Joy Ming (jming@infosci.cornell.edu)
Collective Data, Collective Power with Low-Wage, Frontline Workers

AI and data-driven tools are transforming work, especially for low-wage frontline workers. These workers have used data to resist these changes and collectively improve their working conditions— connecting with other workers, exchanging information, and gathering evidence of employer violations [1, 2, 14]. However, this data-driven advocacy raises concerns around data privacy, technological burden, and limited impact [9, 11, 12]. **My research examines how AI and data-driven technologies can build collective power for low-wage frontline workers while mitigating technological harms.** Investigating these tradeoffs contributes to an understanding of collective data and collective power that is relevant to other community-engaged and participatory data projects—exploring how data can mobilize marginalized workers, equitably represent diverse stakeholder voices, and transform workforce dynamics and policy.

My dissertation investigates the tradeoffs between the potential benefits and harms of data-driven advocacy for home care workers in Upstate New York in partnership with Healthcare Workers Rising and 1199SEIU, the largest union local in the United States. The 2.9 million home care workers in the United States provide crucial and increasingly in-demand services to older adults and people with disabilities, empowering these clients to live and thrive within their own communities [8]. Despite the essential nature of their work, home care workers are often overlooked by the public and in policy [3], facing financial precarity in addition to physically and emotionally demanding working conditions [13]. Moreover, they face increasing datafication and platformization with federally-mandated visit tracking technology [4]. However, prior research shows that technology could help mediate connections and build solidarity between these spatially isolated workers [10].

My research delves into the harms and opportunities of technology in relation to unrecognized and unregulated labor in order to design and deploy technology to acknowledge the contributions of home care workers and protect their labor. First, I delved deeper into how home care workers perceived technology through interviews with thirteen home care workers [6]. I found that employer-mandated technologies enabled additional work, largely because the design did not account for the emotional and relational facets of care work. However, the workers indicated the potential for technology to improve the working conditions if it maintains worker autonomy. To explore this potential, I created design provocations of worker-centered, data-driven tools and discussed them with focus groups with ten workers and five worker advocates [5]. I found these tools could aggregate instances of employer violations and facilitate communication across workers to exchange experiences and build solidarity. Finally, I explored concrete approaches to the challenges raised by designing and piloting an intervention to track working hours and conditions [7]. I found the combination of quantitative and qualitative data collected faithfully represented the workers' experiences and, if scaled up with appropriate organizational resources, had potential to organize workers, hold employers accountable, and promote caring policies.

My research highlights considerations around the nature of home care and informs future design of data-driven technology. My research found empirical evidence of how technology complicates and is complicated by the nuances of the home healthcare work context, including the physical infrastructure of poor internet connectivity where workers changed phones frequently and the occupational infrastructure where last-minute changes led to technological errors in work and wage tracking. These findings inform designs of technology in this context to account for specific cases of invisible work and worker attitudes towards data and technology. Moreover, it theoretically contributes to designing for sociotechnical mechanisms of invisibility, stakeholders with heterogeneous goals, and systemic change.

My dissertation research identified key questions of collective data in worker organizing that I am further exploring. First, there is the question of data management, specifically how sociotechnical systems can support community-owned knowledge and data, exploring participatory data governance models like

data stewards or data cooperatives by drawing on research related to collective knowledge (e.g., community networks) and statistical approaches to privacy (e.g., differential privacy). Then, the second question is how that data and other AI tools can be used to mobilize communities and facilitate solidarity across stakeholders and movements, exploring computer-supported collaboration across physical and circumstantial distances between care workers and care recipients, workers and employers, and the Global North and Global South. The final question is how sociotechnical infrastructures could be built in community-based organizations and governments to support AI and data-driven policy change, such as open data initiatives and building technical capabilities in human services organizations, taking into consideration topics related to AI biases and harms in public interest. I am exploring research directions across geographies beyond New York State and domains beyond home care with researchers and community partners in security and privacy, social work, and global development.

The Computational and Citizen Science Research Workshop is a great opportunity to share my findings on the power of collective data in home healthcare and low-wage workers. My research especially addresses questions relating to "Citizen Science Trust, Equity, Ethics, and Responsible AI" by examining the ethical and social tradeoffs between the potential public good that AI and data-driven tools could bring for workers as well as the potential harms of accuracy, privacy, and bias. This would be applicable to not only future efforts towards public data for public good, but also to developing socially responsible systems. As my research focuses specifically on low-wage, frontline workers, I also contribute to "Broadening Participation at the Intersection of Computing and Citizen Science Research." In the home healthcare work space specifically, this includes workers who are primarily women, people of color, and immigrants who give care to clients who are aging or have disabilities. My research presents perspectives of these marginalized communities around data privacy, data work, and data impact and informs future engagements. Moreover, I look forward to discussing how human-computer collaborations foster and inhibit social change, empowers community-based organizations to effectively leverage new technologies, and improves the lives of essential but precarious workers.

[1] Calacci, D. and Pentland, A. 2022. Bargaining with the Black-Box: Designing and Deploying Worker-Centric Tools to Audit Algorithmic Management. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2 (Nov. 2022), 1–24.
[2] Irani, L.C. and Silberman, M.S. 2013. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2013), 611–620.
[3] Kennedy 2015. Wage Theft as Public Larceny. *Brooklyn law review*. (2015).
[4] Mateescu, A. 2021. *Electronic Visit Verification: The Weight of Surveillance and the Fracturing of Care*. Data & Society.
[5] Ming, J., Gong, D., Ngai, C.S.E., Sterling, M., Vashistha, A. and Dell, N. 2024. Wage theft and technology in the home care context. *Proceedings of the ACM on human-computer interaction*. 8, CSCW1 (Apr. 2024), 1–30.
[6] Ming, J., Kuo, E.F., Go, K., Tseng, E., Kallas, J., Vashistha, A., Sterling, M. and Dell, N. 2023. "I Go Beyond and Beyond": Examining the Invisible Work of Home Health Aides. *Proceedings of the ACM on Human-Computer Interaction*. 7, CSCW1 (2023), Article 59.
[7] Ming, J., Tolera, H., Tu, J., Yitzhaki, E., Ngai, C.S.E., Sterling, M., Avgar, A.C., Vashistha, A. and Dell, N. 2024. Exploring Data-Driven Advocacy in Home Health Care Work.
[8] PHI 2024. *Direct Care Workers in the United States: Key Facts 2024*.
[9] Pierre, J., Crooks, R., Currie, M., Paris, B. and Pasquetto, I. 2021. Getting Ourselves Together: Data-centered participatory design research & epistemic burden. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021).
[10] Poon, A., Hussain, V., Loughman, J., Avgar, A.C., Sterling, M. and Dell, N. 2021. Computer-Mediated Peer Support Needs of Home Care Workers: Emotional Labor & the Politics of Professionalism. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (Oct. 2021), 1–32.
[11] Silberman, M.S. and Irani, L. 2016. Operating an Employer Reputation System: Lessons from Turkopticon, 2008-2015. *Comparative Labor Law & Policy Journal*.
[12] Spektor, F., Khovanskaya, V., Forlizzi, J. and Fox, S. 2024. Data Wagers in Worker Advocacy Research.
[13] Stacey, C.L. 2011. The Caring Self. *The Caring Self*. Cornell University Press.
[14] Toxtli, C., Suri, S. and Savage, S. 2021. Quantifying the Invisible Labor in Crowd Work. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (Oct. 2021), 1–26.