

Algorithmic discrimination at work

Aislinn Kelly-Lyth

Researcher, Bonavero Institute of Human Rights, University of Oxford, Oxford, UK

European Labour Law Journal

2023, Vol. 14(2) 152–171

© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/20319525231167300

journals.sagepub.com/home/ell



Abstract

The potential for algorithms to discriminate is now well-documented, and algorithmic management tools are no exception. Scholars have been quick to point to gaps in the equality law framework, but existing European law is remarkably robust. Where gaps do exist, they largely predate algorithmic decision-making. Careful judicial reasoning can resolve what appear to be novel legal issues; and policymakers should seek to reinforce European equality law, rather than reform it. This article disentangles some of the knottiest questions on the application of the prohibition on direct and indirect discrimination to algorithmic management, from how the law should deal with arguments that algorithms are ‘more accurate’ or ‘less biased’ than human decision-makers, to the attribution of liability in the employment context. By identifying possible routes for judicial resolution, the article demonstrates the adaptable nature of existing legal obligations. The duty to make reasonable accommodations in the disability context is also examined, and options for combining top-level and individualised adjustments are explored. The article concludes by turning to enforceability. Algorithmic discrimination gives rise to a concerning paradox: on the one hand, automating previously human decision-making processes can render discriminatory criteria more traceable and outcomes more quantifiable. On the other hand, algorithmic decision-making processes are rarely transparent, and scholars consistently point to algorithmic opacity as the key barrier to litigation and enforcement action. Judicial and legislative routes to greater transparency are explored.

Keywords

equality, discrimination, technology, employment

Corresponding author:

Researcher, Bonavero Institute of Human Rights, University of Oxford, Oxford, OX1 2JD, UK.

E-mail: akellylyth@llm20.law.harvard.edu

I. Introduction

Biases in automated decision-making systems (ADMS) are now well-documented,¹ and algorithmic management tools are no exception. The debate on how to deal with this phenomenon has developed at speed in the legal and policy literature. A report for Equinet in 2020 noted that the European debate was at an ‘embryonic’ stage,² but a report for the European Commission in the same year found that a ‘majority of countries’ in the Union had witnessed ‘at least some degree of public discussion on issues of AI and discriminatory risks’.³ While the discussion has since matured, the question of how to deal with algorithmic discrimination remains an open one. Legislators in the US and EU have tabled proposals which would tackle the issue tangentially, by imposing new requirements with a view to reducing the risk of bias.⁴ Neither proposal would alter the framework of discrimination law itself. In the UK, a 2018 proposal to deal with discriminatory automated decision-making in the workplace was resisted by the Government on the basis that ‘the Equality Act [2010] already protects workers against direct or indirect discrimination by computer or algorithm-based decisions.’⁵

The employment sector has long been a key forum for the development of equality law and has emerged as a focal point for academic discussion on algorithmic discrimination.⁶ This contribution adopts the same focus to assess the capacity of EU law to respond to algorithmic discrimination: do gaps in the framework exist, and if so, how best should they be filled? Section 2 argues that the concepts of direct and indirect discrimination are already apt to cover many cases of algorithmic bias and explains how some apparently novel doctrinal questions are in fact examples of how algorithmic discrimination ‘shines a new light on...“traditional” problems...of EU gender equality and non-discrimination law’.⁷ Legislative proposals to deal with algorithmic discrimination should therefore aim to reinforce, rather than reform, discrimination law. While the rules on direct and indirect discrimination are not specific to any protected characteristic, employers bear a particular duty with respect to persons with disabilities: the duty to make reasonable accommodations. Section 2 goes on to consider the extent to which that duty can be met in a context of algorithmic

-
1. See Cathy O’Neill, *Weapons of Math Destruction* (Crown Publishers 2016); for a seminal piece in the legal literature, see Solon Barocas and Andrew D Selbst, ‘Big Data’s Disparate Impact’ (2016) 104 California Law Review 671.
 2. R Allen and D Masters, *Regulating for an Equal AI: A New Role for Equality Bodies—Meeting the New Challenges to Equality and Non-discrimination from Increased Digitalisation and the Use of Artificial Intelligence* (Equinet 2020) 30.
 3. Janneke Gerards and Raphaële Xenidis, ‘Algorithmic Discrimination in Europe: Challenges and Opportunities for Gender Equality and Non Discrimination Law’ (Publications Office of the European Union 2020) 90.
 4. HR 6580—117th Congress (2021–2022): Algorithmic Accountability Act of 2022 (2022, March 2), available at <<https://www.congress.gov/bill/117th-congress/house-bill/6580>> accessed 7 December 2022, creating a framework for impact assessments covering elements including algorithmic discrimination; and Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM (2021) 206 final (22 April 2021) (AI Act), referencing the risk of discriminatory algorithms at Recitals 33, 40, 44, and 50 and, *inter alia*, requiring ‘examination in view of possible biases’ as a form of data governance, Art. 10(2)(f).
 5. Data Protection Bill Deb 22 March 2018, cols 209–320.
 6. See, for example, Pauline T Kim, ‘Data-Driven Discrimination at Work’ (2017) 58 William & Mary Law Review 857.
 7. Aislinn Kelly-Lyth, ‘Challenging Biased Hiring Algorithms’ (2021) 41 Oxford Journal of Legal Studies 899; Javier Sanchez-Monedero and others, ‘What Does It Mean to Solve the Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives from the UK on Automated Hiring Systems’ [2020] arXiv:191006144 [cs]. For case study examples from the employment context, see Carsten Orwat, ‘Risks of Discrimination through the Use of Algorithms: A Study Compiled with a Grant from the Federal Anti-Discrimination Agency’ (2020) 30–34.
 7. Gerards and Xenidis (n 3) 9.

decision-making. Section 3 lays out a plan for reinforcement of existing law: the obstacles which information asymmetries pose to litigation must be a critical target for any legislative proposal. Algorithmic impact assessments are proposed as a legislative reform which would enable greater transparency.

2. The robustness of existing law

Computer science scholarship on the unequal effects of ADMS⁸ quickly inspired incisive legal scholarship.⁹ EU law provides that direct discrimination occurs where an outcome is ‘on grounds of’ a protected characteristic, while indirect discrimination can arise when an employer applies a facially neutral provision, criterion or practice (PCP) that puts people with a protected characteristic at a ‘particular disadvantage’.¹⁰ Direct discrimination is outright prohibited in most employment cases, whereas unlawful indirect discrimination does not arise if the use of the PCP can be objectively justified.¹¹

2.1. Indirect discrimination

The use of an algorithm can easily be understood as a provision, criterion, or practice (PCP),¹² and if a job applicant or worker can show that the deployment puts their protected group at a disadvantage—and that they have been, or would be, disadvantaged by its use—then they can make out a *prima facie* case of indirect discrimination, thus shifting the burden to the employer to justify the use. To do so, the employer will first need to identify a legitimate aim which the ADMS has been deployed to achieve. This initial step may well prove uncomplicated. If human decision-making has been automated purely for cost-saving purposes, that is unlikely to be adequate.¹³ However, Hacker points out that

-
8. See, for example, Dino Pedreshi, Salvatore Ruggieri, and Franco Turini, ‘Discrimination-aware data mining’ (2008) Proceedings of the 14th ACM SIGKDD international conference on Knowledge Discovery and Data Mining 560; Cynthia Dwork and others, ‘Fairness through Awareness’ (2012) Proceedings of the 3rd Innovations in Theoretical Computer Science Conference 214.
 9. See, for example, Barocas and Selbst (n 1); Kim (n 6); Crystal S Yang and Will Dobbie, ‘Equal Protection Under Algorithms: A New Statistical and Legal Framework’ (2020) 119 Michigan Law Review 291; Deborah Hellman, ‘Measuring Algorithmic Fairness’ (2020) 106 Virginia Law Review 811; Jason R Bent, ‘Is Algorithmic Affirmative Action Legal?’ (2020) 108 Georgetown Law Journal 803.
 10. Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast) [2006] OJ L204/23 (Gender Equality Directive (recast)) Art. 2; Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation, OJ L 303 (2000) (Employment Equality Directive) Art. 2; Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (Racial Equality Directive) Art. 2.
 11. Direct discrimination can be justified by a genuine occupational requirement, which is a high bar: Gender Equality Directive (recast) Art. 14(2); Racial Equality Directive Art. 4; Employment Equality Directive Art. 4(1). For further discussion of justification and exceptions, see European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European Non-Discrimination Law* (Publications Office of the European Union 2018) 91–109.
 12. See, by analogy, *British Airways v Starmar* [2005] IRLR 862 [17]; *The Government Legal Service v Brookes* [2017] UKEAT/0302/16/RN; *United First Partners Research v Carreras* [2018] EWCA Civ 323.
 13. Case C-243/95 *Hill and Stapleton v Revenue Commissioners and Department of Finance* [1998] ECR I-3739, ECLI:EU:C:1998:298, para 40; Case C-393/10 *O’Brien v Ministry of Justice* [2012] 2 CMLR 25, ECLI:EU:C:2012:110, para 66.

the ‘predictive task of the algorithmic decision-making process itself will often furnish a legitimate aim’.¹⁴ An employer might argue, for example, that the algorithm is being deployed to measure employee success.

