The Need to Update the Equality Act 2010: Artificial Intelligence Widens Existing Gaps in Protection from Discrimination

Tetyana (Tanya) Krupiy*
Newcastle University

Abstract

The use of artificial intelligence (AI) to produce decisions about individuals can result in discrimination. Despite the fact that the employment of AI as part of the decision-making process is growing in the United Kingdom, there is limited literature examining gaps in legal protection in the Equality Act 2010 that the employment of AI gives rise to. This article identifies what assumptions contained within a number of provisions of the Equality Act 2010 result in this legislation having gaps in legal protection in the context of the use of AI. It proposes a number of solutions.

Keywords: Equality Act 2010; discrimination; artificial intelligence; digital; algorithmic; decision-making.

[A] INTRODUCTION

Computer scientists Valentin Hofmann, Pratyusha Ria Kalluri, Dan Jurafsky and Sharese King demonstrated in 2024 that language models operating on artificial intelligence (AI) software exhibit prejudice against individuals who use dialect when speaking English (Hofmann & Ors 2024: 2). While the language models generated statements reflecting positive stereotypes about African Americans when producing a response to a specific user query, they exhibited archaic stereotypes dating to before the civil rights movement when matching individuals to opportunities (ibid 2-3). For instance, AI scored individuals who said, "she been pulling" (ibid 42) as having a lower intelligence quotient score (ibid 46) than individuals who said "she's been pulling" (ibid 42). The researchers found that the intervention of programmers during the

programming stage exacerbated the problem by making it harder to detect prejudices in decision outcomes (ibid 393). It is imperative to address discrimination arising from the deployment of AI. The use of this technology is spreading in the United Kingdom (UK) (Stacey 2023). Yet, Jack Maxwell and Joe Tomlinson point out that it is challenging to apply the Equality Act 2010 in some circumstances to question unequal outcomes associated with the employment of AI (Maxwell & Tomlinson 2020: 353). The example they give is when the application of AI produces accurate predictions in one but not another geographical area (ibid). The Westminster Parliament plans to legislate in order to create obligations for developers of AI (Kyle 2024). One of these initiatives should involve updating the Equality Act 2010.

Currently, there is limited literature scrutinizing the gaps in legal protection from discrimination stemming from the use of AI that exist in the Equality Act 2010. There are few proposals regarding how these gaps can be remedied. Jeremias Adams-Prassl, Aislinn Kelly-Lyth and Reuben Binns are among a few scholars who have demonstrated how existing precedent can be applied to address some of the new ways in which AI brings about discrimination (Binns & Ors 2023: 1856-1857). It is crucial to establish what aspects of the Equality Act 2010 prevent it from achieving its aim to increase equality of opportunity in the context of AI use and to propose solutions. This is particularly the case because many scholars have written on the shortcomings of the Equality Act 2010 relating to protecting individuals from discrimination outside of the AI context (Butlin 2011: 434; Saunders 2020: 27; Connolly 2023: 663-664). For example, Hannah Saunders observes that the Equality Act 2010 protects people with some disfigurements from discrimination but does not capture the full spectrum of disfigurement (2020: 27). Karon Monagham described the Equality Bill that became the Equality Act 2010 as a "wasted opportunity" for many people (2009: 13).

The present article will demonstrate why several key provisions in the Equality Act 2010 do not capture some instances of discrimination when organizations employ AI to make decisions about applicants. It will explain why some of the assumptions underlying these provisions do not hold in the context of AI. These findings will serve as a basis for suggesting what considerations the Westminster Parliament could bear in mind when revising the Equality Act 2010. Due to the limitations of space, this article cannot consider all possible applications of AI and every provision of the Equality Act 2010. The limited scope of inquiry is not problematic. The purpose of this article is to pave the way for further discussion rather than to identify and to provide solutions to all problems.

For the purpose of this discussion, it is necessary to define AI. This stems from the fact that numerous definitions of this technology exist (Martínez-Plumed & Ors 2018: 5180). The UK Department for Science, Innovation and Technology defined AI in 2023 as products and services that are "adaptable" and "autonomous" (2023: 13). Autonomy refers to the fact that AI can produce decisions without ongoing human control; adaptivity denotes that AI can draw inferences from the data based on detecting patterns within the data (ibid 22). The UK Government Data Ethics Framework elaborates that AI entails the use of statistics to "find patterns in large amounts of data" (Central Digital and Data Office 2020). This article uses both documents as a basis for defining AI. This choice stems from the fact that, when used in conjunction, the two documents reveal that the use of statistical techniques and drawing inferences from large volumes of data (ibid; Department of Science, Innovation and Technology 2023: 22) are core characteristics of how AI operates.

The article has the following structure. Section B will define the provisions of the Equality Act 2010 that are the subject matter of discussion in this article. It will explain some of the assumptions that lie at the heart of how the drafters formulated these provisions. Section B also will introduce the central argument relating to why these assumptions create challenges for applying the provisions of the Equality Act 2010 in question to the context of AI use in some cases. Section C will further develop this argument by reference to how AI operates and by reference to specific AI programs. Section D will propose what considerations the legislators can bear in mind when revising the Equality Act 2010. It will put forward that there is a need to reconceptualize the concepts of the protected characteristic and group membership. The article will propose an alternative understanding of these two concepts. It will argue that it is necessary to rethink the relationship between the treatment, the affected individual, the protected characteristic and group membership. Judges need to be able to apply multiple tests, either in isolation or in conjunction, to establish whether the operation of the AI decision-making process gives rise to discrimination. The article outlines the overall framework of the solution without explaining all the details. Fleshing out the details is a subject of follow-up work. This is the case because it is necessary to consult the affected communities in order to draft a legal test that reflects the needs of individuals. Another reason for limiting the scope of inquiry stems from space limitations.

[B] THE NEED FOR CONCEPTS DEFINING DISCRIMINATION TO BETTER ACCOUNT FOR COMPLEXITY

Sandra Wachter states that when AI produces a decision, often, the decision is not causally linked to a particular characteristic or group membership of the subject of the decision-making (Wachter 2023: 199). This state of affairs is due to AI basing the decision on correlations, meaning on the relationships between the data (ibid). A corollary is that there is no causal relationship between the AI decision, the decision-making criteria the AI uses and an applicant's group membership (ibid). What is more, there is no causal relationship between the decision, the decision-making criteria AI employs and a particular characteristic of the applicant (ibid). Wachter concludes that: "The pursuit of mere correlation in AI renders causation disposable." (ibid 200) Although the Equality Act 2010 requires proof of causation for cases of direct discrimination (*Onu v Akwiwu; Taiwo v Olaigbe* 2016: paragraph 30) but not for cases of disability-based discrimination (Equality and Human Rights Commission 2011: paragraph 5.3) and cases of indirect discrimination (*Essop and Others v Home Office* 2017: paragraph 39), the manner in which AI produces decisions nevertheless poses a challenge to the practical application of the provisions of this legislation in some cases.

