October 6, 2022

Chairman Robert White
Committee on Government Operations and Facilities
District of Columbia Council
John A. Wilson Building
1350 Pennsylvania Ave. N.W., Suite 107
Washington, DC 20004

Dear Chairman White,

BSA | The Software Alliance appreciates the opportunity to submit comments on the Stop Discrimination by Algorithms Act (SDAA). These comments are intended to complement our oral testimony during the Council's September 22, 2022, hearing.

BSA is the leading advocate for the global software industry before governments and in the international marketplace.[1] Our members are enterprise software companies that create the technology products and services that power other businesses. They offer tools including cloud storage services, customer relationship management software, human resources management programs, identity management services, and collaboration software.

BSA members are on the leading edge of providing businesses — in every sector of the economy — with trusted tools, including Artificial Intelligence (AI). As leaders in the development of enterprise AI systems, BSA members have unique insights into the technology's tremendous potential to spur digital transformation and the policies that can best support the responsible use of AI.

At the outset, we want to emphasize that BSA would welcome the opportunity to discuss the SDAA with you and your staff. BSA supports the general intent and purpose of the SDAA. We fundamentally agree that the use of high-risk AI should be subject to safeguards. While the adoption of AI can unquestionably be a force for good, we also recognize the significant risks to society if this technology is not developed and deployed responsibly.

Regarding the SDAA, BSA strongly agrees that when AI is used in ways that could unlawfully discriminate or impact access to important life opportunities, the public should be assured that such systems have been thoroughly vetted to identify and mitigate risks associated with unintended bias.

[1] BSA's members include: Adobe, Alteryx, Atlassian, Autodesk, Bentley Systems, Box, Cisco, CNC/Mastercam, CrowdStrike, Databricks, DocuSign, Dropbox, Graphisoft, IBM, Informatica, Intel, Kyndryl, MathWorks, Microsoft, Okta, Oracle, Prokon, PTC, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Unity Technologies, Inc., Workday, Zendesk, and Zoom Video Communications, Inc.

200 Massachusetts Avenue, NW
Suite 310
Washington, DC 20001

P 202-872-5500
W bsa.org

We have concerns, however, with the current legislative language. For example, the bill appears to require third-party audits of AI systems, even though there are no existing standards or professional accreditors for such audits, in contrast with other fields like privacy and cybersecurity. In addition, the bill contains expansive reporting requirements that could be read to require companies to disclose large amounts of data, even though such disclosures may inadvertently create significant security and privacy risks. The bill also requires entities to retain data for a five-year period, which can conflict with data minimization requirements in privacy laws.

Moreover, the bill uses several broad or undefined terms that can create uncertainty about how it would apply. For instance, the bill defines "service providers" in Section 3(7), but also suggests in Section 3(4) that a service provider is a type of "covered entity." This creates significant uncertainty about what types of companies will be considered service providers and what their obligations would be under the bill. Clarifying the role of service providers is a critical component of ensuring the bill can work in practice.

Due to the importance and complexity of this legislation, BSA strongly encourages the Committee to engage directly with a broad range of stakeholders, including members of the technology sector that create and deploy AI systems, as you consider regulating the use of AI by businesses operating in the District of Columbia.

Going forward, we strongly encourage approaching AI regulation by focusing on the need to manage risks across the lifecycle of an AI system.

This has been an area of particular focus for BSA and our member companies over the last several years. In June 2021, BSA released Confronting Bias: BSA's Framework to Build Trust in AI, which sets forth a risk management approach for confronting concerns about bias.[2]

AI is used in so many different contexts that only a flexible, risk management approach will be successful. The BSA Framework is built on three key elements:

1. Identifying the risks of bias through impact assessments across a system's lifecycle;
2. Mitigating those risks through concrete, actionable practices; and
3. Setting forth key corporate governance structures to promote organization accountability.