If the aim is accepted as legitimate, the focus turns to proportionality. The algorithm must be appropriate for achieving the aim, in that it is a suitable means of achieving the stated aim; and its use must be necessary, in the sense that it is the least discriminatory option available and strikes an acceptable balance between the rights and interests of employer and employee.¹⁵

One promising point should be drawn from this brief summary: where a less biased but equally accurate algorithmic system can reasonably be obtained and deployed, then the employer should take that step.¹⁶ The obligation to use the least biased option arguably also means that employers should ensure that reasonable debiasing efforts have been made.¹⁷ On the other hand, scholars have suggested that the nature of algorithmic decision-making means that the proportionality test will be routinely satisfied for indirect discrimination purposes:¹⁸ a concerning possibility which requires consideration of two questions. First, will courts inevitably hold that predictive algorithms are more ‘accurate’ than humans, and therefore an appropriate and necessary means of achieving a legitimate aim? Second, is use of a biased ADMS permitted—or even required—if it replaces a more discriminatory human?

2.1.1. *More effective than the human alternative*

Hacker notes that ‘that there must not exist similarly effective but less discriminating ways of achieving the same result’ but argues that ‘to the extent that the classifier does possess significant predictive accuracy, its effectiveness will likely surpass any alternative ways...based on human decision making.’¹⁹ He suggests that the entity employing the algorithm will therefore ‘likely be able to convincingly claim that there is no similarly effective alternative to machine learning’.²⁰

Assuming that algorithms with high ‘accuracy’ scores are almost always ‘effective’ or ‘appropriate’ is problematic. Say, for example, an employer uses a machine learning model to give job applicants a ‘suitability score’. The algorithm is trained on previously successful applications, and the features it identifies as being correlated with successful performance include speaking slowly,²¹ using active verbs,²² and playing certain sports.²³ Although the employer has not

14. Philipp Hacker, ‘Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under EU Law’ (2018) 55 CMLR 1143, 1161.

15. Case 170/84 *Bilka - Kaufhaus GmbH v Weber Von Hartz* [1986] ECR 1607, ECLI:EU:C:1986:204, para 36.

16. Case C-83/14 *CHEZ Razpredelenie Bulgaria AD v Komisia za zashtita ot diskriminatsia* [2015] ECR I-00, para 128; *Secretary of State for Defence v Elias* [2006] EWCA Civ 1293, [2006] 1 WLR 3213 [302].

17. Kelly-Lyth (n 6) 907.

18. Pablo Martínez-Ramil, ‘Discriminatory Algorithms. A Proportionate Means of Achieving a Legitimate Aim?’ (2022) 4 *Journal of Ethics and Legal Technologies* 3, concluding that ‘[i]n most scenarios, indirect discrimination produced by [machine learning] systems will pass the proportionality test of the CJEU’.

19. Hacker (n 14) 1162.

20. *ibid.*

21. Loren Larsen, ‘HireVue Assessments and Preventing Algorithmic Bias’ (*hirevue.com*, 21 June 2018) <<https://perma.cc/9D2R-U98Q>> accessed 8 July 2022.

22. Jeffrey Dastin, ‘Amazon scraps secret AI recruiting tool that showed bias against women’ (*Reuters*, 11 October 2018) <<https://perma.cc/328A-UJFM>> accessed 8 July 2022.

23. Dave Gershgorn, ‘Companies Are on the Hook If Their Hiring Algorithms Are Biased’ (*Quartz*, 22 October 2018) <<https://perma.cc/7BCD-9SVZ>> accessed 8 July 2022.

analysed why these traits are correlated with success, it continues to use the algorithm because its predictive accuracy is high—in that those who score well go on to receive positive (human) managerial assessments and are quickly promoted.

If it turns out that female applicants speak more quickly, use more tentative language, and do not play the relevant sports, will a rejected female applicant succeed in a case for indirect discrimination? The employer's aim of recruiting competent workers seems to be legitimate, and deploying the algorithm appears to be an effective means of achieving that aim. On the other hand, one might ask what the algorithm's 'predictive accuracy' really tells us. The 'suitability scores' might be correlated with speedy promotion, but the promotion itself occurs in an unequal world: the algorithm has learned the traits of those who do well at the company, but the human managers could be implicitly biased in favour of men. In other words, the algorithm has been fed data from an unequal world and is accurate at predicting unequal outcomes in that world.²⁴

Deeming such predictions 'appropriate' does not accord with the 'fundamental aim' of EU non-discrimination law, which Wachter, Mittelstadt and Russel describe as being 'not only to prevent ongoing discrimination but also to change society, policies, and practices to...achieve substantive rather than just formal equality'.²⁵ In this way, 'EU non-discrimination law...aims to systematically erode inequalities over time'.²⁶ An algorithm may be accurate in predicting the status quo, but if the status quo is not neutral, then does this really count as 'accuracy'?²⁷

The legal framework provides scope for courts to answer that question in the negative. For example, where an employer uses an algorithmic system which predicts 'culture fit',²⁸ the court could take a step back in the analysis to hold that this is simple maintenance of an unequal status quo, which is not a legitimate aim.²⁹ Alternatively, rather than accepting predictive 'accuracy' as an indicator of appropriateness, the courts could require employers to explain why the features identified by the algorithm are appropriate for achieving the given aim.³⁰

On this latter point, Hacker suggests that the decision-maker is 'not under an obligation to specifically explain the factors that led to the algorithmic output', and there does not appear to be a need

24. Sandra Wachter and others, 'Bias Preservation in Machine Learning: The Legality of Fairness Metrics under EU Non-Discrimination Law' (2020) 123 West Virginia Law Review 735, 762–64 raise concerns about the number of fairness metrics in the fair machine learning literature which are 'bias preserving', in the sense that they are satisfied by a perfect classifier; in other words, the status quo is treated as neutral and if the algorithm is a perfect predictor for this status quo then it is deemed fair. For discussion on self-justifying feedback loops, see Jeremias Adams-Prassl and others, 'Directly Discriminatory Algorithms' [2022] Modern Law Review (Early View), 7–9.

25. Wachter and others (n 24) 745.

26. *ibid* 752. Emphasis supplied.

27. A/HRC/49/52 Special Rapporteur on the rights of persons with disabilities, Gerard Quinn, 'Rights of Persons with Disabilities: Report of the Special Rapporteur on the Rights of Persons with Disabilities' para 62 notes that a recruitment algorithm trained on historic data may 'consolidate societal barriers to the hiring of persons with disabilities'.

28. *King v The Great Britain–China Centre* [1991] EWCA Civ 16, [1991] IRLR 513 at [36]; Ifeoma Ajunwa, 'The Paradox of Automation as Anti-Bias Intervention' (2020) 41 Cardozo Law Review 1671, 1715–17.

29. For analogous reasoning, see C-127/92 *Enderby v Frenchay Health Authority* ECLI:EU:C:1993:313, Opinion of AG Lenz, para 49; see also Case C-167/97 *Regina v Secretary of State for Employment, ex p Nicole Seymour-Smith and Laura Perez* [1999] ECR I-623 para 76; for discussion, see Adams-Prassl and others (n 24) 10–11.

