Centre *for* MEDIA, TECHNOLOGY *and* DEMOCRACY

**POLICY BRIEF**

# AI News Audit

## AI, Canadian Journalism, and Paths for Policy Action


**Taylor Owen & Aengus Bridgman**

**Taylor Owen** is the Beaverbrook Chair in Media, Ethics and Communication and Founding Director of the Centre for Media, Technology and Democracy, McGill University.

**Aengus Bridgman** is Associate Director, Research, and Director of the Media Ecosystem Observatory at the Centre for Media, Technology and Democracy, McGill University.



MARCH 2026

AI companies built their products using Canadian journalism without permission and without compensation, and are now delivering that journalism to consumers as their own product. Existing copyright and media policy frameworks were not designed to address this.

In February and March 2026, we conducted the first large-scale empirical audit of how AI models use and distribute Canadian journalism. We ran two studies. First, we tested four major AI models on 2,267 real Canadian news stories in both English and French (18,134 queries in total) to measure what models have absorbed from their training data and whether they attribute it. Second, we enabled web search and asked the same models about 140 specific recent articles from seven Canadian outlets across 3,360 experimental conditions, to measure whether AI models produce viable substitutes for current journalism and whether they credit the source.

When asked about Canadian news events drawn from their training data, ChatGPT, Gemini, Claude, and Grok provide no source attribution 82% of the time. When given web access and asked about specific recent articles, the same models covered enough of the original reporting to substitute for the source in 54 to 81% of cases. Models linked to Canadian news sites in 29 to 69% of responses, but named the originating outlet in the response text in only 1 to 16% of cases. When we named the outlet and asked the same models for citations, attribution rates reached 74–97%. The rules governing how these companies use journalism (who gets credited, who gets compensated, and what obligations attach to those who profit), are being set right now, by default, through inaction. Canada has tools and precedent to act responsibly.

This memo accompanies the technical brief *AI News Audit: How AI Models Use and Distribute Canadian Journalism*, which contains the full methodology, data, and analysis. What follows are the implications of that evidence for Canadian policy

**A note on AI in this research.** This project was an experiment in developing an AI-assisted research methodology. Two senior researchers, Aengus Bridgman and Taylor Owen, designed a pipeline in which AI tools were embedded at each stage of the process, from study design to data collection and response coding to statistical analysis and prose drafting to graphic design, to test what this methodology could produce in a compressed timeline. Claude (Anthropic) was the primary AI tool used throughout. All code and content was reviewed, tested, and verified by Bridgman and/or Owen. The methodology itself is part of what this project set out to test.

## 1. What the Study Shows

AI companies interact with Canadian journalism at three stages (ingestion, production, and distribution), and at each one they extract value without returning it.

***Ingestion: AI companies have absorbed Canadian journalism into their training data at scale.*** Models we tested demonstrated detailed, accurate knowledge of Canadian current events for stories within their training windows: all but Claude answered questions about domestic politics, provincial affairs, and local events with a specificity that points to Canadian journalism as the primary source. Yet among those knowledgeable responses, 92% provided no attribution. The journalism has been converted into training data and the source has been erased.

***Production: AI models synthesize that journalism into substitute products that reduce consumers' need to visit the source.*** When we enabled web search and asked models about 140 specific recent articles from seven Canadian outlets, every model produced responses that covered enough of the original reporting to reduce a consumer's need to visit the sources in 54% to 81% of cases. Models credited the originating outlet in only 1% to 16% of those responses.

***Distribution: AI models are becoming an important channel through which the public encounters news, but they rarely send readers back.*** Models included links to a Canadian website in 29% to 69% of responses, but rarely named the originating outlet in the response text. Links provide a pathway back to the source; a consumer who clicks through can reach the newsroom. But a consumer who simply reads the response will not know which newsroom produced the reporting. The outlets that receive the most AI visibility are a handful of large, free, nationally prominent organizations. CBC, CTV, and the Globe and Mail capture most of the attention. Paywalled and regional outlets, including many that do substantial original reporting, fall well below proportional representation. The Toronto Star received 11 named-as-source mentions across over 18,000 responses. The Montreal Gazette received one.

