

# Algorithms, Apps & Artificial Intelligence: The next frontier in discrimination law

A paper by Robin Allen QC and Dee Masters updated for the Public Law Project session entitled, "AI Justice: Artificial intelligence decision-making and the law" on 16 October 2018



[www.cloisters.com](http://www.cloisters.com)

020 7827 4000

[clerks@cloisters.com](mailto:clerks@cloisters.com)

[@CloistersLaw](https://twitter.com/CloistersLaw)



Robin is a discrimination barrister from Cloisters who has appeared in over 145 reported cases in the Industrial Relations Law Reports (more than any other barrister). These include many path-finding appellate cases at the highest level in the UK and Europe. He has been instructed in over 40 cases in the House of Lords/Supreme Court and undertaken test cases in relation to every protected characteristic. He has undertaken many public interest amicus and intervention briefs, both in the UK and worldwide, for among others the Equality and Human Rights Commission, the Harvard University Law School, Public Concern at Work, and Liberty. He has also given evidence to Parliamentary Committees on the technical aspects of proposed legislation such as the Marriage (Same Sex Couples) Act 2013, and he advised the European Commission on the content of the first two key equality Directives. He is a consultant to HelpAge International on the UN's Open Age Working Group on the proposal for a UN Convention on the Rights of Older Persons and to the European Age Platform.



Dee is also discrimination barrister from Cloisters. Her practice primarily consists of multi-week trials involving multiple and complex claims. Due to her expertise, she regularly advises NGOs and government agencies on the development of discrimination law and she has delivered judicial training on discrimination law at ERA in Trier funded by the European Commission. She has advised the Equality Commission of Northern Ireland (ECNI) and the Northern Ireland Commissioner for Children and Young People (NICCY) on its response to proposed legislation extending the existing prohibition on age discrimination. In August 2015, she was the keynote speaker at a conference attended by politicians and policy makers in Belfast concerning age discrimination. Dee has also been advising AGE PLATFORM, an NGO, on a proposed European Directive which will extend the prohibition on age discrimination. She drafted the Technical Guidance accompanying the ban on age discrimination in goods, facilities and services on behalf of the Equality and Human Rights Commission in GB alongside colleagues in chambers. In 2018, she gave evidence as an expert witness to the Women and Equalities Committee on older people in the workplace.

## CONTENTS

Introduction .....	4
A real problem .....	6
Algorithms and their hidden dangers .....	7
Scope of the Equality Act 2010 .....	10
Commercial context.....	11
Direct discrimination .....	11
The liability of the code provider... ..	13
Harassment .....	14
Indirect Discrimination.....	16
Duty to make reasonable adjustments .....	17
Discouraging diversity through pattern recognition.....	18
Public sphere.....	20
Criminal justice and policing .....	20
Using statistics .....	22
Sentencing decisions .....	25
Human rights perspective.....	27
Conclusion.....	28

## Introduction

1. How often does the media depict the relentless increase in technology as a danger to our health, our children and our security? More recently, commentators have started to identify the ways in which technology discriminates against users because of their race, disability, gender or sexual orientation. Indeed, in the open letter<sup>1</sup> on artificial intelligence ("AI") in The Guardian from 18 June 2018 experts<sup>2</sup> in this field set out their concerns about inequality as a result of the ever-increasing use of AI as follows:

According to the [World Economic Forum]<sup>3</sup> "Global Gender Gap" report,<sup>4</sup> it will take 100 years to unlock the potential of gender equality in terms of health, education, and policies, and 217 years to reach economic parity. Without these commitments and standards, AI will not improve-and could in fact make worse the biases of our societies.

The internet today is inherently unequal, largely because it was created by organizations dominated by men. As we enter the AI revolution that will define our future, women make up less than 30% of research positions worldwide. Across the four largest tech companies - Apple, Google, Microsoft, and Facebook - fewer than 20% of technical roles are held by women.

At the same time machines are teaching themselves from data sets that reflect or even amplify society's past and present biases: "Homemaker" is to "woman" as "programmer" is to "man"; "mother" is to "nurse" as "father" is to "doctor." Committing to diversity in AI leadership and defining the standards by which we will hold all AI accountable will take a society-wide effort-across government and industry.

We are signing this letter to call for a set of standards for AI to make sure it is a force for progress, not an impediment to it. Who's with us?

2. Our aim is to contribute to a developing discussion about the potential for discriminatory effects arising from ill-thought out algorithms, apps and AI, with the hope that you too will be able to spot when to be able to challenge the too ready use of AI without a clear understanding of what it is doing.

---

<sup>1</sup> Equal AI: an open letter, The Guardian print edition 18 June 2018.

<sup>2</sup> Robert LoCascio, Founder and CEO LivePerson, Ariana Huffington, Founder and CEO Thrive Global, Baroness Lane Fox, Founder Doteveryone, Dr Justine Casell, Associate Dean for Technology Strategy and Impact in the School of Computer Science, Carnegie Mellon University, and Jimmy Wales, Founder Wikipedia.

<sup>3</sup> [www.weforum.org/](http://www.weforum.org/)

<sup>4</sup> <https://www.weforum.org/reports/the-global-gender-gap-report-2017>

3. We want to make clear that this is not however going to be a paper saying that these are all bad things which should be stopped at any price. That would be neither wise nor necessary. AI can confer great benefits such as the early diagnosis of skin cancer<sup>5</sup>, dementia<sup>6</sup> or the Zika virus<sup>7</sup>. Nonetheless there are reasons for lawyers interested in public law and equality law to be engaged in this development to make sure that its beneficial effects are not undermined by adverse ones. When it is used to the detriment of an individual, based on biased assumptions or stereotypes, it is discrimination and will often be unlawful. Moreover, since AI may well play an important role in decision-making in a judicial and broader criminal justice setting it is even more important that we start to examine and question that technology now.
4. This issue of discriminatory technology was first raised in Cloisters by Dee Masters last year. We first wrote about this subject at the beginning of the year<sup>8</sup> and since then have had many enquiries and follow ups both here and in Europe. Robin Allen QC recently lectured on it to the European Commission funded European Rights Academy in Trier.<sup>9</sup>
5. It is in any event an issue which is under real scrutiny in Parliament as we set out below. Parliament has been looking at a whole range of issues including the misuse of data, potential breaches of the General Data Protection Regulation (“GDPR”)<sup>10</sup> and the domestic Data Protection Act 2018.<sup>11</sup> These regulatory issues are another dimension which we think you should be aware of and these topics are the subject of a different paper by Robin Allen QC and Dee Masters entitled, “Algorithms, apps & artificial intelligence 2: Can data protection laws be used to challenge discriminatory tech?”<sup>12</sup>.
6. Moreover, commercial organisations are now starting to grapple with ways in which to tackle discriminatory technology. Companies such as IBM, Google, Microsoft and Facebook have released or are soon to release technology which will hopefully allow

---

<sup>5</sup> <https://www.theguardian.com/society/2018/may/29/skin-cancer-computer-learns-to-detect-skin-cancer-more-accurately-than-a-doctor>

<sup>6</sup> <https://imperialbrc.nihr.ac.uk/2018/05/23/artificial-intelligence-ai-improves-stroke-and-dementia-diagnosis-in-brain-scans/>

<sup>7</sup> <https://www.sciencedaily.com/releases/2018/06/180613113804.htm>

<sup>8</sup> See <http://www.cloisters.com/blogs/algorithms-apps-artificial-intelligence-the-next-frontier-in-discrimination-law> and PLC magazine, April 2018.

