University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

Fair AI : Challenges and Opportunities

Feuerriegel, Stefan ; Dolata, Mateusz ; Schwabe, Gerhard

DOI: https://doi.org/10.1007/s12599-020-00650-3

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-188091
Journal Article
Accepted Version

The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:
Feuerriegel, Stefan; Dolata, Mateusz; Schwabe, Gerhard (2020). Fair AI : Challenges and Opportunities. Business Information Systems Engineering, 62(4):379-384.
DOI: https://doi.org/10.1007/s12599-020-00650-3

Bus Inf Syst Eng
https://doi.org/10.1007/s12599-020-00650-3

CATCHWORD

Fair AI

Challenges and Opportunities

Stefan Feuerriegel · Mateusz Dolata · Gerhard Schwabe

Received: 27 December 2019 / Accepted: 1 April 2020
© The Author(s) 2020

Keywords Artificial intelligence · Algorithmic fairness · Decision support · Trust · Information systems

1 Introduction

Information systems (IS) are currently undergoing a fundamental shift: Until recently, decision support was developed upon rule-based and thus deterministic algorithms. However, with recent advances in artificial intelligence (AI), these decision rules have been replaced by probabilistic algorithms (e.g., deep learning; see Kraus et al. 2020). Probabilistic algorithms make inferences by learning existing patterns from data and, once deployed, provide predictions for unseen data under some uncertainty. Owing to this, they are prone to biases and systematic unfairness whereby individuals or whole groups are treated disparately.

The lack of fairness in AI applications has been repeatedly demonstrated in prior research. For instance, decision support systems for credit loan applications were found to favor certain socio-demographic groups in a disproportional way (Hardt and Price 2016; O'Neil 2016). As a consequence, people living in certain areas, those with a specific ethnic background, or women were less likely to obtain a loan from the bank. This can prevent whole neighborhoods from improving their standard of living and cause further economic and societal problems, thus reinforcing existing imbalances. Table 1 lists further examples of situations in which unfairness in AI could lead to discrimination against individuals or whole groups of people.

The term *"fair AI"* refers to probabilistic decision support that prevents disparate harm (or benefit) to different subgroups (Barocas and Selbst 2016). In fair AI, the objective is to provide systems that both quantify bias and mitigate discrimination against subgroups.[1] One might be inclined to think that simply omitting sensitive attributes from a decision support system will also solve fairness issues. However, this is a common misunderstanding: several non-sensitive attributes act as proxies (e.g., salary is a proxy of gender, ZIP code is a proxy for ethnicity, family structure is proxy of race or religion) and, hence, even decision support systems without knowledge of sensitive attributes are deemed unfair.

The objective of this article is to introduce IS practitioners and researchers to "fair AI". As detailed above, there are various areas within IS that are prone to unfairness. In fact, information systems maintain or even reinforce existing unfairness in AI rather than mitigating it. When relying upon such information systems, businesses

Accepted after two revisions by Ulrich Frank.

S. Feuerriegel (✉)
ETH Zurich, Weinbergstr. 56/58, 8092 Zurich, Switzerland
e-mail: sfeuerriegel@ethz.ch

M. Dolata · G. Schwabe
Department of Informatics, University of Zurich,
Binzmuehlestrasse 14, 8050 Zurich, Switzerland
e-mail: dolata@ifi.uzh.ch

G. Schwabe
e-mail: schwabe@ifi.uzh.ch

[1] Many applications that are subject to fairness issues originate from the area of supervised machine learning; however, fairness is also a concern in other areas of AI such as unsupervised learning (Garg et al. 2018) and even rule-based inferences. Hence, this catchword follows the terminology from Russell et al. (2015); that is, we use the term "AI" consistently as it allows us to highlight that the implications of (un)fairness as discussed in this catchword apply to all subareas.

