NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY
VOLUME 25, ISSUE 4: MAY 2024

FAIR ADMISSIONS, FAIR DECISIONS, AND FAIR OUTCOMES: AN ANALYSIS OF ALGORITHMIC BIAS IN EDUCATION, EMPLOYMENT, HEALTHCARE, AND HOUSING

McLean Waters*

Artificial Intelligence ("AI") has surged in popularity over recent years, especially in its accessibility to the public. Increased productivity and the automation of simple tasks has clearly displayed the benefits of AI in everyday life. However, AI has several drawbacks. Since AI algorithms are written by humans, they are influenced by the qualities that humans prioritize, either intentionally or unintentionally. This Article argues that AI shares similarities with other systems built by humans, as it is susceptible to the implantation of bias and heavily influenced by the system's creators. Furthermore, this Article considers examples of this discrimination in systems that utilize AI. Finally, this Article analyzes existing legislation to provide a regulatory framework that prevents the implantation of bias into algorithms, which could lead to widespread discrimination.

TABLE OF CONTENTS

* J.D. Candidate, University of North Carolina School of Law, 2025. The Author would like to thank the staff of NC JOLT, with particular thanks to Austin Kamer, Julia Vaughan-Jones, Hayfa Ayoubi, and McGee Roman for their invaluable feedback. He would also like to thank his incredible wife for her unwavering support.

I. INTRODUCTION

Imagine during the COVID-19 pandemic, a young professional was furloughed from her job. Eventually, her company reinstated the role and asked the candidate to re-apply for the same role for which she was well-qualified. However, the candidate was required to conduct a video interview using an artificial intelligence ("AI") platform, which scores an individual's responses to questions as well as their body language during the interview. While the candidate had all the skills necessary for this role, the AI system scored her body language so poorly that she did not receive the job. The candidate then came to find out that this system was trained on the faces and

voices of white male applicants, leading to consistently lower scores for females and people of color.[1]

For Anthea Mairoudhiou, a make-up artist, this hypothetical scenario was her reality. She was not rehired after HireVue, an AI-screening program, scored her body language poorly.[2] Mairoudhiou's story is not an isolated incident, but one of many stories of individuals experiencing discrimination from AI systems.[3] In the midst of massive adoption of AI in recent years, AI discrimination is increasing at an alarming rate, despite AI leading to improvements and advances in areas such as "speech recognition, natural language processing, translation, . . . computer programming, and predictive analytics."[4] As developments in AI have accelerated exponentially, stakeholders have expressed concerns over how this type of technology will be used.[5] Previous technological advances, such as email, word-processing, and better electronic databases, have all improved efficiency in the workplace by assisting with simple and routine tasks.[6] AI, however, has the capability to replace non-routine cognitive tasks, such as information categorization or sorting information into flexible

---

[1] *See* Ifeoma Ajunwa, *An Auditing Imperative for Automated Hiring Systems*, 34 HARV. J.L. & TECH. 1, 7 (2021) (arguing that auditing AI systems is not only beneficial, but necessary as the technology develops).

[2] *See* Charlotte Lytton, *AI Hiring Tools May Be Filtering out the Best Job Applicants*, BBC (Feb. 16, 2024), https://www.bbc.com/worklife/article/20240214-ai-recruiting-hiring-software-bias-discrimination [https://perma.cc/F4YA-G3B5] (detailing how HireVue has been the focus of many complaints and lawsuits regarding AI discrimination).

[3] *See* Rachel Goodman, *Why Amazon's Automated Hiring Tool Discriminated Against Women*, ACLU (Oct. 12, 2018), https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against [https://perma.cc/3RUC-4BPX] ("In 2014, a team of engineers at Amazon began working on a project to automate hiring at their company. . . . But . . . the project was canned just a year later, when it became clear that the tool systematically discriminated against women applying for technical jobs, such as software engineer positions.").

[4] Alexandre Georgieff & Raphaela Hyee, *Artificial Intelligence and Employment: New Cross-Country Evidence*, 5 FRONTIERS ARTIFICIAL INTEL. 1, 2 (2022).

[5] *See id.*

[6] *See id.*

classifications, using "highly sophisticated algorithmic techniques to find patterns in data and make predictions about the future."[7] AI's technological advances can be both beneficial—by assisting workers with exhausting tasks—and detrimental—by eliminating jobs or "degrading work quality."[8]

AI often operates in these gray areas, where using the technology has clear benefits and harms. To illustrate this idea, consider the world of employment and hiring. A recent study "found that 83% of human resources leaders rely in some form on technology in employment decision-making."[9] While AI tools are beneficial, they need regulations and limitations. For example, the use of AI in hiring can speed up the process and potentially eliminate bias if the algorithms are correctly crafted.[10] On the other hand, AI tools can replace human jobs and, more importantly, the human approach to certain roles. Moreover, AI is only as great as the algorithm used to create it, meaning creators could—either intentionally or unintentionally—introduce their own bias and subjectivity into the algorithm.[11] To prevent hiring bias from appearing in AI, some legislatures are requiring companies and organizations to place limitations on how AI is used during the hiring process and on the underlying algorithms of the technology itself.[12]

Even without ill-intent, bias and discrimination often seeps into AI systems in several ways. Consider what would happen if a

---

[7] *Id.* at 2.

[8] Eva Selenko, et al., Artificial Intelligence and the Future of Work: A Functional-Identity Perspective, 31 CURRENT DIRECTIONS PSYCH. SCI., 272, 272 (2022).

[9] Keith E. Sonderling, *Do Robots Care About Your Civil Rights*?, CHI. TRIB., https://digitaledition.chicagotribune.com/infinity/article_share.aspx?guid=285d3467-3dbe-49b1-810e-014aefee1a3e [https://perma.cc/YP5H-JQHY] (last visited Feb. 4, 2024).

[10] Keith E. Sonderling et al., The Promise and The Peril: Artificial Intelligence and Employment Discrimination, 77 U. MIA L. REV. 1, 4 (2022).

[11] *Id.* at 5.

[12] Jonathan Kestenbaum, *NYC's New AI Bias Law Broadly Impacts Hiring and Requires Audits*, BLOOMBERG L., (July 5, 2023, 5:00 AM), https://news.bloomberglaw.com/us-law-week/nycs-new-ai-bias-law-broadly-impacts-hiring-and-requires-audits [https://perma.cc/Z7LD-2W5B].

company were to input the resumes and applications of all its highest performing employees into an algorithm that was designed to find applicants who matched those profiles. If the company had historically hired mostly men instead of women, then the algorithm could take this data and create a preference for male applicants, when the reality was that women had just been given less opportunities. This was the case for Amazon, which developed an AI hiring algorithm that downgraded female applicants because most of the data used to teach the algorithm came from male applicants.[13] In addition, if all the company's highest performing employees were older, the algorithm may disfavor younger applicants, disregarding the time it took for those high performing applicants to achieve their status. Reliance upon historical data sets for machine learning exemplifies some of the potential issues surrounding AI today.[14]

While the world of employment illustrates some of the pros and cons of AI, it is only one of numerous industries that experience the benefits and drawbacks of using this technology. As AI usage promulgates across the country, the potential for AI to be used in a way that leads to discrimination also grows. As previously mentioned, AI algorithms force creators to make choices about which data sets are preferred and not preferred to achieve desired results. The results of these choices vary across industry and can appear to provide the best candidates for a job, determine the best applicants for housing, or analyze which patients Medicaid should cover. Regardless of the industry, AI can incorporate individual or collective bias into its algorithm, which can lead to discrimination.

This Article demonstrates how AI incorporates unacceptable discrimination. Examples of human and technological discrimination come from analogies to the recent Supreme Court decision in *Students for Fair Admissions v. President Fellows of*

---

[13] *See* Goodman, *supra* note 3 ("[T]he project was canned just a year later, when it became clear that the tool systematically discriminated against women applying for technical jobs, such as software engineer positions.").

