



FAVORABLE

Health and Government Operations Committee

HOUSE BILL 1323

Algorithmic Decision Systems –Procurement and Discriminatory Acts

March 15, 2021

Chair Pendergrass, Vice Chair Pena-Melnyk, and members of the committee, my name is Vandana Janeja and I am a Professor and Chair of the Information Systems Department at UMBC and a resident of Howard County. I am pleased to be able to testify in support of Delegate Hill's [House Bill 1323 Algorithmic Decision Systems –Procurement and Discriminatory Acts](#). Passage of this bill would reflect leadership in the [landscape of efforts](#) across the country considering the societal impacts of using artificial intelligence (AI) systems, particularly in government.

I have been an educator and researcher in data analytics for over 13 years. I submit my views in this testimony as a researcher for data driven decision making and how these systems can or should incorporate ethical viewpoints. I have co-authored community reports for the National Science Foundation (NSF) and co-organized a workshop on embedding ethics in the data science pedagogy. I have also served as an AAAS Science Technology Policy fellow at NSF, in the Computer Information Science and Engineering Directorate's Office of the Assistant Director, where I am also affiliated as an expert in the area of data science.

This testimony is written jointly with James Foulds, an Assistant Professor in the Information Systems Department at UMBC. He is an expert who researches fairness and bias in Artificial Intelligence (AI) and machine learning systems. He has received three National Science Foundation awards on this topic, including the prestigious NSF CAREER award, a grant from the National Institute of Standards and Technology, and a number of research papers in highly selective publication venues.

In this written testimony we reflect on our support for this bill in terms of the increased use of AI, its benefits, pitfalls and responsible AI best practices. As experts, we believe that this bill's purpose, to ensure that systems purchased by the state are not replicating bias through the use of algorithms, is both laudable and possible.

Increased use of AI and Data Driven Systems in Decision Making

With the advances in data collection, capturing and sensing we are living in a constant deluge of data. Several disparate and rich data sources are emerging to help with actionable knowledge discovery. While disparate data proves to be a challenge to work with, it can also be seen as an opportunity to tap into relevant datasets and discover yet unknown patterns and support data driven decision making. Indeed AI and data driven decision making systems are in use in our own backyards in an effort to support our communities (e.g. [MD COVID-19 pandemic response](#), [DHS program to improve flood resiliency](#), [UMBC's Flood Bot project in Ellicott City](#), [filling supply chain gaps with AI](#)). AI and data driven systems are also coming into play into consequential decision making impacting lives and livelihoods (e.g. [MD courts using AI systems for bail decisions](#), [facial recognition in Capital Gazette shooting](#), [ICE facial recognition searches](#)). With technological advances now enabling the incorporation and systematic pattern analysis of data at scales using machine learning - beyond human perception, we must revisit how we view our work as part of a complex ecosystem, filled with feedback loops.

Benefits of using AI and Data-Driven Systems in Decision Making

Algorithmic decision systems, developed using data-driven technologies from artificial intelligence, machine learning, and statistics, are extremely beneficial when done well. These systems can lead to substantial efficiency benefits in terms of the speed, cost, and scalability of decision-making compared to human decision-making, and they often lead to better and more reliable decisions. If implemented appropriately, these systems also have the potential to be more fair than human decision-makers, since the same procedure is applied evenhandedly to everyone, while humans are subject to both implicit and explicit biases.

Pitfalls of AI and Data Driven Systems in Decision Making

Fair and equitable behavior of data-driven algorithmic decision systems is far from guaranteed. Artificial intelligence and machine learning systems aim to encode the patterns that they observe in data, regardless of whether those patterns arise from fair societal processes or justifiable data collection and data preparation choices. Hence, inequities in our society are reflected in data, and in data-driven algorithms. Human prejudice in annotating data with the labels that an AI system aims to predict (e.g. whether an individual “deserves” a particular governmental service or program), will be reflected in the behavior of a data-driven AI system.

The data used to train the algorithm may not be representative of the individuals in society, e.g. by neglecting historically marginalized communities. Even if the data are representative, minority groups will by definition be represented less often in the training data set, potentially biasing the outcomes. Machine learning methods are vulnerable to the problem of *overfitting*, in which a predictive model fits too well to its training data, and hence fails to generalize to the rest of the population outside of the training data. Overfitting can further lead to an amplification of the discriminatory biases already present in the data, leading to inaccurate and discriminatory decisions. If a data-driven algorithm is allowed access to sensitive demographic information

such as an individual's gender, race, ethnicity, nationality, sexual orientation, and age, it will typically learn to encode unwanted discriminatory behavior and stereotypes. Even if sensitive demographics are hidden from the algorithm, other measured attributes which are correlated with them may act as "proxy variables" to the demographic information (e.g., [zip code is correlated with race in the United States due in part to historical segregation policies](#)).

Of course, beyond all of the above subtle issues which may occur in otherwise well-designed systems with otherwise "good" data, any errors in an individual's recorded data are likely to lead to incorrect decisions. [Experts have made the case](#) that if an individual is not allowed to access the data used to make a decision or an explanation of why the decision was made, they will have no recourse to challenge that decision, even when that decision was made in error.

To highlight the overall pitfalls, well-known cases of AI bias include [racial bias in bail and sentencing decisions from criminal recidivism risk prediction systems](#), and [disparities in the accuracy of computer vision systems along lines of gender and skin tone](#) with potentially harmful impacts for individuals impacted and also for society as a whole.

Responsible AI and Data Driven Systems in Decision Making: Best Practices

Mediating the discussions between application stakeholders is becoming ever more important. The vast complexity of data availability and algorithmic decision making tools requires vigilance around access, privacy, provenance, curation, and interoperability, issues of fairness, accountability and transparency. To help navigate increasingly complex decision making systems, there is a movement toward standards and best practices such as a data ethics oath (e.g. [IEEE](#), [ACM](#), [NASEM](#)), [ethically aligned business standards](#), private sector best practices for responsible AI (e.g. [Google](#), [Accenture](#), [Microsoft](#), [PwC](#), [IBM toolkit](#)), [best practices and policies to reduce consumer harm](#). These issues are a natural consequence of deploying solutions in the real world, and must be considered intentionally by all parties involved.

While some may have concerns about the availability of established standards, there are many established best practices. Indeed the conversations in the [state](#), [federal](#) and [private sectors](#) have highlighted the need for more stringent measures in deploying AI and data-driven decision making systems to prevent propagation of societal biases and discrimination.

This bill provides sensible and reasonable requirements which will hold the software vendors accountable and would help to ensure that responsible AI best practices will protect Maryland's citizens and advance the state's efforts to be inclusive and fair.

Considering the far reaching implications of AI and data-driven algorithmic decision systems, which may cause harm if not supported with the right best practices, We appreciate your favorable consideration of HB 1323.