Published online: 11 May 2020

Table 1 Example applications of AI with known fairness issues

Area	Fairness issues
Recidivism prediction	Automated systems such as COMPAS for predicting recidivism, i.e., the likelihood that a prisoner will commit a crime when released, were shown to deny release to people of color more often than to white people (Angwin et al. 2016). Compared against the number of actually committed crimes, the system was shown to have a racial bias, even though was not provided with explicit information about race in the first place, but information on the family structure, ZIP code, or education were available as proxies. (Chouldechova 2017)
Human resources	AI is increasingly used to screen job applications and identify promising candidates. Fairness laws forbid such systems to discriminate – either explicitly or implicitly – by gender, race, or disability. An example at Amazon (Barocas et al. 2018) showed that such information might be not available in an explicit manner yet that the probabilistic algorithms behind AI might use other data as proxy, e.g., a birth place as a proxy for race
Image classification	Algorithms that were trained with, e.g., Google Images have learned to make inferences from mostly white persons and thus are more likely to make errors when classifying pictures of black persons, e.g., by misidentifying them as objects or ignoring them altogether (Zou and Schiebinger 2018). This has implications for the accuracy of face recognition for logging-in to smartphones
Natural language processing	Using neural networks for text representation highlights that existing biases were replicated in computational representations (Garg et al. 2018). As a result, generated texts can include content or words that are generally considered racist or discriminating against minorities

and organizations are exposed to substantial legal risk. In this vein, legislative bodies across the world are implementing laws that forbid disparate treatments in algorithmic decision-making (White & Case 2017); e.g., in the US, fair lending laws penalizes algorithmic biases in risk scoring, while, in the EU, accountability for artificial intelligence is enforced by the General Data Protection Regulation (GDPR). Hence, achieving fair AI is relevant to both the potential victims of discrimination and the institutions that rely on AI within their decision support systems.

Recent reports point out that the adoption of fair AI in businesses, organizations, and governments is lagging behind (AI Now Institute 2018). As we will discuss later, potential reasons for this sluggish progress are located along all dimensions of IS, namely people (e.g., trust), technology (e.g., design principles, economic implications), and organizations (e.g., governance). In the following sections, this article reviews theoretical concepts of fairness, links them to fairness of AI, and derives suggestions for IS research.

2 Background

2.1 Definitions and Origins of Fairness

Fairness, understood as the impartial and just treatment of people, has been dealt with in philosophical and theological discussions for centuries, often in connection with justice (Miller 2017). These discussions have been long dominated by the question of what distribution of what rights is fair. For instance, let us consider Aristotle's example (Cooper 1996) of distributing flutes among a group of musicians, when there are fewer flutes than musicians. Different options emerge: distributing the playing time equally between the musicians, distributing the flutes at random, providing the flutes to the most skilled musicians, holding regular competitions to choose those who will receive the honor of playing the flutes, etc.

In today's understanding of democratic societies, the above example is addressed by the idea that all people with equal gifts should have equal opportunities regardless of their initial position in society (Rawls and Kelly 2003). To this end, fairness refers to the equal distribution of chances for self-advancement as a way to achieve equity in the distribution of goods (i.e., individuals' benefits are proportional to their input). This also implies what is considered unfair: preventing individuals from improving their situation (e.g., by limiting their access to a loan) based not on their contribution to society (e.g., conducting specific work), but rather based on unrelated choices (e.g., neighborhood) or sensitive characteristics (e.g., ethnicity, gender).

Anthropology focuses on the social origins of fairness, arguing that it is an innate aspect of being a human or even a primate (Brosnan 2013). Here prior research suggests that fairness has been developed as an effective strategy in evolution in order to foster collaboration (Hamann et al. 2011). The focus is on interactions between humans, while new forms of collaborations between humans and AI are emerging (Seeber et al. 2020).

Overall, fairness is not a clear-cut concept. However, it is evident that unfairness has a substantial impact on the functioning of societies. Therefore, it is key to build information systems which are capable of detecting unfairness and dealing with it in an adequate manner. Using mathematical notions of fairness can offer a step in this direction.

2.2 Mathematical Notions of Fairness in AI

Different definitions have been put forward that formalize fairness in AI mathematically (cf. Barocas et al. 2018, for overview). These can be grouped into concepts (so-called notions) of fairness across (i) groups or (ii) individuals, as detailed in the following.