[14] *See* Ifeoma Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, 41 CARDOZO L. REV. 1671, 1687 (2020) (arguing that automation does not lead to anti-bias results, but that it instead reinforces existing bias).

Harvard College and through current examples of companies using AI to discriminate or AI use resulting in discrimination. Based on the overwhelming potential for discrimination in AI, this Article sets forth regulations on algorithm creation and continual audits to ensure that AI algorithms and systems do not repeat discrimination mistakes of the past.

This Article proceeds in six parts. Part II explores AI generally and identifies potential ways that discrimination creeps into AI algorithms. Part III exemplifies how simple inputs in decision-making processes can lead to large scale discrimination without the use of technology through a review of the recent Supreme Court decision in *Students for Fair Admissions v. President & Fellows of Harvard College*. Part IV considers the existing federal anti-discrimination framework and discusses the ways that those laws interact with one another. Part V surveys modern-day examples of AI usage resulting in discrimination across numerous different industries. Part VI explores current legislation at both the state and federal level which could provide guidance on which solutions may be best suited for regulating AI. Finally, Part VII of this Article examines the current state of regulation surrounding AI and recommends legislation that should be introduced to properly regulate this emerging area of law.

II. ARTIFICIAL INTELLIGENCE: OPERATING IN GRAY AREAS

While AI has become a common buzzword across numerous industries, it is important to understand how this technology functions in order to comprehend the ways that discrimination can occur. Put simply, algorithms are "a set of instructions that a computer uses to solve a problem."[15] Another important term of art in this field is "machine learning," which has been defined as "using repetition and experience as how humans seem to learn" or "[u]sing software whose operations mimic these methods, employing artificial intelligence techniques to enhance the ability of a machine

[15] Leigh Harvis-Nazzario, *It's Not the Algorithms, It's the People: Preventing Bias in Automated Hiring Tools Starts with Humans*, 49 RUTGERS COMPUT. & TECH. L.J. 138, 140 (2022).

to improve its own performance."[16] In other words, algorithms and machine learning replicate the decision-making process of humans, without humans "having to take in all the information and time needed and sort through it on [their] own."[17]

In order to replicate this human learning, algorithms, at their most basic level, are a collection of if-then statements designed to "very quickly to give you an outcome: [i]f you do A then you get B[,] [i]f you do C instead, then you get D."[18] In addition, machine learning allows algorithms to " 'learn[]' your preferences and take[] them into account during your next use of the if-then statements."[19] When using the internet, AI can "start to learn your behavior (machine learning), so depending on the pages you follow, the ads you click, and the searches you perform, the algorithms will make recommendations based on your preferences."[20]

For a computer to make recommendations or predictions, an engineer must train the underlying algorithms.[21] For example, to train an AI algorithm to recognize an image of a candle, you "feed" that algorithm numerous images of what a candle is so that the algorithm can learn the color, shape, size, and other dimensions of a candle. In the employment sphere, employers could train hiring algorithms to find ideal candidates by "feeding" the algorithm large amounts of data of what an ideal candidate is, such as their education, experience, or credentials.[22]

---

[16] *Machine Learning*, THE LAW DICTIONARY, https://thelawdictionary.org/machine-learning/ [https://perma.cc/H669-E9V6] (last visited Feb. 4, 2024).

[17] Carolyn Lyden, *How Do Algorithms Work? A Basic Primer for Non-Marketers*, SEARCH ENGINE J. (Aug. 31, 2020), https://www.searchenginejournal.com/how-do-algorithmswork/378978/#close [https://perma.cc/A4SK-48WJ].

[18] *Id.*

[19] *Id.*

[20] *See* Harvis-Nazzario, *supra* note 15, at 140.

[21] *See id.*

[22] *See* Miranda Bogen, *All the Ways Hiring Algorithms Can Introduce Bias*, HARV. BUS. REV. (May 6, 2019), https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias [https://perma.cc/4LWC-ADW8] (providing numerous hypothetical and practical examples of algorithms introducing bias the hiring process).

Because algorithms can consume and parse huge amounts of data, they have become integral parts of our lives.[23] In the span of an average day, "algorithms determine the optimal way to produce and ship goods, the prices we pay for those goods, the money we can borrow, the people who teach our children, and the books and articles we read."[24] The ease with which some of these tasks can now be performed clearly demonstrates many of the advantages that AI can provide in our lives. However, the downsides of AI may not always be as clear.[25]

At the simplest level, an AI system engages in confirmation bias; after learning an individual's preferences, it can shape the information a user sees to continue showing an individual only the information the AI system thinks the individual wants to see.[26] While this type of learning may be beneficial for sites such as YouTube and Netflix, the process of showing users only what they want to know can become harmful when it pertains to important information that people need to know. One of the largest examples of confirmation bias occurs with politics, where individuals begin to see only information that "affirm[s] their existing interests and beliefs," which can prevent individuals from listening to the other side and engaging in productive dialogue.[27]

Another problem occurs when individuals assume that AI is "a set of abstract principles manifesting rational objectives," and therefore, is always fully neutral.[28] However, AI is often far from neutral because AI is designed to make predictions in certain ways and have preferences for certain characteristics. The key issue is that "[w]hile algorithmic decision-making may initially seem more reliable because it appears free from the irrational biases of human judgment and prejudice, algorithmic models are also the product of

---

[23] *See* Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 56 (2019) (asserting that regulation for AI should come from private industry instead of public governance).

[24] *Id*.

[25] *See id.* at 56–58.

[26] *See id.* at 57.

[27] *Id.*

[28] *Id.* at 58.

their fallible creators, who may miss evidence of systemic bias or structural discrimination in data or may simply make mistakes."[29] Furthermore, while errors may be unintentional, they "risk reifying past prejudices, thereby reproducing an image of an infinitely unjust world."[30] Because the stakes are so high, there is a clear need in America for greater regulations regarding AI and how it operates.

III. *STUDENTS FOR FAIR ADMISSIONS* AND ITS EXAMPLE OF DISCRIMINATION

Bias in decision-making has the potential to creep into systems of all shapes and sizes, including college admissions decisions. This Part provides an overview of the recent Supreme Court decision, *Students for Fair Admissions v. President & Fellows of Harvard College*[31] ("SFFA" or "*Students for Fair Admissions*") and summarizes how the Court came to find discrimination in that case. In addition, this Part draws comparisons between the discriminatory nature of certain college admissions policies and the potential of AI algorithms to operate similarly and even amplify those concerns.

A. Students for Fair Admissions Reasoning and Analysis

On June 29, 2023, the Supreme Court decided that the consideration of race in the process of college admissions violated the Equal Protection Clause of the Constitution and, therefore, race could no longer be used as a criteria in determining whether a student is admitted.[32] The petitioners, Students for Fair Admissions, brought claims against Harvard College ("Harvard") and the University of North Carolina ("UNC") under the Equal Protection Clause of the Fourteenth Amendment.[33] The petitioners argued that the admissions processes for Harvard and UNC violated the Equal Protection Clause by using race in a way that led to stereotyping and negative impacts on certain students. The Supreme Court held that

---

[29] *Id*. at 59.
[30] *Id.*
[31] Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 600 U.S. 181 (2023).
[32] *Id.* at 230.
[33] *Id.* at 193–94.

