

## Bias does not equal bias

Baumgartner, Renate; Kuhn, Sarah

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

### Empfohlene Zitierung / Suggested Citation:

Baumgartner, R., & Kuhn, S. (2021). Bias does not equal bias. *TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis / Journal for Technology Assessment in Theory and Practice*, 30(2), 69-70. <https://doi.org/10.14512/tatup.30.2.69>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

MEETING REPORT

# Bias does not equal bias

Renate Baumgartner, Center for Gender and Diversity Research,  
University of Tübingen, Brunnenstr.30, 72074 Tübingen, DE  
(renate.baumgartner@uni-tuebingen.de) 0000-0002-3401-1870

Sarah Kuhn, Institute of Sociology, University of Tübingen, Wilhelmstraße 36,  
72074 Tübingen, DE (sarah.kuhn@student.uni-tuebingen.de)

The interdisciplinary conference *Fair medicine and artificial intelligence* was held at the Center for Gender and Diversity Research, University of Tübingen, from 3–5 March. About 70 participants from the social sciences, philosophy, and medical ethics developed socio-technical perspectives on artificial intelligence (AI) and machine based applications as well as deep learning technologies in the medical and healthcare sectors. The big promises for possible future applications of AI in the professional medical field, e. g. diagnosis, prognosis, and therapy recommendations, lead to assumptions that AI will soon play an important role in the health sector and will help to address healthcare disparities, which are currently posing an individual threat, a threat to social justice and a major challenge to the healthcare system. AI could, for instance, reveal human bias, provide more equal treatment to all patients, make health care more accessible and identify possibilities of improvement to develop a more just healthcare system. However, critical voices warn that AI might heighten existing inequalities as technical complexities make them harder to detect.

## Socio-technical perspectives

*Future applications of AI.* E. Detfurth (York University) talked about AI data-driven applications for dementia care. She showed how classification systems of data repositories and brain atlases find their way into AI tools and described chances and limitations of AI-assisted dementia diagnosis. A. K. Kühnen (TU Dresden) focused on the question of representation of BIPOC, concluding that AI might reproduce and exacerbate inequalities between “white” and non-white racial groups. When analyzing racial bias, technological as well as economic, historical and biopolitical aspects need to be considered. C. Bath and S. Samerski (TU Braun-

schweig & Hochschule Emden Leer) spoke about a framework for examining AI-based health apps for diagnosis. They will use feminist concepts of agency from Science and Technology Studies (STS) as framework in their upcoming ethnographic study on classifications and biases in data. K. Napiwodzka and K. Cierszko (Adam Mickiewicz University, Poznań) asked whether AI could provide a safe space for female body politics in the highly contested political environment in Poland. P. Martin and J. Ding (University of Sheffield) presented the repurposing of common drugs for the treatment of rare diseases as a possible use of AI in medical research. The assistance of AI could accelerate and economize the approval process and improve access to therapy. W. Ernst (Johannes Kepler University, Linz) asked if standards of medical research can be questioned with AI and pointed out issues such as: Who can be the representative of who’s body? How are categories envisioned? And to whose benefit?

*Fairness and diversity in medical AI.* C. Kropp and K. Tampe-Mai (University of Stuttgart) considered ‘accessibility’ as a crucial point for social in/justice in AI based smart healthcare systems. The constraints to be considered are financial access, usability and digital health literacy. S. Morais dos Santos Bruss (TU Dresden) used a feminist-decolonial perspective to explore “surrogate”-robotics and the care-revolution. H. Drukarch (Leiden University) showed how a lack of diversity for AI in medicine means erasure, exclusion and silencing of minorities. However, data can also construct (new) normalities when presented as facts and reproduce and naturalize social categories.

*Postcolonial perspectives.* K. Vlantoni and K. Papanastasiou (National and Kapodistrian University of Athens) analyzed expectations concerning the integration of medical AI in Greece against the background of technological enthusiasm and nationalism. S. Mbelu (Erasmus University Rotterdam) reflected on how to design and provide AI enabled health insurance platforms in Nigeria without the pitfalls the Global North has experienced. He concluded that new technologies may have the potential to add important value, e. g. for healthcare universalism, but also exacerbate health disparities to the detriment of the most vulnerable. With the example of Native American Tribes, T. Hendl and T. Roxanne (LMU Munich) pointed out risks of using digital surveillance for racialized minorities during the COVID-19 pandemic. They argued for the inclusion of and respect for indigenous perspectives and indigenous data sovereignty.