One reason why it is challenging to map onto the Equality Act 2010 some instances of discrimination occurring as a result of the use of the AI decision-making process is that this legislation defines the terms "protected characteristic" and "group membership" as having a one-dimensional character. Another reason is that the definitions of direct discrimination and indirect discrimination in the Equality Act 2010 do not conceptualize the relationship between the protected characteristic as a ground of legal protection, the individual experiencing discrimination and group membership in a complex manner. One possible approach to enable the Equality Act 2010 to protect individuals from AI-based discrimination in a more comprehensive manner is to include additional provisions defining prohibited treatment (Binns & Ors: 2023: 1857). Such provisions would need to have definitions of discrimination that capture the complex interrelations and interdependencies that exist between the affected individual, protected characteristic, group membership and harmful treatment/practice. Furthermore, the Equality Act 2010 needs to define the protected characteristic and group membership as having multiple dimensions. In order to lay the groundwork for developing this argument, this section will explain how a number of provisions in the

Equality Act 2010 treat the relationship between the prohibited conduct, the affected individual and the protected characteristic as a ground of legal protection and group membership.

The assumption that there is a direct unidirectional relationship between the protected characteristic as a ground of legal protection, the affected individual and the prohibited treatment can be seen in how the Equality Act 2010 defines a number of key provisions. Section 4 of the Equality Act 2010 defines the grounds for legal protection by reference to the possession of a protected characteristic. It assumes that the prohibited treatment can be directly linked to the possession of a protected characteristic. Section 4 of the Equality Act 2010 contains a closed list of the protected characteristics. These protected characteristics are age, disability, gender reassignment, marriage and civil partnership, pregnancy, maternity, race, religion, belief, sex and sexual orientation.

The Equality Act 2010 contains numerous definitions of how discrimination can occur. According to section 13(1) of the Equality Act 2013, A engages in direct discrimination against person B if "A treats B less favourably than A treats or would treat others" "because of" B's protected characteristic. In *Higgs v Farmor's School No 3*, the court held that it would look at why the defendant acted in a particular way and whether this conduct was "because of or related to" the possession of a protected characteristic (2023: paragraph 82). The terms "because of" and "related to" (ibid) connote the presence of a direct and unidirectional relationship between the prohibited conduct, the affected person and the protected characteristic.

In the case of *EAD Solicitors LLP and Others v Abrams*, the court held that section 13 of the Equality Act 2010 covers cases of direct discrimination where a person suffers adverse treatment due to the protected characteristic of another person (*EAD Solicitors LLP v Abrams* 2015: paragraphs 14-15). An example is when a company suffers adverse treatment due to providing financial support for Islamic education (ibid paragraph 29). Such cases are known as direct discrimination by association (ibid paragraph 10). In this case, the court interpreted section 13 of the Equality Act 2010 as being applicable when there is such a direct relationship between the person suffering the adverse treatment and the person with a protected characteristic that one can substitute the person with a protected characteristic for a person without a protected characteristic. Although this case recognizes that there can be an intervening element between the adverse treatment and the protected characteristic, the relationship between the adverse treatment and the

protected characteristic nevertheless has to be direct and unidirectional. It must be possible to substitute the person with the protected characteristic for the person experiencing adverse treatment in order for the treatment to be due to the possession of a protected characteristic.

The Supreme Court, in the case of *Lee v Ashers Baking Company Ltd*, explained that the prohibition of direct discrimination by association extends to cases where the reason for the less favourable treatment was a proxy for a protected characteristic (2018: paragraph 25). There must be an "indissociable relationship" between the proxy for the protected characteristic and the treatment (ibid paragraph 48). An example is when a hotel rents rooms to married couples but not to civil partners in circumstances when only heterosexual couples can get married (ibid paragraph 25; *Preddy v Bull* 2013). The court said that the connection between the proxy for the protected characteristic and the treatment needs to be closer than having "something to do" with the protected characteristic (ibid paragraph 33). However, the court thought that it would be "unwise" to define what degree of closeness satisfies the requirement of there being an "indissociable relationship" between the protected characteristic and the treatment (ibid paragraph 34). This case confirms that the prohibition of discrimination requires a direct and unidirectional relationship between the person suffering the adverse treatment, the harmful conduct and the protected characteristic. This aspect stems from the requirement that there be an "indissociable" relationship (ibid) rather than a degree of connection (ibid paragraph 33) between the unfavourable treatment and the protected characteristic.

Similarly, the definition of disability-based discrimination in the Equality Act 2010 assumes that there is a direct unidirectional relationship between the affected individual, the protected characteristic and the adverse treatment. Section 15(1)(a) defines discrimination against a person with a disability as treating that person "unfavourably" "because of something arising in consequence" of that person's disability. An example is when an employer dismisses a person due to taking a leave from work that is related to having a disability (Equality and Human Rights Commission 2011: 72 paragraph 5.3). According to section 15(1)(b) of the Equality Act 2010, such treatment should not be a "proportionate means of achieving a legitimate aim". Since section 15(1) of the Equality Act 2010 uses the phrase "because of" to link the adverse treatment to an aspect that is connected to an individual's disability, this provision creates a direct unidirectional relationship between the negatively affected individual, protected characteristic and unfavourable treatment. It follows that, although the definitions of direct discrimination and disability-based

discrimination differ (ibid), the two definitions have the same assumptions and underlying structure. Since courts recognize direct discrimination by association (*Lee v Ashers Baking Company Ltd* 2018: paragraph 25), the difference between the definitions of direct discrimination and disability-based discrimination is not substantial. While the prohibition of disability-based discrimination covers cases of an employer dismissing an employee due to taking a disability-related leave from work (Equality and Human Rights Commission 2011: 72), the prohibition of direct discrimination covers cases of an organization refusing to do business with another organization due to that organization employing individuals with a disability (*EAD Solicitors LLP v Abrams* 2015: paragraph 29). Both the prohibition of direct discrimination and the prohibition of disability-based discrimination address unfavourable treatment that can be directly linked to an individual with a disability.

The definition of the prohibition of indirect discrimination in section 19 of the Equality Act 2010 further consolidates the assumption that there is a direct unidirectional relationship between the negatively affected individual, the protected characteristic and unfavourable treatment. Additionally, section 19 adds group membership to this sequence. The definition of indirect discrimination contains the term group membership because it focuses on preventing harmful effects (Connolly 2018: 127) and on achieving "equality of results" (*Essop and Others v Home Office* 2017: paragraph 25). This approach to defining what constitutes discrimination contrasts with the definition of the prohibition of direct discrimination (Binns & Ors 2023: 1853). The prohibition of direct discrimination focuses on formal equality between two individuals and on the reason for the treatment (ibid).

