



LIMITED-TIME SALE \$0.49/WEEK

OPINION

Canada's new AI data centres could be obsolete before they open

ROBERT DIAB

SPECIAL TO THE GLOBE AND MAIL

PUBLISHED 1 HOUR AGO



Alberta Premier Danielle Smith, centre left, is seen at a Calgary press conference in June, where she announced that Meta will build a sprawling AI data centre north of Edmonton.

Listen to this article

[Learn more](#) about audio

[Log in](#) or [create a free account](#) to listen to this article.

Robert Diab is a law professor at Thompson Rivers University.

Earlier this month, Alberta Premier Danielle Smith announced that Meta Inc. will build a sprawling \$13-billion AI data centre north of Edmonton. It will consume about a gigawatt of electricity – roughly 70 per cent of what the city of Edmonton draws – from a new plant that runs on natural gas. While Meta's project is private, Canada has committed \$925-million over five years to building “sovereign AI data centres” across the country, to keep us competitive as AI transforms the economy.

Ms. Smith's announcement comes as opposition to data centres is growing, with New York last week becoming the first U.S. state to pause new data-centre construction, cities in Ontario and California voting to block them and Manitoba Premier Wab Kinew turning down proposals to build them.

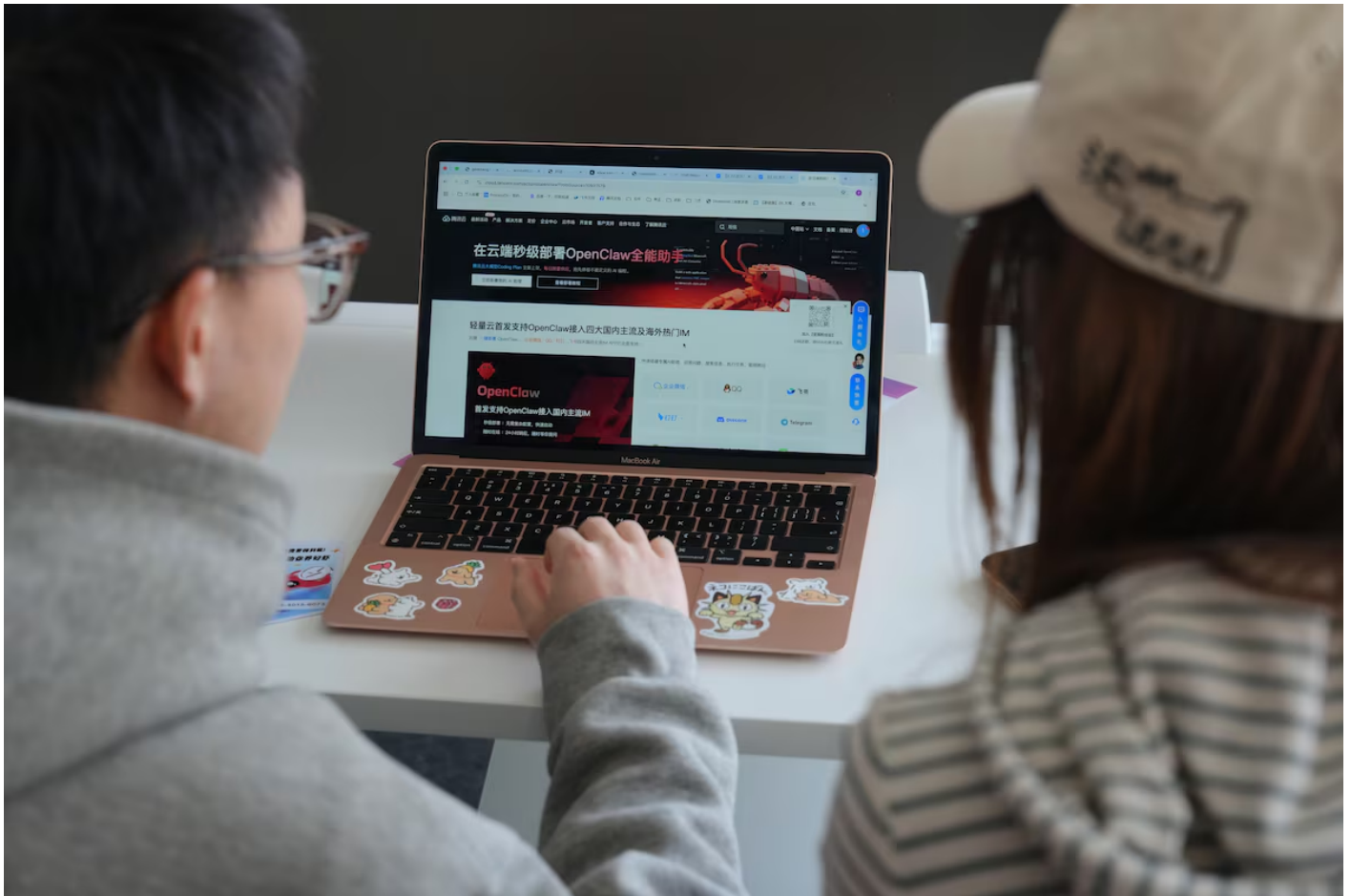
What to know about Ottawa's push for more data centres