<sup>9</sup> The lecture is available on the ERA’s website.

<sup>10</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), see <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679>

<sup>11</sup> See <http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted>

<sup>12</sup> [https://www.cloisters.com/images/machine\\_learning\\_blog\\_July\\_2018.pdf](https://www.cloisters.com/images/machine_learning_blog_July_2018.pdf)

algorithms to be “equality proofed”.<sup>13</sup> Legislation is also being considered in the US to create accountability around algorithms.<sup>14</sup>

## A real problem

7. Sara Wachter-Boettcher in her book, “Technically Wrong: Sexist Apps, Biased Algorithms and other Threats of Toxic Tech” outlines many chilling examples of discriminatory technology from the United States and some from the UK.<sup>15</sup>
8. On 25 May 2018, so only last month, the Science and Technology Committee of the UK Parliament said in the conclusions of its report on “Biometric strategy and forensics services”<sup>16</sup> (“**Fifth Report**”) that:

Facial image recognition provides a powerful evolving technology which could significantly help policing. There are concerns, however, over its current use, including its reliability and its potential for discriminatory bias. We welcome the Government’s assurances that the technology is only being used at the moment for targeting those on ‘watch lists’ rather than as a blanket approach. The technology should not be generally deployed, beyond the current pilots, until the current concerns over the technology’s effectiveness and potential bias have been fully resolved. Ministers and Parliament, rather than the police, should take the final decision on any wider deployment. The Biometrics Strategy should include an undertaking that the House will be given an opportunity to debate and vote on the issue.

9. In giving this advice, the House of Commons, Science and Technology Committee was only repeating the concerns of academics about the effects of machine learning in a raft of different situations where the machines concerned take decisions that affect people’s lives. It had received evidence of these concerns as part of its Inquiry into Algorithms in

---

<sup>13</sup> <https://www.computerworlduk.com/data/how-ibm-is-taking-lead-in-fight-against-black-box-algorithms-3684042/>

<sup>14</sup> <https://www.propublica.org/article/new-york-city-moves-to-create-accountability-for-algorithms>

<sup>15</sup> Published by W. W. Norton & Company, 2017, ISBN10 0393634639, ISBN13 9780393634631

<sup>16</sup> <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/800/80003.htm>

Decision Making<sup>17</sup> and which it set out in its Fourth Report which it published on 15 May 2018 (“**Fourth Report**”).

10. No sooner had we included the above paragraphs in this paper than we learned that Liberty is seeking Crowdfunding to take a case challenging the use of facial recognition data by the police.<sup>18</sup> Indeed, Liberty is expecting to lodge its judicial review in the autumn of 2018.<sup>19</sup>
11. Before discussing what was concerning the Select Committee let us look a little more at what exactly is AI and how it can be biased.

## Algorithms and their hidden dangers

12. At the heart of AI is the “algorithm”. Algorithms are a set of steps created by programmers. They usually perform repetitive and tedious tasks in lieu of human actors. For example, when LinkedIn informs a user that someone within her network is also connected to five people who are her contacts, it is an algorithm – and not a human – that has quickly compared the two networks to find common contacts.
13. The power of an algorithm is often linked to “machine learning” which is a means of refining algorithms and making them more “intelligent”. Here is an extract from “The privacy pro's guide to explainability in machine learning”<sup>20</sup> published by the International Association of Privacy Professionals<sup>21</sup> which explains more:

---

<sup>17</sup> The House of Commons Science and Technology Committee, Report on Algorithms in decision making, the Fourth Report of Session 2017–19, is available here <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pdf>

<sup>18</sup> See <https://www.libertyhumanrights.org.uk/news/press-releases-and-statements/cardiff-resident-launches-first-uk-legal-challenge-police-use> and <https://www.crowdjustice.com/case/facial-recognition/>

<sup>19</sup> The ways in which discriminatory technology can be challenged utilising the Human Rights Act 1998 and judicial review is outside the scope of this paper. However, a useful overview of how public law principles can be used to challenge discriminatory technology is contained in “Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality” (2018) by Marion Oswald, Jamie Grace, Sheena Urwin & Geoffrey C. Barnes, Information & Communications Technology, and which is available here: <https://www.tandfonline.com/doi/pdf/10.1080/13600834.2018.1458455?needAccess=true>.

<sup>20</sup> With thanks to John Higgins CBE, previously Director – General of Digital Europe for suggesting this site.

<sup>21</sup> <https://iapp.org/news/a/the-privacy-pros-guide-to-explainability-in-machine-learning/>

### **What is machine learning?**

Machine learning is a technique that allows algorithms to extract correlations from data with minimal supervision. The goals of machine learning can be quite varied, but they often involve trying to maximize the accuracy of an algorithm's prediction. In machine learning parlance, a particular algorithm is often called a "model," and these models take data as input and output a particular prediction. For example, the input data could be a customer's shopping history and the output could be products that customer is likely to buy in the future. The model makes accurate predictions by attempting to change its internal parameters — the various ways it combines the input data — to maximize its predictive accuracy. These models may have relatively few parameters, or they may have millions that interact in complex, unanticipated ways. As computing power has increased over the last few decades, data scientists have discovered new ways to quickly train these models. As a result, the number — and power — of complex models with thousands or millions of parameters has vastly increased. These types of models are becoming easier to use, even for non-data scientists, and as a result, they might be coming to an organization near you.

14. Algorithms are, of course, code written by humans for human purposes, and algorithms can discriminate on the grounds of protected characteristics when they become tainted by the unconscious assumptions and attitudes of their creators.
15. This is what researchers<sup>22</sup> at the Centre for Intelligent Sensing at Queen Mary University of London<sup>23</sup> said in written evidence to the House of Commons, Science and Technology Committee as part of the background to its Fourth Report (see paragraph 9 above) –

Algorithms themselves enforce the biases present in the algorithm designer and/or the data used to "train" the algorithm. This can have effects as deleterious as the biases that they remove. As a simple example: if past data on employee recruitment is used to "train" an algorithm for making employment decisions, then in many cases it will reproduce the biases that were present in previous practice, unless deliberate effort is made to mitigate that. The removal of such bias is non-trivial and is an ongoing research field. Biases are difficult to uncover if the algorithm is complex; if the algorithm is simple and auditable - ideally, publicly auditable - then there is much greater chance that algorithm-based decisions can be fair. Likewise,

---

<sup>22</sup> Dr Dan Stowell, Dr Emmanouil Benetos, Dr Bob Sturm, and Dr Laurissa Tokarchuk.

