


Current Reflections on EU Anti-Discrimination Law	 ERA european.law
Trier, 11-12th June 2018	

ALGORITHMS, APPS & ARTIFICIAL INTELLIGENCE: THE NEXT FRONTIER IN DISCRIMINATION LAW?

--

ROBIN ALLEN QC¹



www.cloisters.com

All websites visited since 7th June 2018.

¹ Joint – Head of Cloisters’ Barristers Chambers, 1 Pump Court, Temple, London EC4Y 7AA, + 44 (0) 20 7827 4000, ra@cloisters.com

CONTENTS

Introduction	3
A real problem.....	4
Artificial Intelligence	5
What was the Parliamentary Select Committee talking about?.....	6
Direct discrimination by algorithm.....	7
The liability of the code provider... ..	7
Harassment	8
Indirect Discrimination	9
Duty to make reasonable adjustments.....	10
Discouraging diversity by replicating the past	10
Human rights perspective	11
Conclusion.....	12

Introduction

1. It is a great pleasure to be back at the Academy of European Law in Trier. In the 2000s I gave many lectures here on the new Equality Directives and many from my chambers have come to lecture here since. We are always very well looked after. The barristers in my chambers [Cloisters](#) are constantly dealing with the most important issues in discrimination law. I would urge you to visit our website (www.cloisters.com) and to sign on to follow us on Twitter.
2. We will make sure you are alerted to our many equality law blogs [@Cloisters](#), many of our blogs are also picked up by other websites. Please do follow me at [@RobinAllenQC](#) – I promise not to overdo the tweets!
3. This conference on EU Anti-Discrimination Law is a great opportunity for academics and lawyers to discuss what is going on and to reflect on the next steps in this field. I am therefore delighted to be able to make a start on a discussion about some issues which have been causing us in Cloisters to be concerned.
4. So do you think that there could be a problem with discrimination problem inform Algorithms, Apps and Artificial Intelligence and machine learning? You might have thought - as many of my friends from the digital world had done till we discussed this issue - that this is issue has been created by lawyers in search of work!
5. As you will see just as we humans are learning the dangers of stereotyping, machines are taking over! Machine stereotyping confer benefits, and when it does – such as the early diagnosis of skin cancer² – it is useful. When it is used to the detriment of an individual it is discrimination and will often be unlawful.
6. So, my aim this morning 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 want to be involved in it in your own home states.
7. I want to make clear that this is not however going to be a talk saying that these are all bad things which should be stopped at any price. That would be neither wise nor necessary. Nonetheless there are reasons for us as jurists interested in equality law to be engaged in this development to make sure that its beneficial effects are not undermined by adverse ones.
8. I think the time to start that discussion is now. This issue was first raised with me by my colleague in Cloisters Dee Masters³ and we first wrote about this

² <https://www.theguardian.com/society/2018/may/29/skin-cancer-computer-learns-to-detect-skin-cancer-more-accurately-than-a-doctor>

³ <https://www.cloisters.com/barristers/dee-masters>

subject at the beginning of the year.⁴ We have had many enquiries and follow ups. This paper is a development of that work and I wish to acknowledge the considerable contribution she has made to this paper.

A real problem

9. How often does the media depicts 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.
10. 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.⁵
11. This side of the Atlantic is equally affected. For instance just last month⁶ the Science and Technology Committee of the UK Parliament said in the conclusions of its report on “Biometric strategy and forensics services”⁷ -

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.

12. In giving this advice the Parliamentary 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.

⁴ See <http://www.cloisters.com/blogs/algorithms-apps-artificial-intelligence-the-next-frontier-in-discrimination-law>

⁵ Published by W. W. Norton & Company, 2017, ISBN10 0393634639, ISBN13 9780393634631

⁶ On the 25th May 2018.

⁷ <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/800/80003.htm>

13. So before discussing what was concerning the Parliamentary Select committee let us look a little more at what exactly is this Artificial Intelligence or “AI” that we are talking about.

