September 23, 2026

The Honourable Mélanie Joly
Minister of Industry
Innovation, Science and Economic Development Canada
C.D. Howe Building
235 Queen Street
Ottawa, ON K1A 0H5
Canada

Re: BSA Comments on AI Transparency

The Business Software Alliance (BSA)[1] welcomes the opportunity to provide feedback on ways that the Government of Canada can support AI transparency.

BSA is the leading advocate for the global software industry. Our members recognize the importance of privacy and security protections because our companies create the business-to-business products that help companies across the economy innovate and grow. BSA members provide tools including cloud storage services, human resource management software, identity management services, and cybersecurity products. As a result, privacy and security protections are a core part of BSA members' operations. Many BSA members have significant operations in both Canada and the US, with investments and employees in both countries.

As you consider policy approaches to AI transparency, we highlight five important priorities. These address both AI policies broadly and AI transparency specifically.

- **Risk-Based Approach. AI policies should not be one-size-fits-all and must be calibrated to the risk created by particular uses of AI.** Because the utility and benefits of AI are so diffuse, AI-related policies can implicate a wide range of issues. However, policies should avoid imposing requirements that govern all uses of AI, as many low-risk uses, such as enabling enhanced workflow collaboration or better budget forecasting, do not raise the same types of concerns that might arise in other scenarios, such as the management and operation of critical infrastructure or eligibility determinations for important life opportunities. As a result, AI policies should be risk-based and obligations should focus on uses of AI that pose greater risks.

[1] BSA's members include: Adobe, Alteryx, Amadeus, Asana, Atlassian, Autodesk, Avalara, Bentley Systems, Box, Cisco, Cohere, Cohesity, Dassault Systemes, Databricks, Datadog, Docusign, Dropbox, Elastic, EY, Graphisoft, HubSpot, IBM, Kyndryl, MathWorks, Microsoft, Notion, Okta, OpenAI, Oracle, PagerDuty, Palo Alto Networks, PTC, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., TrendAI, TriNet, Veeam, Workday, Zendesk, and Zoom Communications Inc.

- **<u>Role-Based Responsibilities</u>**. **AI policies should create role-based responsibilities, to ensure any new obligations fit a company's role.** The AI supply chain is complex and evolving and includes a variety of companies, including AI developers, integrators, and deployers. In practice, developers have insight into model design and training, integrators into system-level modifications, and deployers into real-world use and user impacts. Creating role-based responsibilities for different AI actors is important for at least three related reasons:

    - First: it ensures companies can implement the obligations assigned to them. Different companies across the AI supply chain will have access to different information and be in position to take different actions. Assigning role-based responsibilities can help avoid making companies responsible for information or actions that are outside of their control.

    - Second: it ensures that consumers and business customers are better protected because the companies that are best positioned to identify and address risks are the entities designated to take relevant steps to protect them.

    - Third: it creates a policy framework that is workable in practice.

    For more information about different companies along the AI supply chain, including the roles of developers, integrators, and deployers, we are enclosing with this submission BSA's primer on Effective Accountability Along the AI Value Chain.[2]

- **<u>Relevant Transparency Disclosures</u>. Transparency requirements should ensure that the appropriate actor in the AI supply chain is providing relevant information to the intended audience and that sensitive information is protected**. We recommend that any transparency obligations recognize four guiding principles:

    - First: the responsibility should fit the role of the company in the AI value chain.

    - Second: the information an AI actor shares should be relevant to the intended audience.

    - Third: trade secrets and confidential proprietary information should be protected.

    - Fourth: companies should maintain flexibility to determine the appropriate mechanisms for communicating relevant information.

    These principles recognize the different goals that can be served by AI transparency measures. In some instances, regulators may want access to non-public information to assess a company's legal compliance. In other instances, policymakers may want to encourage companies to share information with other companies or with the public more broadly. Policies on AI transparency should recognize that different information will be

[2] *See* BSA, Effective Accountability Along the AI Value Chain: Right-Sizing AI Responsibilities for the Role (June 9, 2026), *available at* https://www.bsa.org/policy-filings/effective-accountability-ai-value-chain-right-sizing-ai-responsibilities-role.

relevant to these different audiences — and avoid assuming that the same information should be shared with regulators as with other companies along the AI supply chain or the public more broadly. Indeed, while each of these measures are based in transparency, they have different objectives. Policies on AI transparency should be carefully considered to ensure they meaningfully address transparency goals while protecting trade secrets and other proprietary information. For more information about appropriate transparency along the AI value chain, we are including with this submission BSA's Best Practices for Information Sharing Along the General Purpose AI Value Chain.[3]

- **Content Provenance and Authentication**. **Policies should support the use of content authentication and provenance information to increase the trustworthiness of AI-generated content.** BSA understands that with the advent of generative AI, innovative AI-powered content creation can also raise questions about whether an image, audio, or video file is authentic, which could undermine trustworthiness in the marketplace. As a result, BSA supports obligations for generative AI application developers to leverage open, technical standards, such as the Coalition for Content Provenance and Authenticity (C2PA) standard, to enable machine-readable, non-visible disclosures for AI-generated images, audio, and video. Using such tools helps individuals understand the origin and history of AI-generated content, increasing trust and reliability in the AI ecosystem. However, no policy should require visible labels on AI-generated content, since these can be easily removed, making them an ineffective and less durable method of labeling AI-generated content than machine-readable tools. For more information about BSA's recommendations on content authenticity, we are enclosing with this submission BSA's Primer on Content Authenticity.[4]

