



Business Software Alliance Comments on the National Institute of Standards and Technology's Extended Outline for a Proposed Zero Draft for a Standard on Documentation of AI Datasets and AI Models

October 17, 2025

The Business Software Alliance (BSA) appreciates the opportunity to provide input on the National Institute of Standards and Technology's (NIST) Extended Outline for a Proposed Zero Draft for a Standard on Documentation of AI Datasets and AI Models (Documentation Outline).

BSA is the global trade association of the enterprise software industry, representing companies that are leaders in AI, cybersecurity, cloud computing, and other cutting-edge technologies.¹ We work in over 20 markets in the United States, Europe, and Asia, advocating for policies that build trust in technology so that every industry sector and the public can benefit from innovation. As a result, BSA members have unique insights into AI's tremendous potential to spur digital transformation and the policies that can best support the responsible use and adoption of AI.

NIST's "Zero Drafts" pilot project on pre-standardization research is an important effort in achieving the shared objective of spurring AI adoption because accelerating international standards development plays a critical role in unleashing AI innovation. In particular, interoperable approaches to transparency can enhance responsible use and promote accountability, building trust in the technologies that enables increased adoption. We appreciate NIST's focus on this important issue but are concerned that the current draft may not achieve these goals. We provide recommendations below on how to enhance the Documentation Outline to enable the proposed standard to work most effectively in practice.

Specifically, we recommend that NIST:

- Adopt a tiered approach to transparency documentation;
- Revise the outline to address the need to protect sensitive proprietary information; and
- Ensure the "Zero Draft" proposed documentation standard complements existing transparency efforts and provides an interoperable solution.

I. Adopt a Tiered Approach to Transparency Documentation

We are concerned the Documentation Outline fails to recognize that different types of documentation are appropriate for different audiences.

Transparency is a bedrock principle applied across many different domains, whether the disclosures relate to the scope of a warranty, the potential side effects of medication, or the use of consumers' personal information. Stakeholders have sought to leverage this principle in the AI context to accomplish several

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

different accountability goals: (1) encourage organizations to implement responsible governance practices that enable them to identify relevant issues, consult relevant personnel, and address problems that may arise; (2) enable regulators conducting investigations to assess whether organizations have employed reasonable practices in AI development and use; (3) provide relevant information to other organizations in the AI ecosystem that are integrating AI into their products and services; and (4) enhance public understanding of how AI systems are responsibly developed and used to build trust in the marketplace.

Although each of these goals enhances accountability, relevant information must be shared with the appropriate audience to accomplish each objective. First, regulators conducting law enforcement investigations may have broader access to information to enable a fulsome investigation. Second, other AI ecosystem actors may need information that may not be as extensive as a regulatory inquiry but nevertheless contains sufficient technical information to enable them to integrate, configure, or adapt the technology or assess its appropriateness for certain uses. Third, the public, which will largely lack the technical expertise to understand granular technical details or the ability to preserve the confidentiality of the information, should receive high-level overviews in plain language that broadly enhances their understanding and helps to increase trust.

NIST's Documentation Outline acknowledges that documentation artifacts should be tailored to the needs and uses of interested parties, and that "[p]ublic-facing documentation will typically be less comprehensive than documentation for limited external audiences (e.g., customers, partners, and regulators), ... particularly with respect to sensitive information, such as trade secrets or personal information." Despite explicitly recognizing that different audiences exist in the transparency landscape, and that the information disclosed should be calibrated for the relevant audience, NIST's Documentation Outline ignores this reality in its implementation. NIST designated the information included in the outline for public disclosure, but it does not correspondingly tailor the information to the relevant audience, essentially determining, in practice, that all information is relevant to everyone. Additionally, beyond protecting sensitive information, like trade secrets and confidential information, NIST should consider information that may undermine the security of the dataset or model, like specific details on model weights or training data, to ensure documentation doesn't inadvertently provide a roadmap to adversarial actors to break through those defenses.