Artificial Intelligence

14. One of the best places to start that I know is in “The privacy pro’s guide to explainability in machine learning”, published by the International Association of Privacy Professionals (IAPP).⁸ Here are some extracts with the words and concepts we need to get to grips with in *italics* –

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*.

15. *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.
16. However, algorithms are 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. When they are based on the product of AI as well there is a double

⁸ <https://iapp.org/news/a/the-privacy-pros-guide-to-explainability-in-machine-learning/>

opportunity for bias to come into play. Let's have a look at this from the concern of the recent Parliamentary report.

What was the Parliamentary Select Committee talking about?

17. The basis for their concern was expressed in this way –

As we noted in our recent report on Algorithms,⁹ 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.

18. The research¹⁰ to which they were referring was carried out by Joy Buolamwini and Timnit Gebru and it is fascinating. Their Abstract is as follows –

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

19. As an interesting aside it is worth noting how the authors then went about trying to cure the bias by using males and females from a range of Parliaments from around the world.¹¹

⁹ See Science and Technology Committee, The use of algorithms in decision-making, Fourth Report of Session 2017–19, HC 351.

¹⁰ 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>

¹¹ Ibid. figure 1.

Direct discrimination by algorithm

20. One such algorithm might 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".
21. The problem, was that Maggie Delano is a lesbian and any Valentine's gift she might buy would most likely be for a woman.¹²
22. At a stroke of a line of code, Etsy had 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 Member States – such as my own – where discrimination on the grounds of sexual orientation in service provision is forbidden, this would be direct discrimination and a claim could theoretically be.
23. Another algorithm, also highlighted by Sara Wachter-Boettcher, was utilised by a chain of gyms in Britain called Puregym. In 2015, Louise Selby, a paediatrician, was unable to use her gym swipe card to access the locker rooms.
24. 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.¹³ This is of course ridiculous and very annoying but in the UK and I am sure in many other Member States it would be unlawful too.
25. Service providers need to be completely aware that 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...

26. However the problem does not stop there as in most cases the service provider will not have written the relevant code itself but have bought it from an outside source. 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

¹² Sara Wachter-Boettcher, *ibid.*, page 32-33.

¹³ Sara Wachter-Boettcher, *ibid.*, page 6.

discriminatory assumptions might not be immediately apparent to a purchaser of the software. In itself this raises profound issues of transparency of course.¹⁴

27. 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”. Dee Masters and I advise that they should also insist that the undertaking providing the code indemnifies them against any discriminatory effects that it may have. There is plenty to be worried about.
28. Discrimination claims against the 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.

Harassment

29. There is plenty of scope for technology to lead to harassment. Numerous illustrations are contained in Sara Wachter-Boettcher’s book, for example:
30. In August 2016, Snapchat introduced a face-morphing filter which was “inspired by anime”. In fact, the filter turned its users’ faces into offensive caricatures of Asian stereotypes.¹⁵
31. Smart phone assistants in 2017 nearly all have default female voices e.g. Apple’s Siri, Google Now and Microsoft’s Cortana. This echoes the dangerous gender stereotype that women, rather than men, are expected to be helpful and subservient.¹⁶
32. 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”.¹⁷ 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.
33. In the UK, users who are offended by this type of technology might be able to bring harassment claims against service providers.¹⁸ Although the compensation for injury to feelings in discrimination claims against service providers is often low, 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 then I can foresee

¹⁴ The issue of data management under the GDPR is outside this paper but is of itself very important.

¹⁵ Sara Wachter-Boettcher, *ibid*, page 7.

¹⁶ Sara Wachter-Boettcher, *ibid*, pages 37 – 38.

¹⁷ Sara Wachter-Boettcher, *ibid*, pages 129 – 132.

¹⁸ And perhaps GDPR claims as well.

that the CEOs of Google and other companies will be back in front of the European Parliament. So again precautions at the outset are called for.

