

## References

- Elizabeth S Anderson. What is the point of equality? *Ethics*, 109(2):287–337, 1999.
- Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine bias. *Pro Publica*, 2016.
- Richard J Arneson. Equality and equal opportunity for welfare. *Philosophical studies*, 56(1):77–93, 1989.
- Solon Barocas and Andrew D Selbst. Big data’s disparate impact. *Cal. L. Rev.*, 104:671, 2016.
- Reuben Binns, Michael Veale, Max Van Kleek, and Nigel Shadbolt. Like trainer, like bot? inheritance of bias in algorithmic content moderation. In *International Conference on Social Informatics*, pages 405–415. Springer, 2017.
- Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. In *Advances in Neural Information Processing Systems*, pages 4349–4357, 2016.
- Aylin Caliskan-Islam, Joanna J Bryson, and Arvind Narayanan. Semantics derived automatically from language corpora necessarily contain human biases. *arXiv preprint arXiv:1608.07187*, 2016.
- Gerald A Cohen. On the currency of egalitarian justice. *Ethics*, 99(4):906–944, 1989.
- William Dieterich, Christina Mendoza, and Tim Brennan. Compas risk scales: Demonstrating accuracy equity and predictive parity. *Northpoint Inc*, 2016.
- Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 214–226. ACM, 2012.
- Ronald Dworkin. What is equality? part 1: Equality of welfare. *Philosophy & public affairs*, pages 185–246, 1981.
- Benjamin Eidelson. *Discrimination and Disrespect*. Oxford University Press, 2015.
- Hanming Fang and Andrea Moro. Theories of statistical discrimination and affirmative action: A survey. Technical report, National Bureau of Economic Research, 2010.
- Nancy Fraser and Axel Honneth. *Redistribution or recognition?: a political-philosophical exchange*. Verso, 2003.
- Margaret Gilbert. Collective epistemology. *Episteme*, 1(2):95–107, 2004.
- Sara Hajian and Josep Domingo-Ferrer. A methodology for direct and indirect discrimination prevention in data mining. *IEEE transactions on knowledge and data engineering*, 25(7):1445–1459, 2013.
- Moritz Hardt, Eric Price, Nati Srebro, et al. Equality of opportunity in supervised learning. In *Advances in Neural Information Processing Systems*, pages 3315–3323, 2016.
- Deborah Hellman. *When is discrimination wrong?* Harvard University Press, 2008.
- Elisa Holmes. Anti-discrimination rights without equality. *The Modern Law Review*, 68(2):175–194, 2005.
- Bell Hooks. *Yearning: Race, gender, and cultural politics*. 1992.
- Robert Huseby. Can luck egalitarianism justify the fact that some are worse off than others? *Journal of Applied Philosophy*, 33(3):259–269, 2016.
- Matthew Joseph, Michael Kearns, Jamie Morgenstern, Seth Neel, and Aaron Roth. Rawlsian fairness for machine learning. *arXiv preprint arXiv:1610.09559*, 2016.
- Faisal Kamiran, Toon Calders, and Mykola Pechenizkiy. Techniques for discrimination-free predictive models. In *Discrimination and Privacy in the Information Society*, pages 223–239. Springer, 2013.
- Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. Inherent trade-offs in the fair determination of risk scores. *arXiv preprint arXiv:1609.05807*, 2016.

- Matt J Kusner, Joshua R Loftus, Chris Russell, and Ricardo Silva. Counterfactual fairness. *arXiv preprint arXiv:1703.06856*, 2017.
- Annabelle Lever. Racial profiling and the political philosophy of race. *The Oxford Handbook of Philosophy and Race*, page 425, 2016.
- Kasper Lippert-Rasmussen. *Born free and equal?: a philosophical inquiry into the nature of discrimination*. Oxford University Press, 2014.
- Derek Parfit. Equality and priority. *Ratio*, 10(3): 202–221, 1997.
- Philip Pettit. Responsibility incorporated. *Ethics*, 117(2):171–201, 2007.
- Edmund S Phelps. The statistical theory of racism and sexism. *The american economic review*, 62(4):659–661, 1972.
- John Rawls. *A theory of justice*. Harvard university press, 2009.
- Thomas M Scanlon. *Moral dimensions*. Harvard University Press, 2009.
- Frederick F Schauer. *Profiles, probabilities, and stereotypes*. Harvard University Press, 2009.
- Shlomi Segall. What’s so bad about discrimination? *Utilitas*, 24(1):82–100, 2012.
- Amartya Sen. *Inequality reexamined*. Clarendon Press, 1992.
- Charles Taylor. *Multiculturalism*. Princeton University Press, 1994.
- Larry S Temkin. Inequality. *Philosophy & Public Affairs*, pages 99–121, 1986.
- Jens Damgaard Thaysen and Andreas Albertsen. When bad things happen to good people: luck egalitarianism and costly rescues. *Politics, Philosophy & Economics*, 16(1):93–112, 2017.
- Michael Veale and Reuben Binns. Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2):2053951717743530, 2017.
- Michael Walzer. *Spheres of justice: A defense of pluralism and equality*. Basic books, 2008.
- Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, and Krishna P Gummadi. Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *Proceedings of the 26th International Conference on World Wide Web*, pages 1171–1180. International World Wide Web Conferences Steering Committee, 2017.
- Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. Fa\*ir: A fair top-k ranking algorithm. *arXiv preprint arXiv:1706.06368*, 2017.