JUST TECH

Co-opting or Centering the Human Worker in the Digital Age? An Overview of Human-in-the-Loop Legislation in the United States

By: Mishal Khan

LABOR & ECONOMY

Khan, Mishal. "Co-opting or Centering the Human Worker in the Digital Age? An Overview of Human-in-the-Loop Legislation in the United States." Just Tech. Social Science Research Council. August 5, 2026. DOI: https://doi.org/10.35650/JT.3102.d.2026.

Since the release of ChatGPT in 2023, predictions of widespread job loss have littered headlines. Silicon Valley's tech titans claim they will reshape the global economy with the creation of artificial general intelligence (AGI)—a fantastical technology that will lead to limitless cost savings by eliminating the need for human labor altogether. Economists, too, insist that the new wave of digital technologies differ radically from technologies of the industrial age.[1] Today's technological advancements, we are told, do more than automate tasks requiring physical human labor; they have the capacity to imitate the most fundamental human capabilities: language, reasoning, and decision-making.

Yet, there is still no clear evidence that AI is *replacing* jobs en masse.[2] What we do know is that data-driven technologies—from the sophisticated to the mundane—are increasingly being deployed *alongside* workers in almost every sector of the economy. Nurses are interacting with clinical decision support systems, teachers are using edtech in classrooms, and a whole variety of workers—physicians, public sector workers, and human resources managers—are using algorithmic recommendations to make decisions about everything from healthcare coverage and public benefits to employment.

But what do worker-machine interactions look like on the ground? What should the rules of the road be? How are these technologies impacting both the quality of decision-making and the workers making decisions? Over the last few years, a core regulatory concept has emerged to address these concerns: maintaining a "human-in-the-loop." Keeping a human-in-the-loop is meant to ensure public trust,

safeguard against machine-generated errors, and preserve the “human element,” as digital technologies proliferate throughout our economy.

Drawing on trends emerging from tracking legislation at the intersection of technology and work in the United States over the last few years,[3] I lay out the wide range of human-in-the-loop provisions emerging across legislatures today. Overall, there has been an explosion of legislation in the broader tech and labor arena since 2025.[4] Amid the deluge, sector-specific bills containing human-in-the-loop provisions have emerged as a significant new trend, and this has continued into 2026. While human-in-the-loop provisions are a hard-won victory against the pressures of total automation, the question for the moment is how to ensure that these provisions do not co-opt, but rather actively center, the human worker.

The Stakes for Workers

For unions and labor advocates, ensuring that human-in-the-loop provisions are enshrined in law is pivotal—a vital protection for workers in the face of potential job loss. Even though it is still early days, some of these bills are already making it through legislatures and have been signed into law. For example, this year Washington passed a law (HB 5395) that will require a physician to review any algorithmic decision that would deny a healthcare benefit. Similarly, Rhode Island passed a law (S 2197) specifying that AI cannot be used to replace therapeutic functions but can be used to support the delivery of mental healthcare. Many of these bills are not being directly put forth by labor but are often supported by broad coalitions, including professional associations, civil society organizations, and consumer groups. The enactment of these laws is remarkable in a regulatory environment characterized by heated opposition and intense lobbying against AI regulation of any kind.[5]

However, researchers have been skeptical of the efficacy of keeping a human-in-the-loop as a protective measure, referring to the concept as a “procedural theater”[6] or a “comforting distraction.”[7] Is it simply another version of what researchers Sarah Fox and Samantha Shorey call “augmentation-washing”—that is, a euphemistic slogan meant to obscure the extent to which technologies actively degrade rather than enhance working conditions?[8] Researchers and advocates have long documented the potential and already existing harms that can arise even in the absence of full-scale job automation. Algorithmic systems can supplant a worker’s judgement, leading to deskilling, the erosion of their decision-making authority and scope of practice, and eventually lower pay.

There are other questions, too. Who will be liable if these systems cause harm? For example, if an algorithm incorrectly recommends a course of action with life-threatening impacts on a patient’s health, will the worker be held accountable? Researcher Madeleine Elish calls this dilemma the “moral crumple zone,” describing how workers end up absorbing responsibility for system failures.[9] Is the worker a credentialed specialist and do they have the requisite training needed to understand how the system functions and how to interpret its results? What pressures are on workers to accept the algorithmic recommendation, and do they have the right to override or reject it?

These are complex and still unresolved questions. The guardrails that policymakers establish over the next few years will be critical for workers, and they are currently being negotiated in the legislative arena.

