Preprint for XRDS • WINTER 2016 • VOL.23 • NO.2

Algorithmic Bosses, Robotic Colleagues: Toward human-centered algorithmic workplaces

We already know algorithms can make our lives and our work more efficient, but how can we go beyond that to create trustworthy, fair, and enjoyable workplaces in which workers can find meaning and continuously learn?

By Min Kyung Lee (mklee@cs.cmu.edu)

Algorithms pervade every aspect of our daily lives, more so today than ever. Algorithms decide what we see online, from Google search results and Amazon ads to Facebook News Feed and Netflix recommendations. But they're not just online. Algorithms are increasingly taking on the roles of bosses, managers, and coworkers. Companies use algorithms to find candidates for jobs, evaluate the performance of customer service agents, and connect patients with physicians. Governments use algorithms to determine when to patrol certain areas and screen immigration applicants.

Algorithms have actually helped give rise to an entirely new type of workplace and workforce. In emerging digital work platforms, algorithms flexibly and efficiently match customers with service providers; you've experienced this type of service if you've ever used Uber or TaskRabbit. Robots are slowly but steadily entering the workplace, and working right alongside people. They deliver goods in hotels, care patients and older adults with support from nurses or caregivers, and assemble machinery alongside employees in small businesses. Because of these recent changes, along with an industry push for efficiency, the social impacts of algorithmic technologies are no longer negligible.

My goal is to design such algorithms to better support human values, motivations, and unique capabilities. We already know algorithms can make our lives and our work more efficient, but how can we go beyond that to create trustworthy, fair, and enjoyable workplaces in which workers can find meaning and continuously learn?

This is an opportune and exciting time to address the aforementioned question. Historically we have seen many first-generation technological design choices leave long legacies, even if those choices were less than optimal. The design principles we use to create "smart on-demand transportation," "social delivery robots," or "robotic lawyers and doctors," will likely influence their successors. It is therefore important that academic and industry researchers examine the social, organizational, and ethical implications of these new technologies and incorporate this knowledge into their designs.

To that end, I am studying how people work with algorithmic technologies, in both the real world and the lab, in order to create design principles and systems that will enable more productive, fair, and enjoyable work. In this article, I'll reflect on some of the important issues and questions that have come up in my investigations of algorithmic systems in two different professional roles— algorithmic "bosses" that manage and govern workplace decisions, and robotic "coworkers" that share space with people in a workplace.

PROMISES AND PERILS OF ALGORITHMS

What's so promising about algorithms is they can be used to help people process massive amounts of data and gain insights beyond what the human eye and mind can come up with alone. For example, without utilizing algorithms, a human resources worker or an immigration office employee may only be able to look at and analyze around a hundred cases a day. Each application may not receive enough attention; heuristics and subconscious biases will inevitably play a greater role in the selection process, because humans cannot process massive data in a limited time. But by taking advantage of the pattern matching and processing power of algorithms, people can analyze from a thousand to a million cases a day. As organizations and cities become more technologically advanced, new data and new infrastructure can shine new light on human behaviors. That information can influence how organizations and cities make their managerial and governance decisions, such as allocating resources, hiring new employees, assembling or incentivizing teams of individuals, evaluating performances, and more.

However, we can’t treat data and algorithms as a silver bullet. What makes cities and organizations thrive are the people who comprise them. People working with algorithms would need to trust and cooperate with these managerial and governance decisions for them to be effective. My research suggests simply applying algorithms to a situation won’t automatically result in decisions that elicit cooperation, inspire trust, or feel motivating and fair. I have studied how Uber and Lyft drivers [1] and users of Splididt [2] (a website that applies fair division algorithms to social division tasks) and Snackbot [3, 4] (a social snack delivery robot in a workplace) interact with algorithmic systems. My findings suggest both the logic and working mechanisms of the algorithms themselves, and the interfaces and interaction that surround them, should accommodate diverse human motivations, behaviors, and contexts. This requires new research into human and social aspects of these intelligent systems, because these are not typical parts of computational and mathematical investigations of algorithms. Drawing on my experiences, I below share challenges in and future research questions for creating human-centered algorithmic workplaces.

FAIR, DIVERSE, AND MOTIVATING ALGORITHMIC MANAGEMENT

Algorithms can assign tasks for workers or allocate resources on a larger scale than a human could manage. For example, with Uber and Lyft, the assignment algorithm automates the driver-rider match, allowing just a few human managers to oversee hundreds and thousands of drivers in each city. Many traditional workplaces such as hospitals, retail stores, and construction companies are also using algorithms to schedule and allocate tasks and budgets. How can we make these decisions fair?