30. Tor-Inge Harbo, 'The Function of the Proportionality Principle in EU Law' (2010) 16 European Law Journal 158, 165. For discussion of the need for the measure to be appropriate for achieving its aim, see Joined Cases C-297/10 and C-298/10 *Sabine Hennigs v Eisenbahn-Bundesamt* and *Land Berlin v Alexander Mai* [2011] ECR I-7965, ECLI:EU:C:2011:560; Case C-555/07 *Seda Küçükdeveci v Swedex GmbH & Co KG* [2010] ECR I-365, ECLI:EU:C:2010:21.

to explain why the factors are relevant because the focus is on the ‘effectiveness of the differentiation practice’.³¹ This may be the approach in existing case law, but it is not required by the legislation. Discussing substantively similar US law,³² Selmi argues that unexplained predictive accuracy should not be enough: where an employer’s justification for algorithmic disparate impact centres on the algorithm’s ability to identify characteristics correlated with job success, ‘it is difficult to see a [US] court accepting such an explanation’ without more, ‘particularly when the algorithm may well have been created on biased data’.³³ European case law could similarly evolve to reject the ‘trust us’ defence³⁴ in relation to algorithmic ‘accuracy’, and take a more stringent approach to ‘appropriateness’.

2.1.2. Less biased than the human alternative

If an aim is legitimate and the algorithm is deemed ‘appropriate’ in meeting the aim, then the court will ask whether the algorithmic option is less biased than the human alternative. Although algorithmic bias is now a well-studied problem, algorithmic systems may still be less biased than the humans they replace,³⁵ and employers may well adopt such tools with a view to fostering diversity and equality.³⁶

Imagine, then, two employers: A and B. Both are aware of discrepancies in outcomes between protected groups in the managerial evaluation system,³⁷ and both are considering whether to adopt an algorithmic management tool to measure employee performance in an objective manner. Both weigh up the evidence about the diversity-enhancing potential of people analytics tools against the evidence of potential discriminatory outcomes. On balance, employer A is concerned about its managers’ implicit biases and decides to replace managerial assessments with an algorithmic tool to score employee performance. The scores will be considered alongside sales metrics when making promotion decisions. Meanwhile, employer B is concerned by reports of algorithmic bias and decides to continue relying on human manager reviews. If the tool which A has begun

31. Hacker (n 14) 1161.

32. Public Law 88-352, Civil Rights Act of 1964 (Title VII) § 703 (k), codified at 42 USC § 2000e-2 (k).

33. Michael Selmi, ‘Algorithms, Discrimination and the Law Essays’ (2021) 82 Ohio State Law Journal 611, 643.

34. *ibid* 641.

35. Manish Raghavan and others, ‘Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices’ (FAT* ’20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, January 2020) <<https://arxiv.org/abs/1906.09208>> accessed 8 July 2022; Sendhil Mullainathan, ‘Biased Algorithms Are Easier to Fix Than Biased People’ *The New York Times* (6 December 2019) <<https://perma.cc/D9BR-KF7N>> accessed 8 July 2022.

36. See, for example, ‘Increase diversity and mitigate bias’ (HireVue, undated) <<https://perma.cc/277C-26KY>>; for discussion, see Keith E Sonderling, ‘How People Analytics Can Prevent Algorithmic Bias’ (*International Association for Human Resources Information Management*, 12 June 2021) <<https://perma.cc/DAR6-SSEC>> accessed 8 July 2022. See also Kelly-Lyth (n 6) 908; Selmi (n 33) 615–17; Ashesh Rambachan and Jonathan Roth, ‘Bias In, Bias Out? Evaluating the Folk Wisdom’ in Aaron Roth (ed), *1st Symposium on Foundations of Responsible Computing (FORC 2020)* (Leibniz International Proceedings in Informatics (LIPIcs), Dagstuhl, Germany, Schloss Dagstuhl–Leibniz-Zentrum für Informatik 2020).

37. Note that where employment outcomes are being affected by managers’ implicit biases, an affected individual will have a direct discrimination claim: see *R (on the application of E) v JFS Governing Body* [2009] UKSC 15, [2010] 2 AC 728 [64]. The fact that direct discrimination covers implicit bias has not been recognised as explicitly at the EU level, but the Court has held, for example, that an employer’s failure to explain its decision to hire a particular candidate could be a relevant factor when determining a direct discrimination case brought by another candidate: Case C-415/10 *Galina Meister v Speech Design Carrier Systems GmbH* ECLI:EU:C:2012:217.

to deploy has poorer scores for women on average, but nonetheless exhibits less bias than the human managers at either firm, how would an indirect discrimination claimant fare against each employer?

First, consider a woman working at employer A. She argues that the use of the tool to inform promotions is a practice which puts women at a particular disadvantage. The employer does not deny that the tool produces lower scores on average for women, but argues that its use is objectively justified: the tool was (i) adopted to achieve the legitimate aim of objectively evaluating workers; (ii) is at least as accurate as the human evaluators, in that those scored highly by the algorithmic tool go on to perform well in their new roles;³⁸ and (iii) is the least biased means of achieving the aim. Any indirect discrimination is outweighed by the importance of the aim.³⁹ As explored above, the employer's justification may fail at stage (ii)—but if the appropriateness of the tool is accepted, will the practice be deemed proportionate?

Hacker suggests that if the employer is using the least biased algorithm available, then the proportionality analysis should 'turn on whether the algorithmic decision-making procedure at least significantly and verifiably reduces bias vis-à-vis other types of (non-algorithmic) decision-making procedures'.⁴⁰ If it does, then 'algorithmic discrimination based on biased training data should... be considered proportionate'.⁴¹ That indeed seems to be the most likely conclusion: following this analysis, employer A—who chose the least biased option available—can justify its use of the tool.⁴²

Let us now consider an equivalently situated woman working at employer B. The employer's practice in this case is the use of human managerial evaluations for making promotion decisions. Female employees at the firm receive poorer evaluations on average, and the claimant has not been promoted in several years despite doing well on all other measures.⁴³ The employer argues that selecting the best employees for promotion is a legitimate aim, and that using (human) managerial evaluations is a proportionate means of achieving that aim. In this case, the result will turn on how the court or tribunal assesses proportionality on the facts. The use of managerial assessments may be appropriate, but is it necessary if deployment of the algorithm is a reasonably practical and equally effective alternative? If the employer recognised the human bias and dismissed the algorithmic alternative while taking no further action, then the court may find that it has failed the objective justification test. If so, one might conclude that the employer is effectively required to automate the process.

In one sense, this outcome makes sense: where reasonably practicable, employers should use the least biased means available for reaching their legitimate goals. On the other hand, the conclusion seems jarring when one considers broader equality goals.⁴⁴ Accepting and potentially even (indirectly) mandating the use of a biased algorithmic system seems to entail an entrenchment, rather than

38. The problems with this stage of the analysis are discussed above, section 2.1.1.

39. The proportionality assessment also entails a balancing exercise in which the organisation's need to impose the PCP is weighed against the discriminatory effect: *Homer v Chief Constable of West Yorkshire* [2012] 3 All ER 1287, [2012] UKSC 15 [24] and *Barry v Midland Bank Plc* [1999] 1 WLR 1465 at 870B, both analysing the EU case law. However, Hacker (n 14) 1164 suggests that where the employer is using the least discriminatory option available, 'there is no real conflict between the interests of the decision maker and those of the protected group', implying that the additional 'balancing' exercise is irrelevant.

40. Hacker, *ibid*.

41. *ibid*.

42. Note that an algorithmic tool 'based on biased training data' might be directly discriminatory: Adams-Prassl and others (n 24). Justification will generally not be possible if so.

43. See, by analogy, *Essop v Home Office* [2017] UKSC 27, [2017] 1 WLR 1343.

44. See text around n 25.

an erosion, of inequality. On closer examination, however, one finds that the employer's failure to use the algorithm will usually only be part of the analysis. In reality, other options are available. The employers could mandate implicit bias training for all employers, for example, or adjust the weighting given to the evaluations in decision-making processes. While employer B might be found to be indirectly discriminating if it does nothing at all about its biased human evaluations, it is unrealistic to imagine that a court would 'require' it to use a specific algorithmic tool. Similarly, employer A would be well-advised to consider implicit bias training for its human managers as an alternative to deploying a biased algorithmic tool; and even if its use of the tool is justified on one occasion, the objective justification regime is not static: a practice which starts off being justified may cease to be so.⁴⁵ In this sense, EU law's broader substantive equality goals are not lost.