Section 19(1) of the Equality Act 2010 defines indirect discrimination in terms of a person A applying a provision, criterion, or practice to a person B that is "discriminatory" to person B "in relation" to B's protected characteristic. Section 19(2) clarifies that the provision, criterion or practice "is discriminatory in relation to" B's protected characteristic if it 1) applies to all persons, 2) either "puts, or would put, persons with whom B shares the characteristic at a particular disadvantage" compared to persons who do not share B's protected characteristic, 3) either "puts, or would put, B at a disadvantage" and 4) A cannot show that the provision, criterion or practice is a "proportionate means of achieving a legitimate aim". Section 23(1) of the Equality Act 2010 elaborates that when comparing cases of treatment of A and individuals sharing B's protected characteristic for the purpose of applying section 19, "there must be no material difference between the circumstances relating to each case".

Sections 19 and 23 of the Equality Act 2010 conceive of the relationship between the negatively affected person, the protected characteristic, group membership and disadvantageous treatment as being direct and unidirectional. In the case of *Grundy v British Airways plc*, the judges explained that they identified a pool for comparison that "suitably tests the particular discrimination complained of" (2007: paragraph 27). The pool should enable a comparison of "like with like" (ibid paragraph 28) so that the circumstances of individuals in the pool are "not materially different" (ibid paragraph 33). By way of illustration, in the case of *Natasha Allen v Primark Stores Ltd* the court found that the pool in question related to all individuals whom the employer required to work a late shift on a compulsory basis and excluded individuals who could decline such a request (2022: paragraph 37).

By requiring the court to construct a pool consisting of persons whose circumstances are "not materially different" (*Grundy v British Airways plc* 2007: paragraph 33), section 19 of the Equality Act 2010 conceives of the protected characteristic and group membership as being in a direct unidirectional relationship. In addition to being an attribute of a person, the protected characteristic becomes overlapping with group membership. Moreover, the fact that the Supreme Court held in *Essop v Home Office* that identifying a criterion, provision, or practice "will also identify the pool for comparison" (2017: paragraph 41) points to the fact that the court constructs a direct relationship between the possession of the protected characteristic, group membership and disadvantageous treatment when choosing a pool. This discussion demonstrates that sections 19 and 23 of the Equality Act 2010 conceive of the negatively affected individual, the protected characteristic, group membership and the disadvantageous treatment as being in a direct unidirectional relationship. What is more, the prohibition of direct discrimination, disability-based discrimination and indirect discrimination share the same underlying assumptions and structure. This is despite the fact that the drafters formulated each definition of discrimination in a different way. The next section will demonstrate why this assumption creates challenges in some circumstances for applying the provisions of the Equality Act 2010 to the context of AI use.

[C] THE EQUALITY ACT 2010: AI USE AND GAPS IN PROTECTION

The assumption that there is a direct unidirectional relationship between the affected individual, conduct or practice, the protected characteristic and group membership creates difficulties for applying a number of provisions in the Equality Act 2010 to the context of the employment of AI as part of the decision-making process. Hilde Weerts and colleagues argue that it can be difficult to show a link between the possession of a protected characteristic and treatment in the context of AI use (Weerts & Ors 2024: 2). They maintain that the solution is to measure the extent to which one can derive a protected characteristic from another variable that has a relationship to a person's data and that AI uses to produce a decision (ibid 3-4). One can then measure the extent to which the AI uses a variable from which one can derive an applicant's protected characteristic in order to generate a prediction (ibid). For example, there is discrimination when the AI produces a decision based on inferring a person's race from the postcode (ibid 3). These scholars conclude that the prohibition of direct discrimination can be used to capture the specific way in which AI subjects individuals to unfavourable treatment in such cases (ibid 10). Their work is based on the work of Anya Prince and Daniel Schwarcz (ibid 4). Prince and Schwarcz coined the concept of "proxy discrimination" to denote instances when AI uses data that is predictive of having a protected characteristic to reach a decision about an applicant (2020: 1261).

While Weerts and collaborators explain that it can be difficult to determine what influence some proxies have on the decision outcome (Weerts & Ors 2024: 9), they do not discuss how the Equality Act 2010 can respond to the fact that there are cases when an AI decision-making process penalizes individuals for attributes that are akin to the recognized protected characteristics, but that cannot be linked to protected characteristics. In another article, Binns and colleagues put forward that there is a need to develop an additional category of direct discrimination (Binns & Ors 2023: 1857) that captures cases where the employment of AI disadvantages an applicant even though the developer had no intent to discriminate, there was no individual proxy corresponding to a protected characteristic, and there were no combinations of features that acted as a proxy for the protected characteristic (ibid 1856). They do not clarify how legislators could formulate such a concept.

The following example illustrates that AI can detect correlations in the data (Mittelstadt & Ors 2016: 5) that operate akin to a protected

characteristic but that do not map onto a protected characteristic in section 4 of the Equality Act 2010. Bart Custers explains that the patterns that AI detects in the data can result in AI penalizing someone for having physical attributes, such as shoe size (2023: 2). There is little difference between subjecting someone to unfavourable treatment based on that person's sex and shoe size because both are physical attributes of a person. Yet, in many cases, it will be impossible to link the shoe size to the protected characteristics that the Equality Act 2010 recognizes.

Consider the protected characteristic of sex in section 4 of the Equality Act 2010. In many cases, the shoe size is not a proxy for one's sex. This is the case because individuals of different sexes have both overlapping and different ranges of shoe sizes (Jurca & Ors 2019: 4). Male individuals have feet lengths in the range of 220-300 millimetres on average while female individuals have feet lengths in the range of 210-280 millimetres on average (ibid). There is a large overlap between the feet sizes of individuals of male and female sex. If someone's foot is 240 millimetres in length, one cannot conclude, based on this information, whether this person is of the male or female sex. Additionally, the fact that there are non-binary, multigender and transgender individuals means that it is not meaningful to talk about the shoe size as corresponding to a particular sex. As a result, shoe size is not a proxy for the protected characteristic of sex in many cases. In many cases, there will be no direct unidirectional relationship between the unfavourable treatment, the affected individual and the possession of a protected characteristic when the AI uses shoe size as a basis for generating a prediction, a score, or a decision relating to an applicant.

The lack of a direct unidirectional relationship between shoe size, the affected person and a negative AI decision renders it hard to invoke the prohibition of direct discrimination in the Equality Act 2010 to challenge the decision. The prohibition of direct discrimination in section 13(1) of the Equality Act 2010 requires that the less favourable treatment be "because of" the possession of a protected characteristic. Shoe size is neither a protected characteristic nor can it be associated with having a protected characteristic in many cases. Besides, it can be hard to demonstrate that the AI deployer treated an individual less favourably on the ground of shoe size in cases where the operation of the AI decision-making process disadvantaged individuals with a number of different shoe sizes to different degrees. The prohibition of indirect discrimination in section 19(2)(b) of the Equality Act 2010 is difficult to apply to these types of cases for a similar reason. This provision requires one to demonstrate that the use of AI puts individuals who share the protected

characteristic at a disadvantage in comparison to individuals who do not. However, it is not possible to link shoe size to having a protected characteristic in many cases. The difficulty of applying the prohibition of indirect discrimination is exacerbated by the fact that the use of AI as part of the decision-making process could disadvantage individuals with different shoe sizes to varying degrees. Such cases could present a challenge in identifying the correct pool of candidates in order to undertake a comparison of the treatment of different individuals. Additionally, the fact that the AI could issue a decision based on the shoe size and based on other attributes of multiple applicants (Fawcett & Provost 2013: 107; Broussard 2020) further complicates tracing the decision to belonging to a pool of candidates who have a particular shoe size.