One way is to draw on a line of research in mathematics and economics that investigates fair division problems. Basing the decisions on algorithms and actual data could remove potential biases and ad-hoc decisions, leading to fairer outcomes. However, my study of Spliddit users suggests otherwise, showing that the concept of fairness is socially rather than mathematically constructed [2]. I asked groups of participants to divide household chores or other tasks among themselves. Half of the groups made their decisions through group discussion. In the other half of the groups, each person entered their preferences for each task into the Spliddit website. Mathematically-proven fair division algorithms then used these preferences to assign tasks to each member in a way that maximally satisfied everyone's preferences. After each member completed all the chores or tasks they'd been assigned, we asked them how fair they thought the division decisions had been. When we compared their fairness ratings across these two decision-making methods, the algorithms were thought to be less than fair one third of the time. Interviews helped us understand why. Many participants had multiple concepts of fairness that go beyond what the algorithm assumed—maximizing individual benefit without sacrificing others' benefit. Some groups would have preferred equal distribution in order to minimize social comparison and jealousy; others wanted to sacrifice their own benefit because they thought others' preferences might be stronger. These social and cognitive factors make perceptions of fairness different from economic fairness. Discussion allowed people participate in the process of making the decision, share what fairness meant to them, and adjust their outcomes for the team; algorithmic mediation did not allow such conversations.

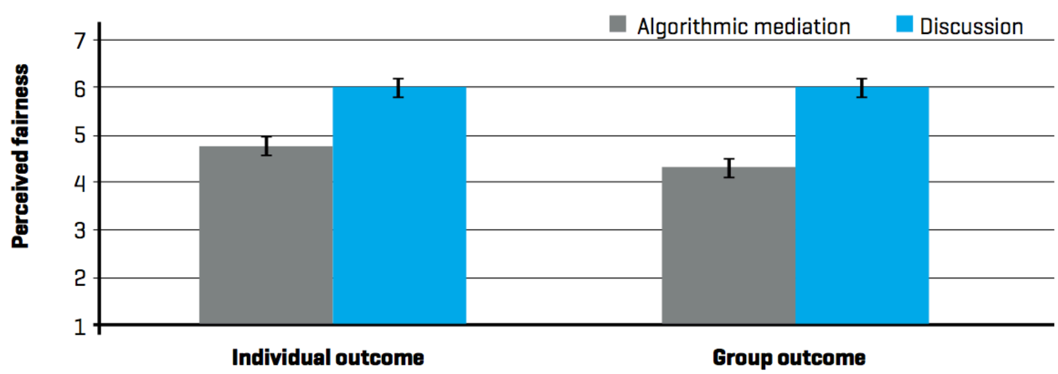

Figure 1. Decisions made by fair division algorithms were perceived as less fair than those made through discussion. How can we create algorithmically-managed workplaces that are fair in not only the mathematical but also the social sense?

The above result has implications for workplace "optimization." Much ongoing work in academia and industry seeks to determine optimal team compositions, communication patterns, traffic flows, and so on. It would be easy to rely on economic and efficiency-centered metrics for optimization, and assume rationality in order to model and predict human behaviors. But there is a danger that we might unintentionally reinforce or impose economic values at the expense of other important values, such as social and altruistic behaviors, which are difficult to define in traditional mathematical and economic terms.

The behaviors of Uber and Lyft drivers around surge pricing suggest that algorithms based on such narrow assumptions won't always motivate people in the real world, particularly when workers have a choice not to cooperate with the algorithm and change their routines instead [1]. Surge pricing dynamically raises fares in areas with predicted high demand, in an effort to balance supply and demand by motivating drivers to come to high-demand areas. Yet half of the drivers who I talked to said they chose not to go to surgepriced areas. Some drove just for fun, or for altruistic reasons, and thought surge pricing was unfair to customers. Others said it changed too fast or in unpredictable ways, so they were unable to meaningfully use the information and incorporate it into their routines.

These studies highlight the challenges in building efficient and optimized yet fair and motivating algorithmic workplaces, and demonstrate the risk of imposing the limited, economic values that algorithms often embody onto social tasks. Further research should investigate how algorithms can accommodate more diverse types of motivations and values, and the emotions that people feel about the decisions that algorithms make.

