To:
Joint Committee on State Administration and Regulatory Oversight
91st General Court of the Commonwealth of Massachusetts

September 25, 2019

Mr. Chairman and Members of the Committee:

Thank you for giving me this opportunity to testify in front of you on this significant bill you are considering today. My name is Karthik Dinakar, and I am a computer scientist at MIT in the fields of machine learning and natural language processing. My research is part of the Ethics & Governance of AI Initiative at MIT and Harvard Law School. I study the interplay between machine learning and decision making in the areas of criminal justice and medicine.

I support the bill S.1876/H.2701 you are considering today. The commission that the bill seeks to establish as well as the issues that it will study is of enormous consequence and import, not just in our commonwealth, but across the entire nation. I sit before you today filled with the spirit of an icon who often appears in my research. Julius Chappelle, the African-American politician who served with great distinction as a member of the Massachusetts State legislature from 1883 to 1891. Back then, private life insurance policies discriminated against black people by differentially charging them a higher premium. They justified this practice by claiming that data showed higher mortality rates for blacks, and hence, black people merited higher premiums. Julius proposed a bill to ban this practice, reframing the future potential of free black people for an equal chance at life unfettered by the bonds of slavery. Julius' vision for the future prevailed against fatalistic clinging of statistics, and his bill passed. The passage of the law in Massachusetts triggered a wave of similar passages in other states across the country. I believe very strongly that the legislation you are considering today is of the same consequence and import.

I have three core issues that I submit for the consideration of the commission that this legislation seeks to establish. Let me highlight these three core issues by giving a concrete example.

I have studied the use of machine learning risk-assessment algorithms in the pretrial criminal justice system. Proponents of risk-assessments algorithms claim that they are a way to reduce bias in judicial pretrial decisions. The risk-assessment algorithm gives each defendant two risk scores. These risk scores must be interpreted as a kind of prediction. The first risk-score is how likely they'd miss their court date and the second is how likely they are to get rearrested while they await trial. Pretrial judges consider both scores while deciding whether to release the defendant pending trial and on what the terms of the release would be. But these risk scores suffer from grave fundamental flaws.

First, these risk-assessment algorithms are trained by statisticians on historical arrest data. It is clear from decades of research that historical arrest data are not equivalent to crime data. They are more a reflection of policing practices than an accurate measure of crime. We know that people of color and underprivileged socio-economic communites are overrepresented in historical arrest data. Because risk-assessment algorithms learn from historical arrest data, they can learn the biases baked in that data. Making these algorithms "fair" and "interpretable" is an active area of research within the AI machine learning community. There are now many technical definitions of fairness. For example, one definition is to ensure predictive parity of risk-scores across different races. It is equivalent to the statistician placing his or her thumb on the scale. But who decides which definition of fairness to encode into the algorithm? And is the concept of fairness even reducible to a computational

interpretation? Research from the state of Kentucky has shown that after the state adopted pretrial risk-assessment algorithms, racial disparities in the pretrial carceral population has increased, despite claims of fairness and transparency by the creators.

Second, risk-assessment algorithms don't accurately measure the risk of new violent criminal activity as they are often touted as doing. Violent crime while on bail is rare. Research has shown that for every hundred defendants given "high-risk" scores for violent crime, only eight go on to commit a crime while awaiting trial. The risk-scores for defendants marked as low, medium, and high-risk are found to have overlapping margins of errors on their risk-scores, with the implication that we cannot identify high-risk defendants from the medium and low-risk ones. An irrational fear of mass violent crime is a contributing factor behind the age of mass incarceration. It should not be driving us to invent mathematical objects that cannot even accurately identify who is likely to commit a crime upon release. Building a data regime using algorithms further cement and ossify this broken frame and failed logic.

The commission that this legislation seeks to establish must consider three core issues of the algorithms and decision making in our state along with those already enumerated by the draft text:

1. **Ensure that the concept of fairness is not implicitly and legally sanctioned as a computational definition.** One cannot claim to be fair by checking a couple of technical definitions. The idea of "actuarial fairness" was invented in the 1980s by the health insurance industry to counter claims of discrimination by grass-roots activists and civil-rights organizations. The framing of fairness in complex mathematical terms made it seem scientific and objective. It was also difficult for non-statisticians to understand. This is widely seen as a reason why the principle of solidarity in health insurance - advocated in many state legislatures, lost the debate to proponents of individualized health risk in the 1980s. This committee can take the lead in making sure that this does not happen again in the age of algorithms.
2. **Ensure that every technical step taken in building algorithms is participatory, including stakeholders from civil society.** Validation of algorithms is not enough. Stakeholders must be involved in every step in the process, from training data criticism, algorithm specification, and evaluation, including stringent periodic oversight.
3. **Make concrete legislative protections for the due process rights of individuals and communities that are not currently addressed by existing jurisprudence of disparate impact and disparate treatment.** There is increasing alarm and concern in legal scholarship that some of the technical issues I have highlighted above and many others are not adequately addressed by the concepts of disparate impact and disparate treatment when it comes to algorithmic predictions. Massachusetts must take the lead in addressing this legal lacunae.

Mr. Chairman, Julius Chappelle was born into slavery in South Carolina, but made invaluable contributions to this august body during his time here in the 1880s. He reframed discriminatory practices of the life insurance businesses by rejecting fatalistic statistics based on biased historical data. He imagined the equal future potential of those that had become free. He codified this ideal into law. That very law triggered a wave of positive change all across our nation. You and this committee are embarking on a similar endeavor. I wish you the very best. **I strongly support this legislation and urge the Committee to give S.1876 and H.2701 a prompt favorable report.**

Sincerely,

Karthik Dinakar