French-language journalism faces a compounded version of this problem. French stories were absorbed into training data at rates comparable to English ones, but French outlets appeared in citations only 10% of the time. Radio-Canada and La Presse dominate the small number of French citations that do appear. The Journal de Montreal, one of Quebec's most widely read newsrooms, is nearly invisible to AI systems. Their content is ingested. Their contribution is erased.

## 2. Implications for Journalism

The pattern our data reveal is not a series of isolated failures. It is a structural shift in how the informational value of journalism is captured and distributed. AI companies have built commercial products that depend, in significant part, on the reporting that Canadian journalists produce. They have done so without compensation, without attribution, and without any obligation to sustain the infrastructure they are drawing from. The result is a system that accelerates the economic decline of the journalism it relies on.

This is not the same problem that social media posed. Platforms captured advertising revenue by aggregating attention around news content. That was a distribution and monetization problem. AI companies are doing something different: they are absorbing the substance of journalism, and delivering it directly to consumers as their own product. The consumer's need to visit the source is not just reduced by algorithmic demotion, as it was with social media. It is rendered unnecessary by the AI's response itself.

The democratic implications extend beyond the economics of newsrooms. When AI models deliver confident answers about Canadian public affairs that are drawn from journalism but stripped of source and context, the public's ability to assess the reliability of the information it receives is diminished. Canadians are increasingly receiving information about their own politics and society through (largely American) systems that neither credit the source nor guarantee accuracy.

## 3. Implications for Policy

Two existing policy instruments are directly relevant, and neither currently reaches what AI companies are doing.

The Online News Act established the principle that technology companies profiting from the work of Canadian

journalists should enter into a fair process to determine the value of this exchange. Our findings show that these same underlying principles are likely also applicable to AI companies. But the Act's definitional architecture, built around entities that index and display news content, does not capture companies that absorb and synthesize it.

The question is whether and how C-18's scope should be extended to a fundamentally different form of intermediation. That is not a simple amendment. C-18's bargaining framework was designed to quantify the value of facilitating access to news content; there is no obvious equivalent metric for what AI companies do when they absorb reporting into model weights and deliver derivative answers to consumers. The $100 million annual Google agreement under C-18 provides a reference point for the scale of compensation at stake, but the mechanism for arriving at obligations for AI companies is genuinely unresolved.

The Copyright Act provides the other potential lever, but its application to AI training remains entirely unsettled. The Act's fair dealing doctrine is limited to enumerated purposes, and whether large-scale commercial AI training constitutes "research" has never been tested in Canadian courts. The government's 2023 consultation revealed deep disagreement and produced no formal position. The publishers' lawsuit against OpenAI, cleared to proceed in Ontario, will eventually produce judicial answers. But courts will resolve individual disputes between specific parties. They will not design a copyright framework suited to the systemic challenge AI poses to journalism. Parliament is better positioned to do that, though the trade-offs involved, between enabling AI innovation and protecting creators' rights, between permissive training regimes and restrictive ones, are real and should not be treated as simple.

A growing number of jurisdictions are exploring a third path. Statutory licensing is structurally different from the bilateral bargaining model that underlies C-18. Rather than requiring AI companies to negotiate compensation with individual publishers, a statutory licence permits use by law while making payment mandatory, with rates set through a centralized mechanism rather than deal by deal. Policymakers in the EU, Indonesia, and Latin America are developing versions of this model for AI training, and the World Intellectual Property Organization has begun examining it as a global framework. The rationale is that it addresses two structural weaknesses of alternative approaches: bilateral deals tend to favor large publishers with the leverage to negotiate, while opt-out regimes are functionally unenforceable once content has been ingested into model weights. For Canadian journalism, where our data show that AI visibility is concentrated among a handful of large, nationally prominent outlets while regional and French-language newsrooms are largely invisible, this distributional question is central to any compensation framework. Whether a collective mechanism would produce more equitable outcomes than bilateral negotiation depends on design choices that have not yet been tested at scale.