Among the unique features of the BSA Framework is that it recognizes these elements need to be followed at all stages of the AI lifecycle: Design, Development, and Deployment and Use phases. Further, there are a variety of AI development and deployment models, and the Framework recognizes that the appropriate allocation of risk management responsibilities will vary depending on the type of system, including who develops the algorithm, trains the model, and ultimately deploys the system.

AI Bias Can Arise Throughout the AI Lifecycle. To combat AI bias, it is essential to understand the many sources of risk and the variety of ways they can manifest in an AI system. While much attention has understandably focused on data as a source of bias, the potential vectors of risk precede data collection efforts and begin at the earliest stages of a system's conception and design.

The initial step in building an AI system is often referred to as "problem formulation." It involves the identification and specification of the "problem" the system is intended to address, an initial mapping of how the model will achieve that objective, and the identification of a "target variable" the system will be used to predict. Because many AI systems are designed to make predictions about attributes that are not directly measurable, data scientists must often identify variables that can be used as proxies for the quality or outcome it is intended to predict.

While the use of proxy target variables can be entirely reasonable, the assumptions underlying the choice of proxies must be closely scrutinized to ensure that it does not introduce unintended bias to the system.

[2] Confronting Bias: BSA's Framework to Build Trust in AI, available at https://ai.bsa.org/confronting-bias-bsas-framework-to-build-trust-in-ai.

200 Massachusetts Avenue, NW
Suite 310
Washington, DC 20001

P 202-872-5500
W bsa.org

The risk that can arise during this process of problem formulation is perhaps best exemplified by a recent study of a widely used healthcare algorithm that hospitals rely on to identify patients in need of urgent care. The research team concluded that the algorithm was systematically assigning lower risk scores to black patients compared to similarly sick white counterparts because it relied on data about historical healthcare costs as proxy for predicting a patient's future healthcare needs. Unfortunately, because black patients have historically had less access to healthcare, the reliance on spending data painted an inaccurate picture and led to dangerously biased outcomes.[3]

The data used to train an AI system is a second major vector for bias. If the data used to train a system is misrepresentative of the population in which it will be used, there is a risk the system will perform less effectively on communities that may be underrepresented in the training data. Likewise, reliance on data that itself may be the product of institutional or historical biases can entrench those inequities in an AI model. The process of "labelling" training data can also introduce bias. Many AI systems require training data to be "labeled" so that the learning algorithm can identify patterns and correlations that can be used to classify future data inputs. Because the process of labeling the data can involve subjective decisions, there is the potential for introducing unintended bias into the training data.

Finally, even a system thoroughly vetted during development can begin to exhibit bias after it is deployed. AI systems are trained on data that represents a static moment in time and that filters out "noise" that could undermine the model's ability to make consistent and accurate predictions. Upon deployment in the real world, AI systems inevitably encounter conditions that differ from those in the development and testing environment. Further, because the real-world changes over time, the snapshot in time that a model represents may naturally become less accurate as the relationship between data variables evolves. If the input data for a deployed AI system differs materially from its training data, there is a risk that the system could "drift" and that the performance of the model could be undermined in ways that will exacerbate the risks of bias. For instance, if an AI system is designed (and tested) for use in a specific country, the system may not perform well if it is deployed in a country with radically different demographics. Bias can also arise if an AI system is deployed into an environment that differs significantly from the conditions for which it was designed or for purposes that are inconsistent with its intended use.

Combatting AI Bias Requires a Lifecycle-Based Approach to Risk Management. Although the challenges of AI bias are significant and without simple solutions, they are not insurmountable. Efforts to combat bias must start by recognizing that the issue requires a lifecycle-based approach to risk management.

Risk management is a process for ensuring systems are trustworthy by design by establishing a methodology for identifying risks and mitigating their potential impact. Risk management processes are particularly important in contexts, such as cybersecurity and privacy, where the combination of quickly evolving technologies and highly dynamic threat landscapes render traditional "compliance" based approaches ineffective. Rather than evaluating a product or service against a static set of prescriptive requirements that quickly become outdated, risk management seeks to integrate compliance responsibilities into the development pipeline to help mitigate risks throughout a product or service's lifecycle.