- **Harmonization**. Because AI systems are developed and deployed globally, AI regulations must work across jurisdictions. As policymakers consider new AI regulatory frameworks, they should align any new policies with established international frameworks and standards rather than creating divergent approaches. For example, policymakers should promote a common definition of AI to ensure that stakeholders have a shared understanding of the technology. BSA supports the OECD's definition of AI, which has been referenced by regulators and legislators worldwide. Harmonizing AI policies with global frameworks will reduce fragmentation and provide greater clarity for organizations operating across multiple jurisdictions.

* * *

[3] *See* BSA, Best Practices for Information Sharing Along the General Purpose AI Value Chain (Sept. 3, 2024), *available at* https://www.bsa.org/policy-filings/best-practices-for-information-sharing-along-the-general-purpose-ai-value-chain.
[4] *See* BSA, Primer on Content Authenticity (March 3, 2026), *available at* https://www.bsa.org/policy-filings/primer-on-content-authenticity.

Thank you again for the opportunity to provide our views. We are attaching several resources that BSA has created on these issues, including:

- Effective Accountability Along the AI Value Chain: Right-Sizing AI Responsibilities for the Role
- Best Practices for Information Sharing Along the General Purpose AI Value Chain
- Primer on Content Authenticity

We welcome an opportunity to further discuss these issues.

Sincerely,

Kate Goodloe
Managing Director, Policy
Business Software Alliance

Effective Accountability Along the AI Value Chain: Right-Sizing AI Responsibilities for the Role

Artificial intelligence (AI) fuels digital transformation in every sector of the economy. Manufacturers use AI to design safe and sustainable products; small businesses rely on AI-based translation tools to reach global customers; health researchers use AI to improve patient care and drive new medical breakthroughs; and companies in all industries can use AI systems to improve the accessibility of their products for people with disabilities. In these areas and countless more, AI creates new opportunities to solve complex challenges.

Successful AI adoption requires trust and confidence in these innovative technologies. As policymakers create frameworks for increasing trustworthiness and accountability, they should recognize a key principle: responsibilities should fit the role of the company.

Creating role-based responsibilities for different AI actors is important for several distinct but related reasons. First, it assigns responsibilities to companies that they can implement based on their access to relevant information and position in the AI supply chain, preventing unworkable obligations for things outside of their control. Second, it ensures that consumers and business customers are better protected because the companies that are best positioned to identify and address risks are the entities designated to take relevant steps to protect them. Third, it creates a policy framework that is workable in practice.

Key Actors in the AI Ecosystem

The AI supply chain is complex and evolving and includes a variety of companies, including AI developers, integrators, and deployers. In many scenarios, a developer may create an AI system, and a deployer will use that system in the real world.

DEVELOPERS design, code, or produce AI systems

EXAMPLE

A company that designs, codes, or produces an AI system, such as a software company that develops an AI system for speech recognition, is a developer.

DEPLOYERS use AI systems

EXAMPLE

A company that uses an AI system, such as a bank that uses an AI system either developed internally or by a third party to make loan determinations, is a deployer.

An organization can act as both a DEVELOPER and a DEPLOYER

EXAMPLE

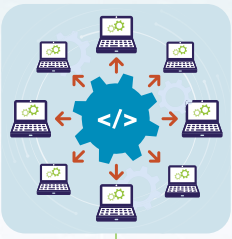
A cybersecurity company develops AI software that monitors network traffic and customer transactions and then uses it on its own systems. The company is both a developer and a deployer.

In other contexts, one company may develop an AI model, a second company may integrate that AI model into an application or platform (AI system), and a third company may use that AI system to perform a business function, such as summarizing documents on a content management platform that is generating insights into a company's financial transaction data.

DEVELOPERS

A model developer can create a generative AI model that summarizes text or generates images.

INTEGRATORS



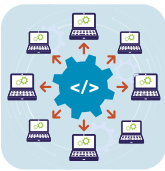
An integrator may incorporate that generative AI model into its own software application, enabling features such as transcribing meeting notes and summaries on a video communications platform or analyzing business contracts using document management software.

DEPLOYERS

A deployer may use the integrator's software application to assist with common business tasks, such as leveraging it to query key business documents to find answers to specific questions or generate summaries.

Developers

Developers include companies that create AI systems, including the underlying models, and sell or license the entire AI system to customers. Examples include human resource management or workflow collaboration software for businesses. In the general-purpose AI context, model developers create AI models, which can be used in a wide variety of different applications. For example, a company may develop a foundation model that can be adapted for many different use cases. The same foundation model can be used to power a range of AI systems including search engines, chatbots, spam detection software, tools that summarize long text documents, and a variety of other applications. Developers will have information about how their AI models or systems have been developed, but they generally lack information about how other companies deploy their AI tools.