2.2. Direct discrimination

The above analysis shows how the indirect discrimination framework is apt to deal with many cases of algorithmic discrimination, despite initial doctrinal concerns. At the same time, indirect discrimination should not be used as a 'conceptual "refuge" to capture the discriminatory wrongs of algorithms'.⁴⁶ Although many European scholars have regarded direct discrimination as comparatively 'less important',⁴⁷ the regime has an important role to play.⁴⁸ Where direct discrimination arises, the practical consequences are significant: while the possibility of justification (and its accompanying complexity) is a crucial component of the indirect discrimination regime, direct discrimination is often straightforwardly unlawful.⁴⁹

There are broadly two types of direct discrimination case which appear in the existing European case law, both of which can be mapped onto examples of algorithmic discrimination.⁵⁰ The first occurs where a criterion used is inherently discriminatory. The CJEU held in 1988 that pregnancy is so closely related to sex, for example, that differential treatment on the basis of pregnancy constitutes sex discrimination even if there was no intention to discriminate against women.⁵¹ If an

45. *IDS Employment Law Handbooks: Vol 4—Discrimination at Work* (Sweet & Maxwell 2022) para 17.15.

46. Gerards and Xenidis (n 3) 9.

47. Raphael Xenidis and Linda Senden, 'EU Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination' in Ulf Bernitz and others (eds), *General Principles of EU Law and the EU Digital Order* (Wolters Kluwer 2020) 19. See generally Frederik J Zuiderveen Borgesius, 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence' (2020) 24 *The International Journal of Human Rights* 1572; Philipp Hacker and others, 'Towards a Flexible Framework for Algorithmic Fairness' [forthcoming] *INFORMATIK2020*; Joshua Simons and others, 'Machine Learning and the Meaning of Equal Treatment' in *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (ACM 21 July 2021) 956; Kelly-Lyth (n 6) 906.

48. Adams-Prassl and others (n 24).

49. The carve-outs are narrow and specific, applying to genuine occupational requirements (Gender Equality Directive (recast) Art. 14(2); Racial Equality Directive Art. 4; Employment Equality Directive Art. 4) and age (Employment Equality Directive Art. 6), for example.

50. *JFS* (n 37) [64]. The CJEU has not been as explicit as the UK Supreme Court, but its case law proceeds broadly on the same lines: see Adams-Prassl and others (n 24) fn 77.

51. Case C-177/88 *Elisabeth Johanna Pacifica Dekker v Stichting Vormingscentrum voor Jong Volwassenen (JV-Centrum) Plus* [1990] ERC I-3941, ECLI:EU:C:1990:383; cf the US case of *General Electric Co. v Gilbert et al.* 429 US 125 (1976) 149, and note the similarities between Justice Stevens' opinion (at 162) and the CJEU decision. For other examples of inherently discriminatory criteria, see C-267/06 *Maruko v Versorgungsanstalt der deutschen Bühnen* ECLI:EU:C:2008:179; Case C-356/09 *Pensionsversicherungsanstalt v Kleist* [2008] ERC I-1757, ECLI:EU:C:2010:703.

algorithm generates a lower score when presented with the name of a girls' school rather than a boys' school, its deployment (it is submitted) will similarly constitute direct discrimination.⁵²

The second type of direct discrimination is subjective. This applies where a protected characteristic affects the decision-maker's decision, whether consciously or not.⁵³ Elsewhere, Adams-Prassl, Binns and I argue that where an algorithm exhibits traits that are equivalent to human subjective bias,⁵⁴ the law must respond in kind. In other words, where an algorithm learns to prefer the traits of a particular group because that group is overrepresented in the training data due to previously biased human decisions, this is equivalent to implicit bias, and must be direct discrimination.

This latter form of discrimination gives rise to a difficult doctrinal question: where does the line fall between direct and indirect discrimination? At section 2.1.1 above, it was suggested that the use of an algorithm which identifies as positive traits slow speech, use of active verbs, and playing of certain sports might not be an appropriate means of achieving a legitimate aim if those traits simply map onto a particular protected characteristic. However, if the algorithm has learned to replicate previous recruiters' biases, then we might instead conclude that this is subjective direct discrimination in an automated form—and the justificatory stage (with its questions of appropriateness) will not even be reached.

2.2.1. Drawing the boundary

Drawing the boundary is difficult because while indirect discrimination looks at the effects of a decision-making process, direct discrimination focuses on reasons.⁵⁵ Rationales for human decisions are often difficult to divine: in the absence of a smoking gun, it will be difficult to prove discrimination on the part of any manager, even if a particular ethnic group consistently performs more poorly in that manager's assessment.⁵⁶ Humans can rationalise their decisions *ex post facto*.⁵⁷ By contrast, the true reasons for algorithmic outputs are often identifiable—and even where an algorithm operates as a black box, its training data can be examined in a way that would never be possible for humans. While automated decision-making therefore provides a promising opportunity to uncover direct discrimination, it also means that the question of whether a decision has been made 'on grounds of' a protected characteristic is likely to become more complex.

Complexities arise even in relation to inherently discriminatory proxies. Which algorithmically identified features will be deemed 'inherently' discriminatory? Gerards and Xenidis note that while a feature such as pregnancy is clearly inherently related to sex, the Court's jurisprudence has not determined whether 'a nearly 100% overlap is required, or [whether] it would enough to show statistically that a certain variable...has a 90% or 80% overlap with a given protected ground'.⁵⁸

52. By analogy, see *Dastin* (n 22).

53. *Nagarajan v London Regional Transport* [2000] 1 AC 501.

54. Ryan Steed and Aylin Caliskan, 'Image Representations Learned with Unsupervised Pre-Training Contain Human-like Biases' (2021) *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* 701.

55. This distinction exists in the case law, although there is no consensus in the literature: Adams-Prassl and others (n 24) 11.

56. It is 'unusual to find direct evidence of racial discrimination' because '[f]ew [decision-makers] will be prepared to admit such discrimination even to themselves': *King v The Great Britain-China Centre* [1991] EWCA Civ 16, [1991] IRLR 513 [36].

57. Ana Carolina, Alfinito Vieira and Alex Graser, 'Taming the Biased Black Box? On the Potential Role of Behavioural Realism in Anti-Discrimination Policy' (2015) 35 *Oxford Journal of Legal Studies* 121.

58. Gerards and Xenidis (n 3) 64.

Similarly, in the context of ‘subjective’ algorithmic discrimination, where the model has apparently learned to replicate implicit bias, it will be necessary to assess whether the features are indeed the result of automated implicit bias or just the ‘ground truth’.⁵⁹ For example, slow speech and active verbs might be genuine indicators of strong sales metrics, but if the employer’s clients prefer men and men use slow speech and active verbs, then the sales success prediction risks collapsing into gender prediction. If the algorithm is effectively predicting who the clients will like, and the clients like men, then using that algorithm for promotion decisions could constitute direct discrimination.⁶⁰ How does one distinguish between merit and societal preference in a systemically unequal world?⁶¹

Finally, existing interpretations of direct discrimination law do not seem to capture all cases in which equivalently situated individuals are treated differently because of an interaction with a protected characteristic: some algorithms simply work less well for certain protected groups because those groups are underrepresented in the training data. If a facial recognition algorithm is trained on photographs which disproportionately show white men, and is consequently worse at recognising black women,⁶² then it seems artificial to deny that impacted black women are being disadvantaged on grounds of race and sex where that algorithm is applied in a real-world context.⁶³

While these conundrums are far from straightforward, they do not challenge the legislative framework of discrimination law in significantly new ways. The ‘100 percent overlap’ question which Gerardis and Xenidis raise in relation to inherently discriminatory proxies has already been a vexed one, even in non-ADMS cases;⁶⁴ and tracking the upstream ‘reasons’ for a decision has always been complex—the difficulty of uncovering true reasons in human cases has simply obscured the challenge. In other words, although ‘algorithmic discrimination blurs the boundaries between the doctrines of direct and indirect discrimination’,⁶⁵ such blurriness has always existed.⁶⁶ Deeper concerns about the continued relevance of the distinction are not unique to algorithmic discrimination.