"[e]liminating racial discrimination means eliminating all of it."[34] Likewise, the Court held that "the Equal Protection Clause applies 'without regard to any differences of race, of color, or of nationality'—it is 'universal in [its] application.'"[35] Under this principle, the Supreme Court held that the admissions programs of these two universities had to comply with strict scrutiny, which requires that the universities "never use race as a stereotype or negative,"[36] and that there be a compelling interest that requires the use of race as a factor in discrimination.[37]

The Supreme Court ultimately concluded that the admissions programs for both Harvard and UNC violated the Fourteenth Amendment and were therefore unlawful.[38] Specifically, the Supreme Court stated that "[b]oth programs lack sufficiently focused and measurable objectives warranting the use of race, unavoidably employ race in a negative manner, involve racial stereotyping, and lack meaningful end points."[39]

In determining that Harvard and UNC violated the Equal Protection Clause, the Supreme Court analyzed the process by which both schools used race in their admissions decision-making processes.[40] At Harvard, race was used in two different ways. First, applicants were assigned numerical scores in six different categories, and application readers "can and do take an applicant's race into account" when scoring an applicant on the "overall" category.[41] Second, students who have been recommended for admission were considered as a large group to make sure there is not a "dramatic drop-off in minority admissions from the prior class."[42] In relying upon this historical data, Harvard attempted to maintain a

---

[34] *Id.* at 206 (quoting Yick Wo v. Hopkins, 118 U.S. 356, 369 (1886)).
[35] *Id.*
[36] *Id.* at 213.
[37] *See id*. at 214–15.
[38] *See id.* at 230.
[39] *Id.* ("We have never permitted admissions programs to work in that way, and we will not do so today.").
[40] *See id.* at 194.
[41] *Id.*
[42] *Id.*

certain racial percentage of its incoming student body year after year, which the court found unconstitutional.[43]

At UNC, race was used in a similar manner as application readers were required to consider race and ethnicity in their decision-making process.[44] While the Court does not discuss whether the results were intentional, the Court did note that "during the years at issue in this litigation, underrepresented minority students were more likely to score [highly] on their personal ratings than their white and Asian American peers."[45] In addition, for students of different races, the results of UNC's admission process could differ dramatically.[46] For example, in the second highest academic decile, 83% of black applicants were admitted while only 47% of Asian applicants were admitted, and in the third academic decile, 77% of black applicants were admitted while only 34% of Asian applicants were admitted.[47] The Supreme Court found that the use of race in the decision-making process, combined with results that seem to favor one race over another, was unconstitutional and discriminatory.

B. SFFA Application and Connection to Artificial Intelligence and Discrimination

Harvard and UNC used race as a plus for individual applicants of certain races or looked at the class as a whole to determine racial consistency over the years. The discrimination and bias concerns underlying race-conscious admissions are present in AI decision-making processes, considering that AI systems use race and other protected characteristics in their algorithms. Consequently, organizations that employ the use of AI systems for decision-making purposes should be wary.

The first potential issue for organizations is when the AI systems use protected characteristics as criterion for making decisions. As demonstrated in the *Students for Fair Admissions* decision, the use

---

[43] *See id.* at 231.
[44] *See id.* at 195–96.
[45] *Id.* at 196.
[46] *See id.* at 197 n.1.
[47] *Id.*

of protected characteristics, such as race, can be an issue that implicates constitutional protections. With AI systems, the ability to use protected characteristics in decision-making opens these systems and their users up to several constitutional violations, which is further discussed in Part V. While protected classes are not often explicitly used in AI algorithms, the *Students for Fair Admissions* decision should put organizations on notice of potential constitutional violations that could come from their system.

The second potential issue for businesses and other institutions comes from a reliance upon historical data sets. In this case, the universities relied upon historical data to keep racial demographics consistent year over year.[48] Similar to how the Court in *Students for Fair Admissions* determined that the university admissions programs' reliance upon historical data did not comply with strict scrutiny,[49] an organization's reliance upon historical data to train AI systems can also be problematic.[50] For example, an algorithm for textual analysis displayed gender bias when it began to associate certain words with different genders, such as connecting the title of "doctor" with men and "nurse" with women.[51] These algorithms were likely trained on historical data where men were more consistently the only ones with the opportunity to pursue medical degrees, thereby linking the term "doctor" with men. Using historical data that contains bias or a previous preference for certain groups over others is a more discrete way for AI systems to discriminate, but it can still lead to unfair outcomes.[52]

---

[48] *See id.* at 194–95.

[49] *See id.* at 213.

[50] *See* Ajunwa, *supra* note 14, at 1686–87 ("[B]ig data may, in fact, contribute to the segregating of individuals into groups because of its 'ability to make claims about how groups behave differently,' an action forbidden by anti-classificationist laws." (citation omitted)).

[51] *See* Madalina Busuioc, *Accountable Artificial Intelligence: Holding Algorithms to Account*, 81 PUB. ADMIN. REV. 825, 825 (2021) ("[N]atural language processing (NLP) algorithms for textual analysis can display recurrent gender biases . . . for instance, associating the word 'doctor' with 'father' and 'nurse' with 'mother.' ").

[52] *See* Zhisheng Chen, *Ethics and Discrimination in Artificial Intelligence-Enabled Recruitment Practices*, 10 HUMANITIES & SOC. SCI.

The third potential issue for organizations occurs when the results of AI systems demonstrate discriminatory preferences which unintentionally lead to disparities between protected classes. The *Students for Fair Admissions* case exemplified this, where the UNC admissions process led to black students being admitted at a disproportionality higher rate than Asian students.[53] AI operates similarly. Even when the algorithm does not use protected characteristics or rely on biased historical data, the outputs from the algorithm can still create discriminatory impacts, also known as a disparate impact.[54] As discussed later in Part V of this Article, the potential for discriminatory impacts has become a focus of regulatory bodies such as the Equal Employment Opportunity Commission ("EEOC").[55]

The *Students for Fair Admissions* decision demonstrates some of the many ways that the use of protected characteristics in decision-making processes can be unconstitutional, and some of these concerns, such as explicitly using protected characteristics and historically biased data or having disparate impacts, are present in AI systems. While the concerns surrounding college admissions programs can affect numerous people, the injection of AI into these processes nationwide broadens the potential for discrimination. Consider that "the impact of one biased human manager is constrained in comparison to the potential adverse reach of algorithms that could be used to exclude millions of job applicants from viewing a job advertisement or to sort thousands of resumes."[56] The potential for these decisions to impact thousands, if not

---

COMMC'NS 567, 568 (2023) ("If the underlying data is unfair, the resulting algorithms can perpetuate bias, incompleteness, or discrimination, creating potential for widespread inequality.").

[53] *See Students for Fair Admissions, Inc*., 600 U.S. at 197 n.1.

[54] *See* Jennifer G. Betts et al., *EEOC Issues New Guidance on Employer Use of AI and Disparate Impact Potential*, OGLETREE DEAKINS, (May 24, 2023), https://ogletree.com/insights-resources/blog-posts/eeoc-issues-new-guidance-on-employer-use-of-ai-and-disparate-impact-potential/ [https://perma.cc/8XJG-2Z5U].

[55] *See id.*

[56] Ajunwa, *supra* note 14, at 1679.

millions, of people demonstrates the need to develop regulations for AI.

IV. BREAKING DOWN THE EXISTING ANTI-DISCRIMINATION FRAMEWORK

Before providing recommendations on potential AI regulations for the United States ("U.S."), it is important to understand the existing state of anti-discrimination law. This Part provides a brief overview of federal anti-discrimination law and how that can relate to the regulation of AI. This Part summarizes different federal anti-discrimination provisions, including (A) provisions relating to employment decisions, like the Civil Rights Act Title VII, the Age Discrimination in Employment Act, and the Americans with Disabilities Act; (B) provisions relating to programs that receive federal funding, such as the Civil Rights Act Title VI and the Affordable Care Act; and (C) provisions relating to housing, such as the Fair Housing Act.

A. Employment Related Anti-Discrimination Laws

In the context of employment decisions, the use of AI "implicates Title VII of the Civil Rights Act of 1964 ("Title VII"), a federal law that protects employees and applicants against discrimination based on race, color, sex, national origin, and religion."[57] Specifically, Title VII states that:

> [i]t shall be an unlawful employment practice for an employer to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin.[58]

An AI algorithm could violate Title VII by screening out individuals who belong to protected classes without sufficient business reasons for why the AI algorithm screened the individuals out.[59] For example, imagine that an AI algorithm was written to prefer candidates within a certain commuting distance of a

[57] Sonderling et al., *supra* note 10, at 6.
[58] 42 U.S.C. § 2000e-2(a)(1).
[59] Sonderling et al., *supra* note 10, at 6.

company's office. If that preference created a "statistically significant disparate impact on certain races or those from a particular national origin, and the employer fails to demonstrate its geographic restriction is job-related and a business necessity, the employer will likely be liable under Title VII."[60] However, Title VII is not the only concern for employers.