In practice, that means companies that develop or use high-risk AI systems should establish a comprehensive approach for performing impact assessments. Impact assessments are widely used in a range of other fields—from environmental protection to data protection—as an accountability mechanism that promotes trust by demonstrating that a system has been designed in a manner that accounts for the potential risks it may pose to the public. The purpose of an impact assessment is to establish organizational processes to guide the development and use of high-risk systems by requiring internal stakeholders to identify the risks that a system may pose, quantify the degree of harm the system could generate, and document any steps that have been taken to mitigate those risks to an acceptable level. By

[3] See Heidi Ledford, Millions of Black People Affected by Racial Bias in Health-Care Algorithms, Nature (Oct. 24, 2019), available at https://www.nature.com/articles/d41586-019-03228-6.

establishing a process for personnel to document key design choices and their underlying rationale, impact assessments are an important transparency and accountability mechanism.

The impact assessment methodology in the BSA Framework includes more than 40 diagnostic statements that should be documented throughout an AI system's lifecycle. Among its key recommendations is for organizations to maintain documentation about:

- The objectives and assumptions of the system, including its intended use cases and its target variable;
- The metrics that will be used as a baseline for evaluating bias in the system;
- The provenance of the data used to train the system, an evaluation of its appropriateness for the intended use case, and the steps that were taken to scrutinize the data for biases;
- The rationale for selecting data attributes and their impact on model performance; and
- The lines of responsibility for monitoring the system following deployment and plans for responding to potential incidents or system errors.

Mitigating AI Bias Requires Diverse, Interdisciplinary Expertise. A common refrain in the BSA Framework relates to the vital role of diversity in AI risk management efforts. Effectively identifying potential sources of bias in data requires a diverse set of expertise and experiences, including familiarity with the domain from which data is drawn and a deep understanding of the historical context and institutions that produced it. Moreover, oversight processes are most effective when team members bring diverse perspectives and backgrounds that can help anticipate the needs and concerns of users who may be impacted by or interact with an AI system.

Because "algorithm development implicitly encodes developer assumptions that they may not be aware of, including ethical and political values," it is vital for organizations to establish teams that reflect a diversity of lived experiences and that traditionally underrepresented perspectives are included throughout the lifecycle of the AI design and development process.[4] To the extent an organization is lacking in diversity, it should consult with outside stakeholders to solicit feedback, particularly from underrepresented groups that may be impacted by the system.

Finally, we want to emphasize the importance of recognizing the different roles that different companies play in developing and deploying AI systems. **It is critical for any AI regulation, including the SDAA, to account for these different roles.**

Risk management is a collective responsibility. Focusing on the need to manage risks throughout the lifecycle of an AI system also facilitates important communication between the multiple stakeholders that may have different roles to play in managing AI risks. In many instances, the risk of bias may emerge at the intersection of system design decisions that were made by the system's developer and downstream decisions by the organizations that may deploy that system. In such a circumstance, risk management responsibilities will necessarily be shared by the system developer and the organization that deployed it. In other situations, the customers may, for privacy or other purposes, not allow the developer to view or assess data that may be used to re-train or fine tune the AI model. In that scenario, the customer that deployed the AI system is best positioned to document features of data sets used to train the model and evaluate the impact of use in different contexts.

While the precise allocation of risk management responsibilities will vary depending on the use case, as a general matter AI developers will be best positioned to provide information about the system's design and capabilities to enable the deployer to make informed deployment and risk mitigation decisions.

[4] Inioluwa Deborah Raji et al., *Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing*, FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (January 2020): 33–44, https://doi.org/10.1145/3351095.3372873.

Thank you again for the opportunity to share BSA's perspective on the SDAA. We look forward to continuing this conversation and to serving as a resource to you and other members of the DC Council.

Sincerely,

Tom Foulkes
Senior Director, State Advocacy
BSA | The Software Alliance

200 Massachusetts Avenue, NW
Suite 310
Washington, DC 20001

P 202-872-5500
W bsa.org