<sup>23</sup> <http://cis.eecs.qmul.ac.uk/>

if the data used to “train” the algorithm is auditable (and anonymised when appropriate), algorithmic biases could be identified.

16. It is worth unpicking why the House of Commons, Science and Technology Committee was so concerned that algorithms can be based on discriminatory assumptions. One practical example which supported their concerns related to facial recognition technology. They said in their Fifth Report (referred to at paragraph 8 above) that –

As we noted in our recent report on Algorithms,<sup>24</sup> research at MIT in the US found that widely used facial-recognition algorithms were biased because they had been ‘trained’ predominantly on images of white faces. The systems examined correctly identified the gender of white men 99% of the time, but the error rate rose for people with darker skin, reaching 35% for black women.

17. The research to which they were referring was carried out by Joy Buolamwini and Timnit Gebru<sup>25</sup> and it is fascinating. The Abstract published at the head of this research states –

Recent studies demonstrate that machine learning algorithms can discriminate based on classes like race and gender. In this work, we present an approach to evaluate bias present in automated facial analysis algorithms and datasets with respect to phenotypic subgroups. [We found that currently widely used] datasets are overwhelmingly composed of lighter-skinned subjects ... and introduce a new facial analysis dataset which is balanced by gender and skin type. We evaluate 3 commercial gender classification systems using our dataset and show that darker-skinned females are the most misclassified group (with error rates of up to 34.7%). The maximum error rate for lighter-skinned males is 0.8%. The substantial disparities in the accuracy of classifying darker females, lighter females, darker males, and lighter males in gender classification systems require urgent attention if

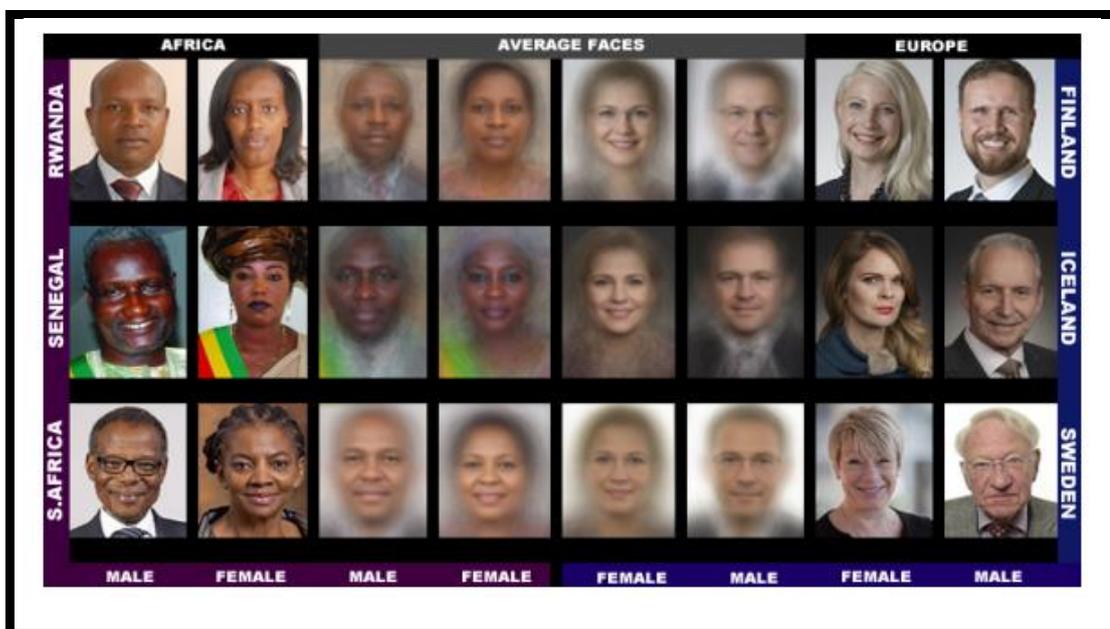
---

<sup>24</sup> See Science and Technology Committee, The use of algorithms in decision-making, Fourth Report of Session 2017–19, HC 351.

<sup>25</sup> See “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification” in the Proceedings of Machine Learning Research 81:1–15, 2018 Conference on Fairness, Accountability, and Transparency; this is available at <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

commercial companies are to build genuinely fair, transparent and accountable facial analysis algorithms.

18. As an interesting aside it is worth noting how the authors then went about trying to cure the bias by creating a new data set which was based on male and female faces from a range of Parliaments with impressive levels of gender parity from around the world. This created a more balance representation of both gender and racial diversity. Their paper identified the range they used pictorially<sup>26</sup> –



19. Using this data set, they concluded that a non-biased selection of faces from which the AI system was to learn was much more successful. In other words, it is possible to create more effective technology by challenging discrimination.

### Scope of the Equality Act 2010

20. In the rest of this paper, we consider the ways in which the Equality Act 2010 (“EA 2010”) in the UK can be used to prevent technology companies from discriminating against customers, potential customers, the public and individuals who interact with the judicial system or policing.

<sup>26</sup> Ibid. figure 1.

21. The prohibition on discrimination in the provision of goods, facilities and services is contained within Part 3 of the EA 2010. The prohibition applies to any service provider concerned with the provision of a service to the public or a section of the public for payment or not (s.29 EA 2010).
22. In very broad terms, a service provider must not discriminate (direct or indirect discrimination), harass or victimise a person as to the terms on which the service is provided, by terminating the service, by refusing to provide the service or by subjecting the person to any other detriment (s.29 (2) – (5) EA 2010). There is also an obligation on the service provider to make reasonable adjustments (s.29 (7) EA 2010).<sup>27</sup> Identical provisions apply to persons exercising a public function (s.29 (6) EA 2010). In other words, the goods, facilities and services provisions in the EA 2010 apply to activities which people probably “don’t want” such as being subject to a “stop and search” by the police, being questioned or arrested or the way in which prisoners are treated within the prison setting.<sup>28</sup>
23. We predict that the rise in technology will impact on the private and public sector in very different ways. Accordingly, in this paper we will examine the ways in which the EA 2010 might be deployed in a commercial context before examining the public sphere.

## Commercial context

### Direct discrimination

24. One algorithm with biased assumptions must have been used by Etsy, an online retailer for unique gifts. It contacted users on Valentine’s Day with a view to encouraging purchases from its site. It appears to have used an algorithm that assumed female users of its website were in a relationship with a man: one customer, Maggie Delano, received the message “Move over, Cupid! We’ve got what he wants. Shop Valentine’s Day gifts for him”.

---

<sup>27</sup> There are some important exceptions to the principle of non-discrimination in the field of goods, facilities and services. It is outside the scope of this paper to outline those exceptions exhaustively. Instead we will outline those exceptions where relevant to our analysis.

