

Work, Meaning, and AI Displacement

PHIL 164: Technology and Human Values

UC San Diego, Summer 2024

Anushka Sheoran

Revisions and Ongoing Questions

The essay below was written in summer 2024. A few things I would say differently now, or am sitting with as open questions:

- The topic of AI safety has become interesting to me, in a way that feels distinct from the jobs question. The breadth of deployment is one reason (modes of unsafe interactions tremendously increased), but another one is how capable the models are getting and what that implies for alignment research. The work on maintaining meaningful human control over increasingly capable systems suggests this is a genuinely hard problem (not one that resolves itself as a byproduct of good engineering), and there is a real window for getting it right that is not obviously indefinitely open. This thinking triggered from reading work from orgs like METR, Redwood Research and AI 2027.
- My view on job displacement has gotten more convoluted and unsure. My current working view is that in the short term the people most at risk are a more specific group than "white-collar workers" broadly: those who cannot upskill within the timeframe of displacement, still need to work (no financial runway to wait it out), and are not in roles with enough strategic leverage to be retained for judgment rather than execution. Middle management, loosely defined, fits this profile more precisely than the aggregate projections suggest, and I think that specificity matters for what interventions would actually help.
- On new job creation: the essay treats the Luddite counterargument as fairly settled, but I am less sure now. Whether new jobs are being created in the aggregate is probably yes. I've had multiple people (specifically professors) clearly assert that jobs are fine and more jobs will be created. Again, haven't dug deep into it again but need to have a working draft opinion. Whether those jobs are accessible to the workers most displaced is a different question, and one I do not think the literature is answering as carefully as it should.
- Since writing this, I reviewed additional papers on AI's labor market impact as part of a graduate agentic AI course, including a preprint by Shao et al. on auditing automation desire across the U.S. workforce (<https://arxiv.org/abs/2506.06576>) [1], which complicates the picture in useful ways. Workers, it turns out, express positive automation desire for a substantial share of tasks, particularly repetitive or low-value ones, while being protective of tasks they find meaningful or identity-adjacent. That sits in some tension with the uniform displacement framing I use here.
- On the force multiplier argument: I still think the basic dynamic is right, but I would add more nuance given my experience in the past few months. The availability of tools like Claude Code and other agentic systems has democratized coding capability in ways that were not obvious even a year ago, and there is a real window right now where a small team or even an individual with the right skills can build things that previously required an engineering organization. Whether that window stays open long enough to allow any meaningful redistribution of capability, or whether the structural advantages of large companies (data, compute, distribution, existing customer relationships) end up consolidating the gains anyway, is something I do not know. Also, my thinking may be limited by being in a "tech bubble" (which some say might burst) since I'm doing my Master in CS and given investors don't know what will be successful yet are pouring large amounts of money into AI ventures. My instinct is still that incumbents win in the long run, and I think the economics support that, but the current moment feels less predetermined than the force multiplier framing in the essay implies, and I wanted to flag that.

- Almost two years later, the discourse has mostly settled into the pattern you might expect: Most mainstream coverage makes this sound maximally alarming, full of large numbers and frightening implications. Most people deep in tech will tell you those numbers miss the point because people will adapt, find other things to do, and end up being capable of more. Neither camp is entirely wrong but both framings are too clean, and the actual uncertainty in the middle is where the more interesting questions sit.
- On the what work provides section, I would want to think about it from a Dharmic lens as well (where svadharm etc. fit in)
- The tone throughout is more assertive than the evidence probably warranted, and a few of the claims function more as assumptions than as conclusions I had actually worked through. I have left the piece as originally written.

1. Framing

The debate about AI and jobs tends to start with scale: how many jobs, how fast, which sectors. The scale projections are striking enough on their own terms that they probably warrant more serious, proactive attention than they are currently getting, because the disruption being projected, if it lands the way some analysts expect, would affect a large number of people's livelihoods in ways that could produce real social and economic chaos. What I kept finding, though, is that the conversation tends to happen in pockets. Technologists think through the capability questions without much fluency in policy or social dynamics, and policymakers try to respond to something they can only partially see. Another core question I'm thinking about is: if employment, as we have organized it, gets restructured faster than most people can follow, what fills the space it leaves?

