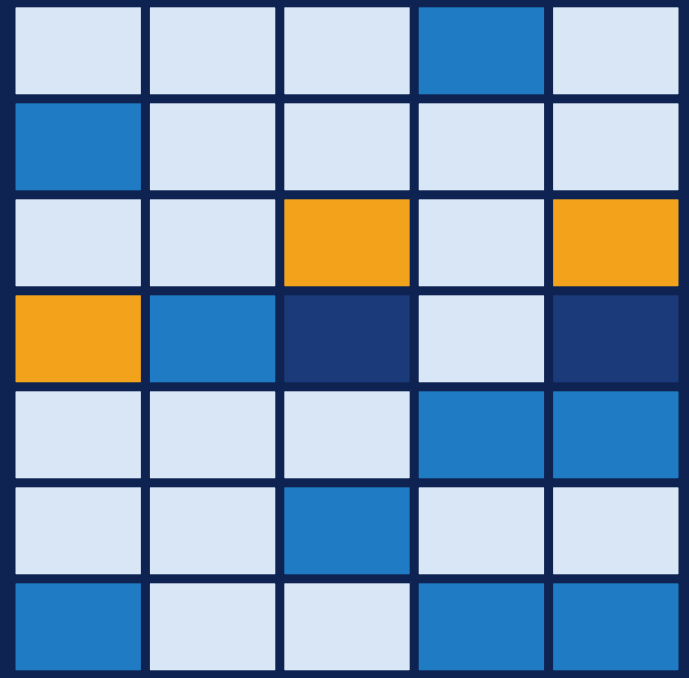
Sovereign AI

Shaping Canada's next digital chapter



October 2026


RBC
Thought
Leadership
Sovereign AI
Shaping Canada's
Next Digital Chapter
JUNE 2026
RBC

The Concept of Sovereignty

Our report defines AI sovereignty as freedom from coercion — the ability to structure dependencies so they cannot be used as leverage

Which models we use

Can we switch foundation-model providers without losing capability, controls or audit history?

Whose hardware runs inference

Are our inference workloads tied to one chip vendor or one hyperscaler region?

Which jurisdiction governs the data

Do foreign-law access regimes reach the data and prompts flowing through our AI systems?

Which providers can be substituted under pressure

If a provider is compelled, sanctioned or restricted, how fast can we route around it?

> *"For Canadian leaders, sovereignty in the AI era should not mean isolation. It means freedom from coercion."*

The starting point: a hyper-concentrated AI supply chain

Canadian firms do not yet play a meaningful role in the global AI supply chain

~85%

of Canada's cloud infrastructure spending goes to three U.S. tech firms

~88%

of enterprise foundation-model usage is served by another three U.S. firms

80–90%

of the advanced AI chip market is held by NVIDIA alone

US$600B

projected global sovereign-AI market by 2030 (McKinsey & Co.)

The trade-off: sovereign infrastructure and sovereign requirements can carry a cost premium and reduce technological competitiveness. More than 70% of global executives, investors and governments nevertheless call sovereign AI an “existential concern” or strategic imperative.

What has changed since we published the report in June

SUMMER — TRADE AND ACCESS UNCERTAINTIES

Jun 4
Ottawa releases AI for All strategy: 850 MW sovereign compute by 2030, public supercomputer by 2031, $700M affordable compute

Jun 12–Jul 1
U.S. Commerce export directive takes Anthropic's Fable 5 and Mythos 5 offline worldwide for 19 days (see next slide)

Jul 1
U.S. declines to renew CUSMA in its current form — annual reviews through 2036; Chapter 19 remains fully binding

Jul 13
Reuters reveals OSFI's Apr 29 email to CROs, CTOs and CISOs warning Mythos-class models compress the time to fix vulnerabilities

Jul 23
ISED opens national AI-transparency consultation (labelling, incident tracking, agent monitoring); closed Sept 23

Aug 22
Canada suspends trade talks; digital and cultural policy among cited disputes

Sep 8
Canadian counter-tariffs covering electronics take effect

FALL — PROGRESS TOWARDS A SOVEREIGN BUILD-OUT

Early Sep
Minister Solomon and FCM launch Canada's Responsible Data Centre Development Principles, a common national framework

Sep 14
Ottawa welcomes Bell AI Fabric expansion in Saskatchewan; exploring up to 6 GW of new data-centre capacity and $100B+ investment

Sep 16–17
ALL IN Montréal: Cohere–Aleph Alpha definitive deal (~$20B, pending approval); Canada and Germany fund LawZero; first Sovereign Technology Alliance meeting