Another relevant federal anti-discrimination law is the Age Discrimination in Employment Act of 1967 ("ADEA").[61] The ADEA states that "[i]t shall be unlawful for an employer . . . to fail or refuse to hire or to discharge any individual or otherwise discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's age."[62] The ADEA prohibits both intentional discrimination on the basis of age as well as practices that "although facially neutral with regard to age, have the effect of harming older workers more than younger workers."[63] Facially neutral practices that harm certain groups more than others are known as disparate impact practices, and the use of AI could lead to a violation of the ADEA.[64] For example, if an AI algorithm was written to screen out candidates who did not possess a certain technological skill that was only recently being taught in college, that algorithm could have a disparate impact on older populations who attended college before the development of this technology.

Along with Title VII and the ADEA, the Americans with Disabilities Act ("ADA") could also be implicated by the use of

---

[60] *Id.*

[61] Questions and Answers on EEOC Final Rule on Disparate Impact and "Reasonable Factors Other Than Age" Under the Age Discrimination in Employment Act of 1967, U.S. EQUAL EMP. OPPORTUNITY COMM'N, [hereinafter Questions and Answers on EEOC Final Rule] https://www.eeoc.gov/regulations/questions-and-answers-eeoc-final-rule-disparate-impact-and-reasonable-factors-other-age [https://perma.cc/L5DU-4NQ3] (last visited Sept. 20, 2023).

[62] Age Discrimination in Employment Act, 29 U.S.C. § 623(a)(1).

[63] Questions and Answers on EEOC Final Rule, supra note 61.

[64] *Id*.

AI.[65] The ADA requires that no employer "shall discriminate against a qualified individual on the basis of disability in regard to job application procedures, the hiring, advancement, or discharge of employees, employee compensation, job training, and other terms, conditions, and privileges of employment."[66] Examples of discrimination under the ADA could be classifying an application based on a disability in a way that adversely affects an applicant's opportunities, not making reasonable accommodations to the physical or mental limitations of an applicant, or using employment tests that tend to screen out applicants with disabilities.[67] Employers could potentially violate the ADA by using AI in a number of ways,[68] such as algorithms that prioritize candidates without disabilities over candidates with disabilities. An ADA violation could occur if an applicant could not lift twenty pounds due to a physical disability and was unable to gain accommodations.[69] When read in conjunction with one another, Title VII, the ADEA, and the ADA prevent workplace discrimination and limit how employers can use AI in hiring processes.

B. Anti-Discrimination for Programs Receiving Federal Funding

For programs that receive federal funding, such as schools and certain health programs, the federal government has created a number of provisions limiting the ways that organizations can utilize protected characteristics in decision-making.[70] For example, Title VI of the Civil Rights Act of 1964 states that "[n]o person in the United States shall, on the ground of race, color, or national origin,

---

[65] The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees, U.S. EQUAL EMP. OPPORTUNITY COMM'N (May 12, 2022) [hereinafter The American With Disabilities Act and the Use of Software], https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence. [https://perma.cc/Z2UK-YYQ3].

[66] Americans with Disabilities Act, 42 U.S.C. § 12112(a).

[67] § 12112(b).

[68] The Americans with Disabilities Act and the Use of Software, supra note 65.

[69] *Id.*

[70] *See generally* 42 U.S.C. §§ 2000d, 18116 (discussing ways that programs which receive federal funding cannot discriminate on the basis of several different protected characteristics).

be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance."[71] Under Title VI, if a recipient of federal funding is found to have discriminated and they are not able to comply with the statute, then the agency that is providing funds should "either initiate fund termination proceedings or refer the matter to the Department of Justice."[72] As with other anti-discrimination provisions, the use of AI could be implicated through either specific use of these characteristics or through disparate impacts upon certain populations.

In addition, the Affordable Care Act, also known as the Patient Protection and Affordable Care Act ("ACA"), has specific provisions which prevent discrimination for any health programs which receive "federal financial assistance."[73] The ACA specifically states that the protections under Title VI of the Civil Rights Act, Title IX of the Education Amendments of 1972, and the ADA of 1975 all apply.[74] This means that individuals in health programs which receive federal funding may not be "excluded from participation in, be denied the benefits of, or be subjected to discrimination" on the basis of race, color, national origin, sex, age[,] or disability."[75] Programs that receive federal assistance could include any hospitals that accept Medicare or doctors who receive Medicaid payments, which includes a large number of health care providers.[76] There are numerous different ways that federal statutes can be implicated by the use of AI in settings like hospitals and schools, which will be discussed in Part V.

---

[71] 42 U.S.C. § 2000d.

[72] *Title VI Of The Civil Rights Act Of 1964 42 U.S.C. § 2000d Et Seq.*, U.S. DEP'T JUS., C.R. DIV., https://www.justice.gov/crt/fcs/TitleVI-Overview [https://perma.cc/FH2R-8APQ] (last updated June 6, 2023).

[73] Patient Protection and Affordable Care Act, 42 U.S.C. § 18116.

[74] *See* 42 U.S.C. § 18116.

[75] *See id.*; *see also* 42 U.S.C. §§ 2000d, 6101; 20 U.S.C. § 1681.

[76] *See Section 1557: Coverage of Health Insurance in Marketplaces and Other Health Plans,* U.S. DEP'T HEALTH & HUM. SERV., https://www.hhs.gov/civil-rights/for-individuals/section-1557/fs-health-insurance/index.html [https://perma.cc/2ZE7-H6ZB] (last updated Aug. 25, 2016).

C. Federal Anti-Discrimination Housing Provisions

Finally, the federal government created protections for individuals in the housing context, specifically through the Fair Housing Act ("FHA").[77] The FHA says that it is unlawful to "refuse to sell or rent after the making of a bona fide offer, or to refuse to negotiate for the sale or rental of, or otherwise make unavailable or deny, a dwelling to any person because of race, color, religion, sex, familial status, or national origin."[78] The FHA also prohibits individuals or organizations from advertising that properties are only available to certain protected classes or to represent to anyone that a dwelling is not available on the basis of a protected characteristic.[79] Housing regulations, like the ADEA, protect individuals from explicit discrimination, as well as the discriminatory effect of housing policies that have a disparate impact on a group based on protected characteristics, even if that discrimination was not intentional.[80] These federal statutes provide a clear picture of the characteristics that cannot be used to discriminate against individuals, including race, sex, national origin, age, disability status, religion, marital status, and others. While the federal government has clearly prohibited discrimination in these areas, some organizations using AI have already faced charges and settlements for discrimination on the basis of protected characteristics.

V. CURRENT EXAMPLES OF AI DISCRIMINATION

This Part provides numerous examples of cases or situations where the use of AI has led to discrimination in violation of the aforementioned federal anti-discrimination statutes. Within this

---

[77] *See* 42 U.S.C. § 3604.

[78] *Id.*

[79] *See id.*

[80] *See* 24 C.F.R. § 100.500 (2023) ("Liability may be established under the Fair Housing Act based on a practice's discriminatory effect, as defined in paragraph (a) of this section, even if the practice was not motivated by a discriminatory intent."); *Id.* ("(a) Discriminatory effect. A practice has a discriminatory effect where it actually or predictably results in a disparate impact on a group of persons or creates, increases, reinforces, or perpetuates segregated housing patterns because of race, color, religion, sex, handicap, familial status, or national origin.").

Part, there are examples of discrimination in employment, education, healthcare, and housing. These cases demonstrate the need for greater regulation for AI and display how discrimination can arise from either algorithm creation or through algorithm use.