Some background worth establishing: AI refers broadly to the capacity of computers to learn from large datasets rather than being explicitly programmed for every step, and what this means for labor is a shift in which kinds of work are actually at risk. The old framework was routine versus non-routine: physical, repetitive tasks could be automated, while complex and judgment-heavy ones could not. AI challenges this at the root, because the more useful question is how clearly defined the steps of a task are and whether a system can learn them from examples rather than prescription. By that framing, a lot of technical work turns out to be automatable, and with generative AI, even roles requiring higher-level cognition, drafting, advising, synthesizing, have become genuinely vulnerable. Goldman Sachs economists estimated that as many as 300 million full-time jobs could be lost or diminished globally by the rise of generative AI [Goldman Sachs, 2], and the workers most at risk are white-collar workers: analysts, writers, paralegals, people who did not expect automation to be their problem.

2. Platform Dynamics and Inequality

Running alongside the job displacement story is a structural design: Data has become one of the most valuable assets in the modern economy, and the companies controlling most of it are pulling away from everyone else in ways that are not obviously reversible. Platform companies offer services to billions of users in exchange for control of the data those users generate, and that data funds better models, which attract more users, which generate more data: a self-reinforcing cycle of collection, storage, analysis, modeling, application, and market domination. Malhotra describes this as a business model that is inherently monopolistic, not accidentally but structurally [Malhotra, 3], and the combination of deep pockets, network effects, and privileged market insight makes it very difficult for new entrants to compete in any meaningful sense. The gap between capitalism's theoretical commitments to individual agency and what platform dynamics actually produce in practice is at least worth sitting with.

The inequality dimension compounds this: The Credit Suisse Global Wealth Report found that the top one percent of people globally already own nearly half the world's wealth [Credit Suisse, 4], and AI, in Malhotra's framing, functions as a force multiplier: it makes the strong stronger and the weak weaker [Malhotra, 3]. Those with technical qualifications and current knowledge are likely to be rewarded with high-paying, stable roles, while everyone else faces something harder. Malhotra describes the potential emergence of what he calls an unemployable class, not people temporarily between jobs but people for whom no viable path back into productive employment exists, and the middle class may not survive this transition intact, though how permanent any of this turns out to be is genuinely uncertain.

3. On the Historical Comparison

When I raise the displacement concern in conversations, someone usually brings up the Luddites (or something similar) as an example. The Luddites were 19th-century English textile workers who smashed industrial machinery because it threatened their livelihoods, and the argument is that they were wrong: technology changes the nature of jobs rather than eliminating them, agricultural workers found manufacturing, factory workers found services, and this time will be no different.

I find this less convincing than it sounds, because the pace of this particular transition feels categorically different from prior ones. The Bain and Company Labor 2030 report projected that worker displacement from AI will occur two to three times faster than it did during the agricultural and industrial transitions [Bain, 5], and earlier disruptions unfolded across generations rather than years. Farmers had time, their children had time to retrain and reorient, and the institutional infrastructure of education and labor markets had time to adapt. That is not the situation we appear to be in, and retraining programs require sustained investment that companies have little structural incentive to provide when displacement moves faster than any training program can track. The breadth of impact of this 'AI revolution' is something to think about as well.

4. What Work Actually Provides

Here is another different but equally relevant lens of thinking about this topic:

Work, in the way we have organized society, is not just an income source... it is an identity source, the structure around which adult life is arranged and evaluated. Hannah Arendt wrote that we live in a "society of laborers" that is about to be liberated from labor but no longer knows of the higher, more meaningful activities that this freedom was supposed to be won for [Arendt, 6]. We spend our youth preparing to work, our adulthood doing it, and our old age recovering from it, and it is hard to even imagine what the organizing principle of life would be if that structure dissolved.