<sup>28</sup> Please see the guidance published by the Equality and Human Rights Commission entitled “Your rights to equality from the criminal and civil justice systems and national security” available here: <https://www.equalityhumanrights.com/sites/default/files/equalityguidance-criminal-civiljustice-2015-final.pdf>

25. The problem, was that Maggie Delano is a lesbian and any Valentine’s gift she might buy would most likely be for a woman.<sup>29</sup>
26. At a stroke of a line of code, Etsy had thus alienated its homosexual client base. Indeed all homosexual clients were at risk of being offended by this ill-considered message and as such there was arguably direct discrimination on the grounds of sexual orientation. In the UK, where discrimination on the grounds of sexual orientation in relation to the provision of a service is forbidden under the EA 2010, this would be direct discrimination and a claim could theoretically be made.
27. Another algorithm was utilised by a chain of gyms in Britain called Puregym.<sup>30</sup> In 2015, Louise Selby, a paediatrician, was unable to use her gym swipe card to access the locker rooms.
28. It transpired that the gym was using third party software which used a member’s title to determine which changing room (male or female) they could access. The software contained an algorithm that the title “Doctor” was coded as “male”. As a female doctor, she was not permitted to enter the women’s changing rooms.<sup>31</sup>
29. This is of course ridiculous and unlawful too. It also received widespread negative publicity.<sup>32</sup> The PR fallout seems to have been relatively well managed it – though at some cost – according to this interview of Puregym’s CEO Humphrey Cobbold by trade press “Health Club Management”<sup>33</sup> –

“There is certainly no intention on our part to be sexist. There are currently more than 200,000 female members of Pure Gym and a large proportion of our staff are female and are absolutely an integral and essential part of our business,” he said.

“This was a software glitch which we take full responsibility for and are working hard to rectify, but we’re not a sexist company at all and actually it’s been heartening to see lots of our members reiterate this in the comment sections of articles.”

---

<sup>29</sup> Sara Wachter-Boettcher, “Technically Wrong: Sexist Apps, Biased Algorithms and other Threats of Toxic Tech” , pages 32-33.

<sup>30</sup> <https://www.information society.co.uk/pure-gym-in-cambridge-sexist-computer-assumed-this-woman-dr-louise-selby-was-a-man-because-she-is-a-doctor/>

<sup>31</sup> Sara Wachter-Boettcher, *ibid*, page 6.

<sup>32</sup> <http://www.itv.com/news/anglia/update/2015-03-18/gym-mix-up-assumes-anyone-with-a-dr-title-would-be-a-man/> and <https://metro.co.uk/2015/03/18/gyms-computer-assumed-this-woman-was-a-man-because-she-is-a-doctor-5110391/>

<sup>33</sup> <http://www.healthclubmanagement.co.uk/health-club-management-news/Pure-Gym-CEO-Were-not-a-sexist-company/314776?source=search>

Cobbold declined to name the provider of the software and said the chain wouldn't be switching as a result of the error. "Ultimately the buck stops with us and it's our responsibility to ensure all components function as they should," he added.

The CEO has personally telephoned Dr Selby to apologise for the error and offered a free six-month membership for her and a friend. He said Dr Selby had accepted his apology and hopes that she will continue to be a member of Pure Gym.

"We had a civilised discussion and I explained how the error came about," said Cobbold. "We've learned that in an organisation such as ours it's absolutely crucial to get even the smallest details right, but ultimately, I'm pleased with the way the team have responded to the issue and we're working hard to ensure it doesn't happen again.

30. This interview highlights several points –

- a. Puregym had a lot of female clients who could have been offended and lost;
- b. The glitch lay elsewhere than with Puregym itself; and
- c. It costs money to restore the goodwill lost by the biased machine.

31. The CEO's interview shows – as might be expected - that the company was not aware that it had been acting in this discriminatory way, but service providers should be aware that ignorance is no excuse. It is irrelevant to the question of liability that the gym did not know and did not intend to discriminate against women. They will normally be fixed with the discriminatory consequences of technology which they use.

#### The liability of the code provider...

32. However the problem does not stop there as in most cases the service provider will not have written the relevant code itself but will have bought it from an outside source. Humphrey Cobbold was coy about this in his interview, but it would seem that his company had indeed bought in the code that thought a "Dr" was always male. What should have happened?

33. Service providers need to manage their exposure. The least they can do is carefully quiz their technology providers to ensure that products have been "equality proofed". We advise that they should also insist that the undertaking providing the code indemnifies them against any discriminatory effects that it may have.

34. At the point of purchase of the code the service provider is in a difficult situation. Algorithms are often closely guarded secrets or so complex that any discriminatory assumptions might not be immediately apparent to a purchaser of the software. In itself this raises profound issues of transparency.<sup>34</sup>
35. Puregym avoided a lawsuit but discrimination claims against a service provider are not merely annoying, they have the capacity to destroy good will and a reputation that has been built up over many years. So having thought about this in advance is at least a partial insurance against that kind of damage. We suspect that Mr Cobbold will have been much more demanding of his IT providers in the future.

## Harassment

36. AI based technology can also easily lead to harassment. Numerous illustrations are contained in Sara Wachter-Boettcher's book.
37. One example she gives concerns Snapchat which in August 2016 introduced a face-morphing filter which was "inspired by anime". In fact, the filter turned its users' faces into offensive caricatures of Asian stereotypes.<sup>35</sup>
38. Smart phone assistants in 2017 nearly all have default female voices e.g. Apple's Siri, Google Now and Microsoft's Cortana. Commentators have said that this echoes the dangerous gender stereotype that women, rather than men, are expected to be helpful and subservient.<sup>36</sup>
39. The Indiana University School of Informatics has been researching the issue for some time. A recent report<sup>37</sup> has found that –

... women and men expressed explicit preference for female synthesized voices, which they described as sounding "warmer" than male synthesized voices. Women also preferred female synthesized voices when tested for implicit responses, while men showed no gender bias in implicit responses to voices.

---

<sup>34</sup> The issue of data management under the GDPR is outside this paper but is of itself very important and will be addressed in a subsequent paper.

<sup>35</sup> Sara Wachter-Boettcher, *ibid*, page 7.

<sup>36</sup> Sara Wachter-Boettcher, *ibid*, pages 37 – 38.

<sup>37</sup> <https://soic.iupui.edu/news/macdorman-voice-preferences-pda/>

40. The University wisely added that the report “suggests that device designers should be aware of reinforcing gender stereotypes.”<sup>38</sup> There does appear to a move away from using female voices in submissive technology<sup>39</sup> but progress is slow.
41. Google Photos introduced a feature which tagged photos with descriptors, for example, “graduation”. In 2015, a black user noticed that over 50 photos depicting her and a black friend were tagged “gorillas”.<sup>40</sup> Of course, Google Photos had not been programmed to tag some black people as “gorillas” but this was the conclusion which the AI at the heart of the technology had independently reached. It is not hard to imagine the degree of offence this must have caused.
42. The strange outcome of this is described by wired.com in this post by Tom Simonite earlier this year<sup>41</sup> under the title “When It Comes to Gorillas, Google Photos Remains Blind”–

In 2015, A black software developer embarrassed Google by tweeting that the company’s Photos service had labelled photos of him with a black friend as “gorillas.” Google declared itself “appalled and genuinely sorry.” An engineer who became the public face of the clean-up operation said the label gorilla would no longer be applied to groups of images, and that Google was “working on longer-term fixes.”