This example illustrates a broader issue. Sandra Wachter notes that the operation of AI decision-making processes can generate many new types of disadvantaged groups (Wachter 2023: 153) due to detecting correlations in the data (ibid 158). Examples include playing video games and being a single parent (ibid 153-154). One cannot map these new groups onto existing protected characteristics (ibid 166). Wachter believes that the legislation should protect these new groups (ibid 203) because they suffer the same type of harm as individuals who currently enjoy protection from discrimination (ibid 195).

The fact that the Equality Act 2010 construes the relationship between the individual, unfavourable treatment, protected characteristic and group membership as being direct and unidirectional gives rise to another problem in the context of AI. Section 4 of the Equality Act 2010 can be difficult to apply even when the operation of AI disadvantages a person based on having a particular protected characteristic. This is the case because individuals with the same protected characteristic or attribute can be affected by the use of the same AI decision-making criterion differently. Consider the case study of bank lending. The Lenddo algorithm uses financial transactions and behavioural data to produce the applicant's creditworthiness score (Bary 2018). This software scores individuals who avoid using one-word subject lines in communication higher on creditworthiness than individuals who do not on the assumption that this behaviour corresponds to caring about details (ibid). The decision-making criterion of using one-word subject lines (ibid) is more likely to negatively affect single parents, individuals with caring responsibilities, individuals who work very long hours, individuals with a disability and individuals who have many children. Such individuals could use a one-word subject line in their communication more frequently due to experiencing greater time pressures.

This discussion corroborates Wachter's claim that the groups that AI generates can combine multiple combinations of protected characteristics (2023: 160-161). Contrary to Wachter, it is not the case that the AI decision-making criterion affects one group more than another group in a particular situation. When formulating this argument, Wachter uses the example of men and women as two groups (ibid). The use of Lenddo could disadvantage women with many children to a greater degree than women who have no children. Similarly, the employment of this software could disadvantage heterosexual women who work long hours to a greater degree than bisexual women who do not work overtime. The use of AI could disadvantage some single parents of different sex to a similar degree. A man with four children could be disadvantaged to a similar degree as a transgender woman with a disability who has two children. It follows that the operation of AI affects numerous individuals with various protected characteristics (ibid), with each individual being affected to a different degree.

Since using the same AI decision-making criterion can disadvantage individuals to different degrees, it could be challenging to identify that the AI decision-making process disadvantages an individual in a particular way based on having a particular protected characteristic. Suppose a woman with four children receives a negative decision and a woman with one child receives a positive decision. In this case, it is difficult to attribute the decision to a woman's sex. Moreover, since individuals with the same protected characteristic are affected differently and individuals with different protected characteristics can be affected similarly, it becomes hard to establish that an AI decision-making criterion affects an individual in a particular way because of having a particular protected characteristic.

The difficulty of capturing how the use of AI affects a diversity of individuals with a particular protected characteristic, such as sex, goes to the heart of how AI operates. AI generates predictions about an individual using data about a group of people (Fawcett & Provost 2013: 107) that AI treats as similar to the individual's data (ibid 24). AI makes sense of each applicant's data based on the correlations it detects between the applicant's data and the data of other individuals (Newell & Marabelli 2015: 5; Mittelstadt 2016: 8). As a result, AI produces a decision based on group characteristics rather than based on those of the applicant (Van Wel & Royakkers 2004: 133; Mittelstadt 2016: 10). Since AI could, for example, allocate individuals who work long hours, some single parents, and individuals with many children to the same group when generating predictions about each person whose data is in that data cluster, it can

be hard for a woman to show that the AI score is based on her sex. It is difficult to argue that the decision is based on a proxy for sex because AI could group male single parents, individuals with a disability and women with no children who work long hours in one group. This discussion shows that it is not always meaningful to talk of having a particular sexual orientation or sex as demarcating group membership in the context of AI. Additionally, the assumption in sections 4, 13 and 19 of the Equality Act 2010 that the treatment can be attached to a protected characteristic of a specific person does not hold in the context of AI, even when the AI decision penalizes the individual for having a protected characteristic.

The inaccurate assumption in the Equality Act 2010 that the negatively affected individual, unfavourable treatment, protected characteristic and group membership have a direct unidirectional relationship interacts with the flawed assumption that the ground of protection can be defined by reference to a discrete protected characteristic and group membership. The interplay between these two assumptions exacerbates the gap in legal protection in the Equality Act 2010 in the context of the use of AI as part of the decision-making process. Consider this example. The website of the company Datrics states that because its AI credit-scoring software detects correlations between thousands of data points, the software "can uncover subtle relationships between seemingly unrelated factors and a person's financial reliability" (Datrics 2024). The website elaborates that: "This score is based on a complex analysis of various factors, including those that may not be immediately obvious, even to financial experts" (ibid). The Datrics software uses credit history, income, transaction analysis, work experience, user behaviour analytics of the applicant and other criteria as a basis for decision-making (ibid).

Since AI can use thousands of data points to generate a decision (ibid), since AI attaches different weights to different data points (Mittelstadt & Ors 2016: 3-4), since one cannot always link the decision-making criterion to a protected characteristic (Binns & Ors 2023: 1856) and since the attributes that the AI uses to generate predictions correspond to the data of different individuals (Fawcett & Provost 2013: 107; Broussard 2020), in some cases the unfavourable treatment will not be based on a discrete attribute, interest or group membership (Wachter 2023: 199). Instead, the decision will be a result of the relationships that the AI detects between the data of different individuals (Fawcett & Provost 2013: 107; Mittelstadt & Ors 2016: 5) in circumstances when such data reflects different aspects of personal identity and circumstances of diverse persons. Zhisheng Chen argues that "statistical discrimination" occurs when AI operates because AI uses historical data about specific

populations to make decisions about particular applicants (Chen 2023: 2). The term "statistical discrimination" refers to individuals using group membership as a proxy to infer missing information, such as someone's work productivity (Tilcsik 2021: 98). Such individuals use their beliefs about a group in order to make predictions about an individual whom they perceive to be a member of that group (ibid).

