



Key Issues on AI in Workplace Safety and Health

Artificial Intelligence technology has the potential to make workplaces safer and healthier when workers are involved in its design, implementation and oversight and when workers can exercise their rights. Examples of beneficial safety technologies include devices that lift patients for nurses, robotic spray booths that shield workers from chemicals, and sensors on vehicles, equipment and in chemical tanks that can detect hazards.

However, technology can also be weaponized against workers. Guardrails are vital to protect workers from AI systems that: promote continuous surveillance of workers instead of safety and health hazards; automate safety procedures requiring human expertise and replace workers' right to use their professional judgment; undermine risk assessments or hazard identifications; enable unsafe scheduling, work pace and retaliation through algorithmic management; or shift the liability of maintaining a safe workplace and verifying the accuracy of AI systems from employers to workers.

- ★ **Algorithmic Management** is becoming more commonplace in workplaces, where workers are managed by AI products instead of human managers. These systems often set unrealistic targets, punish minor delays, and change schedules without human intervention—all resulting in unsafe working conditions. Workers may automatically have their schedules or hours reduced, leading to understaffing and unreasonable productivity quotas—both of which lead to hazards and injuries, such as workplace violence and musculoskeletal disorders. Other examples include AI setting delivery routes without accounting for hazards such as extreme temperatures, extreme weather, traffic or the lack of safe rest break areas. In healthcare settings, AI can assign schedules without accounting for realistic patient loads, staffing or the time it takes to safely use patient lifting devices to prevent injuries.

Human managers should make decisions about work, scheduling, and discipline, rather than algorithms that raise the risk of injury and illness to workers.

- ★ **Risk Assessment and Hazard Identification.** Worker experience and rights, and the expertise of other human professionals, are essential to identifying and fixing workplace hazards, requiring complex, in-person assessments, conversations, and process issues that cannot be fully evaluated by technology. A watered-down risk assessment leads to weak and oversimplified safety programs that lead to unaddressed hazards and risks, putting more workers in danger under false pretenses. Tech companies are developing applications that use AI to identify workplace hazards from a picture alone, missing the point; most hazards and their root causes and contributing factors cannot be identified with a simple snapshot.

Workers and human safety and health professionals need to be in charge of risk assessment and risk management processes and approvals, not machines.



★ **Safety Procedures cannot always be improved with technology.** Industries look for ways to “modernize” safety procedures using technology and AI, but technologies can introduce new hazards or exacerbate the problem, replacing essential and intentional safety procedures. For example, existing federal standards require a “lockout” mechanism that protects workers from machines accidentally turning on during servicing or maintenance. But in manufacturing settings, unions are seeing machines in their facilities that are operated remotely overseas or through machine learning technologies. Instead, onsite workers need to have the control to physically activate, shut down and lock out all power sources to protect them from accidents and injury. Once servicing or maintenance is complete, workers can unlock the machine.

Technology can improve worker safety, but safety procedures require onsite, human control.

★ **Monitoring Systems for constant worker surveillance** are increasingly used. Personal technologies that may use AI, such as wearable monitors and smart helmets, can help with hazard detection and worker safety, but they can also lead to constant, invasive and invisible surveillance of worker behaviors. This often results in employers blaming the worker rather than addressing hazards and fixing broken workplace processes and systems. For example, instead of correcting hazardous conditions and designing the workplace to be safer – such as reducing the need for bending or lifting or keeping a warehouse at a cooler temperature – workers are monitored for bending improperly (ergonomics) or sweating (heat). Keystroke monitoring and AI-enabled behavior detection create constant pressure, stress, fear of retaliation and loss of autonomy. These mental strains also turn into physical risks that force workers into hazardous postures, skipping breaks and rushing tasks. Accounting for human error is a critical part of designing safe work systems and surveillance and monitoring do not necessarily increase productivity, but can increase workplace hazards.

Workplace safety and health tools need to be focused on fixing underlying problems instead of blaming individual workers’ behaviors in an unsafe system.

★ **The Integrity, Accuracy and Safety of AI Systems** are the responsibility of employers and technology companies, yet workers are increasingly burdened with verifying their accuracy and efficacy and held liable for AI errors. Federal law requires employers to maintain a safe workplace, but workers are put in impossible situations when required to work with faulty AI systems. For example, automated monitoring systems have also failed to accurately assess and capture risks for patients, leaving health care workers held accountable for negative health outcomes. Construction workers have been required to use AI systems that recommend unsafe and illegal ladder placements, stopping workers from doing their job safely and being blamed for violating safety rules.

The use of safe AI systems, just as maintaining safe workplaces, must remain the responsibility and liability of employers and technology companies.