2.2.2. *Attributing accountability*

The fact that case law has centred on direct discrimination as a harm perpetuated by biased humans also gives rise to a second doctrinal conundrum: the attribution of accountability where bias becomes automated. Liability for AI-driven harms has already been the subject of policy discussion

59. Here again we encounter the challenges posed by treating the ‘ground truth’ as neutral: Wachter and others (n 24).

60. Client preferences are not genuine occupational requirements; in other words, an employer cannot treat people with a particular protected characteristic worse merely because of their clients’ preferences: Case C-188/15 *Asma Bougnaoui and Association de défense des droits de l’homme (ADDH) v Micropole SA* ECLI:EU:C:2017:204.

61. Rediet Abebe and others, ‘Roles for Computing in Social Change’, (2020) Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* ’20) 252 highlight that ‘exposing the limits of algorithmic notions of fairness has exposed the limits of the underlying legal and philosophical notions of discrimination on which [computational fairness] work has built’.

62. Adams-Prassl and others (n 24) explore this in relation to the algorithms studied in the ‘Gender Shades’ project, Joy Buolamwini and Timnit Gebru, ‘Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification’ (2018) 81 Proceedings of Machine Learning Research 77.

63. To check workers’ identities, for example, Chris Vallance, ‘Legal action over alleged Uber facial verification bias’ (*BBC News*, 8 October 2021) <<https://perma.cc/TE4M-AMRH>>; or to identify potential suspects for arrest, *R (on the application of Bridges) v Chief Constable of South Wales* [2020] EWCA Civ 1058, [2020] 1 WLR 5037.

64. Adams-Prassl and others (n 24) 17–18.

65. Gerardis and Xenidis (n 3) 76.

66. In Canada, see *British Columbia (Public Service Employee Relations Commission) v BCGEU* [1999] 3 SCR 3.

at EU level,⁶⁷ and scholars have begun to question the extent to which ‘current regimes of liability...are adequate to tackle algorithmic discrimination’.⁶⁸ Gerards and Xenidis highlight the ‘variety of different players...involved in the stages of algorithmic decision making’, and suggest that:

[I]f at some point a discriminatory outcome is detected (for instance, because an algorithm systematically suggests that men should be promoted to a certain position rather than women), it may be very difficult for the victim of discrimination...to know whom to hold responsible, liable and/or accountable for that discriminatory outcome...(the developers, the sellers or the end user (...the HR service) of the algorithm).⁶⁹

Where an employment decision is affected by discrimination, the employer seems to be the obvious respondent. The analysis is relatively straightforward for indirect discrimination, where the focus is on effects rather than on reasons. If an employer’s practice of using an algorithmic system to make decisions puts a protected group at a particular disadvantage, this is broadly equivalent to the use of a written test which disadvantages a particular group; both instances would require objective justification.⁷⁰

Case law on direct discrimination does more frequently pinpoint a human wrongdoer. ‘Most employers are corporate and most acts of discrimination are done by individual employees or agents’, suggested the UK Employment Appeal Tribunal in one case, before promptly turning to the scope of employer liability for discriminatory human decisions (*viz* vicarious liability).⁷¹ Under certain legal regimes, the need for a culpable individual may pose serious challenges: direct discrimination under Australian law requires a ‘person’ to engage in discriminatory treatment, and it has been suggested that this requirement cannot be satisfied where an ADMS makes a decision without any input from a natural person.⁷² By contrast, the EU discrimination directives adopt a passive voice: the ‘principle of equal treatment’ means that there ‘shall be no direct or indirect discrimination’ when certain employment-related decisions are made.⁷³ While Union-level law does not specify ‘who should be held liable for discriminatory behaviour’,⁷⁴ Member State

67. European Commission and Directorate-General for Justice and Consumers, *Liability for Artificial Intelligence and Other Emerging Digital Technologies* (Publications Office 2019).

68. Gerards and Xenidis (n 3) 97.

69. *ibid* 46–47.

70. For analogous cases from the UK context, see *The Government Legal Service v Brookes* [2017] UKEAT/0302/16/RN; and *Essop* (n 43). See also *Murray v MaClay Murray & Spens LLP* UKEATS/0004/18/JW [24]–[25], in which the UK Employment Appeal Tribunal noted that ‘[c]ounsel was not aware of any cases about personal liability for an employer’s indirect discrimination’; and that where an employee has no choice in applying a PCP, they ‘they are not likely to be liable for any indirect discrimination’ since ‘the PCP has by definition been created by the employer and applied by him’. This distinction should similarly apply to direct discrimination arising through an algorithm created and applied by an employer: an employer may have directly discriminated even if no individual employee is liable.

71. *CLFIS (UK) Ltd v Reynolds* [2015] EWCA Civ 439, [2015] IRLR 562 [13]. For the position in UK law, see Equality Act 2010 s 109; see similarly the Irish Employment Equality Acts 1998–2015, s 10(1).

72. Natalie Sheard, ‘Employment Discrimination by Algorithm: Can Anyone Be Held Accountable?’ (2022) 45 UNSW Law Journal 617, 634, and see reform suggestion at 638.

73. Employment Equality Directive art 2(1).

74. Directorate-General for Justice and Consumers (European Commission) and others, *A Comparative Analysis of Non-Discrimination Law in Europe 2021: The 27 EU Member States, Albania, Iceland, Liechtenstein, Montenegro, North Macedonia, Norway, Serbia, Turkey and the United Kingdom Compared* (Publications Office of the European Union 2022) 53–54.

legislation generally prohibits discrimination on decisions such as promotions or pay, and it seems unlikely that liability is avoided where such discrimination is automated.⁷⁵ Where an employer's paper-based process involves deducting five points from any individual who attended a girls' school, for example, the employer is liable for direct discrimination, regardless of the scorer's mental state.⁷⁶ In other words, primary liability is directly imposed on the employer. Where an employer directly discriminates via an algorithm rather than via a human, the analysis should be no different.⁷⁷

A straightforward approach holding employers liable where regulated decisions are affected by automated direct discrimination neatly sidesteps questions about whether liability is affected by the employer's relationship to the software (which might be developed in-house, purchased or licenced)⁷⁸ and whether liability should continue to attach if the employer has taken reasonable steps to prevent the discrimination.⁷⁹ The analysis is simplified: if a job applicant or worker has been denied a relevant opportunity by the employer on grounds of a protected characteristic (including where a decision-supporting algorithm has effectively decided on grounds of the characteristic), the employer is liable for direct discrimination. Where no direct discrimination has occurred, but the algorithm disadvantages members of the claimant's protected group, the primary question will be whether the employer can justify the use (within the indirect discrimination framework). The analysis thus promptly returns to the fundamental questions set by the discrimination framework.⁸⁰

2.3. Disability discrimination

The law on direct and indirect discrimination applies broadly to protected characteristics,⁸¹ and the challenges set out above are likely to be encountered in a wide range of cases. Alongside those provisions, EU equality law also imposes a further duty on employers which requires consideration in the context of algorithmic decision-making: the duty to make reasonable accommodations for persons with disabilities.⁸² Employers must take 'appropriate measures' in individual cases to enable persons with disabilities to access, participate, and advance in employment,⁸³

75. cf Gerards and Xenidis (n 3) 117 and 118. Employer liability for discrimination against employees is established throughout the Union: Directorate-General for Justice and Consumers (European Commission) and others (n 74) 38–57, although note variations in the scope of protection in edge cases. This is not to say that liability within the supply chain does not require further unpicking.

76. See, by analogy, *R v Birmingham City Council, ex p Equal Opportunities Commission* [1989] AC 1155.

77. See Dastin (n 22) and discussion in Adams-Prassl and others (n 24) 16.

78. See *Nagarajan v London Regional Transport* [2000] 1 AC 501, in which the UK House of Lords considered the position where different employees have been involved at different stages of the recruitment process.

