Unfair Artificial Intelligence: How FTC Intervention Can Overcome the Limitations of Discrimination Law
Andrew D. Selbst & Solon Barocas

Throughout 2021, the activities of the Federal Trade Commission have suggested that it is poised to take action to address issues of fairness and equity in the commercial use of artificial intelligence (AI). In April 2021, the Commission published a blog post suggesting that it was examining how to use its existing authorities to rein in inequitable AI.[1] In August 2021, Commissioner Slaughter published a white paper focused on economic justice and AI.[2] And if the Washington DC adage that "personnel is policy" is to be believed, the recent nomination of Alvaro Bedoya as well as the hiring of well-known justice-focused scholars and activists as advisors[3] certainly suggests that the FTC is ready to move in this area.

In this paper, we argue that this is a positive and long overdue development. Specifically, the FTC's expansive "unfair and deceptive acts and practices" (UDAP) authority, its investigatory powers, and its inherent advantages as a litigant as compared to individual consumers all mean that FTC intervention could fill notable gaps in the ability of traditional discrimination law to address fairness and equity concerns in the context of AI.

We identify five reasons that FTC action can improve upon the status quo.[4] First, discrimination law as it exists is sector-specific, applying to employment, credit, housing, education, and a few other areas. Some applications of AI can fall outside these contexts and yet still cause significant and disparate harm to consumers. For example, consumer electronics with voice recognition may exhibit systematically worse performance for certain communities, denying these consumers the benefits they've paid for, forcing them to deal with the costs of failure, and likely harming their dignity in the process.[5] No current discrimination law addresses this issue, but the FTC's UDAP authority reaches it. Second, discrimination law is targeted at the choices of decisionmakers, based on an era before decisionmakers regularly outsourced their decisions to the technology they've purchased. Upstream vendors and intermediaries can sell technology to decisionmakers that either contributes to or wholly creates discriminatory results, while

---

[1] Elisa Jillson, *Aiming for Truth, Fairness, and Equity in your Company's Use of AI* (April 19, 2021) https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai

[2] Rebecca Kelly Slaughter, Algorithms and Economic Justice: A Taxonomy of Harms and a Path Forward for the Federal Trade Commission (2021)

[3] *See Fordham Law Professor Olivier Sylvain Named Senior Advisor to Chair of the Federal Trade Commission Lina Khan*, Fordham Law News (October 7, 2021), https://news.law.fordham.edu/blog/2021/10/07/fordham-law-professor-olivier-sylvain-named-senior-advisor-to-chair-of-the-federal-trade-commission-lina-kahn/; Mitchell Clark, *Googler who helped lead 2018 walkout is joining the FTC*, The Verge (Nov 3, 2021), https://www.theverge.com/2021/11/2/22759776/google-meredith-whittaker-ftc-nomination-ai-ethics-regulation.

[4] For a related discussion on AI and the FTC's UDAP authority, see Michael Spiro, *The FTC and AI Governance: A Regulatory Proposal*, 10 Seattle J. Tech. Env. & Innovation L. 26 (2020).

[5] Allison Koenecke et al., *Racial Disparities in Automated Speech Recognition*, 117 Proc. Nat'l Academies Sci. 7684 (2020)

remaining outside the purview of existing discrimination law. The FTC's UDAP authority would allow the agency to address actors other than the ultimate decisionmakers. Third, the FTC possesses advantages as a litigant. The FTC can bring claims based on likely injury to many consumers, rather than having to prove actual injuries in the past. The Commission is not subject to the same standing requirements as individuals. It also has more resources for investigations and will not be bound by contractual concerns, such as arbitration clauses. Fourth, the idea of "unfairness," while not unlimited, is quite expansive and flexible.[6] The UDAP authority can reach concerns that are seen as unfair by many but evade courts' cramped and calcified definitions of "discrimination." Fifth, the FTC's interpretations need not be static. As the Commission and the public learn more about these technologies and the ways in which they discriminate, the UDAP authority is flexible and the FTC nimble enough to try and keep up with a developing understanding of the causes and consequences of unfair AI.

Separate from whether the FTC *should* regulate unfair AI is the question about whether the FTC's UDAP authority permits it. We argue that it does. The FTC Act defines unfair acts as those that meet a three-part test: 1) the act likely results in a substantial injury to consumers; 2) the injury is not reasonably avoidable by consumers themselves; and 3) the costs of the injury are not outweighed by countervailing benefits to consumers.[7] Generally, to be unfair, something need not be otherwise illegal, but the Commission may take inspiration from established public policies.[8] As to the first prong, for AI that discriminates in traditional areas of concern such as employment or credit, a showing of substantial injury will be trivial. More difficult will be milder cases of consumer harm based on disparate ability to enjoy the benefits of one's purchases, but if large swaths of the population are suffering from the same small injury, the FTC can declare that a substantial harm.[9] The second prong asks whether consumers can avoid the injury, and here the case is even easier, as most AI developers and vendors do not even test their systems for bias, let alone disclose their findings, so consumers lack the necessary information to comparison shop. Consumers are even more constrained when the AI system in question has been purchased by companies with whom consumers may have to interact, such as an employer or lender. Finally, the third prong asks the FTC to apply a cost-benefit analysis. Cost benefit analyses are notoriously fraught in the best of cases, but they are simply incoherent when considering concerns with discrimination and distributive justice. Whose costs and whose benefits should be counted and how should they be weighed? The statute does not specify how the agency is expected to go about this analysis. And prior enforcement actions offer little clarity as to how the agency has gone about it in the past. Yet it cannot be that companies are allowed to subject consumers to any conceivable harm, so long as they offer consumers a benefit of greater value. And the benefits that AI technologies offer to the majority of consumers cannot excuse subjecting a minority of consumers to discrimination. Thus, the

---

[6] *See generally* Luke Herrine, *The Folklore of Unfairness*, 96 N.Y.U. L. REV. 431 (2021).
[7] 15 U.S.C. 45(n)
[8] *Id.*
[9] *See, e.g.,* Dennis D. Hirsch, *From Individual Control to Social Protection: New Paradigms for Privacy Law In The Age of Predictive Analytics*, 79 MD. L. REV. 439, 483 (2020) ("[T]he FTC may add together a number of discrete injuries to particular individuals and assess whether the aggregate injury is substantial.")

FTC will have no difficulty finding that their UDAP authority permits an intervention on the issue of unfairly discriminatory AI.

The final question is exactly how the FTC should proceed. After all, the definitions of unfairness and discrimination are essentially contested, and the FTC will likely be no more able to offer a satisfying definition than the many scholars and practitioners who have grappled with the definitional question over the years. Here we look to the FTC's approach to data security for inspiration. This approach is essentially two-pronged, combining guidance for practices to engage in and to avoid, while litigating the worst practices to build up a pseudo-common law.[10] While current academic debates have led to a definitional morass, this approach will allow the Commission to litigate claims that are less controversial, setting a baseline of recognized harm, while building up an understanding over time of what constitutes unfair AI. Though controversial, this approach has proven useful in the data security context and has been upheld by courts that have reviewed it. It provides a useful path forward here as well.

---

[10] *Cf.* Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 Colum. L. Rev. 583 (2014).