A. EEOC Settlement Arising Over AI Use

A recent example of AI discrimination in the employment context arose in the case of *EEOC v. iTutorGroup*,[81] which resulted in one of the first settlements with the EEOC over AI bias.[82] The case, which settled for $365,000, alleged that "iTutorGroup used an AI tool that rejected male applicants over the age of 60 and females over 55."[83] While the defendants denied any allegations and claimed no intentional discrimination, the case shows just how easily bias can creep into AI usage.[84] In the *iTutorGroup* case, the applicant over age 55 first applied to a role using her real age and was rejected from the role.[85] She then subsequently resubmitted her application using a younger age and received an interview request.[86] Additionally, "[t]he EEOC alleged that more than 200 other applicants faced similar age-based rejections."[87] The *iTutorGroup* case demonstrates how the use of AI can violate the ADEA, and if the protected characteristic here were swapped with something like sex or disability status, then it could violate Title VII or the ADA. This case exemplifies the pitfalls of AI usage in the employment context and reflects the need for continued guidance and legislation about AI usage.[88]

---

[81] Complaint, EEOC v. iTutorGroup, Inc., No. 1:22-cv-02565 (E.D.N.Y. 2023).

[82] Bridget Roddy & Francis Boustany, *ANALYSIS: First AI Bias Settlement With EEOC Spotlights Pitfalls*, BLOOMBERG L., (Aug. 24, 2023, 5:00 AM), https://news.bloomberglaw.com/bloomberg-law-analysis/analysis-first-ai-bias-settlement-with-eeoc-spotlights-pitfalls [https://perma.cc/BAA8-MNZZ].

[83] *Id.*

[84] *See id.*

[85] Complaint at 5, *EEOC v. iTutorGroup, Inc.*, No. 1:22-cv-02565.

[86] *See id.*

[87] Roddy & Boustany, *supra* note 82.

[88] *Id.* ("[A] good first step toward liability mitigation is to ensure that tools don't ask certain questions—such as those that would elicit an applicant's age, religion, or other protected characteristic—that could land an employer in court.").

B. Education Admissions and Outdated Systems

The *Students for Fair Admissions* case demonstrated how discrimination can occur in the college admissions process and set the stage for ways that AI could incorporate bias into decision-making processes. In his concurrence, Justice Gorsuch wrote that the analysis of Title VI should be conducted similarly to an analysis under Title VII, which prevents the use of race and other protected characteristics in hiring decisions.[89] In short, Justice Gorsuch argued that Title VI should have applied to the SFFA case, since Title VI "prohibits a recipient of federal funds from intentionally treating any individual worse even in part because of his race, color, or national origin and without regard to any other reason or motive the recipient might assert."[90]

One of the biggest ways that the use of AI can prove to be an issue for college admissions programs occurs when programs use historical data to create admissions algorithms that may have bias imbedded in the data. For example, in 2013, the University of Texas at Austin's ("UT") computer science department began using a machine-learning system to help graduate admissions programs save time.[91] This program, called the Graduate Admissions Evaluator ("GRADE"), was developed by a UT faculty member and a computer science graduate student, and it scored applicants based on how likely they were to be admitted.[92]

The system and the way it was developed presented two major issues. First, the system was trained on historical data, which consisted of a database of past admissions decisions.[93] Reliance on

---

[89] *See* Students for Fair Admissions, Inc. v. President & Fellows of Harv. Coll., 600 U.S. 181, 289 (2020) (Gorsuch, J., concurring).

[90] *Id.*

[91] Lilah Burke, *The Death and Life of an Admissions Algorithm*, INSIDE HIGHER EDUC. (Dec. 13, 2020), https://www.insidehighered.com/admissions/article/2020/12/14/u-texas-will-stop-using-controversial-algorithm-evaluate-phd [https://perma.cc/92VU-PB28] (discussing the development and eventual removal of an algorithm for graduate school admissions).

[92] *Id.*; Austin Waters & Risto Miikkulainen, *GRADE: Machine Learning Support for Graduate Admissions*, 25 PROC. CONF. ON INNOVATIVE APPLICATIONS ARTIFICIAL INTEL. (2013).

[93] Burke, *supra* note 91; Waters & Miikkulainen, *supra* note 92, at 1.

historical data can lead to "signal problems," which occur when "citizens or subgroups are underrepresented due to unequal creation or collection of data."[94] Historically, this program at UT was underrepresented by women and people who are Black and Latinx.[95] Consequently, the use of this historical data only fueled the disparities between races and genders.

Second, the system was never updated or audited, meaning that the preferences of the system were just a snapshot in time from 2013 instead of continually evolving.[96] The creators of the algorithm explicitly stated that their intent was to "replicate what the admissions committee was doing prior to 2013, not to make better decisions than humans could,"[97] which has the potential to continually include biases in the admissions process. A professor from the University of Maryland criticized this system and said that the UT admissions team "built a model that builds in whatever bias [the] committee had in 2013 and [they have] been using it ever since."[98] While the GRADE system was abandoned in 2020, the use of this system in the first place demonstrates the way that AI systems can quickly lead to algorithmic bias.

C. Erroneous AI Algorithm Violating ACA and Title VI

In November 2023, a class action suit was brought against UnitedHealth Group ("UnitedHealth"), which alleged that UnitedHealth used AI "in place of real medical professionals to wrongfully deny elderly patients care owed . . . based on an AI model that Defendants know has a 90% error rate."[99] The plaintiffs alleged that the defendants used an AI model, known as "nH Predict," to predict how much care elderly patients should receive.[100] However, this system was allegedly "rigid and unrealistic" with its predictions for recovery, and the system was

---

[94] Ajunwa, *supra* note 14, at 1686.
[95] *See* Burke, *supra* note 91.
[96] *See id.*
[97] *Id.*
[98] *Id.*
[99] Compliant at 1, Estate of Lokken v. UnitedHealth Group, Inc., No. 0:23-cv-03514 (D. Minn. Nov. 14, 2023).
[100] *See id.* at 3.

used to override "real doctors' determinations as to the amount of care a patient in fact requires to recover."[101] The plaintiffs argued that based on the outputs of the nH Predict algorithm, there were many "inappropriate denial[s] of necessary care prescribed by the patients' doctors."[102] The plaintiffs further contended that they were denied coverage by the nH Predict system because of their age, even though their doctors recommended additional treatment.[103]

The plaintiffs' causes of action arose out of breach of contract, unjust enrichment, and insurance bad faith.[104] While contract and insurance claims are more focused on tort law and not anti-discrimination, there is a possibility that the use of this AI violates federal law protecting individuals receiving healthcare. Under the ACA, individuals may not be excluded from participation in or denied the benefits of health programs on the basis of their age or disability status if the programs receive federal funding.[105] Medicare Advantage, which is the plan that plaintiffs' claims were denied under, is federally funded, and therefore, these claims could be violations of the ACA.[106] If the claims that the algorithm denied care to individuals because of their age or disability status succeed, then this case clearly shows how AI algorithms can violate federal anti-discrimination laws.

D. Evaluating Housing Applicants with Discriminatory Algorithms

In the housing context, the use of AI in determining tenancy has led to recent lawsuits against the algorithm creators as well as management companies that use the AI models to make decisions. In 2022, two individuals in Massachusetts filed suit against SafeRent Solutions ("SafeRent") and Metropolitan Management Group for alleged violations of the FHA and Massachusetts

---

[101] *Id.*
[102] *Id.*
[103] *See id.* at 1–4.
[104] *See id.* at 22–27.
[105] *See* 42 U.S.C. § 18116.
[106] *See How is Medicare Advantage Funded*, MEDICALNEWSTODAY (May 21, 2020), https://www.medicalnewstoday.com/articles/how-is-medicare-advantage-funded [https://perma.cc/54FR-W8KM].

anti-discrimination laws.[107] The suit alleged that "SafeRent's algorithm ha[d] a disparate impact based on race,"[108] which is a violation of the FHA.