79. See, for example, Irish Employment Equality Act 1998 s 15(3).

80. cf Ljupcho Grozdanovski, 'In Search of Effectiveness and Fairness in Proving Algorithmic Discrimination in EU Law' (2021) 58 Common Market Law Review 99, 126–34, proposing that an absence of human agency should be permitted as a defence for respondents in automated discrimination cases. Were such a proposal adopted, questions such as those asked in relation to vicarious liability would likely be introduced into the analysis.

81. Note that there are certain differences in how the characteristics are treated: for example, Article 6 of the Employment Equality Directive permits justification of differences of treatment on grounds of age in certain cases.

82. Employment Equality Directive Art. 5; Convention on the Rights of Persons with Disabilities Art. 5(3).

83. Employment Equality Directive Art. 5.

and this duty applies unless the measures would ‘impose a disproportionate burden on the employer’.⁸⁴

The significance of this duty in the context of algorithmic decision-making remains underexplored in the literature.⁸⁵ On the one hand, the individualised nature of the duty responds to the heterogeneous nature of disability. On the other hand, underrepresentation in datasets can disadvantage outliers, which is concerning where ‘most social structures and institutions treat disabled people...as “atypical” or “abnormal”’.⁸⁶ In this sense, the ‘heterogeneity of disability’ is ‘incongruous with the current paradigm of machine learning’,⁸⁷ in which predictive algorithms learn and apply patterns. This subsection explores how the application of the reasonable accommodations duty in a context of algorithmic decision-making may require system-level design changes as well as individualised adjustments on a case-by-case basis.

Disability is not defined in the Employment Equality Directive and was initially conceptualised by the CJEU as ‘a limitation which results in particular from physical, mental or psychological impairments and which hinders the participation of the person concerned in professional life’.⁸⁸ This approach was based on the medical model of disability,⁸⁹ which focuses on disability as an impairment which is independent of any external or environmental element.⁹⁰ The United Nations Convention on the Rights of Persons with Disabilities (CRPD), concluded some years after the Employment Equality Directive, marked a shift towards the social model, which looks beyond the individual to understand disability as being ‘based on the interactional relationship between people with impairments and the wider environment’.⁹¹ The EU became a party to the Convention in 2010, and is bound by the CRPD obligations to the extent of its competences.⁹² The definitional shift quickly fed into the EU *acquis* via the Court, with the CJEU drawing closely on the CRPD to reframe its definition of disability as a limitation which ‘in interaction with various barriers may hinder the full and effective participation of the person concerned in professional life on an equal basis with other workers’.⁹³

That disability is to be understood contextually is important from the perspective of algorithmic discrimination. Previous studies have identified system-level algorithmic discrimination against persons with disabilities, such as higher rates of facial expression misinterpretation, poorer performance of speech recognition tools, and the labelling of texts which mentioned disability as ‘toxic’ by

84. *ibid.* The duty to provide reasonable accommodations can thus be compared with the objective justification regime in indirect discrimination: Erica Howard, ‘Reasonable Accommodation of Religion and Other Discrimination Grounds in EU Law’ (2013) 38 *European Law Review* 360.

85. A/HRC/49/52 Special Rapporteur on the rights of persons with disabilities, Gerard Quinn, ‘Rights of Persons with Disabilities: Report of the Special Rapporteur on the Rights of Persons with Disabilities’ [19]. Maarten Buyl and others, ‘Tackling Algorithmic Disability Discrimination in the Hiring Process: An Ethical, Legal and Technical Analysis’ in 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’22, New York, NY, USA, Association for Computing Machinery 21 June 2022) 1073.

86. Meredith Whittaker and others, ‘Disability, Bias, and AI’ (AI Now, November 2019) 12.

87. Buyl and others (n 85) 1079.

88. Case C-13/05 *Sonia Chacón Navas v Eurest Colectividades SA* [2006] ECR I-6467, ECLI:EU:C:2006:456, para 43.

89. Lisa Waddington and Andrea Broderick, *Combatting Disability Discrimination and Realising Equality: A Comparison of the UN CRPD and EU Equality and Non-Discrimination Law* (Publications Office of the European Union 2018) 48.

90. *ibid.* 36.

91. *ibid.* 37. The CRPD is the first human rights Convention to be open to ratification by regional integration organisations, a decision that was made so that the EU could conclude the instrument.

92. *ibid.* 31.

93. Joined Cases C-335/11 and C-337/11 *HK Danmark (Ring and Skouboe Werge)*, ECLI:EU:C:2013:222, para 38.

natural language processing tools.⁹⁴ Addressing these impacts is critical, but system-level mitigations which seek to identify and reduce disparities across defined protected groups are not sufficient in the context of disability as currently conceptualised by the Court, for two reasons.⁹⁵

First, every disability is unique, and an individual ‘may fare poorly on an assessment because of a disability...regardless of how well other individuals with disabilities fare on the assessment’.⁹⁶ Second, under the social model, the definition of disability is itself contextual: if disability is seen as a product of a disabling environment, then a conception of disability discrimination which ‘abstract[s] away environmental and social context’ will ‘artificially limit[] categories of disability’ and ‘unintentionally...reinforce both the form and content of the medical model’.⁹⁷ In other words, system-level impacts may not be identifiable if it is the manner and context of the deployment which is ‘disabling’. The US Equal Employment Opportunity Commission (EEOC) points out, for example, that a tool which predicts performance in ‘typical working conditions’ might not accurately predict ‘whether the individual still would experience those same difficulties under modified working conditions’, such as in a setting with reduced sound.⁹⁸ In other words, impacts for persons with disabilities can arise both from the algorithmic management system itself and from the mode of its deployment, and will be impossible to comprehensively assess those impacts at the system level.

Where possible, individualised reasonable accommodations should include the deployment of in-built accessibility functionalities. In other words, the system-level design should be amenable to individual-level response. The EEOC suggests that in the context of algorithmic hiring software, for example, reasonable accommodations might include ‘extended time or an alternative version of the test, including one that is compatible with accessible technology (like a screen-reader)’.⁹⁹ The creation of such functionalities requires consideration at the design stage, and the reasonable accommodations duty can only impose an obligation to procure or develop algorithmic management software with accessibility functions if it can apply anticipatorily.

The UN Special Rapporteur on the rights of persons with disabilities does suggest that the reasonable accommodation obligation ‘may have an anticipatory dimension, in the sense that one should not have to wait for persons with disabilities to present themselves before considering

94. Nicholas Tilmes, ‘Disability, Fairness, and Algorithmic Bias in AI Recruitment’ (2022) 24 *Ethics and Information Technology* 1, 3.

95. Attempts to analyse impacts on other protected groups at the system level can be similarly reductionist because they adopt a ‘limited and essentialist reading of fluid and socially constructed categories of identity, like race and gender’: Whittaker and others (n 86) 11. Note, on the other hand, that A/HRC/49/52 Special Rapporteur on the rights of persons and with disabilities, Gerard Quinn, ‘Rights of Persons with Disabilities: Report of the Special Rapporteur on the Rights of Persons with Disabilities’ para 68 describes the fluidity, heterogeneity and nuance of disability, and the comparatively binary categorisations adopted by developers, as posing challenges which, while challenging, are ‘surmountable’.

96. US Equal Employment Opportunity Commission, ‘The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees’ (EEOC-NVTA-2022-2, 12 May 2022) <<https://perma.cc/LG25-D53T>>.