TRUSTWORTHY ALGORITHMIC WORKPLACES

It's important that workers can trust algorithmically-made managerial decisions; these decisions directly influence the work they do, their satisfaction with their work, and potentially their income and job security. In such cases, transparency (or the lack of it) may play a key role [1]. For example, neither Uber nor Lyft provide drivers or passengers with explanations about how each assignment was made. From a management perspective, this creates a safe backdrop for the company to continuously experiment with the factors that assignment algorithms take into consideration, all behind the scenes. On the other hand, the opaqueness seems to influence workers' attitudes toward the company. For example, one person told me, "Uber is very close-lipped about what actually happens. I mean they say, ‘Oh, we route it to the closest driver,’ or whatever, but who really knows what’s going on behind the scenes. It’s up to whoever engineers their iPhone app.”

The lack of transparency didn’t only influence workers’ attitudes. It also influenced their behaviors around algorithmic decisions. When assignments were undesirable or seemed to make no sense, drivers simply attributed them to errors and rejected them. For example, if a driver receives a request for a 15-minute ride but sees that they’re not actually the closest driver, they may attribute the assignment to an error when the assignment actually could have been made for a legitimate reason. This case exemplifies the importance of algorithmic transparency. Explaining the reasoning behind assignments is important in eliciting cooperation, especially when the assignments are unpopular or exceptional cases.

However, full transparency might not work to the company’s benefit, as workers may use their knowledge to game the system, maximizing individual benefit at the expense of group optimization. For example, many drivers’ least favorite ride is a distant ride request, one that requires driving for more than 15 minutes. Two of our drivers learned from an online forum that the longer a Lyft driver stays online, the wider his or her pick-up radius becomes. They used this knowledge to avoid distant requests by periodically turning driver mode off and on again at traffic signals. This finding raises the question: How do we promote transparency to earn workers’ trust but also prevent workers from gaming the system? With Uber and Lyft, drivers have limited power to refuse incoming requests, and there are financial motivations to accept rides—the more they accept, the more they generally earn. In other contexts with different power structures and incentives, finding the right level of transparency would be even more critical.

AN ALGORITHMIC WORKPLACE THAT PROMOTES LEARNING AND GROWTH

Another important element of a successful workplace is whether it promotes workers’ continuous learning and helps people find meaning in their work. These factors were highlighted by my investigation of two other aspects of Uber and Lyft: algorithmic evaluation and the degree of automation in the ride assignments.

Uber and Lyft evaluate performance using customers' ratings and drivers' rates of accepting algorithmic ride assignments. This automated, data-driven evaluation allows the company to see which drivers are succeeding and which are falling short, and to control service quality. But algorithmic evaluation is not always fair or useful for drivers. Our study found consistent and arbitrary biases in customer's ratings. Drivers felt many uncontrollable factors influence passengers' ratings beyond just their driving and service skills. Passengers sometimes blame drivers for things like arriving late to a destination or having to accept surge pricing. In addition, the metric treats all assignment rejections as service failures, when in reality, not all rejections are the same. For example, some female drivers in our study would not accept male passengers without profile pictures at night due to safety concerns.

It is unclear whether the metrics actually motivated drivers to improve their work in the long run. Once their scores were above a certain threshold of deactivation risk, drivers seemed to develop a detached, indifferent attitude about their ratings. For example, another participant stated: "I used to micromanage my rating, so to speak. I used to sweat and be, 'Oh my gosh my rating is now going down— it's a 4.85,' that kind of thing. Now I don't worry about it. I see there's a lot of error that can take place in the rating." In addition, these quantified metrics did not provide very useful feedback for drivers as to how to improve their performance. Future research should ask: How can we account for nuances and valid exceptions? How can we design useful, quantifiable metrics and rating systems that are fairer to workers?

In applying algorithms, it might be tempting to automate many decisions in the workflow, because automation is easier than thinking about how to model human behaviors in the loop. But ignoring these factors can adversely impact workers' autonomy and learning. For example, Uber and Lyft drivers do not have much control over the types of rides they are assigned, other than to refuse the assignment. But many drivers created workarounds by strategically controlling when and where they turned on driver mode on the app in order to get the types of requests and clientele they wanted.

They turned off driver mode in bad neighborhoods to avoid dangerous situations, or went downtown for successive short rides during the lunch hour.