2.2.1 Group-Level Fairness

Group-level fairness builds upon a predefined sensitive attribute (e.g., race, gender, disability) that describes membership in a protected group s_1. Membership in the protected group should not lead to discrimination. In group-level fairness, discrimination is interpreted by how errors of the prediction model are distributed across groups, in particular within the protected group s_1 vs. outside of it, i.e., the reference group s_2. Note that there is no universal definition of group-level fairness and we thus point to common examples in the following. Furthermore, it is actually mathematically impossible to fulfill all of the following definitions at the same time (Kleinberg et al. 2017). Therefore, the preferred notion of fairness must be chosen by IS practitioners.

Statistical parity represents a simple concept of fairness that is widespread in legal applications. Statistical parity requires the likelihood of events to be equal across groups: to this end, the proportion of affected individuals should be roughly the same inside the protected group and outside of it. Statistical parity focuses only on the predicted outcomes (i.e., the likelihood of paying back the loan), but neither on the actual outcomes (i.e., the fact of paying back the loan) nor on the opportunity due to predictions (i.e., access to a loan).

Other definitions of group-level fairness in AI are tailored to errors in predictions (e.g., Corbett-Davies and Goel 2018; Hardt and Price 2016; Kleinberg et al. 2017). For instance, the so-called equality of accuracy requires that algorithms for AI attain equal prediction accuracies across groups. It relies on the ratio $acc(x)$ of correctly classified individuals over the whole population x. If the accuracy in the protected group, $acc(s_1)$, equals the accuracy in the reference group, $acc(s_2)$, then this algorithm is considered fair according to the accuracy parity metric. In the loan example, this would imply that the same ratio of applicants is classified correctly, independent of whether they belong to the protected group or not. One of the downsides of this approach is that type-I and type-II errors receive the same weight. Hence, alternative definitions have been put forward that specifically focus on these metrics (e.g., equalized odds) or that maintain calibrated class probabilities. This points to an inherent challenge: there is a multitude of different fairness definitions out of which many are mutually exclusive (cf. Corbett-Davies and Goel 2018, for an overview).

The above definitions points towards a key requirement in order to apply group-level fairness: the data must include attributes which allow for the identification of protected groups. However, in many cases, providing such identifiers is forbidden (e.g., in the US, it is often not allowed to ask for ethnic background).

2.2.2 Individual Fairness

Individual fairness is based on the notion that similarly situated individuals should be treated in a similar way (Dwork et al. 2012). Consequently, this approach strives to ensure fairness independent of group membership. Let us consider a classifier f and two individuals x_1 and x_2. Individual fairness would require that the outputs of the classifier be similar for similar individuals, i.e., $f(x_1) \approx f(x_2)$ for $x_1 \approx x_2$. In practice, this relies upon a mathematical definition to measure similarity. Referring to the example of loan applications, this requires that two individuals whose relevant attributes (yearly income, savings, etc.) are equal should be granted equal access to a loan and should be offered the same interest rates.

2.3 Sources of Unfairness in AI

Fairness in AI is violated by so-called biases. In this context, we define *bias* as a systematic deviation of an estimated parameter from true value. Biases can emerge along the complete AI pipeline (Barocas and Selbst 2016), namely with regard to (1) data, (2) modeling, and (3) inadequate applications as discussed in the following.

Data are used for making inferences; however, if data are subject to biases, the same biases are replicated. Examples of biases in this context are similar to biases that can appear in behavioral experiments and stem from data generation or data annotation (Ahsen et al. 2019). For instance, a selection bias occurs when the data are not representative of the wider population or else annotated in a manner that reaffirms the annotators' beliefs or assumptions. This can arise, e.g., when AI is trained for evaluating loan applications with past data for a bank that was restrictive towards young adults, in which case this bias will be subsequently replicated.

Modeling selects relevant features as input and combines them in a meaningful way, though often relying upon correlation rather than causation. Hence, modeling can also be a source of bias due to variables acting as proxies or confounders. For instance, even if race is blinded, an AI can "guess" this value based on where a person lives; the ZIP code, then, can function as a proxy for race or ethnic background.