97. Tilmes (n 94) 5.

98. US Equal Employment Opportunity Commission (n 96).

99. US Equal Employment Opportunity Commission (n 96). The algorithmic predictions themselves might also be adjustable by removing features which are inappropriate in the particular case: Buyl and others (n 85) 1078–79; and see Reuben Binns and Reuben Kirkham, ‘How Could Equality and Data Protection Law Shape AI Fairness for People with Disabilities?’ (2021) 14 *ACM Trans Access Comput* 17:1, 15 on treatment disparity.

what reasonable accommodation might be warranted'.¹⁰⁰ The accepted position under EU law, however, is that the duty is only understood to be 'activated' once the employer has knowledge of a particular person's disability.¹⁰¹

At the same time, system-level design remains relevant when assessing employer liability. The use of a tool without accessibility functionalities might constitute indirect discrimination if its use unjustifiably disadvantages persons with disabilities and there are equally effective alternatives reasonably available;¹⁰² and employers should bear this in mind when selecting or developing algorithmic tools. Moreover, although in-built accessibility functions will go some way to securing disability equality, they will not be sufficient for all individuals. Reasonable accommodation 'might mean providing alternative...tools to accommodate applicants with disabilities'.¹⁰³

In short, although the pattern-identifying nature of machine learning algorithms poses concerns in the context of disability, the EU legal framework already recognises the need for employers to respond to the uniqueness, fluidity, and context-specificity of disability through the reasonable accommodations obligation. This is not to say that the law as it stands is sufficient to ensure equality: reasonable accommodations are contingent on disclosure; assessing the proportionality of accommodation is often difficult; and 'opting out' of an unsuitable assessment may not be a panacea if the alternative process is in practice inequivalent.¹⁰⁴ However, these issues—while pressing—are again not novel to the algorithmic management context.

3. Getting cases to court

A review of the doctrinal challenges across the equality framework reveals that the EU framework is remarkably robust. Where doctrinal issues arise, they are generally either long-standing unresolved questions, or stem from judicial interpretations of discrimination as a human phenomenon. Other challenges highlighted by scholars—such as the personal and sectoral scope of EU equality law—are not particular to algorithmic discrimination, but arise across the board.¹⁰⁵ In short, '[m] any of the issues relating to the discriminatory potential of algorithms are not unique to algorithms

100. A/HRC/49/52 Special Rapporteur on the rights of persons with disabilities, Gerard Quinn, 'Rights of Persons with Disabilities: Report of the Special Rapporteur on the Rights of Persons with Disabilities' para 42 noting that one 'obvious example' of anticipatory accommodation would arise where impacts of screening or interviewing tools can be 'reasonably anticipated', see also paras 33–35 highlighting impacts in employment.

101. Patrick Daly and Darius Whelan, 'Disability in Employment Equality Law: A Reappraisal of the Reasonable Accommodation Duty and Issues Arising in Its Implementation' (2021) 28 Maastricht Journal of European and Comparative Law 744, 749–53. Legislative initiatives to create an explicit anticipatory reasonable adjustment duty have proven unsuccessful: Anna Lawson and Maria Orchard, 'The Anticipatory Reasonable Adjustment Duty: Removing the Blockages?' (2021) 80 Cambridge Law Journal 308, 309.

102. Daly and Whelan (n 101) 750, referencing the principle of Universal Design and noting that 'a failure to design a workplace that is conducive to the needs of all employees is in itself a form of indirect discrimination', irrespective of whether the disability is disclosed. Algorithmic management impact assessments would also provide a framework for considering accessibility: Aislinn Kelly-Lyth and Anna Thomas, 'Algorithmic Management: Assessing the Impacts of AI at Work' (elsewhere in this issue).

103. A/HRC/49/52 Special Rapporteur on the rights of persons with disabilities, Gerard Quinn, 'Rights of Persons with Disabilities: Report of the Special Rapporteur on the Rights of Persons with Disabilities' para 46.

104. Buyl and others (n 85) 1073, 1075.

105. See, for example, Gerards and Xenidis (n 3) 53–62, highlighting, *inter alia*, the hierarchy of protection for different protected characteristics in different domains, and the lack of clarity about non-discrimination protection for platform workers; and 65–66 on intersectional discrimination.

but are problems that have been a staple of antidiscrimination law'.¹⁰⁶ The fact that no new legislation has been adopted at Member State level may well reflect a general consensus that existing legislation 'has a sufficiently wide material scope to cover most examples of algorithmic discrimination'.¹⁰⁷

If the law on the books is adequate, the focus turns to enforcement. Getting cases of algorithmic discrimination before courts is critical for two reasons. First, and most obviously, if equality law is not enforceable, then fundamental rights are not protected. Second, while the doctrinal issues examined above do not pose a novel threat to the framework of EU equality law itself, they do require judicial consideration—and case law will only be updated if cases are brought.

Algorithmic discrimination gives rise to a concerning paradox when it comes to enforceability. On the one hand, automating previously human decision-making processes can render discriminatory criteria more traceable and outcomes more quantifiable. The shift towards algorithmic management should therefore hold significant potential for addressing persistent inequalities in the labour market, including those which stem from implicit bias.¹⁰⁸ On the other hand, algorithmic decision-making processes are rarely transparent, and scholars consistently point to algorithmic opacity as the key barrier to litigation and enforcement action.¹⁰⁹ The 'prevailing non-discrimination model' in EU law is an 'adjudicative adversarial system based on individual litigation', which means that the 'burden of uncovering discrimination...and bringing a case to court lies with the victim'.¹¹⁰ Proving discrimination without access to information is nearly impossible, and cases on algorithmic discrimination remain rare.¹¹¹

In technical terms, some machine learning algorithms do operate in a 'black-box' manner, such that their functioning is impenetrable even to the developers. While technical opacity is often cited as a novel and intractable issue, the true prevalence of black-box algorithms in the employment context remains unclear.¹¹² Moreover, technical opacity does not render the law ineffective.¹¹³ Knowledge of the algorithm's inner workings is not a prerequisite to bringing a successful case: if a claimant who can show that an algorithmic system is putting persons with their protected

106. Selmi (n 33) 630.

107. Gerards and Xenidis (n 3) 111.

108. Buyl and others (n 85) 1079; Binns and Kirkham (n 99) 27–28; Adams-Prassl and others (n 24); Kelly-Lyth (n 6) 928; Gerards and Xenidis (n 3) 122.

109. Kelly-Lyth (n 6); Zuiderveen Borgesius (n 47) 1583; A/HRC/49/52 Special Rapporteur on the rights of persons and with disabilities, Gerard Quinn, 'Rights of Persons with Disabilities: Report of the Special Rapporteur on the Rights of Persons with Disabilities' para 32; Binns and Kirkham (n 99) 9–10; Sheard (n 72) 632, 646, highlighting that the 'main issue with disability discrimination law is its weak enforcement'; Gerards and Xenidis (n 3) 69, 74, 87, 112, 116, citing various national experts; Orwat (n 6) 72, and Mario Martini, 'Algorithmen als Herausforderung für die Rechtsordnung' (Algorithms as a challenge for the legal system) (2017) 21 *JuristenZeitung* 1017, 1024.

110. Gerards and Xenidis (n 3) 75.

111. *ibid* 113–15; although for a recent example of pending litigation, see *Pa Edrissa Manjag v Uber Eats UK* and others UKEAT Case Number 3206212/2021.

112. See Selmi (n 33) 622, for example, suggesting that 'black-box or inscrutable algorithms . . . currently comprise a relatively small portion' of algorithms in use.

113. Grozdanovski (n 80) 108 suggests that black-box opacity 'challenges the basic goal of legal evidence, which is to extract the truth about a dispute from the facts presented by the parties'. While this may be true in theory, the law has long accommodated a lack of transparency in apparently discriminatory employer decision-making: Case C-109/88 *Handels- og Kontorfunktionærernes Forbund i Danmark v Dansk Arbejdsgiverforening ex p Danfoss A/S* [1989] ECR 3199, ECLI:EU:C:1989:383

characteristic at a disadvantage, then they have a *prima facie* case of indirect discrimination;¹¹⁴ and if they can show that a similarly situated individual without their protected characteristic would have received more favourable treatment, for example through a counterfactual simulation, then the two facts together should be sufficient to bring a direct discrimination claim.¹¹⁵ In cases of indirect discrimination, if the employer is unable to explain the algorithm's workings, then a court could find that objective justification is impossible.¹¹⁶ In cases of direct discrimination, the employer's inability to provide a non-discriminatory explanation for the outcome will result in a finding in favour of the claimant.¹¹⁷

The problem is therefore not that technical opacity is an insurmountable challenge, but that potential claimants do not have access to the evidence necessary to bring a case. The burden is initially on the claimant to present facts from which discrimination may be presumed,¹¹⁸ and EU law does not directly impose any disclosure obligations on employers.¹¹⁹