More than two years later, one of those fixes is erasing gorillas, and some other primates, from the service’s lexicon. The awkward workaround illustrates the difficulties Google and other tech companies face in advancing image-recognition technology, which the companies hope to use in self-driving cars, personal assistants, and other products.

WIRED tested Google Photos using a collection of 40,000 images well-stocked with animals. It performed impressively at finding many creatures, including pandas and poodles. But the service reported “no results” for the search terms “gorilla,” “chimp,” “chimpanzee,” and “monkey.”

43. In the UK, users who are offended by this type of technology might be able to bring harassment claims against service providers again under the EA 2010.<sup>42</sup> Although the

---

<sup>38</sup> Ibid.

<sup>39</sup> Amazon’s Polly service in the US added male voices in May 2018. See <https://aws.amazon.com/polly/>.

<sup>40</sup> Sara Wachter-Boettcher, *ibid*, pages 129 – 132.

<sup>41</sup> <https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/>

<sup>42</sup> And perhaps GDPR claims as well.

compensation for injury to feelings in discrimination claims against service providers is often low<sup>43</sup>, it is obvious that a claim brought by a large group of people affected by any such harassment could lead to considerable financial exposure as well as creating a PR disaster. If it is bad enough – and it is difficult to think of anything much more offensive than this - then we can foresee that the CEOs of Google and other companies will be back in front of the European Parliament. So again precautions at the outset must be required.

### Indirect Discrimination

44. Indirect discrimination usually is less of a reputational disaster, but it can be serious. We are clear that the creators of apps (and service providers who purchase them) could also unwittingly expose themselves to indirect discrimination claims by failing to think inclusively about their client base.
45. In 2015, research revealed that of the top 50 “endless runner”<sup>44</sup> games available in the iTunes store which used gendered characters, less than half offered female characters. In contrast, only one game did not offer a male character.<sup>45</sup> Why is that?
46. Whilst there is no necessary connection between a person’s gender and the gender of the character that they would choose within a virtual environment, some research has shown that the majority of users (especially women) will choose an avatar that mirrors their gender identity.<sup>46</sup> It follows that the absence of female avatars will place female users at a particular disadvantage could lead to indirect sex discrimination claims. No doubt a similar analysis could be applied to race.
47. Another problem area is in relation to names. Many services require users to enter their real names. In order to decrease the likelihood of people using false names, algorithms have been developed to “test” entries. This creates barriers for people who have names that are deemed “invalid” by algorithms which have been constructed so as to recognise mostly “western” names.

---

<sup>43</sup> For example, in *Campbell v Thomas Cook Tour Operations Ltd* [2013] Eq. L.R. 658, a disabled woman was awarded £7,500.00 when the Defendant did not make reasonable adjustments at an airport so as to alleviate the difficulties she experienced waiting and queuing over a significant period of time.

<sup>44</sup> These are games where the objective is to keep virtual characters running as long as possible.

<sup>45</sup> Sara Wachter-Boettcher, *ibid*, page 3.

<sup>46</sup> Rosa Mikeal Martey, Jennifer Stromer-Galley, Jaime Banks, Jingsi Wu, Mia Consalvo. “The strategic female: gender-switching and player behavior in online games”. *Information, Communication & Society*, 2014; 17 (3): 286 DOI. This research revealed that within a particular virtual environment, 23% of users who identified as men would choose opposite sex avatars whereas only 7% of women gender-switched.

48. An example highlighted by Sara Wachter-Boettcher is Facebook and a would-be user called Shane Creepingbear who is a member of the Kiowa tribe of Oklahoma.<sup>47</sup> When he tried to register in 2014 he was informed that his name violated Facebook’s policy.



49. Again the algorithm used by Facebook at this point could have been used as the basis of an indirect discrimination claim.
50. Companies will only be able to avoid these risks by thinking broadly about who will use their products and testing products vigorously, with a view to avoiding discrimination, before launching them.

### Duty to make reasonable adjustments

51. We are accustomed to thinking about the duty to make reasonable adjustments in the context of technology. A common example is the feature on many taxi apps whereby a user can ask for a wheelchair adapted car.
52. But there are more subtle ways in which technology can discriminate against disabled users by making assumptions about customer behaviour. Smart weighing scales are an interesting case in point. Sara Wachter-Boettcher writes about a set of scales which tracks basic data about the user which is then stored and used to create personalised “motivational” messages like “Congratulations! You’ve hit a new low weight”.

---

<sup>47</sup> Sara Wachter-Boettcher, *ibid*, pages 54 - 55.

53. The difficulty, as Wachter-Boettcher points out, is that these scales only understood that users would have one goal – weight loss. A user recovering from an eating disorder or in the throes of degenerative disease would likely find these messages counterproductive. Similarly, if they succeed in putting weight on they receive an insensitive message like “Your hard work will pay off [name]! Don’t be discouraged by last week’s results. We believe in you! Let’s set a weight goal to help inspire you to shed those extra pounds”. A simple adjustment like being able to choose your goal would avoid the risk of the manufacturer being in breach of the duty to make reasonable adjustments.

### Discouraging diversity through pattern recognition

54. Technology could also have a worrying impact on diversity as AI becomes more prevalent. As explained, machine learning is based on recognising patterns and “learning” from existing historical data. In the Fourth Report, the Select Committee noted the following evidence of this problem –

A well-recognised example of this risk is where algorithms are used for recruitment. As Mark Gardiner put it, if historical recruitment data are fed into a company’s algorithm, the company will “continue hiring in that manner, as it will assume that male candidates are better equipped. The bias is then built and reinforced with each decision.” This is equivalent, Hetan Shah from the Royal Statistical Society noted, to telling the algorithm: “Here are all my best people right now, and can you get me more of those?” (footnotes omitted)

55. We are aware of some examples of this. For instance Word2vec is a neural network which analyses data so as to understand the semantic relationship between words. The problem is that some of that data will be shaped by historic and continuing direct or indirect discrimination. Research showed, for example, that the system perceived a relationship between being male and a computer programmer whereas women were associated with staying at home.<sup>48</sup> Similarly architects were deemed male and interior designers considered female.<sup>49</sup>

56. Sara Wachter-Boettcher points to a company which decided, in 2016, to utilise this type of software to facilitate recruitment decisions. One way in which the software could be used was to rate CVs so as to identify “matches” between potential employees and

---

<sup>48</sup> Sara Wachter-Boettcher, *ibid*, pages 138 – 139.