The views on what this actually means for human flourishing are not settled. Alfred Marshall believed that human beings degenerate without hard work, that difficulty and effort are not just economically productive but conditions for human flourishing [Marshall, 7], and David Autor has argued along similar lines that idleness is genuinely terrible and that work gives life structure and meaning [Wellisz, 8]. I find both views sympathetic, and also a little uncomfortable, because they imply that the displacement of labor is not just an economic problem but something closer to an existential one.

While the above makes logical 'common sense' to me, current data adds interesting nuance to this reading. Around 70 percent of workers are either disengaged or actively disengaged from their jobs, and nearly 40 percent of people in the UK say they do not believe their work makes a meaningful contribution to the world [Susskind, 9]. If this is what work is actually providing most people, the loss of it turns out to be messier than the headlines suggest: both less catastrophic and differently catastrophic at the same time. The question of whether there exists some third category of activity that could replace paid employment as a source of purpose is one that philosophy has not really resolved. My honest intuition is that people need something to do, and that without it, at scale, things tend to get difficult in ways that go beyond economic hardship and start looking like personal and social breakdown. But I hold that loosely.

5. What Work Actually Provides

Now thinking about the government infrastructure around this:

Governments and policymakers have historically been able to stay agnostic about what "good work" means, with the market defining it and people choosing within it. But if AI displaces enough work fast enough, that agnosticism becomes untenable, because policymakers will have to take explicit positions on what a meaningful life looks like, what kinds of activity deserve support, and what structures replace employment as the organizing principle of adult life. Universal Basic Income and Job Guarantee Schemes are both on the table as partial answers, neither is obviously right, and both require a theory of what we are actually trying to produce in a society where traditional work is scarce.

What I found strange over the course of thinking through this material was how little that conversation actually crosses disciplines. The discourse in technical spaces tends to treat the meaning and social organization questions as downstream problems, things that will get figured out eventually by someone else, and that framing is itself a choice with consequences. The literature on the normative dimensions of work is surprisingly sparse, and scholars in the field note it themselves: relatively little has been written on what work ought to be and what its absence would mean for how we organize society [Susskind, 9].

The technological development is not going to stop. The question is whether the people with the most influence over how it unfolds are willing to engage seriously with the parts of this in way that can mitigate the potential negative effects. Staying neutral, treating AI disruption as a technical problem with technical solutions, is itself a position.

References

- [1] Shao, Y., Zope, H., Jiang, Y., Pei, J., Nguyen, D., Brynjolfsson, E., Yang, D. "Future of Work with AI Agents: Auditing Automation and Augmentation Potential across the U.S. Workforce." arXiv:2506.06576 (2025). <https://arxiv.org/abs/2506.06576>
- [2] Goldman Sachs Global Economics Research. "The Potentially Large Effects of Artificial Intelligence on Economic Growth." (2023).
- [3] Malhotra, R. Artificial Intelligence and the Future of Power: 5 Battlegrounds. Rupa Publications, 2021.
- [4] Credit Suisse Research Institute. Global Wealth Report 2019.
- [5] Harris, K., Kimson, A., Schwedel, A. "Labor 2030: The Collision of Demographics, Automation and Inequality." Bain and Company, 2019.
- [6] Arendt, H. The Human Condition. University of Chicago Press, 1958.
- [7] Marshall, A. Principles of Economics, 8th ed. Macmillan, 1920.
- [8] Wellisz, C. "Late Bloomer." IMF Finance and Development 54(4), 2017.
- [9] Susskind, D. A World Without Work: Technology, Automation, and How to Respond. Allen Lane, 2020.
- [10] Susskind, D. "Work and Meaning in the Age of AI." Brookings Center on Regulation and Markets Working Paper, January 2023. <https://www.brookings.edu/center/center-on-regulation-and-markets/>