

The Civil Rights Implications of Algorithms



A Briefing Report of the
Connecticut Advisory Committee to the
U.S. Commission on Civil Rights

March 2023

Advisory Committees to the U.S. Commission on Civil Rights

By law, the U.S. Commission on Civil Rights has established an advisory Committee in each of the 50 states and the District of Columbia. These Committees are composed of state/district citizens who serve without compensation; they are tasked with advising the Commission of civil rights issues in their states/district that are within the Commission's jurisdiction. Committees are authorized to advise the Commission in writing of any knowledge or information they have of any alleged deprivation of voting rights and alleged discrimination based on race, color, religion, sex, age, disability, national origin, or in the administration of justice; advise the Commission on matters of their state or district's concern in the preparation of Commission reports to the President and the Congress; receive reports, suggestions, and recommendations from individuals, public officials, and representatives of public and private organizations to Committee inquiries; forward advice and recommendations to the Commission, as requested; and observe any open hearing or conference conducted by the Commission in their states/district.

Letter of Transmittal

Connecticut Advisory Committee to
the U.S. Commission on Civil Rights

Members of the Commission

J. Christian Adams
Rochelle Mercedes Garza
Stephen Gilchrist
Gail Heriot
Mondaire Jones
Peter N. Kirsanow
Glenn D. Magpantay
Victoria Frances Nourse

The Connecticut Advisory Committee, as part of its responsibility to advise the Commission on civil rights issues within the state, submits this report, “The Civil Rights Implications of Algorithms.” The report was unanimously approved by the Advisory Committee on March 16, 2023.

Sincerely,

David McGuire, *Chairperson*
Connecticut Advisory Committee

Connecticut Advisory Committee to the U.S. Commission on Civil Rights

David McGuire, *Chair*, West
Hartford

Venesia Hurtubise, *Vice Chair*,
Wethersfield

Ingrid Alvarez-DiMarzo, New
Fairfield

Saud Anwar, South Windsor

Anthony Ciolli, West Hartford

Christine W. Corgel, Westport

Charles Gardner, Sandy Hook

Mark Gius, Southington

Melanie S. Rausch, Bethany

Cheryl Sharp, Cheshire

John M. Tanski, South
Glastonbury

Susan D. Toliver, Stamford

Michael D. Werner, East
Haddam

Olivia S. White, Essex

Acknowledgments

The Advisory Committee extends its appreciation to the panelists at the four briefings to the advocates, government officials, and others for their participation in the briefings. The Committee also extends its appreciation to attorney Spencer Hill for his outstanding contributions to the work of the Advisory Committee. Staff extends appreciation to David McGuire for his unparalleled leadership and to Venesia Hurtubise for her assistance with the report.

Table of Contents

I. Introduction/Background.....6

II. The Growing Use of Automated Decision-Making by the Government6

 A. What are Algorithms? 6

 B. Causes of Discrimination in Algorithms..... 7

 C. Fairness vs. Accuracy as a Source of Algorithmic Bias 9

 D. Use of Algorithms in Government..... 10

 E. Transparency Issues 12

 F. Current Regulatory Approaches 13

III. Summary of the Briefings15

IV. Findings and Recommendations16

I. Introduction/Background

The Connecticut Advisory Committee to the U.S Commission on Civil Rights examined the issue of algorithmic bias in Connecticut. The topic is one of first impression for the Commission and its Advisory Committees. The Committee is examining the civil rights implications in the use of algorithms – a set of instructions for how to solve a problem – by state actors and agencies.

II. The Growing Use of Automated Decision-Making by the Government

A. What are Algorithms?

As complex as they may become, at their core, algorithms are simply a set of instructions for how to solve a problem.¹ The Committee’s work focused on the government’s use of algorithms to drive automated decision-making to perform government functions and business. In this report, algorithms and automated decision-making are often used interchangeably. As more popularly understood, the term “algorithm” is generally used to refer to either artificial intelligence, a subfield of computer science concerned with intelligent behavior, or machine learning, a subfield of artificial intelligence that focuses on computer programs that are able to learn from data.² For example, Netflix uses algorithms to predict which television shows and movies you may want to watch based on your past viewing habits.³ Amazon’s “Alexa” device uses an algorithm to interpret human speech and respond appropriately.⁴

Algorithms have transformed our society, bringing with them a range of benefits and challenges. They can make faster decisions, process more data, and may be more reliable and accurate than a human. Algorithms may allow more reliable predictions in comparison to a human.⁵ These benefits have transformed the way we live and have become so ubiquitous that many people do not realize just how omnipresent algorithms have become in our daily lives.