Furthermore, the barriers that individuals face to accessing opportunities due to their protected characteristics and due to the manner in which social institutions produce social inequality (Torres 2003: 68; Guinier & Torres: 2002) can influence the prediction that the AI generates for an applicant even though the applicant shares either in part or not at all either the protected characteristics or life circumstances of other persons whom the AI treats as being similar (Broussard 2020). Meredith Broussard commented that, most likely, an AI, in part, predicted that a student whose native language was Spanish would fail her International Baccalaureate Spanish examination because it used the historical record of grades from her school as an input (ibid). This student reported that the AI downgraded everyone she knew (ibid). Most of the students who attended this school were racialized and belonged to low-income families (ibid). Broussard's comments point to the fact that, in this case, the AI transferred the social barriers to accessing opportunities that the students of colour from a poor socio-economic background faced onto the applicant who had the advantage of being born to Spanish-speaking parents (ibid).

Since section 4 of the Equality Act 2010 contains a limited list of protected characteristics, it is put forward that this provision does not make it possible to capture the full range of ways in which the attributes (Van Wel & Royakkers 2004: 133) and life circumstances of other individuals that are encoded in the data (Binns & Ors 2023: 1851) can become transferred onto the applicant. For example, section 4 of the Equality Act 2010 does not include socio-economic background. The incident Broussard commented on involved an AI decision-making process penalizing a student for attending a school that had predominantly students from low-income families (Broussard 2020).

The application of direct discrimination in section 13(1) of the Equality Act 2010 to the context of the AI decision-making process poses a challenge because this provision assumes that the applicant is treated negatively solely based on having a protected characteristic. When the adverse treatment is based on the transfer of the attributes and life circumstances of other people onto the applicant (Broussard 2020), the

Amicus Curiae

relationship between the protected characteristic of the applicant and treatment stops being direct and unidirectional. Instead, particular isolated characteristics of the applicant and the characteristics of other individuals may be linked to one another indirectly or not at all. There could be different degrees of relationships between the characteristics of different individuals and the applicant. Such characteristics could include protected characteristics, non-protected characteristics and a mixture of both types of characteristics. Since AI attaches different weights to different correlations between the data (Mittelstadt & Ors 2016: 10), there is arguably a very complex web of relationships between the data of individuals in the same data cluster whom the AI treats as being similar (Fawcett & Provost 2013: 24). Section 13(1) of the Equality Act 2010 arguably renders it challenging to map this social complexity onto the prohibition of direct discrimination in cases where one cannot assign the attributes and life circumstances that are being transferred onto the applicant (Broussard 2020) to a particular protected characteristic.

Another challenge is that the prohibition of direct discrimination by association requires a very high degree of association between the person with a protected characteristic and the person experiencing adverse treatment. Thus, the problem is not confined to that which Weerts and colleagues identify. These scholars posit that the application of the "but for" test gives rise to a challenge in the context of the employment of the AI decision-making process because the relationship between the decision-making criterion and the applicant's protected characteristic can be "elusive" (Weerts & Ors 2024: 5-6). At the heart of the difficulty of applying the prohibition of direct discrimination to the operation of the AI decision-making process is that section 13(1) of the Equality Act 2010 uses the term "because of a protected characteristic" (ibid). Since AI generates predictions about an applicant based on processing the data of a group of individuals (Fawcett & Provost 2013: 107) whom it treats as being similar to the applicant (ibid 24), it is maintained that the prediction will be an outcome of the transfer (Broussard 2020) of an amalgamation of different encodings of life circumstances in the data (Binns & Ors 2023: 1851) belonging to many individuals onto the applicant. It is arguably challenging to trace the relationship between the AI's prediction and the impact of the transfer of the protected characteristics and life circumstances of other individuals onto the applicant (Broussard 2020). According to Binns (2024), the current state of knowledge in computer science makes it impossible to determine how sources of social inequality influenced a particular applicant's ability to obtain a favourable AI decision. The lack of techniques in computer science to map how the life

circumstances of other candidates interact with the sources of societal inequality and AI to lower the applicant's score creates challenges for applying the prohibition of direct discrimination by association.

Multiple decision-making criteria that the AI uses can interact to inhibit access to an opportunity for an applicant (Weerts & Ors 2024: 6). Since there are thousands of data points that the AI uses (Datrics 2024), there are many ways in which the decision-making criteria can interact with the applicant's data and the data of all other applicants to lower the applicant's score. It is suggested that the assumption that there is a unidirectional correspondence between the treatment and a person's protected characteristic in the Equality Act 2010 gives rise to challenges for applying the prohibition of direct discrimination by proxy (Prince & Schwarcz 2020: 1261) to the following cases. In such cases, the treatment is arguably due to the existence of numerous different degrees of association between the applicant's data and the data of other individuals (Wachter 2023: 200) who either possess protected characteristics, belong to new AI-generated disadvantaged groups (ibid 153) or who experience disadvantage due to societal inequality.

Section 14(3)(b) of the Equality Act 2010 exacerbates this difficulty. This provision requires the plaintiff to establish that there was direct discrimination on the basis of each protected characteristic in isolation if the applicant wishes to demonstrate direct discrimination on the basis of a combination of protected characteristics. The focus on the relationship between a single protected characteristic and the unfavourable treatment does not allow one to establish the manner in which unfavourable treatment occurred in the context of the employment of AI as part of the decision-making process. The AI uses decision-making criteria that have a complex relationship with thousands of data points (Datrics 2024) relating to both the applicant and to other individuals (Fawcett & Provost 2013: 107) whom the AI treats as being similar to the applicant (ibid 24). The prohibition of direct discrimination in section 13(1) of the Equality Act 2010 assumes that the less favourable treatment is "because of" the applicant's protected characteristic. Yet, the decision can be due to a complex interaction between the decision-making criteria (Weerts & Ors 2024, 6), the applicant's data and the transfer of attributes of other applicants onto the applicant (Broussard 2020) in circumstances when the decision cannot be traced directly either to the applicant's or to another person's protected characteristic. As Gianna Seglias explains, AI may use a criterion to produce a decision that is related to the protected characteristic without mapping onto that protected characteristic in an exact manner (2021: 66-67). Furthermore, Weerts and colleagues point

to the fact that it is not always possible to establish whether an input acted as a proxy for a protected characteristic (Weerts & Ors 2024: 9).

The fact that the definition of indirect discrimination in section 19 of the Equality Act 2010 requires an applicant to identify a pool of persons with whom that applicant shares the protected characteristic and whose situation is alike (*Grundy v British Airways plc* 2007: paragraph 28) leads to a gap in protection. This occurs arguably in cases where the attributes and life circumstances that the AI transfers onto the applicant (Broussard 2020) either do not correspond to the applicant's protected characteristic or are not very similar to the situation of the applicant. As was shown above, there could be instances when AI groups data of individuals into the same group even though the individuals have different life circumstances (ibid). In such cases, it is suggested that it will be difficult for an applicant to demonstrate that the applicant was in the same situation as another person whose data the AI uses to generate predictions about the applicant (Fawcett & Provost 2013: 107).

