

WITNESS

SUBMISSION TO CALL FOR INPUT

Government of Canada public consultation on AI transparency

SUBMITTED ON: 23rd September, 2026
BY WITNESS

This submission aims to provide input to [public consultation](#) hosted by the Government of Canada's Innovation, Science and Economic Development Ministry to gather public views on strengthening transparency for artificial intelligence (AI) systems and AI-generated outputs. Part of Canada's National Artificial Intelligence Strategy: AI for All, the initiative seeks feedback on key areas such as detecting synthetic content, recognizing AI interactions, improving access to information about AI capabilities and limitations, and tracking AI incidents and agent activities, ultimately informing government steps to ensure safe, responsible, and reliable AI.

The present document addresses Sections present in the consultation formulary, as well as concepts present in the discussion paper: [Enhancing trust in artificial intelligence through increased transparency](#).

Consultation form

Respondent profile

1. **I am responding to this consultation as:** Policy and Advocacy Manager at WITNESS
2. **If you are responding on behalf of an organization, which sector best describes your organization?** Yes, on behalf of a Civil Society Organisation named WITNESS
3. **If you are responding on behalf of a for-profit organization or company, which of the following best describes its size?** NA

Content Related Questions

AI-Generated Content

1. **Would it help you trust what you see online if you could tell whether something was created by AI or by a person? When and for what kinds of content (e.g., image, video, audio) is it most important to know if something was created or modified by AI?**

Yes, distinguishing between AI-generated and human-created content, or the array of possibilities from one of the spectrum to the other, is essential for maintaining trust in digital information across modern information ecosystems. This distinction is particularly crucial for

WITNESS

audio, video, and image content, as these synthetic and manipulated formats are increasingly deployed to misrepresent real-world events, undermine civic integrity, endanger human rights defenders, and proliferate harms such as non-consensual intimate imagery. In an era where digital content can be synthetic by default, public trust relies on the ability to verify authenticity. Establishing clear provenance and transparency mechanisms is therefore critical in contexts involving journalistic reporting, civic documentation, judicial evidence, and democratic discourse, where synthetic manipulation directly threatens freedom of expression, public safety, and accountability.

2. What would best help Canadians determine when content is AI-generated (e.g., invisible or visible watermarks, disclaimers, provenance metadata)?

A multilayered marking approach that encapsulates both encrypted provenance metadata about the content combined with watermarks is necessary, mirroring key provisions of the Code of Practice of Article 50 of the EU AI Act Provenance metadata (such as the C2PA standard) is the most robust mechanism for establishing the chain of custody and authenticity of a file, preserving a complete record of origin and edits without needing to reveal personally identifiable information. While visible/invisible watermarks and disclaimers provide helpful signals, they are often insufficient on their own as they can be easily stripped, cropped out, or overlooked. Provenance metadata ensures the history of a file remains verifiable across platforms, while multi-layered visual or interactive disclosures provide accessible, human-understandable signals for everyday readers.

3. Who in the AI value chain (developers, deployers, or others) should be responsible for providing transparency around AI-generated content? Why?

Pipeline Responsibility is necessary (<https://blog.witness.org/2023/02/building-human-rights-oriented-guidelines-for-synthetic-media/>). Responsibility must be shared across the entire AI value chain, establishing clear accountability at each stage. Developers must prioritize 'authenticity-first' design by default, embedding open, interoperable provenance tools and marking techniques during model training and output generation. Deployers, such as social media platforms and distribution channels, are responsible for ensuring that this provenance information is preserved, verified, and clearly surfaced to users through accessible, interactive disclosures. Finally, governments must play an active regulatory role to enforce these duties, setting binding guardrails and standardizing requirements so that compliance is mandatory industry-wide rather than relying on voluntary, fragmented commitments.

4. Do existing market practices, technical tools, and legal frameworks make it easy enough to know when content is AI-generated? If not, where do gaps remain and what actions (e.g., regulatory measures, guidance and codes of conduct,

WITNESS

standards and technical solutions, research and development, literacy initiatives, procurement requirements) do you think the Government should take?

Significant market gaps persist because existing practices remain fragmented and largely voluntary. To establish meaningful oversight and foster digital trust, we suggest the following strategic actions:

- **Mandate Open and Interoperable Provenance Standards:** Enforce the mandatory adoption of open standards (e.g., C2PA) across all major AI development and distribution platforms to ensure end-to-end verifiability of digital content. Simultaneously, we should push for open standards to meet requirements aimed at reflecting public interest, as legally established. This includes, for example, ensuring that open standards are designed to protect privacy and promote inclusion and accessibility (<https://library.witness.org/product/c2pa-privacy/>)
- **Establish Standardized Labeling Frameworks:** Implement clear, consistent, and user-accessible disclosure guidelines for all synthetic or AI-modified media.
- **Expand Media Literacy Initiatives:** Invest in comprehensive digital literacy programs that equip citizens to critically evaluate and navigate an evolving synthetic media landscape.
- **Fund Infrastructure for Authenticity Verification and detection:** Prioritize research and development funding toward tools that verify authentic, genuine media, recognizing that standalone detection technologies cannot keep pace with the rapid sophistication of synthetic generation. Implementing detection tools across AI models able to read and identify the labels and provenance information embedded in AI generated or manipulated content.