Oct 2
PM Carney launches a National Council on AI to advise on adoption, national champions and sovereign infrastructure

Oct 5–8
Cohere forms a global enterprise-AI alliance with PwC, starting in Canada, and opens a third Canadian office in Ottawa

Oct 13
Ottawa adds $500M to help businesses adopt AI and commercialize Canadian-made technology

Oct 29
Canada to invest up to $1.6B in Toronto-based AI company Cohere

Case study: nineteen days without a frontier model

The June 2026 Anthropic export directive was an unplanned, real-world test of AI substitutability

What happened

Jun 9	Anthropic launches Fable 5 (public) and Mythos 5 (gated)
Jun 12	U.S. Commerce Dept. directs Anthropic to bar all foreign nationals, anywhere, from both models — citing national security and a reported jailbreak that surfaced software vulnerabilities
Jun 12, evening	Unable to verify nationality in real time, Anthropic disables both models for every customer worldwide; other Claude models unaffected
Jun 26	Mythos 5 partially restored to vetted U.S. organizations only
Jun 30 – Jul 1	Controls lifted; Fable 5 returns globally with additional safeguards agreed with Commerce



Concentration risk materialised through policy, not an outage

No system failed. A private letter withdrew a technology globally within hours. Traditional resilience testing does not model this trigger.



Recovery timelines were outside customer control

Restoration depended on a negotiation between a vendor and a foreign government. Neither our contracts nor our SLAs had a seat at that table.