*Discourses and knowledge production.* The third day of the conference started with K. Wiggert’s (TU Berlin) study on data-driven clinical decision support systems for cardiology-related diseases that allow physicians to simulate effects of different treatment strategies. The tools reshape medical reasoning and decision-making. However, physicians who were involved in the process ultimately did not feel represented by the tool. Collaborations between engineers and physicians throughout the development process thus should be based on the needs of physi-

This is an article distributed under the terms of the Creative Commons Attribution License CCBY 4.0 (<https://creativecommons.org/licenses/by/4.0/>)  
<https://doi.org/10.14512/tatup.30.2.69>  
Received: Apr. 15, 2021; revised version accepted: May 06, 2021;  
published online: Jul. 26, 2021 (editorial peer review)



**Illustration 1:** This picture was created by a StyleGAN neuronal network following the input of “diversity” training pictures extracted from Google. Source: Timo Dufner

cians and not only on the ideas of engineers, to improve this aspect. V. Galanos (The University of Edinburgh) opened with an assessment of the discrepancy between AI in public discourses and research. He explored the balance between the accuracy and instability of AI in radiology and proposed the inclusion of contextual reasoning in the development of AI to avoid pitfalls. R. Baumgartner (University of Tübingen) proposed to take on the goal of “health equity” over “fairness” and laid out one of the key challenges in reaching health equity through AI, the “participation vs. privacy dilemma”, concluding that the worth of privacy balanced against the asset of being represented in data-based AI tools is more precarious for minorities than for the majority population.

*Ethical perspectives.* C. Lenk (Ulm University) argued that collection of data variables, such as social determinants of health in patient data, so far is insufficient to consider healthcare inequality. P. Lopez (University of Vienna) presented a new socio-technical typology in data-based algorithmic systems that distinguishes between societal, socio-technical and technical biases. T. Grote (University of Tübingen) talked about the normative relevance of different accounts of algorithmic bias in medical practice. He concluded that in decision support systems, as opposed to automated systems, the assurance of fairness in the final decision is in practice more relevant than algorithmic fairness per se. The last presentation by T. Gremsl and D. Schneeberger (University of Graz) combined ethical and legal perspectives while presenting interdisciplinary commentaries on the proposed European framework of ethical aspects of AI, robotics and related technologies.

*Keynotes.* C. Bath (TU Braunschweig), in the first keynote of the conference, identified algorithmic bias as rooted in discriminatory beliefs of humans whose values and norms are inscribed into the tools. She proposed design methods against discrimination and exclusion informed by gender studies, femi-

nist STS and user-driven approaches to develop non-biased AI. She furthermore suggested starting the development process with the definition of a problem that needs to be solved and not with imagining users, because the later step is prone to mistakes.

In the second keynote, K. Ferryman (New York University, Tandon School of Engineering) presented the *Fairness in precision medicine project* which uses critical medical anthropology and a STS approach to take on the task to center health equity in precision medicine. She also showed how being included in data is important and at the same time a privacy and security risk.

### Lessons learned and outlook

In her summary discussion R. Ammicht-Quinn, director of the hosting center, pointed to topics and questions that kept resurfacing during the conference, in particular issues of categorization, objectification and representation in different health contexts and technologies. Which categorizations are at work and how do they work? Are categorizations a form of objectification? Who can represent (and be representative of) whose body, who is relevant for a specific representation and who counts as standard?

Throughout the conference biases in medical AI had been a focal point on different levels. The question “how does fairness relate to the most desirable bias?” remains – even if we accept that nothing is without bias. Several speakers talked about diversity in all its ambivalences. How is it possible to avoid erasure, exclusion and silencing? How do we deal with the dilemma of either seeking participation or valuing patients’ privacy? The benefits of health AI to facilitate development of treatment of rare diseases was one of the few examples that focused on chances and advantages of AI in medicine. Through the vast majority of topics, the critical analysis of AI pervaded the talks. This seemingly huge gap between the promotion of benefits and fundamental critique was also addressed in both keynotes. AI holds the potential to facilitate and accelerate processes within medicine and healthcare to promote health for all. At the same time, we must be wary of *whose* values and *which* knowledge are inscribed in data and technique during the development processes.