[D] POSSIBLE SOLUTIONS

Given the fact that there are multiple complex interrelationships between the correlations in the data, the affected individual and the disadvantageous outcome, it is desirable to have multiple tests defining what constitutes discrimination in the context of AI. Each of these tests can aim to capture different aspects of the way in which the employment of the AI decision-making process disadvantages individuals. The courts should be able to apply these tests either separately or cumulatively, depending on the situation, in order to address the harm of discrimination. What matters is whether applying each test separately or cumulatively allows the court to better capture the fact that the employment of AI as part of the decision-making process disadvantaged an applicant.

The first test should focus on the process entailed in developing an AI and in AI producing a decision. The second test should focus on the effect of the decision on the applicant or on the type of harm that the prohibition of discrimination is designed to capture. It is put forward that the first test needs to reflect the fact that the subjective decisions that the developer makes when constructing the AI (Mittelstadt & Ors 2016: 2), the target variable that AI is aiming to predict (Barocas & Selbst 2016: 679-680), the process of optimization underpinning the AI (Badar & Ors 2014: 39), the data of the individual (Fawcett & Provost 2013: 24), the data of individuals whom the AI treats as being similar (ibid 107) as well as all the data as a whole (Mittelstadt & Ors 2016: 6) shapes the decision

outcome. The second test needs to focus on the harmful outcome. One should be able to determine whether the harm in question occurred either by reference to a particular person, by reference to a group of persons, or by reference to both.

In addition to embedding these two tests into the Equality Act 2010, it is necessary to rethink and to revise a number of current provisions in the Equality Act 2010. The scholarship of Weerts and colleagues and Wachter points to the fact that one needs to revise how section 4 of the Equality Act 2010 defines a protected characteristic. Weerts and colleagues showed that multiple variables can constitute a complex proxy for a protected characteristic when the AI uses these variables together to produce a prediction (2024: 5). One cannot always know what combination of variables will act as a complex proxy (ibid 8). Moreover, it can be difficult to disentangle the impact of a complex proxy on the decision from the effect arising from the AI using other attributes that are associated with having protected characteristics to generate a decision (ibid 9). Wachter demonstrated that AI can generate new categories of disadvantaged groups (2023: 153) that are not meaningful to a human being (ibid 159).

Custers objects to expanding the list of protected characteristics as a solution to these types of problems (2023: 12). His scepticism is premature. It is necessary to reconceptualize the concept of a protected characteristic and to revise section 4 of the Equality Act 2010 so that it encompasses the possession of any attribute or group of attributes. One can draw inspiration from Article 26 of the International Covenant on Civil and Political Rights 1976. This provision contains a non-exhaustive list of protected characteristics. However, unlike Article 26 of the International Covenant on Civil and Political Rights 1976, the Equality Act 2010 should not have the phrase "on any ground such as" followed by a list of protected characteristics. Rather, it is suggested that it should be sufficient that either an attribute that has predictive value or a group of attributes that have predictive value (Weerts & Ors 2024: 5) reduce an applicant's access to an opportunity by contributing to structural inequality. The concept of structural inequality recognizes that institutions play a role in creating social hierarchies and in producing inequality (Torres 2003: 68; Guinier & Torres: 2002). All such attributes and variables that AI uses to produce a decision should constitute protected characteristics. As a corollary, these attributes and variables, either on their own or in combination, will constitute grounds for legal protection. This discussion points to the fact that it is necessary to reconceptualize the term "protected characteristic"

as having multiple dimensions and as being embedded in a broader social context.

This argument is based on the work of Lily Hu and Shreya Atrey. Hu calls for a recognition that the use of AI can "reinscribe" existing patterns of inequality (Hu 2023: 7). She argues that AI's use of features that are correlated with race should constitute discrimination if these features result in entrenching the subordinate social position of racialized individuals by subjecting them to a "matrix of privileging and subordinating social relations" (ibid 16-17). Hu's proposition should be arguably extended beyond race to encompass any attribute or a group of attributes (Weerts & Ors 2024: 5) that have a role in creating subordinating social relations or in contributing to enacting structural inequality. It is important to include structural inequality in the analysis because Atrey explains that race discrimination "has a clear link with racism" (2021: 2). According to Gerard Torres, the same is the case for gender-based and other types of discrimination (2003: 68). On this approach, the AI's use of particular variable or group of variables (Weerts & Ors 2024: 5) to generate a decision involves the use of a protected characteristic in the following circumstances. The variable or group of variables (ibid) should play a role in the process of creating social hierarchies between individuals who belong to diverse groups (Atrey 2021: 4-5) or in giving rise to "subordinating social relations" (Hu 2023: 17) that disadvantage individuals (Atrey 2021: 9) or in contributing to how social institutions create inequality (Torres 2003: 68) or in giving rise to a new social institution (Krupiy 2020: 2) that gives rise to inequality.

It is crucial to rethink the relationship between the protected characteristic and group membership. The term group membership in the Equality Act 2010 needs to be rethought as having three dimensions. First, it should be possible to define group membership by reference to having either single, multiple, or both single and multiple diverse attributes in common. Second, individuals belonging to the same group can share each attribute to a different degree. It should be sufficient for an association between the attribute that an individual has and the attribute that group members share to have a degree of impact on the decision that is tangible. For example, an association of 30% can suffice. Third, it should be irrelevant that individuals who belong to a group cannot be ordered or partitioned in a coherent and logical manner. What should be relevant is whether an attribute or a group of attributes has a role in sustaining or in creating structural inequality. As was already explained, Hu's scholarship (2023: 16-17) provides some insights into

the circumstances when the use of a particular attribute contributes to structural inequality.

Fourth, there can be shifting and changing intra-group variations within the group that do not deprive the group of its unifying character. The group membership criterion will be satisfied even when the disadvantageous treatment attaches to characteristics of a group that are shifting and do not fall into a sequential or coherent pattern. For example, the core and stable dimension could be the membership of individuals who have natural or prosthetic feet. The shifting and flexible dimension of this group membership could be either having a particular shoe size or falling within a range of shoe sizes that correspond to a greater likelihood of AI generating a negative decision. It should be irrelevant that the pattern of shoe sizes of individuals who are disadvantaged by the AI decision-making process does not fall into a coherent or logical pattern. The judges should be able to draw artificial boundaries between individuals with different shoe sizes and partition them into sub-groups for every different AI application in order to identify individuals with a particular shoe size or variety of non-sequential shoe sizes who are more likely to receive a negative decision in the context of an AI decision-making process on the basis of their shoe size. This is the case because different AI applications could disadvantage individuals with different shoe sizes or combinations of shoe sizes. If the employment of AI disadvantages an individual based on shoe size and another attribute, such as a preference for buying decaffeinated coffee, then these two attributes can count as belonging to a group that has multiple dimensions.

