Anti-Discrimination Law's Cybernetic Black Hole

Marc Canellas

The incorporation of machines into American systems (e.g., crime, housing, family regulation, welfare) represents the peak of evolution, the perfect design, for our devotion to a colorblind society at the expense of a substantively equitable society. Machines increase the speed, scale, and efficiency of operations, while their complex inner-workings and our complex interactions with them effectively shield our consciousness and our laws from the reality that their failures disproportionately affect protected groups.

When investigating alleged discrimination, anti-discrimination law – especially Title VII's protections against employment discrimination – is premised on two flawed assumptions: First, that discrimination always has a single, identifiable, and fixable source; second, that discrimination is exclusively the result of a human or a machine – a human villain with discriminatory animus or a faulty machine with algorithmic bias. These assumptions within anti-discrimination law are completely incompatible with the reality of how humans and machines work together, referred to as "cybernetic systems" in the engineering community. Cybernetic systems are characterized by the properties of interdependence and complexity, meaning that they have numerous, dynamic, uncertain factors that are hard to predict or identify as to how they contribute to performance or failure. The failure of the law and its commentators to understand and address fundamental conflicts between how discrimination is assumed to occur and how discrimination actually occurs within cybernetic systems means that there is no liability for cybernetic discrimination produced from the interaction of humans and machines within organizations.

This cybernetic black hole within anti-discrimination law has set up a system of perverse effects and incentives. When humans and machines make decisions together, it is almost impossible for plaintiffs to identify a single, identifiable source of discrimination in either the human or machine. As machines increasingly mediate human decisions in criminal, housing, family regulation, and welfare systems, plaintiffs will increasingly lose what little protection they had from intentional discrimination (disparate treatment) or discriminatory effects (disparate impact). Decisionmakers who want to reduce liability have no need to reduce discrimination, instead they are incentivized increase adoption of complex, opaque machines to recommend and inform their human decisions – making the burden on plaintiffs as heavy as possible.

There have been endless proposed solutions to the problems of discrimination law and Title VII, but after review the truth emerges that no tweaks to anti-discrimination law or pinch of technological magic will solve the problem of cybernetic discrimination. The only meaningful solution is one that our modern jurisprudence seems to fear most: a strict liability standard where outcomes for protected classes are explicitly used to identify and remedy discrimination.