The more advanced algorithms work by taking a body of data and using it to make predictions according to the programmed goals of its creators. For example, “nearest neighbor” algorithms try to interpret a new input by comparing it to similar data from the past and then making a prediction whether the new input is the same as the others.⁶ This is how, for example, the United States Post Office uses computers to read handwritten addresses on envelopes.⁷ The algorithm compares a letter in new handwriting to examples of handwritten letters of a similar shape, then predicts what the letter in question likely is. The use of this algorithm enables a computer to read an envelope and sort it faster than a human could.

¹ Dr. Philip Thomas, Autonomous Learning Lab Center for Data Science, Briefing before the Connecticut Advisory Committee to the U.S. Comm’n on Civil Rights, September 8, 2022, transcript p. 57 [hereinafter cited as CT September Briefing Transcript].

² *Id.* at 58.

³ *Id.* at 61.

⁴ *Id.*

⁵ *Id.* at 63.

⁶ *Id.* at 59.

⁷ *Id.*

B. Causes of Discrimination in Algorithms

Algorithms are created by humans and operate using information obtained from human society. As such, they are not value free or free from bias. Instead, they reflect the biases of their creators and biases in the data they use to making predictions.

It should first be noted that sometimes algorithms are directed to explicitly discriminate. For example, Facebook created a mechanism by which advertisers could target their ads according to the demographics of individuals, including their race and sex.⁸ For this, Facebook faced lawsuits from multiple organizations, ultimately settling them in 2019, though additional monitoring has led to renewed litigation.⁹

One way algorithms can implicitly discriminate is through reliance on data sets that are themselves tainted by discrimination.¹⁰ For example, Amazon created an algorithm to predict who would make the best employees and then screen applicants based on that criteria.¹¹ In creating the algorithm, it used data from its current workforce over the past ten years.¹² The Tech industry, however, has a long history of sexism and so there were comparatively few women in Amazon's workforce for the algorithm to look at.¹³ As a result, the algorithm predicted that women would not make for good employees and it ended up screening out female applicants based on its historically tainted data set.¹⁴ With no human intervention involved, the algorithm perpetuated the existing patterns of discrimination.

Algorithms can unintentionally mischaracterize or misinterpret the data – even when the data sets are reliable. Google Translate is an algorithm that translates text from one language to another, using scanned texts from one language to choose the most likely correct translation.¹⁵ Not all languages function in the same way, however, which can lead to machines having to make a prediction as to the most closely analogous translation. For example, Turkish only has one pronoun to use for the third-person, “O,” as opposed to “he,” “she,” and “it” used in English.¹⁶ When making a translation, Google scans texts in English and Turkish to select the most appropriate pronoun to use since there is not a clear analogue between the two languages. This can result in the algorithm making sexist translation assumptions based on biased data. For example, all of the pronouns in the following text were selected by Google's algorithm based on its assumptions as to the gender of the sentence's subject:

⁸ Amalea Smirniotopoulous, Senior Policy Counsel at the NAACP's Legal Defense and Educational Fund, *Briefing Before the Connecticut Advisory Committee to the U.S. Comm'n on Civil Rights*, November 7, 2022 [hereinafter referred to as the “CT November Briefing Transcript”], p. 2. According to Ms. Smirniotopoulous, the Department of Housing and Urban Development accused Facebook of enabling advertisers to exclude users based on, “race, color, religion, sex, familial status, national origin, and disability.”

⁹ *Id.* at 6.

¹⁰ Thomas Testimony, *CT September Briefing Transcript*, p. 64.

¹¹ Sruthi Venkatachalam, Yale Law School Media Freedom and Information Access Clinic, *CT September Briefing Transcript*, p. 107.

¹² *Id.* at 106.

¹³ *Id.*

¹⁴ *Id.*

¹⁵ Thomas Testimony, *CT September Briefing Transcript*, p. 67.

¹⁶ *Id.* at 65.

He is a soldier. She is a teacher. He is a doctor. She is a nurse. He is a writer. He is a dog. She is a nanny. It is a cat. He is a president. He is an entrepreneur. She is a singer. He is a student. He is a translator. He is hardworking. She is lazy.¹⁷

Assumptions that a soldier or president must use masculine pronouns while nurses use feminine pronouns are born from Google's library of scanned texts in which the majority of soldiers and presidents are men while the majority of nurses are women. As a result, the algorithm perpetuated those sexist assumptions in its translation.