In some cases, though, these workarounds still don't allow drivers enough control over their work. For example, a former taxi driver now working for Uber mentioned the lack of choice in assignments made it hard for him to create a work strategy. He did not like the Uber assignment system because algorithms made the decisions that he used to make himself, making him feel like he'd lost the agency to enact strategies he'd developed to maximize his income. This could be interpreted as resistance to change, but also raises open-ended ethics questions about the trend in new technology to sacrifice individual control for the sake of overall system efficiency.

ROBOTIC COWORKERS THAT HELP PEOPLE WORK TOGETHER

Another important aspect of successful workplaces is organizational culture. Algorithmic technologies in the form of robots or virtual agents are increasingly taking on the role of coworker, either working directly with people or sharing their physical workspace. The question then becomes: How can we make these new coworkers enjoyable to have around us? Can we leverage them to create better social and organizational culture? For example, how can they help workers socialize better and increase communication within and across teams?

To explore this question, we built a delivery service robot from scratch at Carnegie Mellon and deployed the robot in an office building for a few months in 2012 [3, 4]. Office employees ordered apples, chocolate chip cookies, and other snacks online, and the robot delivered them to their offices. The robot called workers, initiated brief small chats, and then asked them to pick up their snacks.

The results of the field experiment suggest workers built rapport with the robot over time. What was more surprising was the ripple effect that it had on the social dynamics in the workplace. The robot became a common boundary object that participants could easily relate to, creating a topic of conversation and an occasion to socialize, in the way that dogs do in a public park. For example, one person said: "It's usually [...] quiet in my hall.

You know, even if people are in, they might close their door or something. But I think people [were] more likely to be around and laughing and feeling sociable when the robot was there." This suggests a robot in a workplace can have a positive impact on organizational culture.

Figure 2. Snackbot delivered snacks in an office building, built relationshipswith workers, and promoted workers to socialize with each other. How can we create robotic coworkers that are enjoyable to have around and that help us work together better?

However, I am not arguing all robots should be social and chatty. The fact that the robot was social and personalized for different individuals also had unintended consequences. Workers started to compare each other. For example, one person noted she felt jealous when the robot complimented another girl. Interestingly, other people thought the robot was nicer to this participant or even flirting with her. This social comparison may have encouraged people to use the robot more, be nicer to it, and build stronger rapport with it. However, it also created social tensions. Further research would need to investigate how social and how culturally aware a robot should be for different contexts and purposes.

LOOKING FORWARD

Instead of asking, "What will the future of work look like?" I believe we should ask, "What *should* the future of work look like?" Instead of letting ad-hoc, efficiency-centered decisions drive the future, we need to make informed decisions with careful consideration of their organizational and societal impact. I am truly excited about the future these algorithmic technologies can enable. I believe they can usher in more efficient and fair management practices based on the best of both data and human judgment. They may create workplaces where power structures are more equally balanced between workers and managers thanks to the transparency of their decisions. They may even enable workers to make many of these managerial decisions themselves. What the future of work will look like is up to all of us who are living in this critical transition time.

References

[1] Lee, M.K., Kusbit, D., Metsky, E., and Dabbish, L. 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).* ACM, 2015, 1603-1612.

[2] Lee, M.K. and Baykal, S. Algorithmic mediation in group decisions: Fairness perceptions of algorithmically mediated vs discussion based social division. Publication pending for *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW)*. ACM, 2017.

[3] Lee, M.K., Forlizzi, J., Kiesler, S., Rybski, P., Antanitis, J., and Savetsila, S. Personalization in HRI: A longitudinal field experiment. In *Proceedings of 7th ACM/Institute of Electrical and Electronic Engineers (IEEE) International Conference on Human-Robot Interaction (HRI).* IEEE, 2012, 319-326.

[4] Lee, M.K., Kiesler, S., Forlizzi, J., and Rybski, P.Ripple effects of an embedded social agent: a field study of a social robot in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI).* ACM, 2012 695-

704.

Biography

Min Kyung Lee is a research scientist in human-computer interaction at the Center for Machine Learning and Health and the Machine Learning Department at Carnegie Mellon University. Her research examines the social and decision-making implications of intelligent systems and supports the development of more human-centered machine learning applications. Dr. Lee is a Siebel Scholar and has received several best paper awards, as well as an Allen Newell Award for Research Excellence. Her work has been featured in media outlets such as the New York Times, New Scientist, and CBS. She received a Ph.D. in HCI and an M.Des. in interaction design from Carnegie Mellon, and a B.S. summa cum laude in industrial design from KAIST.