

AI Failure Loops in Feminized Labor: Understanding the Interplay of Workplace AI and Occupational Devaluation

Anna Kawakami, Jordan Taylor, Sarah Fox, Haiyi Zhu, Kenneth Holstein

Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA

Extended Abstract

Workplace AI systems often fail in practice. For example, in social services, AI-based decision support tools have been introduced across high-stakes settings, only to be dropped following backlash from workers or the public (Samant et al. 2021). Similarly, in healthcare, researchers have spent decades innovating on AI-based tools to support clinical decision-making, only to find that clinicians ignore them in practice (Yang, Steinfeld, and Zimmerman 2019).

Past research has described a range of challenges that help explain these AI failures. For example, human-computer interaction (HCI) and science and technology studies (STS) literature has identified problems of poor contextual fit, where an AI system’s design clashes with existing worker practices (e.g., (Suchman 1987; Forsythe 1993; Kawakami et al. 2022)). To address these challenges, prior work has proposed a range of new resources to support more responsible or participatory design. Yet, these resources are rarely used in practice, and AI teams continue to design flawed AI systems for the workplace (Delgado et al. 2023).

Existing efforts to improve the design and evaluation of workplace AI tools has overlooked a critical factor: the role of *occupational devaluation*. How might the devaluation of worker expertise interplay with AI design, evaluation, and deployment practices? In our paper, we examine this through the case of *feminized labor*, a particularly extreme form of occupational devaluation (Balka and Wagner 2021). In the U.S., feminized labor often involves care-oriented work, like teaching and nursing. Historically misnomered as “women’s work,” feminized labor is still predominantly performed by women and people of color. While impacts on and of feminized labor have been studied by other disciplines, it remains under-examined in responsible AI research.

Drawing together past scholarship on AI deployments in feminized occupations, we argue that workers in societally devalued occupations are particularly vulnerable to flawed AI deployments, rooted in impoverished understandings of workers’ tasks and expertise. Moreover, we argue that these AI deployments further obscure the visibility of workers’ expertise, triggering a negative feedback loop that further entrenches workers’ devaluation.

To understand this dynamic, we formalize the concept of *AI Failure Loops*: a set of interwoven, socio-technical failures across the development lifecycle of workplace AI systems that can amplify the harms from existing occupational devaluation. *AI Failure Loops* arise at the confluence of *over-claims* about the capabilities of workplace AI systems and *under-recognition* of the complexity of the work that workers perform. We ground an understanding of the driving factors, properties, and impacts of *AI Failure Loops* through three case studies of AI deployments in feminized labor contexts: AI-based risk assessment tools in social work, AI for home healthcare, and AI tutoring systems in K-12 teaching. We conclude with a discussion on future work towards pro-worker AI practice, policy, and organizing.

References

- Balka, E.; and Wagner, I. 2021. A historical view of studies of women’s work. *Computer Supported Cooperative Work (CSCW)*, 30: 251–305.
- Delgado, F.; Yang, S.; Madaio, M.; and Yang, Q. 2023. The participatory turn in ai design: Theoretical foundations and the current state of practice. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–23.
- Forsythe, D. E. 1993. The construction of work in artificial intelligence. *Science, technology, & human values*, 18(4): 460–479.
- Kawakami, A.; Sivaraman, V.; Cheng, H.-F.; Stapleton, L.; Cheng, Y.; Qing, D.; Perer, A.; Wu, Z. S.; Zhu, H.; and Holstein, K. 2022. Improving Human-AI Partnerships in Child Welfare: Understanding Worker Practices, Challenges, and Desires for Algorithmic Decision Support. In *CHI Conference on Human Factors in Computing Systems*, 1–18.
- Samant, A.; Horowitz, A.; Xu, K.; and Beiers, S. 2021. Family surveillance by algorithm: The rapidly spreading tools few have heard of. American Civil Liberties Union (ACLU).
- Suchman, L. A. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
- Yang, Q.; Steinfeld, A.; and Zimmerman, J. 2019. Unremarkable AI: Fitting intelligent decision support into critical, clinical decision-making processes. *Conference on Human Factors in Computing Systems - Proceedings*.