1 DEAD KITTENS

1.1 The fairness of how we manage online systems

In the past several years, perceptions about the internet at large (and the web more specifically) being a public space for all have been challenged by the realities of the need to monitor and moderate content. That pressure has manifested itself in the form of (1) refusing to offer protective service to white supremacists (cloudflare) (2) removing similarly aligned people from dating websites (okcupid), and [al2: okay so some of these examples aren't algorithmic...] (3) demonetizing videos online which discuss topics such as sexual self–identification (LGBTQ videos on YouTube). In some cases, these decisions reflect the conclusions of wandering an algorithmic forest, made up of intermediate decisions and inferences that cannot be retraced; in other cases, [al2: something about ad hoc decisions made unilaterally].

In any event, the notion that these venues are public spaces in the way that we typically think of them has proven untenable.

[al2: Okay I don't know where to take this. I might have gotten this topic out of my system.]

1.2 AI system decisions in publics

Artificially intelligent systems are increasingly making decisions about our lives. Whether they're determining what content we see online [al2: cite Facebook, YouTube, Netflix, ???], or setting how much bail we have to pay to walk free in the time between arraignment and trial [al2: cite Sharad Goel and many other], decisions of wildly variable importance are being made around and about us. [al2: Future risks? Gaydar paper, self–driving cars, etc...?]

A number of researchers have interrogated the introduction of these systems in their corresponding contexts. Lee et al. [2] and Irani and Silberman [1] ask about the decisions that Uber and other on–demand work platforms make on behalf of workers. [al2: Bots make decisions about whether contributions made to Wikipedia are being made in good faith, and ...Stuart Geiger(?) has asked about the fairness of these systems] [al2: Airbnb might be facilitating racism/redlining, and researchers have asked about the decisions that are being made on those sorts of platforms] [al2: Goel and others have sought ways to make algorithmic bail decisions easier to interrogate and manipulate in the interest of justice.] [al2: I want more variety of examples here!]

These efforts to make sense of the algorithmic decisions being made through these systems have struggled to situate AI systems in a broader sociological context. To a large extent, we discuss AI systems as akin to forces of nature, beyond the control of those who designed them. [al2: I really feel like I've seen this in a citable format somewhere — I just need to find it.] This fundamentally removes agency not only from the organizations and people that were involved in their conception, but also from the society which largely informed the values and worldviews which influenced their design. A more robust framework for thinking about AI systems — especially those that make socially, economically, and politically consequential decisions — can give us more [al2: something something something]

OKAY HERE'S THE PAPER UNCHANGED

Researchers have documented some of the ways that algorithmic management systems fall short, and specifically the ways in which people work around these systems to achieve desirable ends. Whether the scholarship discusses Uber drivers toggling availability [2], Nurses finding ways to work around the constraints of electronic medical records (EMRs) [], or developing "folk theories" about how systems work, and working on those theories to achieve desirable status or outcomes []; we're seeing as a community a growing tendency to work around the algorithms that act upon all of us.

These questions drift along the undercurrent of deeper questions: what roles will we have as algorithmically trained systems increasingly manage us? What roles are these systems supplanting, or are they inventing new ones? When and how will algorithmic managers take into consideration the needs of the people they manage, if at all?

All of these questions place us in murky waters, where we lack not only the vocabulary and the structure, but indeed important landmarks, to navigate effectively. In an effort to situate ourselves more effectively to make sense of these phenomena and the questions that they bring, we suggest a framing of algorithmic systems at instantiations of bureaucracies to talk and reason about not only the roles of these systems in situ, but also the roles that these systems play in our workplace and indeed in our society. We will demonstrate that this framing provides us with a reasonable grounding from which to make meaningful progress toward better understanding the roles that algorithms play in our lives; more compellingly, we'll show that this framing unlocks insights to questions that designers of algorithmic management systems must face in a world increasingly run by machines.

We'll show first that bureaucratic institutions suffer from many of the same inherent biases from which algorithms suffer; next, we'll discuss some of the parallels in social power dynamics among those that motivate both systems, those that execute both systems, and those upon whom the system is executed. Then, crucially, we'll explore some of the ways that bureaucracies *work*; namely, the ways that bureaucrats exercise discretion in their interaction with the end–users of bureaucracies. Finally, we'll examine the characteristics that make this operating style — that of undocumented and sometimes conflicted discretion — possible in the first place.

With these thoughts, we'll return to the quandary of algorithmically managed systems and the people upon whom these systems execute the agendas of their designers. We'll show how these parallel origins and characteristics have led to experiences involving bureaucracies strikingly reminiscent of the experiences which people have reported interacting with algorithmic systems. Then, importantly, we'll bring some of the realities of bureaucracies — the discretion, the bending of rules, and the individual empathy of human bureaucrats — to contrast with the realities of AI systems. In the process, we'll attempt to highlight important but insufficiently explored areas of research which we argue must become cornerstones of scholarship on AI in society; for example, the design of discretion on the part of algorithmic systems to deviate in tasks in order to support broader organizational goals.

algorithmic decisions take place under conditions of limited time and information. algorithms typically are constrained by the costs of obtaining information relative to their resources, by their capacity to absorb information, and by the unavailability of information. however, algorithms work with a relatively high degree of uncertainty because of the complexity of subject matter (people) and the frequency or rapidity with which decisions have to be made.

REFERENCES

[1] Lilly C. Irani and M. Six Silberman. 2013. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, 611–620. https://doi.org/10.1145/2470654.2470742
[2] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with Machines: The Impact of Algorithmic and Data–Driven Management on Human Workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, 1603–1612. https://doi.org/10.1145/2702123.2702548