<sup>49</sup> Ditto.

existing successful employees.<sup>50</sup> The dangers should have been obvious. This type of software is more likely to identify new employees who have similar experiences, backgrounds and interests as the current workforce. Any inbuilt stereotyping will mean that new recruits are far more likely to be the same gender and race as existing employees.

57. In such a scenario, an applicant who was rejected because they were “different” to existing employees might be able to bring an indirect discrimination or even perhaps a direct claim. Equally, statistics showing that a workforce lacks diversity might be used by other claimants to boost allegations of discrimination.
58. Understandably the Select Committee in its Fourth Report were very concerned about all this and they noted a further diversity aspect being the lack of a diverse population of professionals working on these issues –

Dr Adrian Weller from the Alan Turing Institute told us that algorithm bias can also result from employees within the algorithm software industries not being representative of the wider population. Greater diversity in algorithm development teams could help to avoid minority perspectives simply being overlooked, by taking advantage of a “broader spectrum of experience, backgrounds, and opinions”. The US National Science and Technology Council Committee on Technology concluded in 2016 that “the importance of including individuals from diverse backgrounds, experiences, and identities [...] is one of the most critical and high-priority challenges for computer science and AI”. Dr Weller also made the case for more representation. TechUK told us: More must be done by Government to increase diversity in those entering the computer science profession particularly in machine learning and AI system design. This is an issue that TechUK would like to see the Government’s AI Review exploring and make recommendations on action that should be taken to address diversity in the UK’s AI research community and industry. (paragraph 43, footnotes omitted)

59. You may also be aware of systems using forms of machine learning to decide on dismissal decisions. The so called Bradford Formula or Factor<sup>51</sup> for calculating the significance of staff absence has been around for some time and is now being sold by many providers as a personnel management tool.

---

<sup>50</sup> Sara Wachter-Boettcher, *ibid*, pages 139 - 140.

<sup>51</sup> Developed by the Bradford University School of Management in the 1980s.

60. There are various formulations of this offered by commercial providers. One – Active Absence<sup>52</sup> - describes the product it offers as -

Calculating an employee’s Bradford Factor is an automated absence management tool already built into Activ Absence. Whether you are in HR or an employee’s Line Manager, you can monitor Bradford Factor scores automatically using configurable absence trigger alerts and management reports.

61. The use of the Bradford Factor/Formula has already been the subject of some litigation and there are growing concerns about the extent to which it is consistent with disability discrimination legislation.<sup>53</sup> In 2016, in *Lowmoore Nursing Home Limited v Miss C Smith*, the Employment Tribunal held that the application of the Bradford Factor/Formula was a provision, criterion or practice (**‘PCP’**) within the meaning of the EA 2010 which placed the disabled claimant at a disadvantage. Moreover, a reasonable adjustment would have been to “shelve” the use of the Bradford Factor process.<sup>54</sup> Similarly, in *Gibbs v Westcroft Health Centre*, the employer rated a pregnant unfavourably using the Bradford Factor/Formula because she had required time away from work for pregnancy related illnesses. The Employment Tribunal concluded that a mechanistic application of the Bradford Factor/Formula amounted to unfavourable treatment contrary to s.18 EA 2010.<sup>55</sup>
62. These cases highlight how AI systems can serve, but must not be allowed to determine, human resource management decisions, if discrimination is to be avoided.

## Public sphere

### Criminal justice and policing

63. Discriminatory technology might also arise in the public sphere, in particular around criminal justice, judicial decision making and policing.

---

<sup>52</sup> <http://www.activabsence.co.uk/bradford-factor-calculator2/>

<sup>53</sup> We should say that we do not know of any misapplication of Active Absence’s tool.

<sup>54</sup> These matters are recorded in the EAT judgment: see *Lowmoore Nursing Home Limited v Miss C Smith*, UKEAT/0239/15/JOJ, 21 June 2016.

<sup>55</sup> Case number: 3400583/2014. The judgment is available on Westlaw. The relevant passage is at paragraph 5.16.

64. One development which requires careful monitoring is the use of facial recognition technology. As already explored at paras 16 to 18 above, there is research which reveals that the accuracy of facial recognition technology falls dramatically when considering non-white individuals. Despite this, facial recognition technology has started to be used by some police forces in the UK. According to Liberty, cameras equipped with automated facial recognition (AFR) software scan the faces of passers-by, making unique biometric maps of their faces. These maps are then compared to and matched with other facial images on bespoke police databases. On one occasion – at the 2017 Champions League final in Cardiff – the technology was later found to have wrongly identified more than 2,200 people as possible criminals.<sup>56</sup>
65. If facial recognition technology fails to adequately identify certain protected characteristics, such as individuals from a particular race, then that racial group is always at a greater risk of being incorrectly identified and as such there is the potential for a direct race discrimination claim under the EA 2010 against the organisations which utilise the technology if a person is then subjected to a detriment. Importantly, direct race discrimination can *never* be justified under the EA 2010.
66. The use of predictive technology in the criminal justice sector is also worrying. In 2017, Durham Constabulary started to implement a Harm Assessment Risk Tool (HART), which utilised a complex machine learning algorithm to classify individuals according to their risk of committing violent or non-violent crimes in the future.<sup>57</sup> This classification is created by examining an individual’s age, gender and postcode. This information is then used by the custody officer, so a human decision maker, to determine whether further action should be taken.<sup>58</sup> In particular, whether an individual should access the Constabulary’s Checkpoint programme which is an “out of court” disposal programme.<sup>59</sup>
67. The reader will immediately see the potential for numerous claims here. A direct age discrimination could be brought by individuals within certain age groups who were scored negatively. Similarly, direct sex discrimination claims could be brought by men, in so far as their gender leads to a lower score than comparable women. Finally, indirect race discrimination or direct race discrimination claims could be pursued on the basis that an individual’s postcode can be a proxy for certain racial groups. Only an indirect race

---

<sup>56</sup> <https://www.libertyhumanrights.org.uk/news/press-releases-and-statements/cardiff-resident-launches-first-uk-legal-challenge-police-use>.

<sup>57</sup> “Machine Learning Algorithms and Police Decision-Making: Legal, Ethical and Regulatory Challenges”, Alexander Babuta, Marion Oswald and Christine Rinik, RUSI, September 2018. Available here: [https://rusi.org/sites/default/files/20180329\\_rusi\\_newsbrief\\_vol.38\\_no.2\\_babuta\\_web.pdf](https://rusi.org/sites/default/files/20180329_rusi_newsbrief_vol.38_no.2_babuta_web.pdf).

<sup>58</sup> “Innocent Until Predicted Guilty? Artificial Intelligence and Policing Decision-Making”, Alexander Babuta, March 2018. Available here: <https://rusi.org/publication/newsbrief/innocent-until-predicted-guilty-artificial-intelligence-and-police-decision>

<sup>59</sup> See foot note 19.

discrimination claim would be susceptible to a justification defence in these circumstances.