Another way that algorithms can inject bias is by replicating the biases and assumptions of their creators. Sometimes this is deliberate, as with Microsoft's Tay AI chatbot program whose users intentionally trained it to make anti-Semitic statements.¹⁸ Other times, the creators may train their algorithms on themselves, leaving the algorithm to assume those working in the tech industry are representative of all humans. For example, one algorithm unintentionally started identifying pictures of Black people as gorillas.¹⁹ The algorithm had been created by very few Black programmers and so unintentionally associated dark skin with apes. In another example, software designed to detect cheating by using student's webcams during exams failed to work on darker skinned students since the program had been tested predominantly on lighter-skinned subjects.²⁰

More worrying is when programmers make assumptions about their data without understanding the biases behind it. In healthcare, a developer created an algorithm to help people get access to programs assisting those with chronic health diseases.²¹ In doing so, the developer coded the algorithm to look for people who paid the most into the healthcare system on the assumption that filtering the data this way was the best way to identify people with chronic issues. The programmer did not realize that health disparities and access to affordable healthcare treatment is heavily influenced by race and, as a result, the program kicked out every low-income patient of color.²²

It's important to note that these algorithms generally do not overtly consider legally protected characteristics like sex or race when looking at data. What more typically happens is that the algorithms look at other variables that are often tied to those characteristics because of past patterns of systemic discrimination in our society. It does this through two ways: making inferences based on past behaviors and by using what are called proxy variables.

Inferences are when a program uses past behaviors to make a prediction that becomes associated with a protected characteristic.²³ For example, algorithms in the financial sector started associating difficulties with paying back loans with people whose names sounded ethnic – Jose, Juan, or Marquetta, as examples.²⁴ As a result, people with similar sounding names were charged higher

¹⁷ *Id.* at 65.

¹⁸ *Id.* at 70.

¹⁹ Dr. Nicol Turner, Senior Fellow of Governance Studies and Director of the Center for Technology Innovation at the Brookings Institute, *CT September Briefing Transcript*, p. 39.

²⁰ Dr. Suresh Venkatasubramanian, Brown University, *CT September Briefing Transcript*, p. 172.

²¹ *Id.* at 40.

²² *Id.*

²³ Turner Testimony, *CT September Briefing Transcript*, p. 8

²⁴ *Id.*, p. 9.

interest credit card rates because of an inference the algorithm made that people with similar names had trouble with their credit loans.²⁵ These inferences can become so complex that the people who created them may not know why inferences are being made. For example, algorithms by some financial companies started looking at whether a user had an Apple computer or a PC to determine the ability to make pay back a loan.²⁶ Without any intention to discriminate and without having the algorithm look at any protected characteristic, the algorithm inadvertently ended up discriminating.

Proxy variables are perhaps more easily understood. Proxy variables are categories of data that are not explicitly based on protected classes but are so intertwined with them in our society that they are functionally equivalent.²⁷ For example, the company Corelogic created a tenant screening program for use by landlords.²⁸ While the program did not explicitly look at race, it did look at arrest records which, due to disparities in the criminal justice system, disproportionately impact Black and Hispanic communities.²⁹ Importantly, the system did not look at conviction records but only at arrest records.³⁰ As a result, the system disproportionately screened out Black and Hispanic applicants without directly intending to do so. Programs like this are used by approximately 90 percent of landlords screening potential tenants.³¹

Another example of the use of proxy variables is with algorithms that help determine appropriate bail for people awaiting trial. The program does not look explicitly at race but does look at: length of employment, total years of education, prior criminal history, history of substance use, criminal activity in the neighborhood the individual lives in, the criminality of the individual's family members or of those within their social networks, their educational attainment, and their immigration status.³² None of these factors are explicitly determined by race but all of them are factors that have been heavily influenced by racial disparities throughout American history.

C. Fairness vs. Accuracy as a Source of Algorithmic Bias

Another way that bias can be built into algorithms has to do with the inherent tension between accuracy and fairness as well as which definition of fairness is used. Accuracy and fairness are separate goals that are not always perfectly aligned.³³ Maximizing accuracy can therefore lead to unfair or discriminatory predictions while maximizing fairness typically leads to less accuracy.³⁴ That tension between the two means some degree of predictive bias will be introduced into the system for the sake of accuracy.

²⁵ *Id.* at 9.