In order to implement this proposition, one could modify the existing position that judges identify a pool for comparison that "suitably tests the particular discrimination complained of" (*Grundy v British Airways plc* 2007: paragraph 27). It is put forward that judges need to be able to define protected characteristics and corresponding group membership by reference to whether such protected characteristics have a role in disadvantaging an individual by giving rise to or contributing to structural inequality. The requirement that individuals should be in a position that is not "materially different" in order to constitute a group (ibid para 33) should be revised for the context of the AI decision-making process. In addition to focusing on the attributes of the applicants, the judges need to focus on evaluating the process entailed in the AI decision-making process, how the AI decision-making process interacts with the social context and the outcome. This is particularly the case because the operation of the AI decision-making process will give rise to biases due to

being in feedback loops with the dynamic cultural and social environment (Friedman & Nissenbaum 1996: 336; Whittaker & Ors 2018: 27-28).

Furthermore, the Equality Act 2010 needs to be amended to recognize that varying degrees of relationships between the negative AI decision disadvantaging the candidate and the protected characteristic suffice. The same is the case for group membership. Contrary to Weerts and colleagues, the threshold that there be a "significant influence" on the decision (2024: 6) is too high. It should be sufficient for the influence to be material, meaningful and not spurious. The judges will need to be able to hear from experts from different fields in order to aid them in carrying out this analysis. The computer science community will need to create tools that facilitate the conduct of this type of inquiry.

The proposed approach differs from the current assumptions underlying the provisions of the Equality Act 2010. This legislation assumes that a protected characteristic corresponds to group membership that can be neatly subdivided into subgroups using ordering criteria, such as a sequential pattern. By way of example, section 9 of the Equality Act 2010 arguably assumes that one can define group membership, such as race, by partitioning the group into sections or subgroups sequentially by reference to the shade of skin colour (UK Government 2024: paragraph 50). Section 9(1) of the Equality Act 2010 states that the term race includes skin colour. Section 9(2)(a) of the Equality Act 2010 states that "a reference to a person who has a particular protected characteristic is a reference to a person of a particular racial group". These two provisions make it clear that race corresponds to belonging to a group and that this group can be subdivided into subgroups. The key phrase denoting this meaning in section 9(2)(a) of the Equality Act 2010 is "a person of a particular racial group".

The validity of this interpretation can be gleaned from the explanatory notes to the Equality Act 2010. The explanatory notes state that, "Colour includes being black or white" (UK Government 2024: para 50). The explanatory notes confirm that the Equality Act 2010 recognizes that all individuals are on a spectrum of having varying skin colours and that the term race applies to individuals of all skin colours (ibid). However, section 9(2)(a) of the Equality Act 2010 artificially draws boundaries between different shades of skin colour. It positions different skin colours as existing on a spectrum and as belonging to particular sub-groups. In order to demonstrate disadvantage, the individual needs to show to which subgroup that individual belongs under section 9(2)(a) of the Equality Act 2010 based on that person's shade of skin colour.

The assumption that groups are stable and can be understood using logic in relation to one another in the Equality Act 2010 makes it difficult to account for groups that have both a dimension of stability and a dimension of being shifting. Consider an AI decision-making process that disadvantages individuals with multiple different shoe sizes. Individuals with shoe size 37 could get a 5% lower performance prediction for a job role. The AI could predict that individuals with a shoe size 40 are 10% less likely to be successful employees, that individuals with a shoe size 38 are 12% less likely to be suitable for the role and that individuals with a shoe size 44 are 10% less likely to perform well. The current logic in the Equality Act 2010 that sub-groups can be ordered sequentially on a spectrum poses challenges in this case. Defining everyone with shoe size 40 as a distinct stable sub-group does not account for the fact that individuals with shoe size 44 experience the same degree of disadvantage. Defining individuals with shoe sizes 40 and 44 as the same sub-group can lead to a situation where one allocates individuals with shoe sizes 37 and 38 to the same group even though these individuals experience different degrees of disadvantage. This discussion points to the need to examine the assumptions underlying the provisions of the Equality Act 2010 that have a relationship with how these provisions conceive of the protected characteristic and group membership. All provisions that either define, elaborate on, or incorporate the concepts of the protected characteristic and group membership need to be revised in order to enable the Equality Act 2010 to capture the manner in which the operation of the AI decision-making process gives rise to disadvantage.

Finally, the new definitions of discrimination that policymakers formulate would need to build on aspects of the prohibition of direct discrimination and the prohibition of indirect discrimination. The relationship between the decision, the affected individual, the attribute that is an input into the AI decision-making process and the data of other individuals that AI uses to generate a prediction (Fawcett & Provost 2013: 107) can be complex and non-unidirectional. As a result, it is arguably necessary to evaluate how individual correlations between any two data points, the correlations between data points of individuals whom AI allocates to the same group (ibid) and correlations between all data points in the model as a whole impacted on the ability of the applicant to access an opportunity (Mittelstadt & Ors 2016: 6). This approach would result in individuals being protected from discrimination based on the attributes and group membership of all individuals whose data the AI uses to generate a prediction.

This argument builds on the work of Linnet Taylor. Taylor proposes that the right to privacy needs to be extended from the individual to all individuals wherever they may be located in the context of the use of digital technology (2018: 105). In the context of the prohibition of discrimination, one arguably needs to think about how the processing of all the data of individuals whom the AI treats as being similar (Fawcett & Provost 2013: 107) disadvantages an individual, contributes to sustaining structural inequality (Hu 2023: 7) or creates new forms of disadvantage (Binns & Ors 2023: 1856; Wachter 2023: 153). Additionally, one needs to evaluate how the correlations between all data within the AI model interact to disadvantage an individual, contribute to sustaining structural inequality (Hu 2023: 7), or create new forms of disadvantage (Binns & Ors 2023: 1856; Wachter 2023: 153).

About the author

Tetyana (Tanya) Krupiy *is a lecturer at Newcastle University. Her research addresses how English law and norms of public international law can respond to technological innovation. She has expertise in AI and robotics. In the past, she received funding from the Social Sciences and Humanities Research Council of Canada to carry out a postdoctoral fellowship at McGill University in Canada. Oxford University Press, University of Melbourne and the European University Institute in Florence, among other publishers, feature her work.*

Email: tanya.krupiy@newcastle.ac.uk.

References

Atrey, Shreya. "Structural Racism and Race Discrimination." *Current Legal Problems* 74 (2021): 1-34.

Badar, Altaf & Ors. "Study of Artificial Intelligence Optimisation Techniques Applied to Active Power Loss Minimisation." *IOSR Journal of Electrical and Electronics Engineering* (2014) 39-45.

Barocas, Solon & Andrew Selbst. "Big Data's Disparate Impact." *California Law Review* 104 (2016) 671-732.

Bary, Emily. "How Artificial Intelligence Could Replace Credit Scores and Reshape How We Get Loans." Marketwatch 29 October 2018.

Binns, Reuben & Ors. "Legal Taxonomies of Machine Bias: Revisiting Direct Discrimination." Paper presented at ACM FAccT Conference, Chicago, 12 June 2023.

Binns, Reuben. 'Private conversation with Tetyana Krupiy.' University of Oxford, 14 May 2024.