AI Interaction

- 1. When and why is it most important to know that you are interacting with an AI system? Are there contexts where you don't need to know that you're interacting with AI?**

We trust that it is most critical to know you are interacting with an AI system in contexts where that interaction affects a natural person's autonomy, access to information, or fundamental rights. This includes direct interactions such as customer support, personalized content feeds, political campaigning, or public interest discourse, where synthetic influence could manipulate perception or behavior. The goal is to safeguard trust and ensure people are not misled into believing they are engaging with a human when they are not.

Throughout WITNESS work, we have stated that the disclosures are more critical when attached to visual content made or altered with Artificial intelligence tools such as deepfakes, videos and so on. In these cases, we defend that the disclosures should also carry the "recipe of the content", informing users how and where that piece of information was made or crafted.

- 2. What factors make a disclosure of AI use meaningful and effective for the user?**

WITNESS

Meaningful disclosure requires moving beyond a simple "AI vs. Human" binary, which often fails to capture the "recipe" of how content is created, edited, and distributed. Effective disclosure should be:

- Layered and Contextual: Users should have access to an easily discoverable "more information" icon that provides graduated levels of detail based on their specific needs (e.g., from a casual reader to a regulator).
- Standardized and Neutral: We advocate for a single, neutral, interactive icon (similar to an information symbol) rather than a complex taxonomy of labels, which risks confusing users and driving "label fatigue".
- Persistent: Disclosures must be technically robust, supported by provenance standards (like C2PA) that ensure transparency signals remain attached even as content is shared, remixed, or modified across different platforms.

3. Who in the AI value chain (developers, deployers, or others) should be responsible for providing transparency around AI interactions? Why?

At WITNESS we stand for a pipeline of responsibility across the AI value chain, meaning that responsibility fits across the different types of stakeholders involved such as developers and deployers. In this sense, we highlight the following:

1. Developers must prioritize "authenticity-first" design, embedding provenance and marking techniques (e.g., digital signatures, watermarks) during the training and inference stages.
2. Deployers (such as social media platforms) must ensure this information is surfaced to the end-user in a clear, accessible, and interactive manner, ensuring signals are not stripped or hidden.
3. Governments play the critical role of setting regulatory guardrails to ensure these standards are interoperable, rights-respecting, and adopted industry-wide, preventing fragmented and unequal transparency practices.

4. Are existing market practices and legal frameworks sufficient to support transparency around AI interactions? If not, where do gaps remain and what actions (e.g., regulatory measures, guidance and codes of conduct, standards and technical solutions, research and development, literacy initiatives, procurement requirements) do you think the Government should take?

Existing frameworks are largely fragmented and voluntary, leaving significant gaps in accountability and technical interoperability. Current practices often shift burdens onto end-users and fail to preserve transparency signals across the content lifecycle. The Government of Canada should:

WITNESS

1. **Mandate Interoperable Standards:** Require the adoption of open, interoperable provenance standards (e.g., C2PA) to ensure the "recipe" of content creation is verifiable.
2. **Protect Rights-Based Transparency:** Ensure that transparency measures are privacy-preserving by design, focusing on the "how" (agent/content) rather than the "who" (PII) to prevent surveillance and potential weaponization.
3. **Invest in Digital Literacy and Research:** Prioritize public funding for digital literacy initiatives and research into authenticity verification tools, rather than focusing solely on the "arms race" of detection tools, which struggle to keep pace with rapid synthetic media evolution.
4. **Collaborate on Governance:** Foster independent, multi-stakeholder advisory mechanisms to review the effectiveness of transparency implementations and ensure they do not disproportionately burden civil society or community media

Information about AI Systems

1. **What kinds of information about AI would you like to know in order to help you make informed choices about when and how to use AI? When and for which types of AI products is this information most important to you?**

We advocate for transparency regarding the "recipe" of content creation. Users should have access to clear information about whether AI was involved in the creation, editing, or manipulation of content, and the nature of those changes. This is most critical for multimodal systems—such as audio, video, and image generators—that directly influence public discourse, civic documentation, and human rights evidence. Knowing the provenance—the chain of custody of a digital file—is essential to maintaining trust in shared information.

2. **What kind of information about AI do businesses need to adopt AI safely and responsibly? Do businesses currently have access to the information that they need? How can the Government best balance the need to increase transparency (including input transparency) against the need to protect confidential business information?**

Businesses need standardized, interoperable provenance information to verify the integrity of the content they use and share. Currently, the market is fragmented and voluntary, which impedes consistent adoption. The Government should mandate open, interoperable standards (like C2PA) that businesses can integrate into their workflows. To balance transparency with confidential business information, regulation must distinguish between system-level data (tools, models, operations, sequence), which should be public to enable accountability, and sensitive personal or trade-secret data, which should be protected by default.