Our findings contribute two pieces of empirical evidence to this policy debate. First, the scale of ingestion is not incidental: AI companies have systematically absorbed the back catalogue of Canadian journalism. Second, the outputs function as market substitutes for the original reporting. Both findings are relevant to the fair dealing analysis, particularly the market-impact factor that Canadian courts consider. But translating empirical evidence into statutory language requires choices about where to draw lines, and reasonable people will disagree about where those lines belong.

While the C-18 and copyright questions will take time to resolve, our data point to an area where intervention may be more immediately achievable. The 74–97% attribution rate we observed when users named the outlet and asked for citations demonstrates that the technical infrastructure for meaningful source identification already exists. A regulatory expectation that AI systems identify journalistic sources by default, in the response

text rather than buried in metadata, would not resolve the underlying economic questions. But it would restore a basic norm of credit that has been abandoned, it would give consumers the ability to assess and verify the information they receive, and it would make visible the full scale of AI's dependence on Canadian journalism, which in turn strengthens the basis for the larger policy responses. Any such standard would also need to account for the demonstrated disadvantage facing French-language outlets, whose content is absorbed at comparable rates but credited at a fraction of the rate of English-language journalism.

## 4. Next Steps

The governance gap we have documented is real and it is widening. The most consequential new form of value extraction from Canadian journalism sits outside the scope of every existing regulatory instrument. There is no compensation obligation, no attribution requirement, and no transparency about what has been ingested.

We do not pretend that the policy solutions are straightforward. Extending the Online News Act raises genuine questions about how to define and measure the value AI companies extract from journalism. Reforming the Copyright Act requires balancing the legitimate interests of creators against the potential costs of restricting AI development. Even attribution standards, which appear simple, involve design choices about what counts as a source and how far the obligation extends. Our findings have implications for four areas of Canadian policy:

1. **The Online News Act.** Our data demonstrate that AI companies are extracting value from Canadian journalism at a scale and in a manner that C-18 was not designed to address. The central question is whether the Act's scope can or should be extended to encompass AI companies that synthesize and distribute journalism, and if so, what obligation structure is appropriate for a form of intermediation that differs fundamentally from social media.
2. **Copyright and fair dealing.** The statutory review of the Copyright Act will need to address AI training directly. Our findings are relevant to the fair dealing analysis, particularly the market-impact factor: the substitution rates we document raise the question of whether large-scale commercial ingestion of copyrighted journalism falls within the doctrine's enumerated purposes. Whether transparency requirements for training data are warranted is a related but distinct question.
3. **Statutory licensing and international frameworks.** A growing number of jurisdictions are developing collective licensing models for AI training. Our data on the concentration of AI visibility among a small number of large outlets suggest that any compensation framework will need to account for the distributional question: whether bilateral negotiations or collective mechanisms are better suited to ensuring that regional and French-language outlets, which lack individual bargaining power, are not excluded from whatever arrangements emerge.
4. **Attribution standards.** The 74–97% attribution rate we observed when users named the outlet and asked for citations suggests that the technical capacity for meaningful source identification already exists. Whether a regulatory expectation of default attribution could be implemented through existing or forthcoming authority, and whether it could function as a practical step that does not depend on resolving the larger economic questions, is perhaps the most immediately tractable question our findings raise.

None of these questions have easy answers, and the trade-offs involved are genuine. But the cost of deferring them is also real. In the absence of deliberate policy choices, the terms of AI companies' relationship to Canadian journalism are being set by corporate design decisions made outside Canadian jurisdiction. The evidence we present here makes the scale of that relationship visible. What democratic institutions do with that evidence

is a political choice, not a technical one.

---

*This memo accompanies "AI News Audit: How AI Models Use and Distribute Canadian Journalism," a Media Ecosystem Observatory Technical Brief, March 2026. The full technical brief and methodology are available at github.com/MEOMcGill/ai-news-audit-brief.*