The argument against these large AI data centres rests mainly on environmental grounds and doubts about their economic benefits. The centres use massive amounts of water and electricity, as North America sets new heat records every year. Project advocates wave away these concerns, promising to use closed-loop cooling or clean hydro power. Opponents point to the lack of independent environmental assessments.

But governments have an additional reason to be cautious about funding or approving large AI data centres – and it has less to do with the environment than with the technology itself.

AI data centres may turn out to be overbuilt, or even obsolete, by the time they are finished, because more of our AI processing is shifting from the cloud to

our personal devices. The language models behind popular tools like ChatGPT, Claude, and Gemini vary widely in size. And smaller, open-source models you can run for free on your laptop, with no connection to the internet, are growing more powerful by the day.



People install the RedClaw AI software on a laptop at an office in Beijing, in March.

ANDY WONG/THE ASSOCIATED PRESS

These small models are typically distilled from larger ones, then released free to download. They read, edit and summarize documents well, but do less well at multi-step reasoning or non-trivial coding. And as these small-language models (SLMs) become more popular, the explosive demand we see now for large-language models (LLMs) may soon crest, or begin to fall, before the centres they rely on are running at full capacity.

Across Canada, the fight against artificial intelligence goes offline

Until recently, small open-source models were too limited to be more than a hobbyist's curiosity – the ham radios of AI. But in June, when Google released

the 12-billion-parameter version of its open-source Gemma 4 model – which handles text, images and audio, and is small enough to run on a laptop – we crossed a meaningful threshold. Tests show it performs about as well as the smaller cloud models that leading AI firms were charging for as recently as 2024, like GPT-4o mini.

Two years is a long time in AI, and today's frontier models from Anthropic and OpenAI are far more capable. But Gemma 4 12B is hardly the upper limit for SLMs. Larger open-source versions are already available, giving a law firm or a government office enough capacity for much of what they do while running on hardware that costs roughly \$5,000 to \$10,000.

For many uses, the advantages of shifting from the cloud to local AI are becoming hard to ignore. Since nothing leaves the premises, your data stays private, including your prompts and results. There are no metering fees and no limits on how much you run.

This may sound speculative. But it is about to go mainstream.

This fall, Apple will roll out its newest operating systems for the iPhone and Mac. Their updated AI assistant will handle many everyday requests on the device itself – surfacing data from e-mails and texts, adding calendar dates, and drafting or editing documents – without sending them to the cloud.

For more demanding queries, including the new conversational Siri, Apple will still rely on a large model running in the cloud. But the trend is clear. Your phone and tablet will run more of your AI locally, for privacy and speed. Many Canadians will soon be doing much of their AI on their devices.

Of course, we are unlikely to give up the large frontier models in the cloud. For those, we will need data centres. But how many we need, and how big, remains uncertain. Demand may be high now, but much of AI development is moving the other way – toward models that are smaller, cheaper and closer to the user. This should give governments pause.

Sign up for the Opinion Newsletter. Keep your opinions sharp and informed with a helpful summary of The Globe’s most provocative columns, covering politics, world affairs, business, arts, health and the way we live

+ SIGN UP | EXPLORE NEWSLETTERS
Report an editorial error

Report a technical issue

Editorial code of conduct

Terms & Conditions

Community Guidelines

Privacy Policy

Disclaimer

Subscribe

Business services

Contact us

Reader services

About us



© Copyright 2026 The Globe and Mail Inc. All rights reserved.

351 King Street East, Suite 1600, Toronto, ON Canada, M5A 0N1

Andrew Saunders, President and CEO