Integrators

Other companies—"integrators"—may integrate an AI model into a particular application, for use by other companies. Some integrators may simply connect the AI model to a specific AI system or application, while other integrators may fine-tune or modify the AI model before including it in an AI system or application. For example, a company may act as an integrator when it develops an AI application that incorporates one or more third-party AI models. Integrators will generally have information about any changes they made to the AI model, but they usually do not have direct insight into the model's initial development or the later use of the AI system by deployers.

Deployers

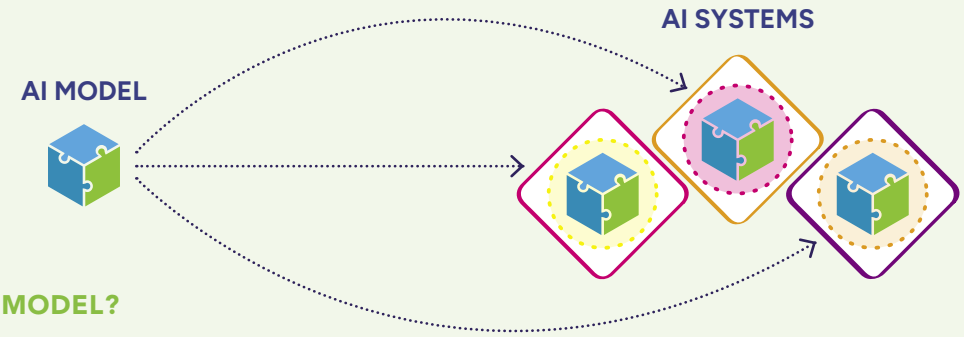
Companies that use an AI tool for a specific purpose are often called deployers. These companies decide when and how to use a particular AI technology, so they will have insight into the facts of a particular use case. But deployers often obtain AI tools from other companies, so they typically lack direct insight into the initial training of the AI tool. Deployers are often the companies that have direct relationships with consumers.

Effective policy and accountability frameworks must recognize these different types of companies to avoid one-size-fits-all requirements and ensure responsibilities fit the role. This is important because the various companies along the AI supply chain will have access to different information and be in different positions to take the necessary actions to identify and address relevant risks to protect consumers and business customers.

For example, a developer of a foundation model would be able to describe issues related to the design and training of its model, but it generally would not have insight into how an integrator customizes the model for a specific application or how a deployer subsequently uses that AI system in the real world. Instead, the deployer using the AI system is generally best positioned to understand how the AI system is being used, whether that use aligns with its intended use, how to incorporate human oversight in its operation, review the outputs generated from the AI system, receive any complaints from end users, and assess real-world factors that may affect the system's performance.

Like other contexts, any legislation or policy framework that creates obligations for companies that design and use AI systems should reflect these different roles and assign responsibilities accordingly.

WHAT IS THE DIFFERENCE BETWEEN AN AI MODEL AND AN AI SYSTEM?



WHAT IS AN AI MODEL?

An **AI model** is a piece of software trained on large amounts of data to, for example, recognize patterns, make predictions, or generate content.

Think of it like:

- A **trained expert** who has learned from many examples, or who continues to learn from experience, or
- A **brain** that can process language, images, or numbers.

Examples of what models can do:

- Predict the next word in a sentence
- Recognize objects in an image
- Generate text, images, or code

On their own, models are usually **not user-friendly.** They don't have user interfaces, buttons, screens, or workflows.

WHAT IS AN AI SYSTEM?

An **AI system** is a complete tool or product that uses one or more AI models to perform a task or solve a problem.

Think of it like:

- A **car** that includes an engine (the model), plus steering wheel, seats, and controls.

AI applications may include a chat interface, buttons, menus, or voice input, or rules, safeguards, and integrations (e.g., with email, calendars, databases).

Examples:

- A chatbot that answers customer questions
- A photo app that enhances images
- A writing assistant that helps draft emails

PRECEDENT FOR ROLE-BASED POLICY APPROACHES

Establishing role-based obligations is considered best practice in privacy and security legislation around the world. For example, privacy laws in the United States and around the world distinguish between the differing roles of controllers that determine how and why data is processed, and processors that handle data on behalf of a controller and according to its instructions. Similarly, in various jurisdictions, cybersecurity legislation generally differentiates between companies and their service providers.

Just as privacy and security laws distinguish between different types of companies that handle consumers' personal data, distinguishing among the different companies along the AI supply chain ensures that legal frameworks appropriately assign obligations to a company based on its role in the AI ecosystem, enabling them to meaningfully fulfill those obligations and better protect consumers.

Best Practices for Information Sharing Along the General Purpose AI Value Chain

BSA members are advancing trust and ethics in artificial intelligence (AI) and we support policies that ensure AI systems are developed and used responsibly. To be effective, AI policies and corporate practices must reflect that different entities have different roles in the AI ecosystem, and therefore will have different obligations based on those roles. For example, an AI developer, an AI deployer, and other parties within the value chain will have different information about characteristics, capabilities, and limitations of an AI system and how an AI system was developed or operates. As a result, these different companies will have different abilities to share information with other actors.

