Bias, Fairness, and Justice

Readings on the normative concepts of bias, fairness, and justice and how they are relevant considerations in data science.

Title	Citation
Are Algorithms Value-Free?	Johnson (2023)
Algorithmic Bias: on the Implicit Biases of Social Technology	Johnson (2020)
Algorithmic Bias: Senses, Sources, Solutions	Fazelpour & Danks (2021)
Fairness	Vredenburgh (2024)
Just Machines	Castro (2022)
Big Data and Compounding Injustice	Hellman (2023)
On the Site of Predictive Justice	Lazar & Stone (forthcoming)
Proceed with Caution	Zimmermann & Lee-Stronach (2021)
Semantics Derived Automatically from Language Corpora Contain Human-like Biases	Caliskan, Bryson, & Narayanan (2017)
The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems	Creel & Hellman (2022)
Facebook has been Charged with Housing Discrimination by the US government	Brandom (2019)

References

Brandom, R. (2019). *Facebook has been charged with housing discrimination by the US government*. https://www.theverge.com/2019/3/28/18285178/facebook-hud-lawsuit-fair-housing-discrimination; Oxford University Press.

Caliskan, A., Bryson, J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*, 183–186. https://doi.org/10.1126/science.aal4230

Castro, C. (2022). Just machines. *Public Affairs Quarterly*, *36*(2), 163–183. https://doi.org/10.5406/21520542.36.2.04

Creel, K., & Hellman, D. (2022). The algorithmic leviathan: Arbitrariness, fairness, and opportunity in algorithmic decision-making systems. *Canadian Journal of Philosophy*, *52*(1), 26–43. https://doi.org/10.1017/can.2022.3

Fazelpour, S., & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, *16*(8), e12760. https://doi.org/10.1111/phc3.12760

Hellman, D. (2023). Big data and compounding injustice. *Journal of Moral Philosophy*, *21*(1-2), 62–83. https://doi.org/10.1163/17455243-20234373

Johnson, G. M. (2020). Algorithmic bias: On the implicit biases of social technology. *Synthese*, *198*(10), 9941–9961. https://doi.org/10.1007/s11229-020-02696-y

Johnson, G. M. (2023). Are algorithms value-free? *Journal Moral Philosophy*, *21*(1-2), 1–35. https://doi.org/10.1163/17455243-20234372

Lazar, S., & Stone, J. (forthcoming). On the site of predictive justice. *Noûs*. Forthcoming. https://doi.org/10.1111/nous.12477

Vredenburgh, K. (2024). Fairness. In *The Oxford Handbook of AI Governance*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.8

Zimmermann, A., & Lee-Stronach, C. (2021). Proceed with caution. *Canadian Journal of Philosophy*, (1), 6–25. https://doi.org/10.1017/can.2021.17