Configuring the Role of the Human: The Legislative Response

Human-in-the-loop provisions have surged in state legislatures across the country—and some at the federal level too. These bills are emerging on a sector-by-sector basis, mostly in healthcare, education, the public sector, journalism, critical infrastructure, and criminal justice.

There is significant variation, however, in how the human-in-the-loop concept is articulated. Does having a human-in-the-loop mean that a human—i.e., a worker—is available by default or only on request? Does it mean that a human reviewer will be part of the appeal process or that they are making the decision from the get-go? To complicate matters further, human-in-the-loop provisions are often buried in confusing regulatory language. For example, some bills state that businesses are "prohibited from solely relying" on AI to make decisions. Others state that a human must be the "final" decisionmaker, while still others that businesses must "preserve the use of in-person assistance." There are bills requiring "oversight" or "significant oversight," "review," "meaningful review," or "affirmative review," "involvement" or "intervention." Beneath this mélange of regulatory language, lawmakers and advocates are attempting to establish the contours of the human-machine relationship in law.

To make sense of this considerable—perhaps overwhelming—diversity, it is useful to divide these bills into several categories.

1. Simple Human Review

In the first camp are *simple human review* bills. For example, in 2025 Virginia passed a law (HB 1642) stating that critical criminal justice decisions cannot be made by AI without some kind of "human involvement." The same year, Kentucky passed a public sector law (SB 4) that would make human review of an algorithmic decision not necessary, but rather dependent on the level of risk or the use-case in question. These laws are prime examples of the leanest articulation of the human-in-the-loop concept. Weaker versions of these bills require companies to simply report if and to what extent a human reviewer is part of a decision-making process, rather than making it an active requirement. Another iteration of these bills state that any kind of human review, of a healthcare document for example, eliminates a business's requirement to disclose that AI was used. In practice, even a cursory review can become an escape clause.

2. **Worker-in-Command**

A second set of bills articulate the concept of *worker-in-command*. These bills require workers not only to review outputs but rather be in control of the technologies they are working with. Many of these bills have language specifying that workers have the right to "reject, modify, and override" the machine. These bills have often been supported by unions and professional associations insisting that workers on

the frontlines have instincts that go beyond the data inputs and predictive capabilities of the tech. For example, the 2025 Federal Right to Override Act would establish that healthcare workers have the right to override a clinical decision-making support system output if, in their professional judgement, doing so is in the best interest of their patient. The original version of a 2025 law in New York (SB 8451), similarly, would have given media workers the right to approve, deny, or modify any AI-generated content used in the newsroom.

Additional concepts have emerged during the 2026 legislative session. A California bill (AB 2575) would protect healthcare workers from being held liable when an AI system causes harm. Another California bill (SB 1011), covering critical infrastructure, would give workers "unimpeded authority" to override an AI model. An earlier version of the bill went further, containing language prohibiting companies from designing systems or implementing performance metrics that may pressure workers into accepting the model's recommendations.

3. Enhancing the Weight of Worker Evaluation

A third category of bills enhance the *weight of worker evaluation* in making critical decisions. Many of these bills also contain worker-in-command provisions. For example, a 2026 bill in California (SB 947) states that employers can only use algorithmic recommendations to assist in making employment-related decisions about workers, such as hiring, firing, or promotion. But even more importantly, that the human reviewer must independently corroborate and compile their own evidence about a worker, shifting the emphasis away from the algorithmic recommendation toward a human evaluation of the worker's broader employment record. Similarly, in the education context, a 2026 bill in Washington (SB 5956) prohibits schools from making disciplinary decisions about students based on algorithmically generated predictions or scores, without independent human investigation and consideration of context.

4. Protecting Essential Functions, Tasks, and Decisions

The final category of bills explicitly lays out the specific *functions*, *tasks,* and *decisions* that employers are prohibited from replacing with digital technologies. For example, a 2026 bill in Illinois (SB 3702), covering nursing, names the discrete tasks that cannot be performed by AI, including nursing assessments, nursing care plans, or triage, admission, discharge, or transfer determinations. Similarly, a 2025 law in Texas (SB 815) outright prohibits the use of algorithms in denying healthcare benefits. As digital technologies spread, these bills in effect carve out domains that are so fundamental to public safety or the public good that they cannot be replaced by AI. By doing so, they protect not only workers' jobs, but the right to perform the most essential elements of their work.



Moving Forward

While tech giants are promulgating grand visions of a world free of human workers, unions and labor advocates, together with civil society and the public, are pushing back, insisting on a different future.[10] Battles are being fought in legislatures across the country to outright prohibit employers from replacing certain jobs—teachers, court reporters, therapists, college faculty, physicians, nurses, and public sector workers—with AI. As a regulatory concept, human-in-the-loop provisions too, especially the stronger forms detailed above, can potentially offer important protections for workers. Yet, whether the language and details they contain are sufficient to manage the full range of harms that can occur as digital technologies become further integrated into workplaces across the economy, remains to be seen.

