
**Hearing on**

**"Beyond I, Robot: Ethics, Artificial Intelligence, and the Digital Age"**

**United States House of Representatives Committee on Financial Services**
**Task Force on Artificial Intelligence**

**October 13, 2021, at 12:00 p.m.**

**Testimony of Aaron Cooper**
**Vice President, Global Policy**
**BSA | The Software Alliance**

**Testimony of Aaron Cooper**
**Vice President, Global Policy, BSA | The Software Alliance**
**Hearing on "Beyond I, Robot: Ethics, Artificial Intelligence, and the Digital Age"**

**Before the United States House of Representatives**
**Committee on Financial Services**
**Task Force on Artificial Intelligence**

**October 13, 2021**

Good afternoon Chairman Foster, Ranking Member Gonzalez, and members of the AI Task Force. My name is Aaron Cooper. I am Vice President of Global Policy for BSA | The Software Alliance (BSA).

BSA is the leading advocate for the global software industry.[1] Our members are at the forefront of developing cutting-edge, data-driven services that have a significant impact on US job creation and growing the global economy. I commend the Task Force for convening today's important hearing, and I thank you for the opportunity to testify.

Enterprise software services, including artificial intelligence (AI) are accelerating digital transformation in every sector of the economy. Artificial intelligence is not just about robots, self-driving vehicles, or social media. It can be used by businesses of all sizes to improve their competitiveness, enhance their value proposition, and increase their capacity to make data-informed decisions.

BSA represents the perspective of the enterprise software companies that help make this possible. Our members create the technology products and services that help other businesses innovate and grow. In that capacity, BSA members are on the leading edge of providing businesses in every sector of the economy with the trusted tools they need to leverage the benefits of AI.

The promise that AI may one day impact every industry is quickly turning into a commercial reality and driving the digital transformation. For instance, Autodesk brings the power of AI to product design, helping American manufacturers improve the performance of their products while reducing their costs and environmental impact. In one recent collaboration, Autodesk worked closely with engineers at General Motors to explore how AI-enabled generative design could help the company optimize its design and manufacturing processes.[2] As an initial proof-of-concept, the two companies set out to improve GM's approach to designing and manufacturing the brackets that secure seatbelts and seats to a car's floor. The partnership yielded immediate benefits, enabling GM to identify a new design that is 40 percent lighter and

[1] BSA's members include: Adobe, Atlassian, Autodesk, Bentley Systems, Box, CNC/Mastercam, DocuSign, IBM, Informatica, MathWorks, Microsoft, Okta, Oracle, PTC, Salesforce, ServiceNow, Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, Twilio, Workday, Zendesk, and Zoom Video Communications, Inc.

[2] General Motors | Generative Design in Car Manufacturing | Autodesk

20 percent stronger than its previous approach. This resulted in a savings of material costs, simplification of the supply chain, and a reduction in assembly time.

Splunk is helping the financial services sector leverage AI to take a bite out of the more than $40 billion that is lost to fraudulent transactions each year. Splunk's software powers a suite of enterprise fraud management capabilities that allow banks to identify transaction anomalies in real time, reduce the frequency of false positives, and better protect consumers from identity theft.[3]

While the adoption of AI can unquestionably be a force for good, it can also create real risks if not developed and deployed responsibly. We commend the Task Force for convening today's hearing to examine the role that frameworks for ethical AI can play in ensuring the responsible use of this technology. This is an area of particular focus for BSA and our member companies are leaders when it comes to responsible AI practices.[4] We recently produced a detailed framework that sets forth a risk management approach for confronting concerns about bias. As the Task Force explores the use of these tools, we offer our perspective on how they can be used to address the risk of bias, which we hope will also inform the broader conversation at the hearing today.

As this Task Force is aware, the data-driven nature of AI makes it susceptible to unintentional bias. Because AI is trained on data from the past, there is a risk that AI systems may replicate and potentially further entrench historical inequities. As AI is integrated into business processes that can have consequential impacts on people's lives, there is a risk that "biased" systems will perform less accurately or unfairly disadvantage members of historically marginalized communities.