### Using statistics

68. It might be argued that predictive technology like HART is not objectionable since its power is based on a statistical analysis which suggests that there are legitimate correlations between certain protected characteristics and particular behaviours. An example of this type of argument is highlighted in a recent paper by the Royal United Services Institute for Defence and Security Services (RUSI) <sup>60</sup>:

The issue of algorithmic bias and discrimination is further complicated by the fact that crime data is inherently 'biased' in a number of ways, because certain classes of people commit more crimes than others. For instance, men commit crime at significantly higher rates than women, are more likely to be involved in violent offences, and are more likely to reoffend. This gender imbalance has been described as 'one of the few undisputed facts "facts" of criminology. Therefore, a crime prediction system that is operating correctly will assign many more male offenders to the 'high-risk' category than female offenders. This can be described as 'fair biases', an imbalance in the dataset that reflects real-world disparities in how a phenomenon is distributed against different demographics. (footnotes removed)

69. The first point to note is that there are commentators who are skeptical that protected characteristics and certain behaviours can be linked in such a concrete way. One recent paper identified that predictive algorithms might actually simply predict who is most likely to be arrested rather than who is most likely to commit a crime. Moreover, there is always a risk that predictions become a "self-fulfilling prophecy" as human actors, like the police, act on algorithms.<sup>61</sup>

70. The second point is that from a legal perspective these types of arguments - that there is no discrimination because statistics reveal differences between protected characteristics - have been rejected in relation to gender and race by the CJEU. This is worth unpicking. Historically it was common to differentiate between men and women in relation to insurance products on the basis of actuarial factors which revealed that their "risk profile"

---

<sup>60</sup> Ibid.

<sup>61</sup> See footnote 58.

was different. For example, women were said to be less likely to be in in car accidents which was said to be relevant to car insurance premiums but more likely to seek medical treatment which was said to be relevant to health insurance premiums. Article 5(1) of Directive 2004/113/EC required member states to ensure that for all new contracts concluded after 21 December 2007, sex was not used as the basis to charge different insurance premiums. Article 5(2) provided that member states could decide *before that date* to permit proportionate differences in such premiums and benefits where the use of sex was a determining factor in the assessment of risk based on relevant and accurate actuarial and statistical data. A Belgian consumer rights association brought an action in the CJEU challenging Article 5(2) on the basis that it was incompatible with the principle of non-discrimination in relation to gender. AG Kokott and the CJEU in *C-236/09 Association belges des Consommateurs Test-Achats ABSL and others v Conseil des ministers* [2012] 1 WLR 1933 had little hesitation in finding that the principle of equal treatment is infringed if actuarial or statistical data is used as the basis of differential treatment. The principle of equality requires men and women to be treated the same in so far as they are in a comparable situations and generic risk profiling did not stop men and women from being comparable. Article 5(2) was accordingly found to infringe the principle of equal treatment. On that basis, we consider it likely that the court would conclude that the use of technology like HART infringes the principle of equal treatment contained in the EA 2010 in so far as less favourable treatment is occurring because of the protected characteristics of gender and / or race.

71. The position in relation to age is more nuanced. Unlike gender and race, there is always scope to justify direct age discrimination under the EA 2010. Accordingly, it is theoretically possible that the users of technology like HART could justify their actions in so far as different age groups are treated less favourably. However, cogent evidence would be required that HART was a proportionate means of preventing crime.

72. At this point, we should refer to the exception to the principle of non-discrimination contained in s.29 EA 2010 in relation to decisions concerning criminal proceedings. The relevant part is para 3 in Schedule 3 of Part 1 of the EA 2010 which reads as follows:

(1) Section 29 does not apply to:

(a) ...

(b) ...

(c) a decision not to commence or continue criminal proceedings;

(d) anything done for the purpose of reaching, or in pursuance of, a decision not to commence or continue criminal proceedings.

73. This provision is rather awkwardly drafted in that it is not immediately obvious if the exception covers *all* prosecutorial decisions or simply decisions *not* to commence or *not* to continue criminal proceedings. Thankfully, the position is much clearer when one examines the predecessor legislation which was essentially consolidated within the EA 2010 since it reveals that only *negative* prosecutorial decisions are exempted from the principle of non-discrimination. Hence, the Disability Discrimination Act 1995 contained the following provision:

21 C Exceptions from section 21B(1)<sup>62</sup>

(4) Section 21B(1) does not apply to –

- (a) a decision not to institute criminal proceedings;
- (b) where such a decision is made, an act done for the purpose of enabling the decision to be made;
- (c) a decision not to continue criminal proceedings; or
- (d) where such a decision is made –
  - (i) an act done for the purpose of enabling the decision to be made; or
  - (ii) an act done for the purpose of securing that the proceedings are not continued.

74. Interestingly, the appropriateness of an exemption in relation to negative prosecutorial decisions and disability discrimination was considered by the High Court in *R (B) v Director of Public Prosecutions (Equality and Human Rights Commission intervening)* [2009] 1 WLR 2072. In that case, the court concluded that the rationale for the exemption was “not hard to see” since prosecutors should be entitled to take into account, when reaching decisions about the reliability of evidence, that a disabled witness might not be able to provide reliable evidence in consequence of their disability (para 58). There are numerous criticisms<sup>63</sup> which can be made of that analysis but for current purposes it is sufficient to note that an algorithm which relies on race, age or gender to inform a prosecutorial decision in a *positive* way (i.e. a decision to pursue to continue criminal proceedings) can

---

<sup>62</sup> This provision specified that it was unlawful for a public authority to discriminate against a disabled person in carrying out its functions.

<sup>63</sup> See, for example, “Monghan on Equality Law”, Second Edition, para 11.56 – 11.62.

be a breach of the EA 2010 since the exemption in para 3 in Schedule 3 of Part 1 of the EA 2010 does not apply.<sup>64</sup>

### Sentencing decisions

75. In the US, algorithms are also being used in relation to sentencing decisions. The most famous example relates to an algorithm used within software called Compas. This is used in some states by judges to inform sentencing decisions.<sup>65</sup> This has led commentators such as journalists working for Propublica to analyse whether the Compas software creates discriminatory outcomes.<sup>66</sup> Propublica concluded that black defendants were twice as likely to be incorrectly labelled as high risk offenders by Compas.<sup>67</sup>

76. Whilst this type of technology is not yet being used in the UK, it is important to note that would probably not infringe the EA 2010. That is, a further exception to the principle of non-discrimination contained in s.29 EA 2010 pertains to judicial functions. The relevant part is para 3 in Schedule 3 of Part 1 of the EA 2010 which reads as follows:

(2) Section 29 does not apply to:

- (e) a judicial function;
- (f) anything done on behalf of, or on the instructions of, a person exercising a judicial function;
- (g) ...
- (h) ...

(3) A reference in sub-paragraph (1) to a judicial function includes a reference to a judicial function conferred on a person other than a court or tribunal.