Such an approach not only contradicts NIST's own description of its objectives, but it also undermines the ability to provide meaningful information in a manner that is suitable for public comprehension. For example, the Documentation Outline refers to detailed data curation, preprocessing, enhancement, and annotation measures. Aside from the proprietary concerns, discussed below, this information will likely have limited utility to the public and do little to build the trust the marketplace requires. A tiered approach to transparency that delineates relevant audiences allows organizations to share information that is consistent with the technical fluency of the audience and the purpose for the disclosure. Failure to acknowledge these important differences can be both over- and under-inclusive. It risks sharing too much or an inappropriate type of information that can detract from the goal of providing simplified public transparency, and it could limit the disclosure of relevant information to downstream actors where the information provided is not appropriately calibrated to the need of the relevant organization.

Recommendation: NIST should adopt a tiered approach to AI documentation that recognizes different information is appropriate for different audiences. It should ensure any public-facing documentation advances the goal of enhancing the public's understanding of issues material to the use or impact of the model.

II. Revise the Documentation Outline to Protect Sensitive Proprietary Information

We are also concerned that the Documentation Outline fails to ensure that companies can protect trade secrets or other sensitive proprietary information. NIST acknowledges, as referenced above, that transparency documentation should account for the need to protect trade secrets and other sensitive

information. However, the detailed categories of information that the Documentation Outline includes undermines this goal and necessarily requires the disclosure of sensitive proprietary information. Indeed, NIST proposes disclosure of detailed information about data sets, from data curation decisions prior to inclusion in the data set to data enhancement and annotation. It also recommends detailed disclosures about training protocols.

The protection of sensitive proprietary information is critical with respect to sharing information to all relevant audiences, but as mentioned above, some information not shared with the public may be shared, as appropriate, with downstream actors in connection with non-disclosure agreements or to regulators with legal authority to protect the confidentiality of the information. The public release of this information runs afoul of the broader aim of unleashing AI innovation, as it risks undermining the competitiveness of different companies in the AI ecosystem and jeopardizing the security of AI, network, and information systems. It also contravenes policy priorities reflected in laws across jurisdictions that seek to protect sensitive proprietary information to ensure a healthy marketplace.

By comparison, even the EU's extremely prescriptive and expansive implementation of the EU AI Act still allows for the protection of trade secrets. In particular, the requirement to provide public information about the content used to train general purpose AI models at least acknowledges, with respect to non-public licensed or private datasets, that a general overview of the data modality type, nature of the content, and linguistic characteristics suffices. As a result, the Documentation Outline's failure to even acknowledge the increased sensitivity associated with non-public licensed or private datasets goes beyond the EU's approach, which itself poses potential challenges even with respect to publicly accessible datasets.

In short, the approach that NIST proposes does not adequately protect the confidentiality of sensitive proprietary information and, consequently, is not feasible for companies to implement in practice, undermines the competitiveness of companies developing AI models, thereby harming, not unleashing, AI innovation, and does not align with existing norms.

Recommendation: Revise the template in the Documentation Outline to provide more general descriptions of sensitive information related to data and training protocols to protect the confidentiality of proprietary information.

III. Ensure the Zero Draft Proposed Documentation Standard Complements Existing Transparency Efforts and Provides an Interoperable Solution

The AI transparency ecosystem includes several documentation artifacts, including data, model, and system cards, company transparency notes, disclosures required by certain jurisdictions, including California, and required templates, such as in the EU. We encourage NIST to complement and harmonize, not duplicate or complicate, this already fragmented landscape. In particular, we urge NIST to consider whether and how its proposed standard differs from existing norms and requirements, and how to leverage the common elements of different approaches in a way that both promotes interoperability and advances adoption of approaches that are best suited to enable continued AI adoption. To this end, NIST should consider mapping the proposed standard to other existing requirements and best practices so that it can implement streamlined and harmonized practices instead of adding duplicative or infeasible responsibilities.

Recommendation: Ensure the proposed AI documentation standard leverages existing best practices that are feasible to implement and facilitate interoperability.

We appreciate NIST's significant contributions to the AI policy dialogue and look forward to providing additional input on the forthcoming "Zero Draft" documentation concept paper.

Sincerely,

Shaundra Watson
Senior Director, Policy