Broussard, Meredith. "When Algorithms Give Real Students Imaginary Grades." *New York Times*, 8 September 2020.

Butlin, Sarah. "The UN Convention on the Rights of Persons with Disabilities: Does the Equality Act 2010 Measure up to UK International Commitments?" *Industrial Law Journal* 40(4) (2011): 428-438.

Central Digital and Data Office. "Guidance Data Ethics Framework: Glossary and Methodology" (nd).

Chen, Zhisheng. "Ethics and Discrimination in Artificial Intelligence-enabled Recruitment Practices." *Humanities and Social Sciences Communications* 10 (2023): 1-12.

Connolly, Michael (2018) "Easy Cases Making Bad Law: The English Judiciary, Discrimination Law and the Statutory Interpretation." PhD diss, Faculty of Laws, University College London.

Connolly, Michael. "Perceived Disability Discrimination and the Deficient Equality Act: Interpretive and Legislative Remedies." *Industrial Law Journal* 52(3) (2023): 635-664.

Custers, Bart. "Reconsidering Discrimination Grounds in the Data Economy: An EU Comparison of National Constitutions." *Computer Law and Security Review* 50 (2023): 1-13.

Datrics. "AI Credit Scoring: the Future of Credit Risk Assessment" 2024.

Department for Science, Innovation and Technology, and Office for Artificial Intelligence. "AI Regulation: A Pro-innovation Approach." Norwich: His Majesty's Stationery Office, 2023.

Equality and Human Rights Commission. "Equality Act 2010 Code of Practice: Employment Statutory Code of Practice." Norwich: The Stationery Office, 2011.

Fawcett, Tom & Foster Provost. Data Science for Business. Sebastopol: O'Reilly Media Inc, 2013.

Friedman, Batya & Helen Nissenbaum. "Bias in Computer Systems." *ACM Transactions on Information Systems* 14(3) (1996): 330-347.

Guinier, Lani & Gerard Torres. *The Miner's Canary: Enlisting Race,*

Resisting Power, Transforming Democracy. Cambridge: Harvard University Press, 2002.

Hofmann, Valentin & Ors. "Dialect Prejudice Predicts AI Decisions about People's Character, Employability and Criminality." arXiv:2403.00742 (2024): 1-54.

Hu, Lily. "What is 'Race' in Algorithmic Discrimination on the Basis of Race?" *Journal of Moral Philosophy* 21 (2023): 1-26.

Jurca, Ales & Ors. "Analysis of 1.2 million Foot Scans from North America, Europe and Asia." *Scientific Reports* 9 (2019): 1-10.

Krupiy, Tetyana (Tanya). "A Vulnerability Analysis: Theorising the Impact of Artificial Intelligence Decision-making Processes on Individuals, Society and Human Diversity from a Social Justice Perspective." *Computer Law and Security Review* 38 (2020): 1-25.

Kyle, Peter. "AI Opportunities Action Plan." Written Questions, Answers and Statements, UK Parliament, 26 July 2024.

Martínez-Plumed, Fernando & Ors. "The Facets of Artificial Intelligence: A Framework to Track the Evolution of AI." Paper presented at IJCAI'18: Proceedings of the 27th International Joint Conference on Artificial Intelligence, Stockholm, 13–19 July 2018.

Maxwell, Jack & Joe Tomlinson. "Proving Algorithmic Discrimination in Government Decision-making." *Oxford University Commonwealth Law Journal* 20(2) (2020): 352-360.

Mittelstadt, Brent & Ors. "The Ethics of Algorithms: Mapping the Debate." *Big Data and Society* 3(2) (2016): 1-21.

Monagham, Karon. "The Equality Bill: A Sheep in Wolf's Clothing or Something More?" *European Human Rights Law Review* 4 (2009): 512-537.

Newell, Sue & Marco Marabelli. "Strategic Opportunities (and Challenges) of Algorithmic Decision-making: A Call for Action on the Long-term Societal Effects of 'Datification'." *Journal of Strategic Information Systems* 24(1) (2015): 3-14.

Prince, Anya & Daniel Schwarcz. "Proxy Discrimination in the Age of Artificial Intelligence and Big Data." *Iowa Law Review* 105 (2019): 1257-1318.

Saunders, Hannah. "Difficult Distinctions in Anti-discrimination Law: Disfigurement, Appearance and Disability." *International Journal of Discrimination and the Law* 20(1) (2020): 21-44.

Seglias, Gianna. "Bias and Discrimination in Opaque Automated Individual Risk Assessment Systems: Challenges for Judicial Review under the Equality Act 2010." *Oxford University Undergraduate Law Journal* (2021): 53-79.

Stacey, Kiran. 2023. "UK Risks Scandal over 'Bias' in AI Tools in Use across Public Sector." *The Guardian* 23 October 2023.

Taylor, Linnet. "On the Presumption of Innocence in Data-driven Government: Are We Asking the Right Question?" In *Being Profiled: Cogitas Ergo Sum,* edited by I Baraliuc & Ors, 101-107. Amsterdam: Amsterdam University Press, 2018.

Tilcsik, Andras. "Statistical Discrimination and the Rationalisation of Stereotypes." *American Sociological Review* 86(1) (2021): 93-122.

Torres, Gerard. "The Ecology of Justice: The Relationship between Feminism and Critical Race Theory." In *Directions in Sexual Harassment Law,* edited by C MacKinnon & R Siegel, 67-73. New Haven: Yale University Press, 2003.

UK Government. "Equality Act 2010: Explanatory Notes." 2024.

Van Wel, Lita & Lamber Royakkers. "Ethical Issues in Web Data Mining." *Ethics and Information Technology* 6 (2004): 129-140.

Wachter, Sandra. "The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-discrimination Law." *Tulane Law Review* 97(2) (2003): 149-204.

Weerts, Hilde & Ors. "Unlawful Proxy Discrimination: A Framework for Challenging Inherently Discriminatory Algorithms." Paper presented at ACM FAccT Conference, Rio de Janeiro, 5 June 2024.

Whittaker, Meredith & Ors. "AI Now Report 2018." New York: AI Now Institute, 2018.

Legislation, Regulations and Rules

Equality Act 2010

International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171

Cases

EAD Solicitors LLP and Others v Abrams [2015] IRLR 978, [2015] UKEAT/0054/15

Essop and Others v Home Office [2017] All ER 551, [2017] UKSC 27

Grundy v British Airways plc [2008] IRLR 74, [2007] EWCA Civ 1020

Higgs v Farmor's School No 3 [2023] IRLR 708, [2023] EAT 89

Lee v Ashers Baking Company Ltd [2018] IRLR 1116, [2018] UKSC 49

Natasha Allen v Primark Stores Ltd [2022] IRLR 644, [2022] EAT 57

Onu v Akwiwu and Taiwo v Olaigbe [2016] IRLR 719, [2016] UKSC 31

Preddy v Bull [2014] 1 All ER 919, [2013] UKSC 73