Nonetheless, the volume and variety of these bills highlight two critical conclusions: (1) There is a growing consensus that human workers remain indispensable to the future envisioned and demanded by the public at large and (2) legislating in this area is as challenging as it is urgent. The potential to hollow out the concept is ever-present. Businesses can win over the public by claiming that a human is somewhere in the decision-making process while pushing for workers to simply monitor the flow of automated processes, not even "in" but "on" the loop.[11] The rapid rollout of AI agents is bringing up additional challenges and questions.[12]

Moving forward, perhaps the strongest protections will come out of detailed regulations tailored to the specific needs, concerns, and features of individual sectors of the economy. Collective bargaining agreements offer crucial templates as unions are carving out strong human review provisions for their members, taking into account the messy reality of human-machine interactions in their workplaces.[13] In fact, we have seen human-in-the loop bills become increasingly detailed and sophisticated over time. As policymakers and advocates continue to grapple with the growing complexities of the digital workplace, it is more important than ever to ensure that legislative interventions are robust and meaningful rather

than performative, with only a shallow nod to the workers whose labor and skills power our economy.

Footnotes

1 "We Must Act Now: A Statement on AI's Transformation of the Economy," Stanford Digital Economy Lab, https://www.wemustactnow.ai/.

2 TechEquity, *White-Collar Bloodbath or AI-Washing? Understanding the AI & Labor Displacement Research Landscape* (TechEquity, 2026), https://techequity.us/wp-content/uploads/2026/05/AI-Labor-Displacement-Research-Landscape-FINAL.pdf.

3 Mishal Khan and Annette Bernhardt, *The Current Landscape of Tech and Work Policy in the US: A Guide to Key Laws, Bills, and Concepts* (UC Berkeley Labor Center, 2025), https://laborcenter.berkeley.edu/tech-and-work-policy-guide/.

4 Mishal Khan and Annette Bernhardt, "Evaluating Trends and Challenges in State Regulation of Workplace Technologies," *Tech Policy Press*, November 3, 2025, https://www.techpolicy.press/evaluating-trends-and-challenges-in-state-regulation-of-workplace-technologies/.

5 Jeremia Kimelman, "Tech Giants are Spending more than Ever to Shape California Politics. See How Much," CalMatters, March 27, 2026, https://calmatters.org/politics/2026/03/meta-google-ai-regulation-elections/.

6 Nicolas Spatola, "AI Efficiency Can Undermine Accountability Even with Humans in the Loop," *Tech Policy Press*, May 5, 2026, https://www.techpolicy.press/ai-efficiency-can-undermine-accountability-even-with-humans-in-the-loop/.

7 Uri Maoz, "Why Having 'Humans in the Loop' in an AI War is an Illusion," *MIT Technology Review*, April 16, 2026, https://www.technologyreview.com/2026/04/16/1136029/humans-in-the-loop-ai-war-illusion/.

8 Sarah E. Fox and Samantha Shorey, "How Augmentation-Washing Hides Labor Automation," *Tech Policy Press*, December 22, 2025, https://www.techpolicy.press/how-augmentationwashing-hides-labor-automation/.

9 Madeleine Clare Elish, "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction," *Engaging Science, Technology, and Society* 5 (2019), https://estsjournal.org/index.php/ests/article/view/260.

10 Mishal Khan and Kung Feng, *A First Look at Labor's AI Values: An Analysis of Recent Statements about Technology by Unions and Other Worker Organizations* (UC Berkeley Labor Center, 2025), https://laborcenter.berkeley.edu/a-first-look-at-labors-ai-values/.

11 Hoag Levins, "In the Loop or On the Loop: The Conundrum of AI Clinical Decision Support," Penn Leonard Davis Institute for Health Economics, September 22, 2025, https://ldi.upenn.edu/our-work/research-updates/in-the-loop-or-on-the-loop-the-conundrum-of-ai-clinical-decision-support/.

12 Samir Passi and Ranjit Singh, *The Oversight Fallacy: Why AI Agents Require More than Humans-in-the-Loop* (Data & Society, 2026), https://doi.org/10.69985/VWCK1626.

13 Lisa Kresge, "Negotiating Tech: An Inventory of US Union Contract Provisions for the Digital Age," UC Berkeley Labor Center, last updated July 2, 2025, https://laborcenter.berkeley.edu/negotiating-tech/.