Indirect Discrimination

34. Indirect discrimination usually is less of a reputational disaster, but it can be serious. Dee Masters and I 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.
35. In 2015, research revealed that of the top 50 “endless runner”¹⁹ 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.²⁰ Why on earth is that?
36. 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.²¹ 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.
37. 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.
38. 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.²² When he tried to register in 2014 he was informed that his name violated Facebook’s policy. Again the algorithm used by Facebook at this point could have been used as the basis of an indirect discrimination claim.

¹⁹ These are games where the objective is to keep virtual characters running as long as possible.

²⁰ Sara Wachter-Boettcher, *ibid.* page 3.

²¹ 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.

²² Sara Wachter-Boettcher, *ibid.* pages 54 - 55.

39. 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

40. 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.
41. 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”.
42. The difficulty, as Wachter-Boettcher points out, is that these scales only understand 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 by replicating the past

43. Technology could also have a worrying impact on diversity as AI becomes more prevalent. As already alluded to, some technology is based on recognising patterns and “learning” from existing historical data.
44. 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.²³ Similarly architects were deemed male and interior designers considered female.²⁴
45. 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

²³ Sara Wachter-Boettcher, *ibid.* pages 138 – 9.

²⁴ Ditto.

software could be used was to rate CVs so as to identify “matches” between potential employees and existing successful employees.²⁵ 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.

46. 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.

Human rights perspective

47. Finally, it is important not to overlook the potential human rights implications of the rise in technology. I 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.
48. 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.²⁶
49. 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.²⁷ 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.

²⁵ Sara Wachter-Boettcher, *ibid.* pages 139 - 140.

²⁶ “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).

²⁷ “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

50. 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 that²⁸ –

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.

51. 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

52. Dee Masters and I 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.
53. 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. Where technology is discriminatory, the sheer number of possible claimants may mean that group actions will become prevalent and very
54. 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.
55. For many years the CJEU has been astute to prevent stereotypical ideas about women from causing then discrimination.²⁹ So while AI may discover that there

²⁸ See [61].

²⁹ See for instance Case C-345/89 *Stoeckel* [1991] ECR I-4047, and in particular the Opinion of AG Tesauero.

are in some situations correlations which are typical their thoughtless applications by businesses and employers is likely to lead to trouble.

ROBIN ALLEN QC

7th June 2018

© 2018 All rights reserved

CLOISTERS' award winning employment team is

"...simply the go-to set if you want to win an Equality Act case'...It fields 'counsel at all ranges of the spectrum, all of whom are accessible, down to earth and able to put clients at ease'..." (Legal 500).

"...consistently involved in the most high-profile employment disputes..." (Chambers and Partners 2014).

Cloisters barristers are highly sought after for their astute practical approach to all aspects of employment disputes across all industry sectors.

OUR CLIENTS

We represent employers and employees in multinationals, SMEs, non-government organisations, charities, government departments, regulators, local authorities, associations and individuals

OUR SERVICES

We have a well-earned reputation for fearless advocacy and the highest of standards. Over three quarters of our employment team are ranked leaders. Many of our barristers achieve ground-breaking results and appear in almost all major employment litigation and landmark cases.

We regularly make submissions to government about development of the law. Our pioneering team includes three past Chairs of the Employment Law Bar Association, Chair and immediate past Chair of the Industrial Law Society, past Chair of the Discrimination Law Association. Many of our members sit on these committees

OUR TRAINING SERVICE

Cloisters offers CPD-accredited training courses, workshops and seminar programmes. Cloisters training programmes are highly regarded and well attended. Our members can deliver bespoke in-house training to law firms and professional bodies at their premises.

CLOISTERS DIRECT

Many of our members are qualified Public Access barristers especially trained to provide legal advice and litigation support directly to organisations.

If you are interested in discussing our work for employers, our training programmes, or wish to subscribe to our free news alerts please email clerks@cloisters.com

Phone - 020 7827 4000 /out of hours 020 7827 4040

Email - clerks@cloisters.com

Follow us on Twitter: @CloistersLaw

Post: Cloisters Chambers, 1 Pump Court, Temple, London EC4Y 7AA

Office opening hours: 8.30 am and 6.30 pm, from Monday to Friday.