²⁶ *Id.*

²⁷ Venkatachalam Testimony, *CT September Briefing Transcript*, p. 107.

²⁸ *Id.*

²⁹ *Id.* at 108.

³⁰ *Id.*

³¹ Smirniotopoulos, *CT November Briefing Transcript*, p. 3.

³² Dr. Vincent Sutherland, Assistant Professor of Clinical Law, Director of the Criminal Defense and Reentry Clinic, and Co-Faculty Director at the Center of Race Inequality at New York University School of Law, Briefing Before the Connecticut Advisory Committee to the U.S. Comm'n on Civil Rights, September 29, 2022 [hereinafter referred to as the Second September Briefing], transcript pg. 4.

³³ Thomas Testimony, *CT September Briefing Transcript*, p. 68.

³⁴ *Id.*

This leads to the important definitional question of what programming for “fairness” means. No algorithm will predict an outcome with perfect accuracy and therefore every prediction will, to some degree, predict people with certain common characteristics act in one way and those without them act in another. For example, an algorithm predicting the grade point averages of students based on their entrance exam was found to routinely predict higher GPAs for male students as compared to female students.³⁵ Correcting for that prediction means introducing a bias into the algorithm to give male students lower scores and female students higher scores even though the prediction for any one student may not have been inaccurate in the first place.

This raises the question of what it means to be “fair” – is it fair to have the algorithm make a biased prediction based on real data or is it fairer to alter the prediction to reflect societal values even if the results are less consistent with the actual data? Once that question is resolved, then there are considerations as to short term fairness versus long term fairness and what to do when those do not align. These are difficult questions that need to be answered to account for bias in algorithms.

Accuracy and fairness create another issue for algorithms in potentially introducing an additional source of bias called “differential validity.” This is when a model may be less accurate at assessing people from different groups.³⁶ By making accurate predictions about one group, that group can be favored in decision making. For example, an algorithm might recommend an equal number of White and Black applicants for a position. In making those recommendations, it may be more accurate when assessing the credentials of the White applicants. When those candidates are later interviewed, the White applicants would be found to be more suitable for the position than Black candidates due to the algorithm’s failure to accurately match Black candidates for the role. As a result, a White candidate will be selected even though there may be a more suitable Black candidate that the algorithm failed to accurately identify.³⁷ In prioritizing fairness but not accounting for a corresponding potential drop in accuracy, the outcome of the algorithm is to perpetuate bias in hiring.

D. Use of Algorithms in Government

While an algorithm’s potential to perpetuate discrimination is troubling in the private sector, it is all the more concerning when used by the government. A few examples will show the danger. Police departments throughout the nation are turning to artificial intelligence to predict areas of high crime so they can send more officers to those locations to prevent criminal behavior.³⁸ However, the algorithm is trained using high-crime data from the recent past, and so it creates a feedback loop: the algorithm will predict an area to have more crime so more officers go to the area which leads to more arrests being made there as opposed to other locations.³⁹ The new arrest data is fed back into the algorithm which then uses that data to predict even higher crime rates in that area and so sends more officers there. This feedback loop could thus create more problems of over-policing and put more people of color into the criminal justice system.

³⁵ *Id.* at 73.

³⁶ Smirniotopoulos, *CT November Briefing Transcript*, p. 3.

³⁷ *Id.*

³⁸ Anjana Samant, ACLU’s Women’s Rights Project, *CT September Briefing Transcript*, p. 128.

³⁹ *Id.*

This is not a hypothetical issue. In New York City, police officers stopped and frisked over five million people over the past decade.⁴⁰ During that time, Black and Latino people were nine times more likely to be stopped than their White counterparts.⁴¹ As a result, predictive policing algorithms trained on data from that jurisdiction will over predict criminality in neighborhoods with predominantly Black and Latino residents. Up to one third of U.S. cities are either using or considering the use of predictive policing tools, including Hartford, Connecticut.⁴²

Once defendants are in the criminal legal system, algorithms again step in. Judges have started to turn to algorithms to make pre-trial detention and sentencing decisions based on assumptions the program makes for how likely a defendant is to commit another crime in the future.⁴³ One such program, the Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS, has been adopted by several jurisdictions.⁴⁴ When the COMPAS program was tested against real world data, it was found to label Black defendants as more likely to reoffend than they actually did and to label White defendants as lower risk though they would later commit another crime.⁴⁵ Not only did an algorithm increase the chance of a person getting put into the criminal justice system, but they made it more difficult for that individual along the way.