Inadequate applications of the model might occur in dynamic settings with drifts or non-stationarities in the underlying population. Here the data from the training population differs from that of the population after deployment. For instance, an AI for evaluating loan applications might have been trained on applicants from one country, but, when used in another country, might not recognize that socio-demographic variables are distributed differently and thus provides incorrect assessments. Furthermore, such an information system is not able to improve if it does not receive data concerning the actual repay rate of those individuals because their loans were denied a priori (see literature on reject inference; e.g., Li et al. 2017).

All in all, IS practitioners must be aware that biases arise at various steps within the AI pipeline and can have multiple sources. The above shows that removing humans from decision support systems does not necessarily prevent biases but, on the contrary, might even reinforce them. Many of the sources of unfairness are not straightforward to identify but instead require thorough domain knowledge.

2.4 Algorithms for Fair AI

Algorithms for fair AI have different objectives, aiming at measuring fairness, designing fair predictions, or modeling fair decisions. (1) Measuring fairness in AI commonly builds upon an in-depth analysis of the prediction performance, where type-I and type-II error rates are critically compared across subgroups. To facilitate this, inequality metrics for algorithmic assessments have been developed (Barocas et al. 2018). (2) Designing fair predictions is commonly achieved by reducing the prediction performance of the majority group so that it approaches the (lower) prediction performance of the minority group (e.g., Hardt and Price 2016; Haas 2019). This can occur in different ways (Friedler et al. 2019), namely via preprocessing techniques, modifications of the underlying classifier, or postprocessing techniques. (3) Modeling fair decisions is approached in different ways, often requiring tailored approaches that carefully model feedback loops.

For IS practitioners, there are a few (proprietary) tools that have been recently developed for ensuring fairness, such as IBM's *AI Fairness 360*. However, these mostly provide programming libraries only, whereas key questions related to the IS design – namely people, technology, and organizations – remain unsolved.

3 Challenges and Opportunities for IS Research

Hitherto, research on fair AI has been primarily conducted by researchers from computer science. However, as discussed above, fair AI has the potential to radically change the nature of decision support in away that exposes all domains to the risk of discrimination. Hence, fair AI has serious social, technological, and organizational implications, which require a holistic, scientific approach. Given the multidisciplinary background of IS, researchers from this field seem ideally suited to explore the capabilities and implications of fair AI. Table 2 summarizes existing gaps with respect to IS research, which are detailed in the following.

3.1 People

Extensive research is required to study user perceptions of fair AI. For instance, a better understanding is needed of which attributes are regarded as sensitive. In practice, sensitive attributes are likely to vary with the underlying use case. For instance, some attributes seem obvious (e.g., race), while other attributes are defined more vaguely (e.g., Christian or American), or are domain-specific (e.g., physically attractive).

Fair AI is related to the wider problem of value alignment: fairness is an important value for humans, one which needs to be taught to AI in decision support systems. IS has the chance to make a lasting impact in this area by specifying models for translating human values identified in philosophy or the social sciences to actionable design principles.

Trust represents the primary prerequisite for an IS ecosystem to succeed (Hurni and Huber 2014). In traditional IS studies, users transfer trust from people or institutions to an IT artifact. Yet AI challenges the traditional conceptualization of trust, since the logic behind its reasoning can often be barely understood. To this end, future research should investigate how fair AI can help in building trust.

3.2 Technology

Several challenges exist when adapting fair AI to applications in practice. For instance, regulatory initiatives such as the GDPR enforce transparent algorithms, yet further research is required to reconcile transparent decision support with fair AI. Statistical approaches for modeling causality (Pearl 2013) are regarded by some as a way to implement fair AI that is tailored to specific uses cases. IS is equipped with the means to develop said casual models (e.g., structural equation models) and make a distinctive contribution to both practice and research in IS. However, this relies upon the premise that the philosophical concept of causality can be described in mathematical language.