For BSA members, earning trust and confidence in the AI and other software they develop is crucial, so identifying ways to reduce the risk of bias is a priority. BSA therefore set out to develop real, credible, and actionable steps to guard against the potential of AI systems producing unintended disparate impacts. The resulting framework – Confronting Bias: BSA's Framework to Build Trust in AI – was released in June and is built on a vast body of research and informed by the experience of leading AI developers.[5]

The Framework outlines a lifecycle-based approach for performing impact assessments to identify risks of AI bias and corresponding best practices for mitigating those risks. The foundation of the Framework is its detailed methodology for performing impact assessments that help ensure that critical decisions are documented and that an organization's product development team, its compliance personnel, and senior leadership are aligned on the appropriate steps for mitigating risks of bias when they are identified. The Framework is intended to scale with risk and recognizes that inherently low-risk systems—for example, a system used to predict the type of fonts being used on a document—may not require a full

---

[3] Detecting Credit Card Fraud Using SMLE | Splunk; Splunk at TransUnion | Splunk

[4] See, e.g., Adobe - Adobe unveils new AI ethics principles as part of commitment to responsible digital citizenship; IBM - 3 lessons from IBM on designing ethical AI technology | World Economic Forum (weforum.org); Microsoft - Our approach to responsible AI at Microsoft; Salesforce - Salesforce Debuts AI Ethics Model: How Ethical Practices Further Responsible Artificial Intelligence - Salesforce News; Workday - Building Trust in AI and ML Through Principles, Practice, and Policy (workday.com).

[5] Confronting Bias: BSA's Framework to Build Trust in AI

impact assessment. But for systems that pose heightened risks, a robust impact assessment is essential to mitigating potential harms.

The Framework is ultimately a playbook that organizations can use to enhance trust in their AI systems through risk management processes that promote fairness, transparency, and accountability, three of the key principles for responsible and ethical AI. The full Framework, with more than 50 actionable diagnostic statements, is attached to my testimony and can be found at ai.bsa.org. Below, I share a few key insights from the Framework.

**Overview**

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. 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 of spending data painted an inaccurate picture and led to dangerously biased outcomes.[6]

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 a 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.

But, what does that all mean in practice?

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

[6] Millions of black people affected by racial bias in health-care algorithms (nature.com)

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 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.

- ***Risk Management is a Collective Responsibility***

The documentation created and maintained as part of an impact assessment also facilitates important communication between the multiple stakeholders that may have roles to play 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.

For instance, some AI developers provide general-purpose AI functionality, such as text analytics tools, that their customers can access and integrate into their own products and services via an API. 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.

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.

- ***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.[7] 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.

**Policy Recommendations**

Public trust is an essential component of a thriving digital economy. While the responsibility for managing the risks of AI falls squarely on the organizations that develop and use AI systems, government can help foster public trust through policies that enhance the benefits of the technology while safeguarding against its potential risks. In the near term, we would advise Congress and the Administration to focus on the following lines of effort.

(1) Ensure consumer and civil rights protections remain fit-for-purpose in the digital age. Decisions that would otherwise be unlawful should not avoid liability simply because they may now involve the use of an AI system. To that end, we have encouraged efforts to audit federal agencies' existing consumer protection authorities to assess whether technological innovation is impeding their ability to enforce the law.[8] And we wrote to this Task Force last year about concerns that a rulemaking at HUD may exacerbate the risk of bias and discrimination in the housing market.[9]

(2) Establish a requirement for organizations to perform impact assessments prior to deploying high-risk AI systems. The BSA Framework can be one useful roadmap for new legislation.

(3) Promote international alignment around AI policy. Given the inherently global nature of the technology ecosystem, it is vital for the US to engage with our trading partners to forge consensus approaches for tackling shared challenges. There is an emerging global consensus that AI regulation should be risk-based and context specific. The EU recently introduced comprehensive legislation along these lines. The US should look for opportunities to drive these conversations, including through NIST's development of an AI risk management framework.

(4) Continue to emphasize privacy and security. Ethics and issues of bias are part of the trust formula, but privacy and data security laws are also essential.

---

[7] 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.

[8] *See* BSA Comments on Office of Management and Budget's Guidance on AI Regulation | BSA | The Software Alliance

[9] US: BSA Letter to the House Financial Services Committee Regarding Equitable Algorithms Hearing

## Conclusion

Digital transformation across industry sectors is creating jobs and improving our lives. But industry, civil society, academia, and the government must work together on guidelines and laws that will ensure companies act responsibly in how they develop and deploy AI.

We appreciate the Task Force's strong focus on issues of ethics and bias. Confronting Bias: BSA's Framework to Build Trust in AI is our attempt to contribute meaningfully to this discussion. Thank you again for the opportunity to testify.