It is important to keep in mind that an algorithm's prediction is not about what *will* happen, but a prediction of probabilities that something *may* happen. This point is demonstrated by looking at another example of algorithms in use by the government: child welfare. Child welfare agencies have started to use algorithms to predict the likelihood that a child may suffer death or serious injury.⁴⁶ These tools work by examining various proxy variables to come up with a prediction for harm. These variables include the family's use of public benefits, involvement in the foster care or criminal justice systems, past housing instability, and neighborhood characteristics like arrest rates and proximity to foreclosed properties.⁴⁷ Like other proxy variables, these data points are heavily influenced by past and present racism, leading to inaccurate predictions.

More troubling, however, is that these predictions impact the ability of child welfare agencies to properly identify instances of real danger and allocate their limited resources appropriately. For example, one of these algorithms in use in Chicago flagged over 4,100 child welfare cases as having a high risk of death or injury.⁴⁸ As a result, caseworkers were overwhelmed with high-risk cases and were unable to investigate all of them. In Los Angeles, 95 percent of cases flagged by a similar system did not result in either severe injury or death, showing just how burdensome reliance on these programs can be.⁴⁹

⁴⁰ Dr. Vincent Sutherland, Assistant Professor of Clinical Law, Director of the Criminal Defense and Reentry Clinic, and Co-Faculty Director at the Center of Race Inequality at New York University School of Law, Briefing Before the Connecticut Advisory Committee to the U.S. Comm'n on Civil Rights, September 29, 2022 [hereinafter referred to as the Second September Briefing], transcript pg. 3.

⁴¹ *Id.*

⁴² *Id.* at 4.

⁴³ Venkatachalam Testimony, *CT September Briefing Transcript*, p. 103.

⁴⁴ *Id.*

⁴⁵ *Id.*

⁴⁶ Samant Testimony, *CT September Briefing Transcript*, p. 118.

⁴⁷ *Id.* at 122.

⁴⁸ David DesRoches, Quinnipiac University, *CT September Briefing Transcript*, p. 164.

⁴⁹ Samant Testimony, *CT September Briefing Transcript*, p. 119.

There are, however, uses of algorithms in government that shift the focus from the people being impacted by decisions to the decision makers. One jurisdiction examined the decisions of judges in setting bail or determining whether individuals should be released.⁵⁰ The AI determined that judges were more likely to set bail or hold people in custody more often in the hour prior to lunch than they would in the early morning or the afternoon.⁵¹ Through the use of an algorithm, a pattern of behavior the judges were not aware of was identified and could be corrected.

In another jurisdiction, the charging decisions of prosecutors were examined and it was found that one group overcharged individuals facing drug charges with additional paraphernalia possession charges.⁵² These individuals were not conscious of the patterns of their decisions, but by turning the lens of algorithms on decision makers, meaningful reforms could be made to eliminate discriminatory outcomes.

E. Transparency Issues

One of the biggest issues involved is the use of algorithms is transparency. This is particularly important when it comes to the government's use of AI as oversight laws like the Freedom of Information Act (FOIA) require transparency. FOIA starts with the premise that all government records are public records open to the public and can only be withheld if a statutory exception applies.⁵³

As an initial point, most of the general public are unaware of an algorithm's use in governmental decision-making at all.⁵⁴ With such little public understanding of whether and when an algorithm is being used, it is difficult for the public to take any meaningful steps to hold the government accountable.

The lack of awareness is compounded by the fact that many government employees are also unaware of whether and when they're using AI.⁵⁵ Part of this has to do with the ambiguity over what the terms "algorithm," "AI," and "machine learning program" actually mean and whether they are subject to public disclosure. For example, the European Union is currently debating laws about artificial intelligence and has spent months trying to define what exactly constitutes "artificial intelligence" and what does not.⁵⁶

Even when agencies are aware of their algorithms and can define them, there is then the further question of whether they are disclosable to the public. Algorithms are not specifically named in most Freedom of Information legislation, leading agencies to believe they are not covered by those

⁵⁰ Sutherland, 7

⁵¹ *Id.*

⁵² *Id.* at 11.

⁵³ Colleen Murphy, Executive Director of the Connecticut Freedom of Information Commission, Briefing Before the Connecticut Advisory Committee to the U.S. Comm'n on Civil Rights, December 19, 2022 [hereinafter referred to as the December Briefing], transcript pg. 7.