BSA's best practices for information sharing along the general purpose AI (GPAI) value chain are intended to support transparency between these different actors.[1]

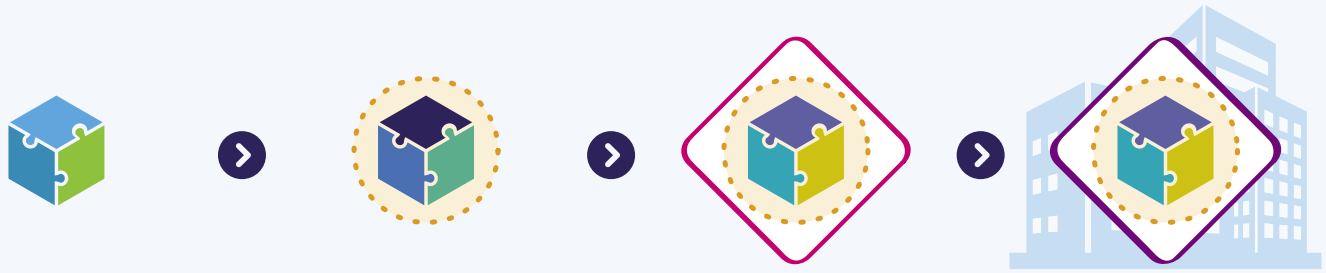
BSA'S BEST PRACTICES:

1. **Identify the entities that should share information.**
2. **Identify the audience for such information.**
3. **Identify methods for sharing that information.**
4. **Set out the types of information to be shared by different actors along the value chain.**

Because different companies have different roles in creating and providing AI systems, the type of information to be shared will vary based on a company's role and the availability of information shared in particular value chains.

[1] General purpose AI (GPAI) models can be used for a wide variety of tasks and may be integrated into a variety of downstream AI systems. As a result, there are a range of different entities along the GPAI value chain. For this reason, BSA's best practices focus on the GPAI value chain and identify the information that different entities should provide to others.

SHARING INFORMATION ALONG THE GENERAL PURPOSE AI VALUE CHAIN



Developer of GPAI Model

Developers of GPAI models should share information about the model with downstream providers.

Downstream Providers

Downstream providers may modify a GPAI model as they integrate it into an AI system. They should share information about the underlying model and their changes to that model with other actors farther downstream the AI value chain.

Deployment

Companies often deploy AI systems that use GPAI models, and have been modified by downstream providers to fit their use case.

1 ENTITIES SHARING INFORMATION SHOULD INCLUDE[2]

- Providers that develop a GPAI model for use by other businesses
- Downstream providers that integrate GPAI models developed by others into an AI system for use by their business customers
- Companies that make available GPAI models to their business customers, without integrating the GPAI model into an AI system[3]

2 AUDIENCE FOR INFORMATION

- Relevant actors situated farther downstream in the GPAI value chain

3 POTENTIAL METHODS FOR SHARING INFORMATION

- Model or system cards
- Public disclosures with standard template
- Standardized technical documentation
- Incorporating links to publicly available information, when available

4 INFORMATION TO BE SHARED

As a general matter, information sharing obligations will depend on what is reasonable and appropriate based on the risks and capabilities of a GPAI model or AI system, the different roles of different actors in the AI value chain, and the working relationship between the various actors, including business customers using AI-related services.

---

[2] A company may be involved in one or more of these activities.

[3] These companies would only provide information about the underlying GPAI model and describe the lack of changes to the model, as the information about a standalone AI system would not apply.

BEST PRACTICES FOR SHARING INFORMATION ALONG THE AI VALUE CHAIN

PROVIDERS OF A GPAI MODEL should share the following information with downstream providers:

- GPAI model name
- Developer name
- Date of release
- Date(s) of revision, if any
- Modality/format (e.g., text, video, image)
- Intended use
- Known limitations
- Model characteristics
- Related software and hardware
- Information reasonably necessary to integrate the GPAI model
- An overview of the training data, including the type and provenance of data and curation methodologies
- Information on data used for testing or validation of the GPAI model, if applicable
- Acceptable use policy, including any prohibited uses
- Contact information to report concerns/issues with the GPAI model

DOWNSTREAM PROVIDERS that integrate a GPAI model into an AI system for use by business customers should share some or all of the following information with relevant actors situated further downstream in the AI value chain:

1. **Information about the underlying GPAI model(s), including:**
   - GPAI model name(s)
   - GPAI developer name(s)
   - Available documentation from GPAI model provider (as set out above)
2. **Information about relevant changes made to the GPAI model by the downstream provider, namely:**
   - Statement summarizing relevant changes by the downstream provider. These may include, as applicable:
     - Changes to the intended use of the GPAI model
     - Changes that alter the model weights (e.g., retraining model, fine tuning model, response learning from human feedback (RLHF)) and a summary of data used to retrain or finetune the model, if applicable
     - Changes that customize the GPAI model without altering model weights (i.e., retrieval augmented generation (RAG), prompt optimization, use of supporting models)
     - Changes to privacy or security controls of GPAI model
     - Information about changes to performance or known limitations if the downstream provider has put the model to use after the relevant changes
     - Estimated cadence of changes
   - Statement of no relevant changes by the downstream provider.[4] These may include:
     - Statement that the provider did not make changes that alter the model weights (e.g., translating the model into a different language without retraining it)
     - Statement that the provider did not customize the GPAI model

[4] These statements may be particularly relevant for companies that provide business customers the ability to access a GPAI model but do not change the model itself.