The plaintiffs, who relied on federal housing vouchers to cover the majority of their rent, argued that SafeRent's algorithm relied heavily upon credit score and credit history and did not take into account other assets.[109] This reliance upon credit history "disproportionately affects Black and Hispanic tenants, in addition to tenants who hold housing vouchers, because Black and Hispanic consumers have a lower median credit score than White consumers."[110]

The District Court of Massachusetts considered whether this utilization of an AI algorithm could amount to a violation of the FHA in a motion to dismiss from the defendants.[111] In denying the motion to dismiss the FHA claims, the court held that it was plausible that SafeRent's algorithm was effectively controlling "the decision to approve or reject a rental application," because the algorithm had "sole control over how scores [were] calculated."[112] The court then held that since these determinations "may disqualify otherwise qualified rental applicants and, as alleged, results in a disparate impact on protected groups, SafeRent is subject to the FHA."[113] The district court's decision, which denied the dismissal of the FHA claims, represents one of the numerous ways that algorithms can be

---

[107] Louis v. SafeRent Sols., LLC, No. 22-CV-10800-AK, 2023 WL 4766192 (D. Mass. July 26, 2023); *see also SafeRent Solutions Accused of Illegally Discriminating Against Black and Hispanic Rental Applicants*, COHENMILSTEIN (May 25, 2022), https://www.cohenmilstein.com/saferent-solutions-accused-illegally-discriminating-against-black-and-hispanic-rental/ ("A lawsuit filed today in U.S. District Court for the District of Massachusetts against SafeRent Solutions, LLC alleges that the national tenant screening provider has been violating the Fair Housing Act and related state laws for years.") [https://perma.cc/A2ZF-D8KT].

[108] *SafeRent Solutions Accused of Illegally Discriminating Against Black and Hispanic Rental Applicants*, *supra* note 107.

[109] *Louis v. SafeRent Sols*., 2023 WL 4766192, at *2.

[110] *Id.*

[111] *Id.* at *8–10.

[112] *Id.* at *9.

[113] *Id.*

used in the context of housing to discriminate against protected classes, even if it is unintentional.

VI. KEEPING AI IN CHECK

While AI usage has advanced in almost every sector of our lives,[114] most states in the U.S. are lagging in their efforts to regulate this emerging technology. This Part considers two pieces of legislation, one currently in force and one that has been introduced in Congress, which provide a framework for how the U.S. should regulate AI.

A. NYC Local Law 144

On July 5, 2023, New York City began enforcing NYC Local Law 144 of 2021 ("Local Law 144"),[115] which requires bias audits to be conducted on automated employment decision tools ("AEDT") prior to their use each year.[116] A bias audit is defined as "an impartial evaluation by an independent auditor" to determine an algorithm's potential for bias.[117] While employers can look for other potential biases, the minimum required by the law is that the "independent auditor's evaluation . . . include calculations of selection or scoring rates and the impact ratio across sex categories, race/ethnicity categories, and intersectional categories."[118]

Local Law 144 requires employers to be transparent with the results of the bias audits by publishing that data "on the employment

---

[114] *See* Brian Kennedy, et al., *Public Awareness of Artificial Intelligence in Everyday Activities*, PEW RSCH. CTR. (Feb. 15, 2023), https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/ [perma.cc/BWQ4-BVSY].

[115] Rules of the City of New York tit. 6, § 5-301 (2021); N.Y.C., N.Y., Admin. Code § 20-871 (2021).

[116] N.Y.C. DEP'T OF CONSUMER & WORKER PROT., AUTOMATED EMPLOYMENT DECISION TOOLS: FREQUENTLY ASKED QUESTIONS 1, 6 (2023).

[117] *Id.*

[118] *Id.*

section of their website."[119] Specifically, the law requires employers to publish:

> [The] date of the most recent bias audit of the AEDT and a summary of the results, which shall include the source and explanation of the data used to conduct the bias audit, the number of individuals the AEDT assessed that fall within an unknown category, and the number of applicants or candidates, the selection or scoring rates, as applicable, and the impact ratios for all categories.[120]

The guidance posted by the New York City Department of Consumer and Worker Protection ("DCWP") defines an AEDT as "a computer tool that: [u]ses machine learning, statistical modeling, data analytics, or artificial intelligence [and] [h]elps employers and employment agencies make employment decisions [and] [s]ubstantially assists or replaces discretionary decision-making."[121] Additionally, Local Law 144 applies to any stage of the hiring process, such as initially screening employees, not only the final hiring decisions.[122] Ultimately, if companies use an AEDT "to substantially help them assess or screen candidates at any point in the hiring or promotion process, they must comply with . . . [Local Law 144's] requirements before using an AEDT."[123]

If an organization qualifies under the criteria provided above for using an AEDT, then the organization must conduct a bias audit of its system and publish the findings.[124] Under Local Law 144, organizations and companies using AEDTs must have independent auditors evaluate the AI and underlying algorithms to provide a score based on potential bias in "sex categories, race/ethnicity categories, and intersectional categories."[125] The auditor cannot work for the employer that uses the AEDT or the vendor that distributes the AEDT, and they must be "someone who exercises

---

[119] *See* tit. 6, § 5-303. ("[A]n employer or employment agency in the city must make the [audit results] publicly available on the employment section of their website in a clear and conspicuous manner.").

[120] *Id.*

[121] *See* N.Y.C. DEP'T OF CONSUMER & WORKER PROT., *supra* note 116, at 1.

[122] *Id.* at 2.

[123] *Id.*

[124] *Id.*

[125] *Id.*

objective and impartial judgment in the performance of a bias audit."[126] Finally, once an independent audit has been completed, companies are required to publicly share the results of the bias audit.[127] The audit results must include the date of the audit, the source of the data audited, the number of individuals assessed by the AEDT within certain categories, and the selection or scoring rates among the focus categories of sex, race, and ethnicity.[128]

Local Law 144 combats bias and discrimination in AI use in a few different ways. First, Local Law 144 requires analysis of the underlying algorithm and historical data used to make hiring decisions through this bias audit. The analysis conducted through these bias audits can help resolve red flags in the criteria used to make decisions, and it can reveal potential areas for discrimination before it occurs. Second, Local Law 144 requires annual audits, which allows the auditors to catch disparate impact cases as they arise, not just cases where the potential for discrimination arises from the terms of the algorithm. Finally, this law provides external motivations in the form of third-party auditors conducting the audit and the requirement to publish the results of the audit. Ideally, these external factors encourage companies to take this law seriously and invest time and resources into ensuring that their systems are free from bias.

B. Algorithmic Accountability Act

On a broader scale, several senators have submitted a bill before Congress called the Algorithmic Accountability Act of 2023 ("AAA"), which is designed to increase insight and visibility into the way that organizations are currently using algorithms.[129] This is

---

[126] *Id.* at 5.

[127] *See id.* at 3.

[128] *See id.*

[129] *See* Press Release, Ron Wyden, U.S. Sen. For Or., Wyden, Booker and Clarke Introduce Bill to Regulate Use of Artificial Intelligence to Make Critical Decisions like Housing, Employment and Education (Sept. 21, 2023), https://www.wyden.senate.gov/news/press-releases/wyden-booker-and-clarke-introduce-bill-to-regulate-use-of-artificial-intelligence-to-make-critical-decisions-like-housing-employment-and-education [https://perma.cc/9U2Y-P8RE].

the third time that the AAA has been introduced, as it was first introduced in 2019,[130] and again in 2022.[131] The AAA attempts to create similar requirements to Local Law 144 but applies to many more industries across the country.[132] Specifically, this bill requires any covered entity to conduct an "impact assessment" for efficiency metrics and bias detection.[133] The covered entities are those that are above a certain economic threshold and deploy any form of augmented critical decision process, which is a process that affects a consumer's education, employment, healthcare, housing, utilities, legal services, and others.[134] The impact assessments are designed to cover a number of different topics (e.g., efficiency, privacy, and negative impacts on consumers) and areas of improvement (e.g., fairness, meaning areas of potential bias, or non-discrimination).[135]

There are several key aspects that make the AAA a step in the right direction. First, the bill applies broadly to several key sectors where bias and discrimination arise from AI use, such as education, employment, healthcare, and housing. Additionally, the AAA requires companies to "[m]aintain and keep updated documentation of any data or other input information used to develop, test, maintain, or update the automated decision system."[136] Keeping track of the data used to develop an automated decision system could be influential in removing biased data in the development of AI algorithms, which could decrease instances of discrimination.