⁵⁴ Turner Testimony, *CT September Briefing Transcript*, p. 17.

⁵⁵ DesRoches Testimony, *CT September Briefing Transcript*, p. 167.

⁵⁶ Venkatasubramanian Testimony, *CT September Briefing Transcript*, p. 200.

laws.⁵⁷ Some agencies have said they believe algorithms are covered by FOIA requests but do not believe they have any responsive data as the algorithms are run by private corporations.⁵⁸

Other agencies believe algorithms are covered by FOIA requests and they have responsive documentation in their control but then believe algorithms are exempt as protected “trade secrets” of the private companies that created them.⁵⁹ Trade secrets are defined by Connecticut statute as:

Information, including formulas, patterns, compilations, programs, devices, methods, techniques, processes, drawings, cost data, customer lists, film or television scripts or detailed production budgets that (i) derive independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from their disclosure or use, and (ii) are the subject of efforts that are reasonable under the circumstances to maintain secrecy; and... Commercial or financial information given in confidence, not required by statute.⁶⁰

Citing to this statute, Connecticut state agencies have claimed that providing the source code or supporting documentation for an algorithm would violate the developer’s rights to keep trade secret information from the public.⁶¹

While this is a developing area of law, this argument has found support in court. For example, the trade secrets exemption was cited in a challenge to the use of algorithms in sentencing decisions when defendants were asking to see what factors the program relied on to help determine an appropriate sentence.⁶² In that situation, the court found that not only was the algorithm protected as a trade secret, but that there was no due process violation to keep it a secret since the programs were only used in an advisory capacity; that it was the judge making the final decision and therefore the algorithm advising the judge could remain a secret.⁶³

F. Current Regulatory Approaches

There are several regulatory initiatives focused on algorithms that are already in place or drafted and under review. Independent organizations with interests in data privacy and policy, such as the Electronic Privacy Information Center (EPIC), are investigating how federal agencies and actors assess governmental compliance with Title VI, which prohibits federal funding of programs that discriminate on the ground of “race, color, or national origin.”⁶⁴ The National Institute of Standards and technology (NIST) is creating a framework for how to mitigate bias in

⁵⁷ DesRoches Testimony, *CT September Briefing Transcript*, p. 167.

⁵⁸ Murphy Testimony, *CT December Briefing Transcript*, p. 5.

⁵⁹ DesRoches Testimony, *CT September Briefing Transcript*, p. 167.

⁶⁰ Connecticut General Statutes Section 1-210(b)(5).

⁶¹ Murphy Testimony, *CT December Briefing Transcript*, p. 5.

⁶² Venkatachalam Testimony, *CT September Briefing Transcript*, p. 104.

⁶³ *Id.*

⁶⁴ Electronic Privacy Information Center, <https://www.epic.org> (last visited Mar. 25, 2023).

sociotechnical systems.⁶⁵ The Federal General Accounting Office has published its own framework focusing on the interdependent roles of governance, data management, system evaluation, and monitoring.⁶⁶ The EEOC and HUD have both put out guidance for how AI can perpetuate discrimination.⁶⁷ Various states such as New York and Utah have taskforces created to build guidelines around the government’s procurement and use of AI applications.⁶⁸ Illinois passed the Biometric Information Privacy Act in 2008 to limit the ways that AI can collect and use biometric data such as facial recognition.⁶⁹

One of the most prominent proposals relating to algorithms is President Biden’s “The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People,” which was released in October of 2022.⁷⁰ This plan focuses on five areas: creating safe and effective systems; establishing algorithmic discrimination protections; data privacy; notice and explanation; and ensuring there are human alternatives, consideration, and fallback.⁷¹ Each of these five areas is an aspect common to the other regulatory approaches cited above.

“Safe and effective systems” means that algorithms should be proactively designed to prevent foreseeable yet unintended harms created through the use of the system.⁷² A key aspect of how this can be achieved is by bringing in stakeholders that will be impacted by the algorithm during its design phase. Too often, these tools are created without this input, creating blind spots in their design that are hard to remediate after the fact.⁷³

“Algorithmic discrimination” means putting in place or clarifying existing civil rights protections to ensure that AI does not violate the law by treating individuals differently based on a protected class or by having a disparate impact on protected classes either directly or indirectly through proxy variables.⁷⁴ These protections include putting in place equity assessments and audits both before the system is deployed to the public and after on a regular basis.⁷⁵ The regular, post-deployment audits testing for disparate impacts are particularly important as machine learning will lead to changes in how a system will operate without human intervention.⁷⁶

“Data privacy” is the principle that individuals “should be protected from abusive data practices via built-in protections,” giving individuals agency in the use of that data.⁷⁷ This includes restricting use of personal data beyond reasonable expectations. This means that personal data collected by one agency for one purpose should not be used by other agencies or for other purposes

⁶⁵ *Id.* at p.177.