3. Information about the AI system into which the GPAI model is integrated:

- AI system name
- Downstream provider name
- Date of release
- Date of revision(s), if any, or cadence of system updates
- Modality/format (e.g., text, video, image)
- Intended use
- Known limitations
- AI system characteristics
- Related software and hardware
- Information reasonably necessary to integrate the AI system
- Information regarding testing or evaluation of the AI system, if applicable
- Acceptable use policy, including any prohibited uses
- Contact information to report concerns/issues with the AI system

The level of detail for this information may appropriately vary based on the risks presented by foreseeable uses of the AI system.

In sharing this information, companies should respect confidentiality obligations and share information in line with protections for trade secrets, intellectual property, or other business confidential information.

ADDITIONAL INFORMATION THAT MAY BE APPROPRIATE IN SPECIFIC SCENARIOS.

Further information may be shared by either the GPAI model provider or the downstream provider, as appropriate and applicable, including:

- Instructions for use for the GPAI model or AI system, as applicable
- Information on data sets, including any datasheets for data sets
- Employee training guidance
- Certifications, if applicable
- Related research

HOW SHOULD INFORMATION BE SHARED?

Information may be shared between different entities in the GPAI value chain in different ways, depending on the AI model and AI system at issue. For example:

- Information may be made publicly available, such as publishing information on the date a model is released or significantly revised. In these scenarios, publishing the information in a manner that downstream providers may link to or otherwise pass to other actors is helpful.
- Information that is not public may also be shared between different actors in the course of a business relationship.
- Providers may also develop standardized ways of notifying others in the value chain when they update such information.

Annex 1
TEMPLATE FOR USE BY: GPAI MODEL PROVIDER

The template below can be completed by the provider of a GPAI model to share information along the AI value chain, consistent with BSA's best practices. This template may also be further developed by companies to document their information-sharing practices.

Information About GPAI Model	
GPAI model name	
Developer name	
Date of release	
Date(s) of revision, if any	
Modality/format (e.g., text, video, image)	
Intended use	
Known limitations	
Model characteristics	
Related software and hardware	
Information reasonably necessary to integrate the GPAI model	
Overview of training data, including the type and provenance of data and curation methodologies	
Information on data used for testing or validation of the GPAI model, if applicable	
Acceptable use policy, including any prohibited uses	
Contact information to report concerns/ issues with the GPAI model	
Additional Information—May Be Appropriate in Specific Scenarios	
Instructions for use for GPAI model	
Information on data sets, including any datasheets for data sets relevant to GPAI model	
Employee training guidance	
Certifications, if applicable	
Related research	

Annex 2
TEMPLATE FOR USE BY: **DOWNSTREAM PROVIDERS**

The template below can be completed by a downstream provider integrating a GPAI model into an AI system to share information along the AI value chain, consistent with BSA's best practices. This template may also be further developed by companies to document their information-sharing practices.

Information About Underlying GPAI Model	
Name of GPAI model(s) integrated into AI System	
Developer(s) of GPAI model(s) integrated into AI system	
Further information on GPAI model, as provided by the GPAI model provider	
Information About Relevant Changes to the GPAI Model by Downstream Provider	
Statement of relevant changes by the downstream provider to the GPAI model. These may include: » Changes to the intended use of the GPAI model; » Changes that alter the model weights (e.g., retraining model, fine tuning model, response learning from human feedback (RLHF)) and a summary of any data used to retrain or finetune the model, if applicable; » Changes that customize the GPAI model without altering model weights (e.g., retrieval augmented generation (RAG), prompt optimization); » Changes to privacy or security controls of GPI model; » Information about changes to performance or known limitations if the downstream provider has put the model to use after the relevant changes; and » Estimated cadence of changes.	
Statement of no relevant changes by the downstream provider to the GPAI model. These may include: » Statement that the provider did not make changes that alter the model weights (e.g., translating model into other languages without retraining it), or » Statement that the provider did not customize the GPAI model	

Information About AI System Into Which GPAI Is Integrated	
AI system name	
Downstream provider name	
Date of release	
Date of revision(s), if any, or cadence of system updates	
Modality/format (e.g., text, video, image)	
Intended use	
Known limitations	
AI system characteristics	
Related software and hardware	
Information reasonably necessary to integrate the AI system	
Information regarding testing or evaluation of the AI system, if applicable	
Acceptable use policy, including any prohibited uses	
Contact information to report concerns/issues with the AI system	
Additional Information—May Be Appropriate in Specific Scenarios	
Instructions for use for AI system	
Information on data sets, including any datasheets for data sets relevant to AI system	
Employee training guidance	
Certifications, if applicable	
Related research	