WITNESS

3. Who in the AI value chain (developers, deployers, or others) is best placed to provide what information about AI systems, how they are created, and their capabilities and limitations?

We champion a "pipeline of responsibility" where each stakeholder bears proportionate responsibility:

1. Developers are best placed to provide technical information on model capabilities, training data lineage, and inherent limitations (e.g., bias, propensity for hallucinations). They must also embed provenance and marking techniques at the training stage.
2. Deployers (e.g., social media platforms) are responsible for surfacing this information to end-users in a clear, interactive, and consistent manner, ensuring transparency signals are not stripped or hidden.

4. Are existing disclosure practices and legal frameworks sufficient to provide the information that users and other stakeholders need? If not, where do gaps remain and what actions (e.g., regulatory measures, guidance and codes of conduct, standards and technical solutions, research and development, literacy initiatives, procurement requirements) do you think the Government should take?

Current frameworks are insufficient as they are largely voluntary and fragmented. To address this, the Government should:

1. Mandate the adoption of open, interoperable provenance standards (e.g., C2PA) to verify content integrity.
2. Require public disclosure of compliance framework summaries and testing/monitoring results.
3. Move away from binary 'AI vs. Human' labels, which are misleading, and toward neutral, interactive 'more-information' icons that explain the 'recipe' of the content.

Broader considerations

1. Given the areas discussed here, when is AI transparency most important, and are there any areas where action would be premature? Are there other areas that should be considered for action to enhance AI transparency?

Transparency is most critical where AI influences public opinion, civic participation, or human rights evidence. This includes the deployment of synthetic media in news, political campaigning, and public interest discourse. Action is arguably "premature" if it focuses solely on proprietary, static detection-based "arms race" solutions, which are ineffective and easily circumvented. Instead, Canada should prioritize "authenticity-first" infrastructure—such as C2PA provenance standards—which remains durable even as AI technology evolves. Other areas for action include ensuring transparency measures do not negatively impact marginalized communities, who may face "false positives" from moderation tools, and protecting human rights defenders from surveillance masked as "transparency."

WITNESS

2. How should any measures account for the circumstances of small and medium-sized enterprises?

Regulations must avoid imposing monolithic compliance costs that only large tech giants can afford. The Government should:

Promote Open Standards: Encourage the use of open-source, interoperable tools (like C2PA) to lower the barrier to entry for smaller players to adopt best practices. Regulation should also facilitate adoption of these standards by providing pertinent resources.

Provide Technical Guidance: Offer accessible, practical resources and training rather than forcing SMEs to navigate complex, proprietary regulatory landscapes.

Tiered Obligations: Calibrate transparency requirements based on risk profile and scale rather than a "one-size-fits-all" approach, ensuring small organizations aren't paralyzed by administrative burdens.

3. How should Canada ensure coherence between any federal measures and provincial, territorial, and international frameworks? What approaches in other jurisdictions could best inform a Canadian consideration of this issue?

Coherence is best achieved by aligning with internationally recognized, rights-respecting standards (e.g., the Coalition for Content Provenance and Authenticity - C2PA) rather than creating fragmented, jurisdiction-specific technical requirements that would be difficult for global platforms to maintain. Canada should participate actively in international forums to advocate for interoperable provenance standards. Regarding other jurisdictions, Canada can learn from the European Union's AI Act (particularly its risk-based approach) while ensuring that Canadian implementation prioritizes civil society engagement and does not replicate the "label fatigue" or privacy risks observed in other implementations.

4. How can Canada design measures to best account for the ongoing evolution in AI technologies and their use?

To remain future-proof, we suggest that Canada focuses on the following:

Enable transparency across the AI value chain: Embedding provenance and enabling detection tools able to demonstrate to users a content was AI generated or manipulated, and support the work of verification.

Focus on Outcomes: Regulate the impact of AI (e.g., misinformation, likeness misappropriation, systemic rights violations) rather than specific current architectures or technical models, which evolve rapidly.

Create Multi-Stakeholder Governance: Establish standing, independent advisory bodies that include civil society, human rights experts, and academic researchers. These bodies should regularly review the effectiveness of transparency measures, allowing for iterative updates as technology changes.

WITNESS

Prioritize Resilience: Invest in tools that empower users to verify the origin and history of information (provenance) rather than trying to detect the nature of content, which becomes an obsolete strategy as generation tools advance.

Demographic questions

1. Where are you currently resident, or if you are responding on behalf of an organization, institution, or business, where is your organization, institution, or business currently based?

Replying on behalf of WITNESS, global civil society organisation focused on making truth undeniable. We are based in the United States, but our work is globally focused.

2. If you are responding as an interested Canadian or resident of Canada, what is your age group?(Choose the category that applies)

Not applicable

3. If you are responding as an interested Canadian or resident of Canada, do you identify as any of the following?(Select all that apply)

Not applicable