities, and consortia currently operate at smaller scale, including in research compute and specialized regulated-sector hosting. Notably, Queens University and Simon Fraser University have signed a partnership on AI compute.

Implications for finance, healthcare, and national security

Financial services, healthcare, and defence and critical infrastructure are the Canadian sectors that face the highest stakes in the convergence of AI and digital sovereignty. Executives in these sectors are likely considering some of the same architectural decisions in the coming years. What follows is a frame of some of the major policy and technology choices in front of them.

Financial services

Canadian banks have moved past the question of whether to deploy AI at scale, with AI adoption moving from approximately 30% in 2019 to 50% in 2023, with 70% expected by the end of 2026, according to joint Office of the Superintendent of Financial Institutions (OSFI) and the Financial Consumer Agency of Canada (FCAC) data. In April 2026, the Canadian Financial Sector Resiliency Group (CFRG), a public-private partnership led by the Bank of Canada, convened on Mythos, signalling that frontier AI is now treated as a financial stability and cybersecurity concern, rather than simply a technology one. The practical consequences are concrete. First, model risk management built for credit and market models is the floor for AI governance, not the ceiling. Second, risk frameworks need to consider the whole pathway for inference data (where and what data queries are being processed by AI models) not simply training. And third, AI-enabled cyber scenarios are among the top threats for financial stability at a systemic level.

What the Big Six and other major financial service firms procure in the next 24 months may determine whether the Canadian sovereign cloud ecosystem reaches meaningful commercial scale. Financial services are not only a subject for Canadian AI policy, but they are also among the largest forces shaping it.

Healthcare

The Pan-Canadian Health Data Strategy continues to advance through Health Canada and the Canadian Institute for Health Information, but the infrastructure for moving health data across provincial boundaries remains underdeveloped. Quebec constrains cross-border health data transfers as a matter of binding provincial law, while Alberta, Ontario and British Columbia have parallel systems that overlap unevenly. The Connected Care for Canadians Act (Bill S-5), advancing through Parliament in 2026, includes provisions on health data interoperability that intersect with AI training data in ways the legislation leaves largely undefined.

Canada has a patchwork of laws on AI and data, which could get clearer, or even more tangled as the federal government reforms the Privacy Act, PIPEDA, and other digital rules. Further, the current provincial patchwork of laws is unlikely to be resolved anytime soon. Meanwhile, procurement cycles for AI scribing, triage support, and administrative automation are already moving without clear sovereign requirements to procure against.

Healthcare leaders and hospital executives likely have options when it comes to digital sovereignty and protecting Canadians' data—that might include federated learning, workload partitioning, deliberate data-flow design—but future leaders may want to invest in solutions that can ensure the data sovereignty of Canadians' health data. Either choice is defensible. But it is crucial to determine the choices through governance and technology architecture decisions, rather than leaving them up to international vendors.

Defence and critical infrastructure

Canada's Defence Industrial Strategy and new NATO commitments create growth opportunities for dual-use firms. Cloud infrastructure for defence presents three options: build within the U.S. ITAR compliance, build outside it for Canadian/allied demand, or build both. Canadian-classified data may require sovereign, separately operated infrastructure, while Five Eyes data likely needs more interoperable infrastructure. For critical infrastructure operators in energy, telecommunications, transportation, and water, the AI policy questions that financial services firms are working through are relevant. It will also be important to monitor the supervisory framework being considered in Bill C-8, an act respecting cyber security, amending the Telecommunications Act and making consequential amendments to other acts. The operators must decide whether to adopt AI risk management proactively or defer implementation until regulatory requirements are finalized.

Sources

Linked sources are publicly available. Government documents, regulatory guidelines, and major reports cited in the brief are listed by topic. Disclaimer: not all sources may be precisely accurate nor are former legal opinions or forward guidance.

Trade and U.S. policy

- CUSMA Chapter 19 — Digital Trade (Government of Canada) — full text of the digital trade chapter, including Articles 19.4, 19.11, 19.12, and 19.16.
- CUSMA Chapter 32 — Exceptions and General Provisions — National Security Exception (Article 32.2).
- USTR 2026 National Trade Estimate Report — USTR release identifying Canadian digital and sovereign-computing measures of concern.
- CCIA Submits Comments in USMCA Review, Urges Preserving Digital Trade Chapter - CCIA — The Computer & Communications Industry Association

AI sovereignty and Canadian capacity

- Sovereign by Design: Strategic Options for Canadian AI Sovereignty (Munk School, March 2026) — five dimensions of digital sovereignty; three procurement models.
- Canadian Anti-Monopoly Project (CAMP)—Parting Clouds: Creating a Competitive Marketplace for Compute
- Canadian Sovereign AI Compute Strategy (ISED) — $2B Sovereign AI Compute Strategy within a $2.4B Budget 2024 AI package; SCIP, AI Compute Access Fund, AI Compute Challenge.
- Government of Canada finalizes investment in Cohere (March 2025) — $240M federal contribution to the Cambridge AI compute facility.
- AI for All National Strategy (June 2026)

Financial regulation and AI risk

- OSFI Guideline E-23 — Model Risk Management (2027) — enterprise-wide model risk management framework, in force May 1, 2027.
- PwC Canada — 2026 Trust in AI Report — readiness gap analysis; 72% prioritize responsible AI but 36% lack governance.

Defence and dual use

- Canada's Defence Industrial Strategy: Security, Sovereignty and Prosperity (February 2026) — full strategy document.
- Prime Minister Carney announcement (February 17, 2026) — official launch announcement; BOREALIS, BDC Defence

Privacy, data governance, and provincial frameworks

- **Quebec Law 25** (Loi modernisant des dispositions législatives en matière de protection des renseignements personnels). Commission d'accès à l'information du Québec — cai.gouv.qc.ca.
- **PIPEDA** Personal Information Protection and Electronic Documents Act. Office of the Privacy Commissioner of Canada
- **EU Adequacy Decision for Canada (2024 renewal)** European Commission Justice and Consumers

Cyber threats and critical infrastructure

- National Cyber Threat Assessment 2025–2026 Canadian Centre for Cyber Security
- Critical Cyber Systems Protection Act framework (Bill C-8) Public Safety Canada
- Anthropic Mythos/Project Glasswing Globe and Mail — Canadian bank execs, regulators meet to discuss risks raised by Anthropic's new AI model - The Globe and Mail
- NGEN Funding BNN —NGen Announces $79.5 Million in New AI Projects to Help Canadian Manufacturers Compete Globally — ≈$29.2M new federal funding plus ≈$50.3M industry co-investment across 20 projects.

Contributors

Jaxson Khan, Senior Fellow at the Munk School of Global Affairs & Public Policy, University of Toronto

Sabreena Shukul, Research Associate, RBC Thought Leadership

Nora Bieberstein, Director, Strategic Programs
Yadullah Hussain, Managing Editor
Caprice Biasoni, Design Lead

Published by

RBC Thought Leadership

June 2026