Annex 3
COMPARISON: BSA BEST PRACTICES AND EU AI ACT OBLIGATIONS FOR GPAI PROVIDERS

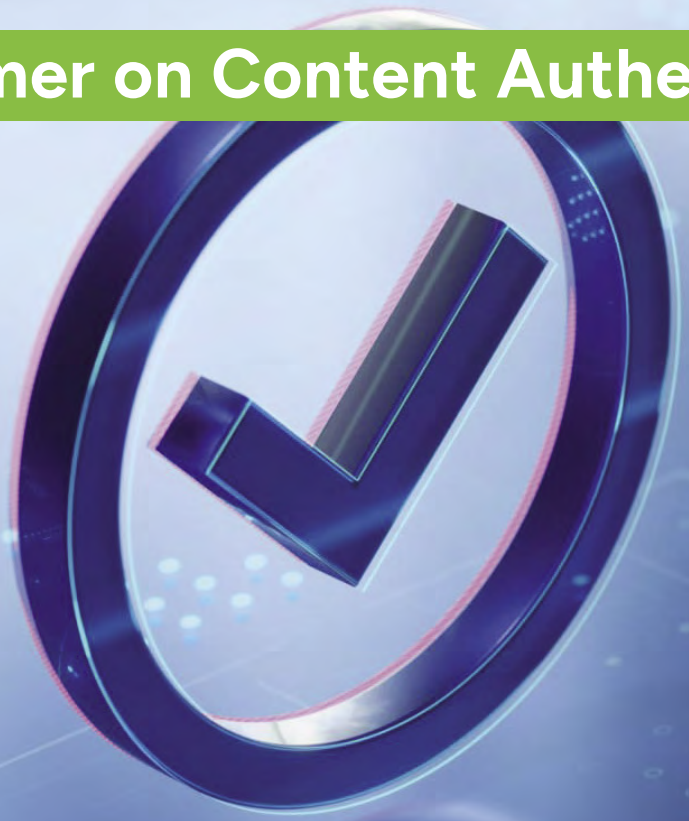
This table compares BSA's recommended best practices for GPAI model providers with the obligations for such providers to share information under the EU AI Act. Article 53(1)(b) of the AI Act requires providers of GPAI models to give downstream providers that integrate the model into their AI system certain documentation, set out in Annex XII.[5]

BSA BEST PRACTICES Information to Be Shared by GPAI provider	EU AI ACT Annex XII Obligations for GPAI Providers to Share Information With Downstream Providers
GPAI model name	
Developer name	
Date of release	1.c: The dates of release and methods of distribution
Date(s) of revision, if any	1.c: The dates of release and methods of distribution
Modality/format (e.g., text, video, image)	1.g: The modality (e.g., text, image) and formats of inputs and outputs
Intended use	1.a: The tasks that the model is intended to perform and the type and nature of systems into which it can be integrated
Known limitations	1.h: The license for the model
Model characteristics	1.f: The architecture and number of parameters 2. A description of the elements of the model and the process for its development, including: (b) the modality (e.g., text, image, etc.) and format of the inputs and outputs and their maximum size (e.g., context window length, etc.)
Related software and hardware	1.d: How the model interacts, or can be used to interact, with hardware or software that is not part of the model itself, where applicable 1.e: The versions of relevant software related to the use of the GPAI model, where applicable
Information reasonably necessary to integrate the GPAI model	2. A description of the elements of the model and the process for its development, including: (a) technical means (e.g., instructions for use, infrastructure, tools) required for the GPAI model to be integrated into AI systems
Overview of training data, including the type and provenance of data and curation methodologies	2. A description of the elements of the model and the process for its development, including: (c) information on the data used for training, testing and validation, where applicable, including the type and provenance of data and curation methodologies
Information on data used for testing or validation of the GPAI model, if applicable	2. A description of the elements of the model and the process for its development, including: (c) information on the data used for training, testing and validation, where applicable, including the type and provenance of data and curation methodologies
Acceptable use policy, including any prohibited uses	1.b: The acceptable use policies applicable
Contact information to report concerns/ issues with the GPAI model	

[5] Outside of the EU, California is considering draft regulations that would apply more broadly than GPAI. The draft regulations would require companies that make or train automated decision-systems to provide information to other downstream actors, including a "plain language explanation of any requirements or limitations that the business identified as relevant to the permitted use." California's draft regulations are expected to undergo public comment in late 2024.

Business Software Alliance

Primer on Content Authenticity



Transparency about AI-generated content is an important part of promoting responsible AI.

The Business Software Alliance supports the development and deployment of reliable content authentication and provenance mechanisms that can help users identify the history and origin of AI-generated content. This can help consumers know when content is human- or AI-generated and can help to address misinformation and disinformation.

www.bsa.org

Across the globe, policymakers are confronting important questions about how consumers can tell if the pictures and videos they see online are real. To do that, consumers need to know about tools that already exist today, like watermarks, digital fingerprints, and secure metadata. These tools can show who made a picture or video—and whether digital tools like AI have modified it. Policymakers should support the use of these and other tools to help consumers understand if content was made by humans or by AI.

BSA's primer on content authenticity:

- Defines content authenticity and content provenance, which are foundational to AI transparency policies.
- Describes existing tools that can identify who created an image or video and whether that image or video has been altered. These include machine-readable watermarks, digital fingerprints, and secure metadata.
- Describes the different actions that different types of companies can take to help support the use of content authenticity and provenance tools.
- Explores the policy issues raised by deepfakes.
- Explains how different AI policies can help consumers tell fact from fiction online.

What Are Content Provenance and Content Authentication?

Content Provenance

Content provenance information describes where content came from and how it changed. It tracks the origin, ownership, and modification history of a digital file, like an image, video, or audio clip. Example: A photograph includes cryptographically signed metadata that shows the date and location in which a photo was taken, the type of camera used to take the photo, and any software editing programs used to edit the file.