IS practitioners demand design principles for implementing fair AI. Here IS as a discipline has the means to

Table 2 Suggested areas of IS research to advance fair AI

People	– Perceptions of fair AI
	– Value alignment between AI and humans
	– Trust towards fair AI
Technology	– Algorithms for fair AI
	– Design principles for IS with fair AI
	– Economic implications of fair AI
Organization	– Business models with respect to fair AI
	– Governance of AI to ensure fairness
	– Policy-making for fair AI

derive and test the design principles that guide the decision-making of practitioners (e.g., in choosing a definition of fair AI that is effective for the relevant domain application). Altogether, these efforts can result in information systems where the AI achieves "fairness by design".

Fair AI has direct implications for economics of IS. This is because fair AI is subject to a fairness-performance trade-off: fairness is achieved at the cost of lowering the prediction performance for certain subgroups (Haas 2019). However, the economic implications have been overlooked, despite the fact this this represents a key prerequisite for management decisions and thus industry adoption.

3.3 Organization

Fair AI is likely to have an impact on businesses and organizations. For instance, it can render new business models with regard to decision support systems feasible that would have been otherwise restricted by fairness laws. Building upon this, IS practitioners require a better understanding of how fair AI is linked to value propositions, value chains, and revenue models.

Organizational aspects of fair AI are strongly linked to governance. It has been argued that internal governance structures are failing at assuring the fairness of AI (AI Now Institute 2018). Hence, the effectiveness of different governance structures for the management of fair AI and their relation to other ethics-oriented processes should be investigated. IS, given its interest in governance of change and technology, has the potential to establish a new management framework with the goal of achieving fair AI.

Both governance and business models rely upon the legal frame offered by policy-makers. Given the increasing call for the regulation of AI applications in public and private spheres, various regulatory bodies have initiated discussions regarding the ethical and practical aspects of AI (European Commission 2018). In this context, IS research, thanks to its real-world impact and expertise in industry, has the opportunity to shape policies.

4 Outlook

A recent BISE editorial specifically has called for *"data science without prejudice"* (van der Aalst et al. 2017) and there is much interest in understanding the real-world implications of fair AI (Martin 2019). Our article provides a starting point for IS researchers pursuing fair AI. For IS researchers and practitioners, this endeavor is of direct relevance: First, AI is expected to become more powerful and pervasive, thus also raising concerns among the public. Second, fairness in decision support systems will soon be enforced by various legal initiatives and, hence, appropriate tools for fair AI must be developed. Third, businesses and organizations without a clear strategy for achieving fair AI run various risks: violating fairness laws poses immense financial and reputational risks.

Needless to say, fair AI introduces unprecedented opportunities for people, organizations, and society. Developing mathematical notions for this purpose allows practitioners to statistically quantify the level of fairness in their information systems and to monitor the effectiveness of fair AI in decision support systems over time. Finally, fair AI promises to reduce discrimination over the status quo: human decision-making is subject to biases, whereas algorithms can be derived to be fair by design.

Acknowledgements Stefan Feuerriegel acknowledges support from the Swiss National Science Foundation (SNSF) via a Digital Lives Grant (183149).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ahsen ME, Ayvaci MUS, Raghunathan S (2019) When algorithmic predictions use human-generated data: a bias-aware classification algorithm for breast cancer diagnosis. Inf Syst Res 30(1):97–116. https://doi.org/10.1287/isre.2018.0789

AI Now Institute (2018) AI Now report 2018. https://ainowinstitute.org/AI_Now_2018_Report.pdf. Accessed 14 Aug 2019

Angwin J, Larson J, Mattu S, Kirchner L (2016) Machine bias. ProPublica https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Accessed 14 Aug 2019

Barocas S, Selbst AD (2016) Big data's disparate impact. Calif Law Rev 104:671–732. https://doi.org/10.15779/Z38BG31

Barocas S, Hardt M, Narayanan A (2018) Fairness and machine learning. http://www.fairmlbook.org. Accessed 01 Apr 2020

Brosnan SF (2013) Justice- and fairness-related behaviors in nonhuman primates. Proc Natl Acad Sci USA (PNAS) 110(Supplement 2):10,416–10,423. https://doi.org/10.1073/pnas.1301194110

Chouldechova A (2017) Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. Big Data 5(2):153–163. https://doi.org/10.1089/big.2016.0047

Cooper JM (1996) Justice and rights in Aristotle's "politics". Rev Metaphys 49(4):859–872