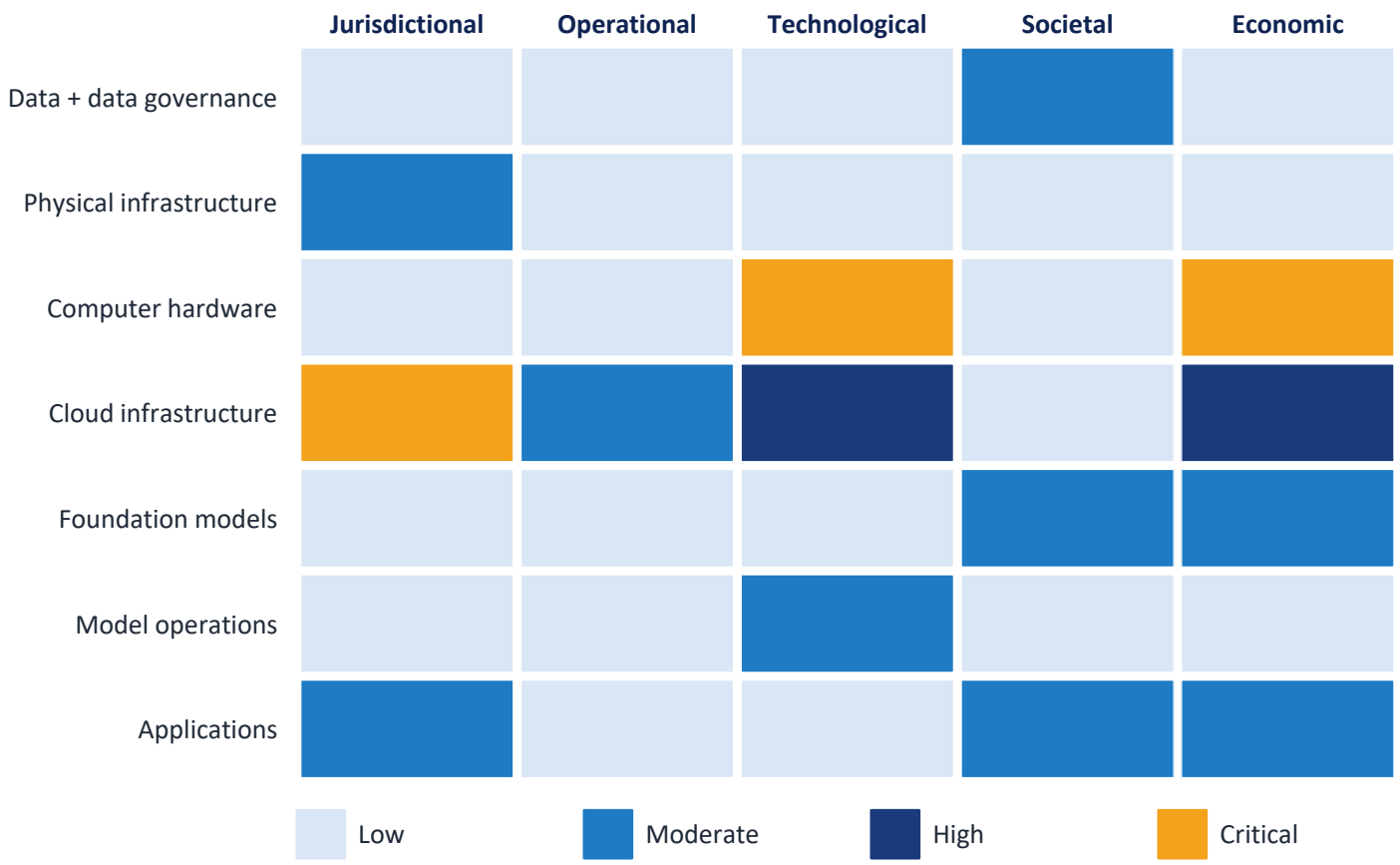


Substitution has to be pre-built

Firms with model-agnostic architecture and a second qualified provider absorbed the gap; others waited. Euro-zone banks were left out of Mythos's initial rollout entirely.

Where the exposure sits: the sovereignty risk heat map

Severity of foreign dependency, supplier concentration and substitutability at each layer of the stack
(Sovereign by Design, Mullin & Khan, Munk School, 2026)



Potential Risk Considerations

Critical — hardware and cloud jurisdiction. Chip supply and the legal reach over cloud infrastructure are the true chokepoints: no near-term domestic or allied substitute.

High — cloud technology and economics. Heavy reliance on a small number of foreign providers; slow and costly to route around.

Moderate — models, applications, data governance. Real but manageable: Canadian and allied alternatives exist, and this is where substitution planning pays off fastest.

CUSMA Chapter 19 narrows the room to manoeuvre

Chapter 19 remains fully binding after the July 1 review — and is now renegotiated annually through 2036

Four provisions that constrain data and AI sovereignty

Art. 19.4

Non-discrimination

No less-favourable treatment of foreign digital products

Art. 19.11

Cross-border data flows

Limits on blocking data transfers for business conduct

Art. 19.12

Computing-facility localization

Cannot require domestic data centres as a condition of operating — and, unlike CPTPP, no “legitimate public policy” exception

Art. 19.16

Source code and algorithms

Limits on compelled disclosure; model weights and training data not expressly addressed — now a live interpretive dispute

The narrower channels that remain

- National security exception (Art. 32.2)
- Federal procurement exclusion — the basis for Ottawa's sovereign-cloud approach to government data
- Prudential carve-outs in financial services to preserve system stability — the channel most relevant to us
- Over a dozen targeted defence exceptions under development

Pressure is rising: USTR's 2026 National Trade Estimate names Canada's sovereign computing initiative and digital rules as trade barriers; CCIA has asked U.S. negotiators to target the Online Streaming Act and Online News Act. Washington increasingly frames AI regulation itself as commercial friction.

Plan for three regulatory scenarios, not one

The CUSMA outcome is unknowable; the architecture decisions we make now should hold across all three

Continuity

Status quo on digital trade

Chapter 19 disciplines persist unchanged through annual reviews.

What it means

Sovereignty remains voluntary. Value from AI is realised on existing cloud; sovereign options are used selectively for the most sensitive workloads.

Modest tightening

Side letters and interpretations

AI-specific interpretation on source code and model weights; sovereign-cloud framework for government data; privacy interoperability commitments.

What it means

Audit-access and explainability demands rise. Documentation and validation evidence for third-party models becomes a hard requirement.

Material renegotiation

Chapter reopened

Localization and data-flow rules re-cut amid tariff escalation; sovereignty requirements extended to regulated sectors.

What it means

Rapid migration of regulated workloads to Canadian-jurisdiction infrastructure. Firms that partitioned early move; others rebuild under pressure.

Domestic uncertainty compounds the trade variable. The Critical Cyber Systems Protection Act (Bill C-8, Royal Assent June 16) is phasing in, while privacy reform (Bill C-36, replacing PIPEDA's private-sector rules) and the Connected Care for Canadians Act (Bill S-5) are still before Parliament. OSFI made frontier AI the focus of its October risk outlook and has issued bulletins on generative, agentic and frontier AI — a finalised supervisory framework for frontier models is still ahead.

The Canadian sovereign stack is maturing

Entities operating, contracted or rapidly building — the procurement options available today

Cohere

Frontier model capability

Only Canadian-HQ'd frontier LLM builder; ~US$20B valuation; Aleph Alpha merger signed Sept 16, pending approval; $240M federal funding; anchor tenant of the Cambridge, Ont. facility. VPC and on-prem deployment. RBC is a national partner and user of North.