Content Authentication

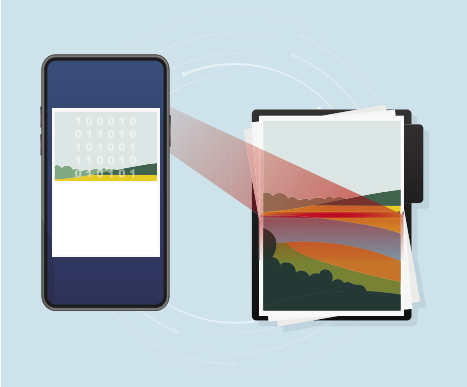
Content authentication helps indicate whether the content and its provenance information are trustworthy and whether the metadata, watermarks, or credentials for a piece of content have been tampered with. Example: a content authentication tool can check a video's embedded signature to confirm it matches the issuer's public key and that the file hasn't been modified since publication.

What Tools Can Help Identify AI-Created Content?

A range of technologies can identify how content was created or modified, including content made through AI tools. Policymakers should not require use of a single technique but should ensure companies can use a range of technologies to mark content as AI-generated or not.

Machine-Readable Watermarks

Machine-readable watermarks are added directly into a file, by embedding a small amount of information that can be detected by a specialized tool. A watermark can be embedded in a file's pixels or data stream, so the origin of the file can be verified even if it is later stripped of metadata. Example: an invisible watermark inside an image can tie that image back to a creator or authenticity platform.

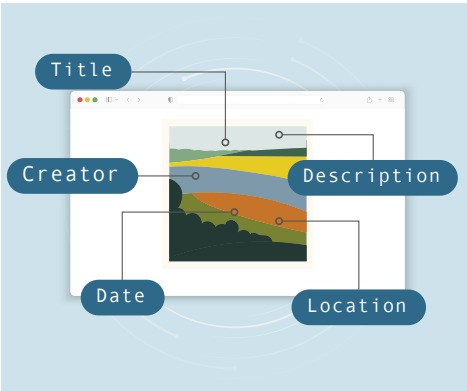


Digital Fingerprints

Digital fingerprints generate a unique identifier (or "fingerprint") from the content's core features, such as its color patterns, structure, or encoding. A digital fingerprint does not modify the file. Instead, it is a unique mathematical signature, such as a hash, that can be stored separately from the file. A file's digital fingerprint can be checked against a database of digital fingerprints, to confirm if a file matches an original or has been altered. Example: a video can be broken into frames to create a fingerprint from visual and audio cues; that signature can be checked against a database of digital signatures to confirm if the video was modified or altered.

Secure Metadata

Secure metadata can be used to store information about who created a file, what tools were used to create it, and how the file was edited. That information establishes the file's content provenance. Example: an image can include metadata that records the creator's identity, editing history, and software used; this can be digitally signed and embedded in the file so that viewers can verify its authenticity and origin.

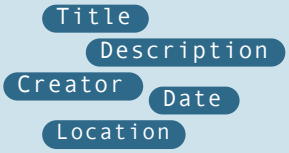


A Robust Approach: Combining Multiple Tools

These technologies can work together to support a robust approach to content authenticity, giving individuals the tools to know how content was created and whether it has been altered. All three processes act as layers of verification.

WATERMARKING
embeds authenticity

FINGERPRINTING
identifies content



METADATA
can document context and authorship

While implementing all three processes may be ideal in many scenarios, it is not always necessary to use all three techniques in every instance. Rather, policies should encourage the use of technical methods that are mature enough for adoption today while remaining flexible enough to accommodate a range of technical and operational uses.

A Bad Fit for AI: Visible Watermarks

Visible watermarks are a low-tech method used for decades to label content, including to clearly identify documents as 'Draft' or 'Confidential.' In the AI context, requiring visible watermarks is not practical, since they can be easily removed. That undermines their effectiveness and creates a false sense of security.

IN FOCUS

AI GENERATED TEXT

There are a range of tools to label images and audio-visual content, to help consumers know if that content is authentic or AI-generated. In contrast, any requirements to label AI-generated text raise a host of practical concerns. To ensure consumers know about text-based AI generated content, we recommend focusing on requirements that help individuals know when they interact with an AI system.

IN FOCUS

COALITION FOR CONTENT PROVENANCE AND AUTHENTICITY (C2PA) STANDARD

The C2PA standard can be used by anyone to incorporate digital provenance information into products and processes. It combines several technologies and is expected to be approved by the International Standards Organization as a global standard.

HOW DOES IT WORK?

1. **CREATING A FILE.** When a person creates a new file, she can use a tool that generates information about the file's provenance—including how it was created and any edits.
2. **CONTENT CREDENTIALS.** The provenance information is encoded into a content credential.
3. **CRYPTOGRAPHIC SIGNING.** The content credential is signed with the private key of the software or hardware used to create it, ensuring its authenticity. A public key is made available for authentication.
4. **EMBEDDING AND/OR WATERMARKING.** The content credential can be stored in several places, including embedding it within the file or storing it in a separate database where it is associated with an invisible watermark or digital signature.
5. **VERIFICATION.** To check the authenticity of a file, an individual can use a tool that checks content credentials to confirm if the file is authentic or has been modified.
6. **PRESENTATION.** Content that is verified as authentic can be marked with a clear indicator, like a badge or icon.

