

Reading list

Michele Loi, University of Zurich.

Essential:

- Angwin, Julia, and Jeff Larson. 2016a. "Machine Bias." Text/html. ProPublica. May 23, 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- . 2016b. "Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say." ProPublica. December 30, 2016. <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say>.
- Chouldechova, Alexandra. 2016. "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments." *ArXiv:1610.07524 [Cs, Stat]*, October. <http://arxiv.org/abs/1610.07524>.
- Berk, Richard, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. 2017. "Fairness in Criminal Justice Risk Assessments: The State of the Art." *ArXiv:1703.09207 [Stat]*, March. <http://arxiv.org/abs/1703.09207>.
- Heidari, H. et al. (2018) A Moral Framework for Understanding of Fair ML through Economic Models of Equality of Opportunity, *ACM Conference on Fairness, Accountability, and Transparency (ACM FAT*) 2019*. Available from <https://arxiv.org/abs/1809.03400v2>

Additional:

- Hardt, Moritz, Eric Price, and Nathan Srebro. 2016. "Equality of Opportunity in Supervised Learning." *ArXiv:1610.02413 [Cs]*, October. <http://arxiv.org/abs/1610.02413>.
- Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan. 2016. "Inherent Trade-Offs in the Fair Determination of Risk Scores." *ArXiv:1609.05807 [Cs, Stat]*, September. <http://arxiv.org/abs/1609.05807>.
- Chouldechova, Alexandra, and Aaron Roth. 2018. "The Frontiers of Fairness in Machine Learning." *ArXiv:1810.08810 [Cs, Stat]*, October. <http://arxiv.org/abs/1810.08810>.