While the AAA is a strong starting point for AI regulation, improvements need to be made before it can be considered a comprehensive regulatory plan for AI. First, the AAA lists numerous other categories, such as the efficiency of the system and the impact on stakeholders, before it considers bias and discrimination as a reason for and part of the impact assessments.[137] While for-profit

---

[130] Harvis-Nazzario, *supra* note 15, at 159–60.
[131] Algorithmic Accountability Act of 2022, S. 3572, 117th Cong. (2022).
[132] *See* Algorithmic Accountability Act of 2023, S. 2892, 118th Cong. (2023).
[133] *See id.*
[134] *See id.*
[135] *See id.*
[136] *Id.*
[137] *See id.*

companies are likely to consider the efficiency and costs of the systems that they use on their own, they may be less likely to consider how their algorithms are negatively impacting different groups. Therefore, the AAA should put a greater emphasis on requiring companies to dig deeper into how their systems could be negatively impacting protected classes. This emphasis could be achieved by making bias results the first reporting requirement for the impact assessments, as well as enforcing penalties for companies who fail to include information relating to bias results in their impact assessment.

Second, the AAA does not seem to provide a clear definition of what it considers to be bias or protected groups.[138] To amend this issue, the bill should take a page from the ACA, which borrowed definitions of discrimination from other statutory provisions, instead of creating a new definition that is specific to healthcare. In the context of the AAA, the bill could incorporate provisions from Title VI, Title VII, the ADEA, the ACA, the FHA, and other anti-discrimination provisions to leverage their definition of disparate treatments and disparate impacts.

Finally, the bill requires the covered entities (i.e., the companies themselves) to conduct the assessments.[139] Unlike Local Law 144, the AAA requires covered entities to conduct self-regulated assessments and send the results to the governing agency. While this system may work, it is unclear how many organizations would invest their own time and energy into developing systems to assess every automation tool they use. Instead, the AAA should follow the example of Local Law 144 and mandate assessments by external auditors. Both Local Law 144 and the AAA represent steps in the right direction towards a proper regulatory framework for AI, but there is still more that needs to be done.

---

[138] *See id.*
[139] *See id.*

VII. RECOMMENDATIONS FOR A FEDERAL ANTI-DISCRIMINATION FRAMEWORK FOR AI

AI's potential for discrimination on a massive scale is a significant cause for concern, especially considering that "any bias introduced in the system will be magnified and multiplied, greatly dwarfing the impact of any prejudice held by any one [decision-maker]."[140] Due to the potentially disastrous consequences of allowing AI to advance unregulated, there is a definite need to provide a regulatory framework for the production and maintenance of AI algorithms. Previous scholarship in this area has either focused on preventing disparate treatment causes of action under Title VII or disparate impact causes of action as separate goals.[141] This Part provides a two-prong framework for regulating AI in an alternative way that leverages existing legislation and regulations to address both causes of action. To do so, it is necessary to create federal regulations that require certain criteria to be complied with in the creation of AI algorithms. Further, it is imperative that additional regulations are put in place for continual maintenance and auditing of AI systems.

A. Regulating the Creation of Algorithms

One unanswered question surrounding AI regulation is whether the liability for AI returning biased results should fall upon the creators of the algorithm or the user of the algorithm.[142] The first (and preferred) option is to create regulations that place some responsibility upon the algorithm creators themselves. Experts have advocated for a fiduciary theory to be applied to the creators of the algorithms, which has been referred to as an "information fiduciary" theory.[143] This theory argues that "information fiduciaries 'have special duties to act in ways that do not harm the interests of the people whose information they collect, analyze, use, sell and

[140] Ajunwa, *supra* note 14, at 1679–80.
[141] *See id.* at 1678–79.
[142] *See id.* at 1720.
[143] *Id.* at 1720–21.

distribute.' "[144] The concept of having an information fiduciary places a specific responsibility upon creators of systems that utilize and leverage data to do so in a responsible and ethical way. While this idea is a strong basis for why AI creators should be regulated, relying upon the fiduciary relationship alone may not be sufficient to create regulations regarding AI creation.

Instead, the path towards regulation for AI creators must follow in the footsteps of other regulatory agencies that require certain safety standards to be met before their products can be sent out into the market. The Consumer Product Safety Commission ("CPSC") is an agency that sets out requirements for producing numerous different consumer products.[145] The CPSC requires products to go through third-party testing programs before they can be sold and distributed to consumers.[146] Additionally, the CPSC incorporates safety standards from an external organization, the American Society for Testing and Materials, to establish the requirements for product safety.[147] Some of these safety standards regulate how the products can be developed, while others regulate what materials the products can be made from.[148]

In addition, the Food and Drug Administration ("FDA") requires intensive testing on drugs before they are able to be sold to the public.[149] The Center for Drug Evaluation and Research ("CDER") sets specific requirements for drug testing to ensure that these drugs "work correctly and that their health benefits outweigh their known risks."[150] The CDER also receives the results of the testing and sends them to a group of unbiased evaluators to determine that the drug's health benefits outweigh the risks.[151] Both the CPSC and the FDA

---

[144] *Id.* at 1721 (citing Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183, 1209 (2016)).

[145] *See* 16 C.F.R. § 1112.15 (2013).

[146] *See, e.g.*, 16 C.F.R. § 1216.2 (2022).

[147] *See id.*

[148] *See id.*

[149] *See Development & Approval Process: Drugs*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/drugs/development-approval-process-drugs [https://perma.cc/5T3K-EREA] (last updated Aug. 8, 2022).

[150] *Id.*

[151] *See id.*

provide specific regulations for the creation of different products in order to protect the public.

To hold creators of algorithms accountable in some fashion, AI regulation must mimic the paths of agencies like the CPSC and the FDA. The idea of "An FDA for Algorithms"[152] has been briefly discussed before by scholars, but a regulatory body for AI should follow more than just the FDA—it should follow other agencies as well. In the context of AI, the first step is to choose which agency would begin to regulate this area. The Federal Trade Commission ("FTC") is the obvious choice, as the AAA recommends the FTC as the best agency to regulate AI.[153] The FTC should follow the structure of other consumer protection agencies to develop safeguards for AI. Instead of simply acting like another FDA and mostly focusing on review of the testing process, the FTC should instead create specific guidelines for how AI algorithms are created and what criteria they are allowed to use. These rules must proscribe specific criteria that cannot be used in algorithm creation, including all the protected characteristics previously discussed, unless there is a definitive business reason to use these characteristics. Additionally, these rules should follow the structure of the CPSC and incorporate the requirements from third-party organizations, such as the Institute for Internal Auditors,[154] to give guidance on how algorithms are to be designed.

Laws regulating the creation of algorithms would be beneficial for multiple reasons. First, they would provide clarity and consistency in the requirements for how algorithms are to be created and what information needs to be avoided at all costs. Second, and most importantly, regulating the creation process would give a clear cause of action for individuals who are harmed by these algorithms. Currently, it is extremely difficult for plaintiffs to bring successful discrimination claims against organizations like their employers because it is incredibly challenging to prove both the impact and the cause.[155] By regulating the creation process instead of waiting until

---

[152] Andrew Tutt, *An FDA for Algorithms*, 69 ADMIN L. REV. 83, 118 (2017).
[153] Algorithmic Accountability Act of 2023, S. 2892, 118th Cong. (2023).
[154] *See* Ajunwa, *supra* note 1, at 44.
[155] *See id.* at 27.