What Are Content Credentials?

The latest C2PA standard relies on content credentials, which combine multiple techniques.

- Content credentials contain metadata about a file, including the date and time it was created and the technology used to create it.
- Watermark identifiers and digital signatures can be added into the content credential, which already includes the metadata.
- The package of credentials is digitally signed and uniquely connected to that file.
- The content credentials are both embedded into the content and stored separately, such as in a database of content credentials.

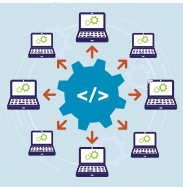
The AI Supply Chain: Different Roles and Responsibilities

The AI supply chain is evolving and includes a variety of companies. Lawmakers must recognize these different types of companies as they create policies on content authenticity—since different types of companies will have access to different information and be in a position to take different actions to protect consumers.

For example, one company may develop an AI model, a second company may integrate that AI model into an application, and yet another company may use that application to create new content.

Model Developers

These companies create AI models, which can be used in a wide variety of different applications. For example, a company may build a foundation model that can be adapted for many different use cases. The same foundation model can be used to power search engines, chatbots, spam detection software, tools that summarize long text documents, and a variety of other applications. The model developer will have information about how that foundation model is developed—but generally lacks information about how other companies deploy the model for specific uses.



Integrators

These companies may integrate an AI model into a particular application, for use by other companies. Some integrators may simply connect the AI model to a specific service or application, while other integrators may fine tune or modify the AI model before including it in a service or application. These companies will generally have information about any changes they made to the AI model, but they usually do not have direct insight into the model's initial development or the particular circumstances in which other companies use the resulting AI application. For example, a company may act as an integrator when it develops an AI application that incorporates one or more AI models.

Deployers

Companies that use an AI tool for a specific purpose are often called deployers. These companies decide when and how to use a particular AI technology—so they will have insight into the facts of a particular use case. But deployers often obtain AI tools from other companies, so they typically lack direct insight into the initial training of the AI tool.

Any policies on content authenticity need to reflect these different roles.

If policies apply one-size-fits-all requirements to these very different types of companies, it will not effectively help consumers. For example, developers of AI applications that generate content are in the best position to apply content provenance information to content generated by that AI application. In contrast, the developer of a foundation model would lack the technical ability to apply watermarks or other digital markings to content that is generated through an AI application developed and used by other entities.

Ensuring laws remain fit over time. What constitutes state of the art in ensuring solutions for content provenance will evolve over time; policymakers should ensure that any legislative framework accommodates such developments. Embracing open standards like C2PA can help in this effort.

Combatting Deepfakes

Tools like invisible watermarks, digital fingerprints, and secure metadata can help to identify authentic content—so that individuals can know when the images, videos, and audio clips they interact with are authentic.

Together, this helps combat the problem of deepfakes. While no bad actor is likely to label their deepfakes as fake, it is important to give individuals tools to understand when content is authentic. For this to work, these tools need to be adopted across devices and platforms. People also need to be educated about these tools, so they know to watch for information about the provenance of an image or a video, while maintaining a healthy skepticism about information they encounter online.

Bad actors will continue to find novel ways to exploit technologies, including AI, for deceptive purposes. But tools like secure metadata and the C2PA standard can be crucial in helping good actors prove the authenticity of their content—helping all of us tell fact from fiction online.

As policymakers consider the most effective way to address deepfakes, they should also consider the different roles and functions of different types of companies. This is crucial since different types of companies will be positioned to address different issues, due to key service-level, technical, functional, and user-based distinctions.

Policymakers should also recognize the different risks involved in distinct types of services. For example, business-to-business software services pose limited risk to user safety and public order, given the size of their user base and the fact that they do not provide services directly to consumers, which may create fewer deepfake-related concerns.

How AI Policies Can Help Consumers

When policymakers want to ensure consumers can tell fact from fiction online, it is important to rely on existing tools that can identify if content was created by AI or not.

POLICY GOAL	SOLUTION
 **Tell consumers when they interact with AI systems.**	When companies provide AI systems that interact directly with consumers they should tell consumers they are interacting with an AI system, unless it is obvious.
 **Help consumers know if content is AI-generated.**	Focus obligations to use content provenance and authentication tools on consumer-facing audio, visual, or video content. These requirements should not apply to text or in business-to-business settings. This is important because using content authenticity mechanisms doesn't make sense for text—like when a consumer uses AI tools to edit the text of an email or translate a document. In business-to-business settings, companies can create separate methods to address content authenticity, based on their specific uses.
 **Support the use of leading global technologies to recognize AI-created content.**	Ensure requirements to identify AI-created content can be satisfied by leading global tools, including open standards like the Coalition for Content Provenance and Authenticity (C2PA). A range of countries are considering regulations on content provenance; adopting country-specific standards can limit consumers' ability to use widely accepted standards that are updated over time.
 **Facilitate transparency about AI-generated content.**	Obligations to apply content provenance information to AI-generated content are usually best placed on developers of AI applications—since AI applications are used to generate content. Companies can use tools like machine-readable watermarks, digital fingerprints, and secure metadata to identify the origin of that content.
 **Make sure content provenance information remains available to users.**	Prohibit stripping content provenance information such as machine-readable watermarks and secure metadata, absent security concerns.