3.1. Judicial provision of transparency

Once again, there is scope for judicial reinterpretation to go some way towards addressing this issue. Gerards and Xenidis suggest that the principle of effectiveness in EU law could provide a basis for national courts to 'consider the refusal to disclose information [about algorithmic performance] as contributing to the establishment of a *prima facie* case of discrimination, thus shifting the burden of proof to the respondent.'¹²⁰ The transparency issue is not new: the absence of any right to information has long created problems for claimants in discrimination cases, and one might ask whether the jurisprudence on this point will change in response to automation.¹²¹ The harm of opacity is, however, pronounced in the algorithmic context, as discrimination is easier to detect in technical terms but less likely to be perceived from the perspective of the victim.¹²²

One avenue to enhanced transparency could be through a 'joint reading' of EU non-discrimination law and data protection law.¹²³ Grozdanovski argues in favour of such a reading, relying on Article 22 of the General Data Protection Regulation (GDPR), which provides for safeguards against solely automated decision-making with significant effect. Where an employer decides that making a significant decision—such as the denial of a bonus payment or selection for redundancy—on a solely automated basis is necessary for the performance of the employment contract,¹²⁴ then certain 'suitable safeguards' must be provided, including a data subject right to 'obtain human intervention' and 'contest the decision'.¹²⁵

114. Kelly-Lyth (n 6) 906.

115. Adams-Prassl and others (n 24).

116. As explained at section 2.1.1 above.

117. Adams-Prassl and others (n 24).

118. European Union Agency for Fundamental Rights and Council of Europe (n 11) 234–35.

119. Case C-104/10, *Patrick Kelly v National University of Ireland (University College, Dublin)* [2011] ECR I-6813, ECLI:EU:C:2011:506.

120. Gerards and Xenidis (n 3) 144.

121. See, for example, Case C-415/10 *Galina Meister v Speech Design Carrier Systems GmbH*, ECLI:EU:C:2012:217.

122. Andrea Bonezzi and Massimiliano Ostinelli, 'Can Algorithms Legitimize Discrimination?' (2021) 27 *Journal of Experimental Psychology: Applied* 447.

123. Grozdanovski (n 80) 121.

124. GDPR Art. 22(2)(a).

125. GDPR Art. 22(3).

In many cases, however, employment decisions will not be solely automated: managers may consider algorithmically generated scores as one factor within a broader evaluation, for example, and Article 22 will not apply in such cases. Moreover, although transparency is a principle of data protection law and the GDPR has provided new access rights, data protection rights are highly individualised and will not furnish data subjects with the comparative information necessary to ground a data protection claim.¹²⁶ Class action lawsuits based on mass data access requests might appear to provide an alternative in theory, but face various barriers in practice.¹²⁷ In short, while stronger judicially created access rights for individual claimants would be helpful, they would not be a panacea.

3.2 Legislative reform

Access to information and evidence is critical to ensure that litigation at the individual and collective level can take place, and that public bodies can be empowered to act.¹²⁸ Expanded rights for potential claimants to access information would be helpful, but proactive sharing of information is necessary for the existing regime to be truly effective.¹²⁹ Legislation specifying standards for published materials would ensure that individuals and group bodies, such as trade unions, can understand and compare the information. By analogy, Sheard notes that there is an ‘urgent need for evidentiary standards identifying which types of statistical steps, tests or counterfactuals should be run to unearth discrimination’.¹³⁰

In short, the legislator should facilitate enforcement of existing equality law by putting information into the prospective litigator’s hands: that is, by requiring the publication of information on algorithmic management tools in a specified format and to a specified standard. No current proposal under EU law meets that imperative. The proposed AI Act classifies algorithmic management systems as high risk and requires that certain documentation be created about the functioning and performance of such systems,¹³¹ but does not furnish job applicants or workers with access to this documentation.¹³² The proposed Platform Work Directive is stronger, providing platform workers (but not job applicants) with access to information about algorithmic management systems, including the ‘categories of decisions that are taken or supported’ by such systems and the ‘main parameters’ that the systems ‘take into account’.¹³³ The proposal does not contain any specific provisions on information about equality outcomes, however, and the absence of standards means that the information might not be sufficient for claimants to ground a case of discrimination.

One of the most promising legislative innovations in the context of algorithmic equality is the algorithmic impact assessment.¹³⁴ Employers could be required to publish impact assessments

126. Kelly-Lyth (n 6) 914, 921; Orwat (n 6) 71–72; Gerards and Xenidis (n 3) 116; Jake Stein and Dan Calacci, ‘Workers Collective Data Access Rights: Adding Context to Worker Data Protection’ (June 2022 preprint).

127. Kelly-Lyth (n 6) 921, and discussed by representatives of National Humans Rights Bodies and National Equality Bodies (‘When rights and tech collide: Protecting human rights and equality in the age of AI’, webinar session by Equinet and the Commonwealth Forum of National Human Rights Institutions, 6 July 2022).

128. Enforcement by public bodies is crucial: Kelly-Lyth, *ibid* 918–23; Gerards and Xenidis (n 3) 145; Orwat (n 6) 70.

129. On the importance of facilitating collective litigation, see Martini (n 109) 1024–1025.

130. Sheard (n 72) 646.

131. AI Act generally, and in particular Arts. 11 and 29, and Annex III, para 4.

132. Aislinn Kelly-Lyth, ‘The AI Act and Algorithmic Management’ [2021] *Comparative Labor Law & Policy Journal* Dispatch No 39, 7–8.

133. Proposal for a Directive of the European Parliament and of the Council on improving working conditions in platform work, COM (2021) 762 final (9 December 2021) (Platform Work Directive) art 6(2).

134. For in-depth review, see Kelly-Lyth and Thomas (n 102).

which provide the information necessary to identify and (where relevant) ground a claim of direct or indirect discrimination. The data protection impact assessment (DPIA) obligation—which is imposed by the GDPR and applies to many algorithmic management systems—provides a solid foundation for such a duty, but the DPIA is subject to no transparency requirements and does not set any concrete standard or format for evaluating rights impacts.¹³⁵ Meanwhile, the Platform Work Directive would require digital labour platforms to ‘monitor and evaluate’ the impact of algorithmic management systems, but the draft Directive does not particularise the format and publication of the results.¹³⁶

Given that algorithmic discrimination does not pose any insurmountable challenges to the statutory framework of EU equality law, legislative reforms should be aimed at securing effective enforcement. Mapping out the requirements for an adequate impact assessment must therefore be a priority for policymakers. Elsewhere in this issue, Anna Thomas and I propose an algorithmic management impact assessment obligation which would build in transparency via stakeholder consultation and publication of redacted assessments.¹³⁷

4. Conclusion

Reference was made at the start of this article to a proposal which would have amended UK non-discrimination law—the framework of which largely reflects EU equality law—to respond to algorithmic discrimination. In response to the Government’s assertion that the law ‘already protects workers against direct or indirect discrimination by computer or algorithm-based decisions’,¹³⁸ the Opposition contended that the law was not translating into practice: the Minister’s description of the labour market was described as ‘Panglossian’.¹³⁹ While the amendment proponent’s critique took aim at the ‘sharp[ness]’ of the law,¹⁴⁰ the real problem is its enforceability. While ADMS brings new opportunities to uncover and challenge unlawful discrimination, man-made opacity is stymying successful litigation. It is this latter problem on which legislators should fix their attention. While the material scope of the Directives is flexible enough to allow for updating via careful judicial interpretation, a strong policy response is required to buttress these provisions with enhanced transparency. Such transparency should be achieved through the mandating of algorithmic impact assessments in the employment context.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article: I acknowledge funding from the European Research Council under the European Union’s Horizon 2020 research

135. GDPR art 35.

136. Platform Work Directive art 7.

137. Kelly-Lyth and Thomas (n 102).

138. Data Protection Bill Deb 22 March 2018, cols 209–320.

139. *ibid* col 320.

140. *ibid*.

and innovation programme (grant agreement No 947806). I am grateful to Jeremias Adams-Prassl, Halefom Abraha, Michael ‘Six’ Silberman, Johanna Wenckebach, Daniel Pérez del Prado, Tobias Müllensiefen, Dan Calacci, Simona de Heer, Lukas Hondrich, and Krishna Gummadi for discussions which informed this work.