⁶⁶ *Id.*

⁶⁷ *Id.* at p.178.

⁶⁸ *Id.* at p.180.

⁶⁹ *Id.*

⁷⁰ White House Office of Science and Technology Policy, “Blueprint for an AI Bill of Rights,” available online at <https://www.whitehouse.gov/ostp/ai-bill-of-rights/#discrimination>.

⁷¹ *Id.*

⁷² *Id.*

⁷³ Samant Testimony, *CT September Briefing Transcript*, p. 130.

⁷⁴ Turner Testimony, *CT September Briefing Transcript*, p. 40.

⁷⁵ *Id.* at p. 15.

⁷⁶ DesRoches Testimony, *CT September Briefing Transcript*, p. 166.

⁷⁷ “Blueprint for an AI Bill of Rights,” *supra* note 70.

without an individual's knowledge or consent.⁷⁸ For example, a local public housing authority with facial recognition software at their entrance to track who can enter should not be sharing that data with local law enforcement for scanning by its facial recognition software as that would be an unexpected use of the data.

“Notice and Explanation” refers to regulations that inform people that an automated system is being used in a given decision making process.⁷⁹ These requirements force developers and organizations that use algorithms to provide clear documentation on the role automation plays in particular decisions as well as what factors and data sets are used by the algorithm in creating its output.⁸⁰

“Human alternatives, consideration, and fallback” is an area of regulation allowing individuals to opt out of having an algorithm impact decisions relating to them and, where feasible, for there to be human alternatives to make those decisions without the use of an algorithm.⁸¹ This is particularly necessary as all systems will eventually fail. For example, during the COVID-19 Pandemic, individuals in many states were required to obtain benefits using third-party identity verification systems.⁸² The system relied in part on facial recognition software that had difficulty scanning darker-skinned people which led to those who were affected having to wait for a human to triage their case.⁸³ This had not been prepared for by the vendor, leading to wait times of up to ten hours.⁸⁴ Regulations requiring these alternatives be put in place will ensure there is a viable alternative for those who cannot or choose not to use algorithm-based systems.

III. Summary of the Briefings

The first panel, held on September 8, 2022, included subject matter experts in computer science and government transparency. Speakers included Dr. Nicol Turner Lee, Senior Fellow of Governance Studies and Director of the Center for Technology Innovation at the Brookings Institute; Dr. Philip Thomas, Co-Director of Autonomous Learning Lab Center for Data Science at the Manning College of Information & Computer Sciences at UMass; Sruthi Venkatachalam, representative of the Yale Law School Media Freedom and Information Access Clinic; Attorney Anjana Samant, Senior Staff Attorney with ACLU's Women's Rights Project; Professor David DesRoches, professor and Director of Community Programs at Quinnipiac University; and Dr. Suresh Venkatasubramanian, Professor of Data Science and Computer Science at Brown University. The speakers focused on three areas: (1) what algorithms are and how they can be used to create and perpetuate discrimination, (2) transparency issues with algorithms used by the government, and (3) regulatory approaches that have been or are being created.

The second panel, held on September 29, 2022, featured Dr. Vincent Sutherland, Assistant Professor of Clinical Law, Director of Criminal Defense and Reentry Clinic, and Co-Faculty

⁷⁸ Samant Testimony, *CT September Briefing Transcript*, p. 132.

⁷⁹ “Blueprint for an AI Bill of Rights,” *supra* note 70.

⁸⁰ *Id.*

⁸¹ *Id.*

⁸² Venkatasubramanian Testimony, *CT September Briefing Transcript*, p.183.

⁸³ *Id.*

⁸⁴ *Id.*

Director on the Center of Race Inequality and the Law at New York University School of Law. The briefing's focus was the importance of using a racial justice lens to inform the design, development, implementation, and oversight of algorithms before, during, and after their deployment.

The third briefing was held on November 7, 2022, and featured NAACP's Legal Defense and Educational Fund's Senior Policy Counsel Amalea Smirniotopoulos. This briefing provided additional insight into how algorithms are currently being utilized in both the private and public sectors as well as recommendations for regulatory approaches.