Corbett-Davies S, Goel S (2018) The measure and mismeasure of fairness: a critical review of fair machine learning. arXiv:1808.00023

Dwork C, Hardt M, Pitassi T, Reingold O, Zemel R (2012) Fairness through awareness. In: Conference on innovations in theoretical computer science (ITCS). https://doi.org/10.1145/2090236.2090255

European Commission (2018) Fair and unbiased algorithmic decision making: current state and future challenges. JRC digital economy working paper 2018-10

Friedler SA, Scheidegger C, Venkatasubramanian S, Choudhary S, Hamilton EP, Roth D (2019) A comparative study of fairness-enhancing interventions in machine learning. In: Conference on fairness, accountability, and transparency (fat*), pp 329–338. https://doi.org/10.1145/3287560.3287589

Garg N, Schiebinger L, Jurafsky D, Zou J (2018) Word embeddings quantify 100 years of gender and ethnic stereotypes. Proc Natl Acad Sci USA 115(16):E3635–E3644. https://doi.org/10.1073/pnas.1720347115

Haas C (2019) The price of fairness: a framework to explore trade-offs in algorithmic fairness. In: International conference on information systems (ICIS)

Hamann K, Warneken F, Greenberg JR, Tomasello M (2011) Collaboration encourages equal sharing in children but not in chimpanzees. Nature 476(7360):328–331. https://doi.org/10.1038/nature10278

Hardt M, Price E (2016) Equality of opportunity in supervised learning. In: Advances in neural information processing systems (NIPS). http://papers.nips.cc/paper/6373-equality-of-opportunity-in-supervised-learning

Hurni T, Huber T (2014) The interplay of power and trust in platform ecosystems of the enterprise application software industry. In: European conference on information systems (ECIS). https://doi.org/10.7892/boris.43332

Kleinberg J, Mullainathan S, Raghavan M (2017) Inherent trade-offs in the fair determination of risk scores. In: Conference on innovations in theoretical computer science (ITCS)

Kraus M, Feuerriegel S, Oztekin A (2020) Deep learning in business analytics and operations research: models, applications and managerial implications. Eur J Oper Res 281(3):628–641

Li Z, Tian Y, Li K, Zhou F, Yang W (2017) Reject inference in credit scoring using semi-supervised support vector machines. Exp Syst Appl 74:105–114. https://doi.org/10.1016/j.eswa.2017.01.011

Martin K (2019) Designing ethical algorithms. MIS Q Exec 18(2):129–142

Miller D (2017) Justice. In: Zalta EN (ed) The stanford encyclopedia of philosophy. Stanford University, Stanford. https://plato.stanford.edu/info.html

O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy, 1st edn. Broadway Books, New York

Pearl J (2013) Causality: models, reasoning, and inference, 2nd edn. Cambridge University Press, New York

Rawls J, Kelly E (eds) (2003) Justice as fairness: a restatement, 3rd edn. Harvard University Press, Cambridge

Russell S, Dewey D, Tegmark M (2015) Research priorities for robust and beneficial artificial intelligence. AI Mag 36(4):105. https://doi.org/10.1609/aimag.v36i4.2577

Seeber I, Bittner E, Briggs RO, de Vreede T, de Vreede GJ, Elkins A, Maier R, Merz AB, Oeste-Reiß S, Randrup N, Schwabe G, Söllner M (2020) Machines as teammates: a research agenda on ai in team collaboration. Inf Manag. https://doi.org/10.1016/j.im.2019.103174

van der Aalst WMP, Bichler M, Heinzl A (2017) Responsible data science. Bus Inf Syst Eng 59(5):311–313. https://doi.org/10.1007/s12599-017-0487-z

White & Case (2017) Algorithms and bias: what lenders need to know. https://www.whitecase.com/sites/whitecase/files/files/download/publications/algorithm-risk-thought-leadership.pdf. Accessed 06 Aug 2019

Zou J, Schiebinger L (2018) AI can be sexist and racist: it's time to make it fair. Nature 559(7714):324–326. https://doi.org/10.1038/d41586-018-05707-8