AI developers violate other anti-discrimination provisions, plaintiffs would be better-suited to recover if developers skipped specific steps or used restricted data sets in algorithm creation. If companies are required to provide more transparency into the inputs and proxies they use in their algorithms, there will be fewer disparate treatment discrimination cases and when they do occur, they will be easier to pinpoint.

Finally, although many critics claim that these regulations will delay the development of AI and technology and harm the economy,[156] regulation is essential for safe and effective algorithms. There is a possibility that given the rapid pace at which algorithms develop, we will "progress from 'too soon' to regulate algorithms to 'too late' in the blink of an eye."[157] AI is advancing at an incredibly fast pace, and it is important that regulations are established before AI goes too far. Because of the FDA's regulations surrounding drug creation, Americans "benefit from having access to the safest and most advanced pharmaceutical system in the world."[158] The goal of creating the safest and most advanced AI system in the world, like the FDA has done with the pharmaceutical industry, should be the focus of regulating agencies like the FTC moving forward.

B. Auditing Algorithms for Accountability

While regulating and reviewing the creation of algorithms will help prevent discrimination in many cases, it will not cover every case where discrimination arises. Thus, there is a need to also hold AI users accountable through tools like bias audits. Bias audits would be a useful part of the regulatory framework for AI because of their ability to fill in the gaps left behind by regulations that only focus on the inputs of AI algorithms. Whereas the previous recommendations are helpful for finding and preventing cases of disparate treatment, the incorporation of bias audits would help to locate causes of disparate treatment and unintentional

---

[156] *See* Katyal, *supra* note 23, at 60 ("Technologists have, and often rightfully so, framed legal regulation . . . as outdated, outmoded, and unnecessarily impeding innovation.").

[157] Tutt, *supra* note 152, at 119.

[158] *Development and Approval Process: Drugs*, *supra* note 149.

discrimination. Additionally, bias audits would help avoid issues surrounding trade secrecy, which can lead to companies' unwillingness to share the proprietary information of their algorithms. "[A]udits by an independent auditing body would serve to allay any fears as to the misuse of proprietary information" and could help provide information for disparate impact claimants.[159]

To implement these bias audits, the government should look to existing or proposed legislation for guidance. Both Local Law 144 and the AAA require either a bias audit or an impact statement, designed to locate potential biases developed in the system. The structure of Local Law 144, which requires scoring of potential bias in certain categories,[160] should be replicated to create consistent standards across industries and AI systems. In addition, if users score low enough in a certain category, regulations could require the user to take action to eliminate the causes of bias or discrimination.

Although one agency may be able to handle the regulation of algorithm creation, it would be burdensome for that same agency to review bias audit results. Instead, the process of reviewing bias audits should be the responsibility of the agency already regulating that industry. For example, the EEOC could regulate the use of AI in employment contexts, the U.S. Department of Housing and Urban Development could regulate AI in housing, and the Office for Civil Rights[161] could regulate AI usage in education. These agencies are the most familiar with what violations of federal anti-discrimination law look like in their own industries, so they are the best voices to define what would need to change in algorithms to prevent bias. Finally, these agencies could provide causes of action for plaintiffs who are harmed by algorithmic discrimination if the organizations were aware of the bias after a bias audit and failed to address the issue. This cause of action would mirror the way that the CPSC

---

[159] Ajunwa, *supra* note 1, at 27.

[160] *See* N.Y.C. DEP'T OF CONSUMER & WORKER PROT., *supra* note 116, at 2.

[161] *How Title VI Legally Prohibits Discrimination in Education*, JUSTIA, https://www.justia.com/education/discrimination-in-education/title-vi/ [https://perma.cc/VP6W-B92T] (last updated May 2023).

provides liability for manufacturers and could help with the issues regarding difficulties in bringing disparate impact claims.[162]

Some scholars have advocated for voluntary methods of incorporating audits into the process of using AI, but the rate at which AI is evolving requires a firmer stance on the process of auditing. For example, some scholars have argued that AI auditing should be subsumed by the idea of Corporate Social Responsibility ("CSR"), which has seen growth in recent years.[163] This idea would be very beneficial if implemented, but the pace at which AI is advancing may move faster than corporations' CSR departments can keep up with, and companies may choose to value their efficient technologies over CSR initiatives if there is no external pressure to change the systems. Second, others have argued that the best path to accountability is through self-regulation, such as conforming to specific codes of conducts or following ethical principles set forth by professional organizations such as the "Association for the Advancement of Artificial Intelligence."[164] Self-regulation only works if companies agree to certain standards, so this recommendation should not be accepted because of the same reason that regulation is required in the first place. Even if just one company were to not abide by these standards or codes of conduct, the widespread influence and impact of AI could have massive impacts on numerous populations. Since AI has such a long reach, it is important to incorporate regulations now, before damage can be done to too many people.

Critics claim that bias audits would be too costly or inefficient to implement and should not be a part of the regulatory framework,[165] but those concerns can be appeased in several ways. First, scholars have argued that "anti-discrimination laws do not

---

[162] *See, e.g.*, *Bad Boy Enterprises Agrees to $715,000 Civil Penalty for Failing to Report Defective Buggies*, U.S. CONSUMER PROD. SAFETY COMM'N (Sept. 23, 2011), https://www.cpsc.gov/Newsroom/News-Releases/2011/Bad-Boy-Enterprises-Agrees-to-715000-Civil-Penalty-for-Failing-to-Report-Defective-Buggies [https://perma.cc/AEX6-JQ6E].

[163] *See* Ajunwa, *supra* note 1, at 41.

[164] Katyal, *supra* note 23, at 109.

[165] *See* Ajunwa, *supra* note 1, at 63.

require shareholder value maximization" and that they instead are designed to serve a different goal.[166] Second, even when cost-savings are taken into account, AI is already a cost-saving measure so companies should have additional funds to use, and "automated hiring system[s] [have] already been designed in such a way to retain and easily produce the information needed for [bias] audits."[167] Finally, new technologies, such as IBM's WatsonX.Governance, are being developed to "drive responsible and ethical decision-making across organizations."[168] As additional technologies like this are developed, it will become cheaper and easier for companies to better control their AI systems and easily produce the data necessary to comply with audit requirements. The regulation of both the creation of AI algorithms and the requirement of audits on these systems should empower the federal government to prevent AI discrimination.

VIII. CONCLUSION

As the world continues to move toward a greater reliance on technology, it is important to ask the question of when society should continue to forge ahead and when it should slow down. Technological advancements have continued to fuel innovation and impact lives in a positive way, whether it is through increasing the efficiency of work, connecting individuals across the globe, or empowering better decision-making. However, as society continues to advance technologically, it is necessary to consider whether these new advancements are helping or hurting society at large.

Whether it is healthcare, employment, or education, no industry is immune from the potential pitfalls of AI and the ways that bias can creep into those systems. In light of this, the federal government must create a regulatory framework to prevent both disparate treatment and disparate impacts on protected classes. Regulating

---

[166] *Id.* (quoting Charles Sullivan, *Employing AI*, 63 VILL. L. REV. 395, 398 n.12 (2018)).

[167] Ajunwa, *supra* note 1, at 63.

[168] *'Break Open the Black Box' with AI Governance*, IBM, https://www.ibm.com/products/watsonx-governance [https://perma.cc/F2CM-HFWV] (last visited Feb. 25, 2024).

algorithm creation and requiring periodic bias audits are the proper first steps to ensuring that AI is used in a manner that prioritizes people over profit and prevents widespread harm from being embedded in systems used on a daily basis.

When Dr. Martin Luther King Jr. said, "[i]njustice anywhere is a threat to justice everywhere,"[169] he was in no way, shape, or form referring to AI. However, when you consider the impact that a small amount of bias can have on an AI algorithm that is used by millions, Dr. King's words still ring true. A small amount of injustice, combined with a system that can transmit that injustice anywhere, is truly "a threat to justice everywhere."[170]

[169] MARTIN LUTHER KING JR., LETTER FROM BIRMINGHAM JAIL 1 (1963).
[170] *Id.*