The fourth and final briefing was held on December 19, 2022, and featured Colleen Murphy, Executive Director of the Connecticut Freedom of Information Commission. This briefing centered on transparency issues surrounding government use of algorithms. In particular, the briefing highlighted how algorithms under current Freedom of Information laws are treated and discussed some issues that come with adapting these laws to complex computer programs.

The briefings each concluded with questions and comments between Advisory Committee members and panelists, illustrating the need for a comprehensive approach prioritizing transparency and equity to be implemented by the state for its agencies and municipalities.

IV. Findings and Recommendations

The Committee's investigation revealed that there is growing recognition of the civil rights implications of governmental use of algorithms. At the end of 2022, the White House noted that "among the great challenges posed to democracy today is the use of technology, data, and automated systems in ways that threaten the rights of the American public."⁸⁵ In issuing its "Blueprint for an AI Bill of Rights" it noted:

Too often, these tools are used to limit our opportunities and prevent our access to critical resources or services. These problems are well documented. In America and around the world, systems supposed to help with patient care have proven unsafe, ineffective, or biased. Algorithms used in hiring and credit decisions have been found to reflect and reproduce existing unwanted inequities or embed new harmful bias and discrimination. Unchecked social media data collection has been used to threaten people's opportunities, undermine their privacy, or pervasively track their activity—often without their knowledge or consent.

The Blueprint for an AI Bill of Rights is a good start towards addressing the civil rights concerns of the government's use of algorithms. The Legislature should incorporate these principles in an overarching Connecticut AI Bill of Rights that includes guardrails for the development, use and monitoring of algorithms to minimize the potential for bias and disparate impact on protected classes as well as a process to review and address the use of algorithms.

⁸⁵ "Blueprint for an AI Bill of Rights," *supra* note 70.

The following findings and recommendations are directed to the state legislature and state agencies.

1. **Finding:** Algorithms can create or perpetuate discrimination through reliance on data sets that are historically biased, consideration of proxy variables for race, differential accuracy rates between groups, and more. Sometimes this bias is intentional, but more often it is a result of unintentional bias on the part of programmers, historical biases in the data, or the unintentional consequence of giving the program specific goals that fail to account for disparate impacts.

Recommendations:

- a. Include people from the protected classes most adversely impacted by the use of automated decision-making in the monitoring and assessment of algorithms used by the government.
 - b. Implement a public education campaign designed to bring awareness of the existence, limitations, and dangers of automated decision-making that is aimed both at the public and state personnel.
2. **Finding:** Algorithms are already in use by many state agencies throughout Connecticut to fill a variety of functions including in hiring and decision-making processes. Despite their use and utility, there is little to no statewide oversight of algorithms during procurement, implementation, or monitoring.

Recommendations:

- a. Validate system designs prior to implementation to minimize built-in bias.
 - b. Algorithms used by government agencies should include an internal audit and ongoing evaluation to ensure that algorithmic outcomes are not discriminatory.
3. **Finding:** There is little transparency for the public to understand when automated decision-making is being used. Even when the public is aware of the use of an algorithm in a particular decision process, there is little transparency about what data the program is relying on in making its decision, how the algorithm functions, whether the data relied on is being used by the algorithm's maker for other purposes, and more.

Recommendations:

- a. Require the government to create and maintain a publicly available dashboard that lists which agencies are using automated decision-making.
- b. Implement frequent independent audits of algorithms with publicly available reporting in the form of algorithmic impact assessments, including disparity testing results and information about mitigation efforts.
- c. Provide an opt-out option and an appeal process that includes human decision-makers for people who believe that they have been negatively impacted by individualized automated decision-making algorithms used by the government or provide a public explanation for why such an opt-out option is not possible.

d. Prohibit sharing or sale of data and provide meaningful consequences for transferring personal information without permission.

4. **Finding:** Current government transparency laws such as the Freedom of Information Act (FOIA) are often stymied when it comes to algorithms. In part this is because state agencies are not aware of what programs they use and whether they rely on artificial intelligence to work. In other part, exceptions to FOIA such as the trade secrets exception prevent disclosure of information about algorithms that is necessary to evaluate their public impact.

Recommendations:

- a. Revise Connecticut Freedom of Information laws to explicitly provide the public access to data regarding state agencies' use of algorithms, including any reviews or analysis of an algorithm's outcome.
- b. Require disclosure of any publicly available data sources used by algorithms relied upon by government agencies.