---

<sup>64</sup> We should stress for completeness that we consider it unlikely that a decision by a custody officer would be classed as a judicial function and therefore fall under a different exemption in the EA 2010 which is discussed further below.

<sup>65</sup> <https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html>

<sup>66</sup> It is denied by Compas' makers that its technology is discriminatory.

<sup>67</sup> <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say>

77. There is no definition of “judicial function” within the EA 2010 beyond this provision. However, there are some related sources of information which suggest that the “judicial function” exception is intended to capture merits based decisions reached by judges and persons in a similar position. In particular:

- a. The Explanatory Notes that accompany the EA 2010 explain that: “A decision of a judge on the merits of a case would be within the exceptions in this Schedule. An administrative decision of court staff, about which contractor to use to carry out maintenance jobs or which supplier to use when ordering stationery would not be”.
- b. There is further guidance from the Equality and Human Rights Commission in its document entitled, “Your rights to equality from the criminal and civil justice systems and national security”<sup>68</sup> where the distinction between a judicial function and related decisions is unpicked. The following passage is material:

Equality law does not apply to what the law calls a judicial act. This means something a judge does as a judge in a court or in a tribunal case. It also includes something another person does who is acting like a judge, or something that they have been told to do by a judge.

**For example:** A father, who is a disabled person who has a visual impairment, applies to court for a residence order in respect of his child. The court refuses his application. He believes that this is because of his impairment. As the decision of the court is a judicial act, he may be able to appeal against the decision, but he cannot bring a case against the judge under equality law.

...

If the disabled person feels that he or she has been treated unfavourably in subsequent dealings with the Crown Prosecution Service or, in Scotland, the Procurator Fiscal’s office, for example if they refuse to call him as a witness because they think he will not present well to the jury because of his learning disability, or if the CPS only offers to meet him a place which is inaccessible to him without making reasonable adjustments, then they may well be able to bring a claim for unlawful discrimination under equality law.<sup>69</sup>

---

<sup>68</sup> <https://www.equalityhumanrights.com/sites/default/files/equalityguidance-criminal-civiljustice-2015-final.pdf>

<sup>69</sup> Page 26.

78. On this basis, technology like Compas could be utilised in the UK without falling foul of the EA 2010.<sup>70</sup> In light of Propublica’s research, this is an area which is likely to consider urgent consideration in the near future if algorithms start to be used in the UK’s legal system in relation to judicial decisions like sentencing.

## Human rights perspective

79. Finally, it is important not to overlook the potential human rights implications of the rise in technology. We suspect that you will be familiar with press stories explaining how robotics will help employers to plug gaps in the labour market. Robotic carers for older and vulnerable people appears to be gaining particular momentum.

80. There is a positive side to increased automation as assistive devices and robots can compensate for physical weaknesses by enabling people to bath, shop and be mobile. Tracking devices can also promote autonomy by allowing people to be remotely monitored. Some human rights instruments have gone as far as enshrining a right to assistive technology. For example, the UN Convention on the Rights of Persons with Disabilities states that assistive technology is essential to improve mobility.<sup>71</sup>

81. However, there are possible negative consequences as identified recently by the UN’s Independent Expert on the enjoyment of all human rights by older people in her report.<sup>72</sup> For example, consent to use assistive technologies might not be adequately sought from older people especially as there is still a prevalent ageist assumption that older people do not understand technology.

82. Overreliance on technology could lead to infantilisation, segregation and isolation. There is also a risk that artificial intelligence might replicate prejudice and discrimination. The report echoes the concerns identified by Sara Wachter-Boettcher when it states that –

---

<sup>70</sup> Please note that Karon Monaghan QC, however, argues with reference to the Human Rights Act 1998 and that the “judicial function” exception would not apply to merits based decisions where an individual would have no other means of challenging discriminatory behaviour. See “Monaghan on Equality Law”, Second Edition, para 11.48.

<sup>71</sup> “States Parties shall take effective measures to ensure personal mobility with the greatest possible independence for persons with disabilities, including by ... (b) Facilitating access by persons with disabilities to quality mobility aids, devices, assistive technologies and forms of live assistance and intermediaries, including by making them available at affordable cost ...” (*Article 20*).

<sup>72</sup> “Report of the Independent Expert on the enjoyment of all human rights by older people”, 21 July 2017, available here:

[https://ec.europa.eu/eip/ageing/library/report-un-independent-expert-enjoyment-all-human-rights-older-people\\_en](https://ec.europa.eu/eip/ageing/library/report-un-independent-expert-enjoyment-all-human-rights-older-people_en)

There is some evidence that artificial intelligence could reproduce and amplify human bias and as a result automated machines could discriminate against some people. Biased datasets and algorithms may be used in judicial decision-making, medical diagnoses and other areas that have an impact on older person's lives. Auditing machine-made decisions, and their compliance with human rights standards, is therefore considered necessary to avoid discriminatory treatment. (para 61)

83. This all indicates that businesses and public contractors and organisations, in the rush to create technological solutions to pressing social needs, must always assess carefully the products that they use bearing in mind the capacity that they have to be a source of discrimination and breaches of human rights, because in the right circumstances, individuals can rely on human rights instruments in litigation against service providers.

## Conclusion

84. We believe that technology is the next frontier for discrimination law. There is infinite scope for novel legal arguments about the application of anti-discrimination provisions and applicable human rights instruments.

85. Whilst most valuable litigation has been confined to the employment field, it is possible that claims will more become prevalent in the goods, facilities and services sphere and in a public setting. Where technology is discriminatory, the sheer number of possible claimants may mean that group actions will become prevalent and very expensive.

86. For many years the CJEU has been astute to prevent stereotypical ideas about women from causing them discrimination.<sup>73</sup> So while AI may discover that there are in some situations correlations which are typical, their thoughtless applications by businesses, public authorities and employers is likely to lead to trouble.

87. In January 2018 the Government announced the creation of a new Centre for Data Ethics & Innovation.<sup>74</sup> The search for Board appointees is under way as we write. The proposal is –

---

<sup>73</sup> See for instance Case C-345/89 *Stoeckel* [1991] ECR I-4047, and in particular the Opinion of AG Tesouro.

<sup>74</sup> See <https://www.gov.uk/government/news/search-for-leader-of-centre-for-data-ethics-and-innovation-launched>

... to make the UK the best place in the world for businesses developing artificial intelligence (AI) to start, grow and thrive.

[The centre] will advise on the measures needed to enable and ensure safe, ethical and innovative uses of data-driven technologies.

The UK already benefits from a world-class regulatory regime and the centre will build on this by making sure we understand and respond to the rapidly evolving ways in which data is impacting our lives.

The centre will make sure we have a governance regime which fully supports both ethical and innovative uses of these technologies. It will deliver its work through extensive engagement with industry, regulators, civil society and the public.

88. It could do well to start with tackling the interface between discrimination law and AI, apps and algorithms.

ROBIN ALLEN QC

DEE MASTERS

3 October 2018

© 2018 All rights reserved