Bell AI Fabric

Sovereign cloud and compute

$2B+ investment anchored on NVIDIA. June 18: Merritt, B.C. facility to double its 6.5 MW by early 2027 to run Cohere's models on Hypertec-built hardware via BUZZ HPC's sovereign cloud.

TELUS Sovereign AI Factory

Carrier-grade reliability

Rimouski, Que., operating since 2025; proposed B.C. cluster with Ottawa targets 60,000+ GPUs by 2032. Two telco-anchored options mean real procurement competition.

Sovereign Cloud Consortium

Critical mass for regulated workloads

ThinkOn, Hypertec, Aptum, eStruxture. ThinkOn was selected (June 25) to build Canada's Sovereign AI Platform for Shared Services Canada — scale without single-vendor lock-in.

Vector, Mila, AMII

Research and talent

The three CIFAR-funded institutes behind Cohere, Element AI and Canada's senior AI bench. Talent retention is itself a sovereignty variable.

AI for All (June 4, 2026)

Federal commitments, execution pending

- Infrastructure “under Canadian control and Canadian law”
- 850 MW sovereign-compute partnerships by 2030, scaling toward 2.3 GW
- Public supercomputer by 2031; $700M affordable compute for smaller firms

Four considerations for firms

1. Plan for three regulatory scenarios, not one

Build AI and data strategy so it holds under continuity, tightening or renegotiation — and under a domestic patchwork that may get clearer or more tangled.

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

What the Big Six procure in the next 24 months may decide whether a Canadian sovereign cloud reaches commercial scale. Risk shapes that procurement.

3. Get the talent and IP question right

Model weights, fine-tuning data and prompt libraries are accumulating IP; people with deep AI skill are strategic assets. Both need protection in vendor terms.

4. Act collectively

No firm can move these questions alone. The CBA, CFRG and the Business Council are the vehicles; firms that have not defined their asks should do so now.

Sources and notes

Report

RBC Thought Leadership, Sovereign AI: Shaping Canada's Next Digital Chapter (June 2026). Contributors: Jaxson Khan (Munk School), Sabreena Shukul, Nora Bieberstein, Yadullah Hussain, Caprice Biasoni. Heat map: Mullin & Khan, Sovereign by Design (Munk School, March 2026).

Updates since publication

- ICPA, USMCA Joint Review — Digital Trade tracker (verified Sept 6, 2026): July 1 non-renewal and annual reviews to 2036; Aug 22 suspension of talks; Sept 8 counter-tariffs; AI-transparency consultation to Sept 23; Article 19.16 dispute.
- Fortune (Jun 13, 2026); Harvard Law Review blog (Jun 2026); Bloomberg Law and Sovereign Magazine (Jul 1, 2026): Commerce Dept. export directive on Fable 5 / Mythos 5 and restoration.
- Reuters via Yahoo Finance / U.S. News (Jul 13, 2026): OSFI April 29 email on Mythos-class cyber risk; Globe and Mail (Jun 2026) on Canadian exposure.
- The Logic, Globe and Mail, Bell/Cohere/Hypertec/BUZZ HPC release (Jun 18, 2026): Merritt, B.C. expansion.
- OSFI Guideline E-23 Model Risk Management (final Sept 11, 2025; effective May 1, 2027) — Torys, McMillan, Protiviti summaries.

Updates, September–October 2026

- ISED (Sept 2026): Responsible Data Centre Development Principles; ALL IN 2026 announcements (LawZero funding, Sovereign Technology Alliance, Mila–Mozilla–Hypertec open-source consortium); Bell AI Fabric investment in Saskatchewan.
- Prime Minister's Office (Oct 2, 2026): National Council on AI. FedDev Ontario (Oct 2026): expanded $500M support for AI adoption and commercialization.
- BetaKit and AI Weekly (Sept 2026): Cohere–Aleph Alpha definitive agreement; PwC (Oct 5, 2026): Cohere alliance; BetaKit (Oct 2026): Cohere Ottawa office; BNN Bloomberg (Oct 15, 2026): Cohere–Aisle.
- Fact-check updates: Public Safety Canada (Jun 16, 2026) on Bill C-8 Royal Assent; Bill C-36 (introduced Jun 15, 2026); OSFI Semi-Annual Risk Outlook (Oct 2026); ThinkOn (Jun 25, 2026); Torys, Investing in Canada (Fall 2026) on the TELUS B.C. cluster; The Logic (Sept 11, 2026) on Cohere's funding; BNN Bloomberg (Oct 25) and Globe and Mail (Oct 26) on trade talks.
- The Next Web on Cohere and sovereign ownership; Open Canada, “The human dimension of Canada's AI strategy.”