When AI Writes The Safety Policy, Who Owns The Mistake?

A drug manufacturer recently gave safety and compliance leaders a preview of a problem that’s going to reach far beyond pharmaceuticals.

In April 2026, FDA issued a warning letter to Purolea Cosmetics Lab after inspectors found significant current good manufacturing practice violations. One part of the warning letter has drawn special attention. FDA said the company used AI to create drug product specifications, procedures, and master production or control records, but failed to review those AI-generated documents to ensure they were accurate and compliant. FDA also noted that the company’s overreliance on AI was documented during the inspection.

That’s the part every safety leader should read twice.

The issue wasn’t that the company used AI. The issue was that AI appeared to become part of the compliance system without the controls, review, expertise, and accountability that compliance work requires. In regulated environments, a polished document isn’t enough. Someone competent still has to know whether the document is right, whether it matches the law, whether it applies to the operation, and whether it can stand up to inspection.

Workplace safety has the same problem, only the consequences can be more immediate.

An employer can ask AI to write a fall protection policy, a forklift safety talk, a lockout procedure, a WHMIS or HazCom refresher, a confined space entry checklist, or a violence prevention plan. Within seconds, the organization has something that looks structured, professional, and usable. It may have headings. It may have steps. It may use the right vocabulary. It may even sound more polished than the internal document it’s replacing.

But looking like a safety procedure isn’t the same thing as being one.

A real safety procedure has to match the law, the task, the equipment, the work environment, the worker group, the hazards, the supervision model, and the employer’s actual controls. A real training program has to do more than transfer information. It has to prepare workers to recognize hazards, follow procedures, ask questions, stop unsafe work when needed, and prove they understood what they were taught. A real compliance system has to create records that show what was done, when it was done, who reviewed it, who received it, and how the employer verified that it worked.

AI can help with parts of that process. It can’t own it.

That distinction matters because workplace safety law doesn't judge employers by whether they generated a document quickly. It judges them by whether they took reasonable steps to protect workers. In Canada, CCOHS describes due diligence in occupational health and safety as the reasonable precautions employers take under the circumstances to prevent workplace harm. CCOHS also explains that due diligence requires employers to identify workplace hazards and take corrective action to prevent accidents or injuries. In the United States, OSHA maintains extensive training requirements across standards, and its training guidance makes clear that employers remain responsible for compliance with the applicable standards.

That's why AI misuse in safety isn't just a technology issue. It's a due diligence issue.

Imagine a maintenance supervisor at a mid-sized manufacturing company. The company has an older piece of equipment that jams regularly. Workers clear the jam by opening a guard, reaching into the machine, and pulling material free. Everyone knows it isn't ideal, but it's been done that way for years.

After a near miss, the safety manager asks AI to draft a lockout procedure. The prompt is simple: "Create a compliant lockout procedure for clearing jams on industrial machinery." The output looks good. It explains shutdown, isolation, lock application, verification, and restart. It includes a short quiz. It even reminds workers to notify affected employees.

The safety manager formats it, adds the company logo, and assigns it through the LMS.

Two months later, a worker is seriously injured. During the investigation, the company discovers the AI-generated procedure didn't identify all energy sources. It didn't account for stored pneumatic pressure. It didn't match the manufacturer's instructions. It didn't describe the specific control points on the machine. It didn't include a hands-on verification step. It wasn't reviewed by maintenance. Supervisors were never trained on how to enforce it. The workers completed the course, but nobody confirmed they could apply the procedure at the machine.

On paper, the company had training.

In reality, **the company had a document-shaped gap in its safety system.**

That's the danger.

AI-generated safety content often fails in ways that aren't obvious at first. The most dangerous failures aren't always wild hallucinations or laughably wrong statements. The more serious risk is content that's 80 percent right, written confidently, and missing the 20 percent that matters most.

A generic fall protection talk may describe harness use but miss the jurisdiction's trigger height, rescue planning duties, equipment inspection requirements, or site-specific anchor limitations. A forklift refresher may describe safe travel but fail to reflect the employer's actual pedestrian traffic patterns, loading dock hazards, battery charging area, attachments, ramps, or blind corners. A violence prevention policy may sound legally responsible but fail to match the province, state, sector, reporting process, investigation duty, or required worker consultation. A confined space checklist may name atmospheric testing but miss rescue capability, permit criteria, isolation controls, attendant duties, or the difference between a confined space and a permit-required confined space under the applicable regime.

That's how organizations get into trouble. They don't always use AI recklessly. They use it efficiently, then mistake efficiency for control.

The common errors are already easy to see.

The first error is asking AI broad compliance questions without defining the

jurisdiction. “Write a compliant safety policy” is not a serious compliance instruction. Compliance with what? Federal OSHA? Cal/OSHA? Ontario? Alberta? British Columbia? A federally regulated Canadian workplace? Construction? Manufacturing? Warehousing? Healthcare? Transportation? The answer changes.

The second error is failing to identify the task and hazard in enough detail. Safety training that doesn’t reflect the actual job becomes generic awareness content. Awareness has value, but it’s not the same as task qualification, procedure training, or competency verification.

The third error is treating AI output as if it were reviewed because it was edited for grammar. A manager may clean up the language, improve the formatting, and add a logo, but none of that answers the compliance question. Was the legal requirement checked? Was the procedure compared against the manufacturer’s manual? Did a competent person review the hazard controls? Did supervisors confirm the steps can be followed during real work?

The fourth error is relying on AI to summarize laws or standards without confirming the source. AI may combine old and new rules. It may mix U.S. and Canadian terminology. It may apply a federal rule where a state or provincial rule is stricter. It may invent a refresher interval because similar training topics have one. It may give a correct-sounding statement that collapses under inspection.

The fifth error is confusing training completion with worker competence. This may be the biggest risk for safety teams. A completed module proves that someone opened, progressed through, and perhaps passed a quiz. It doesn’t automatically prove the worker can inspect fall protection equipment, lock out a machine, recognize heat stress symptoms, enter a confined space safely, operate a lift truck, or respond to workplace violence.

The sixth error is letting AI-generated content enter the safety system without version control. If the company later changes the document, where’s the original? Who approved it? What changed? Why was it changed? Which workers received which version? Which version was in effect on the day of the incident?

The seventh error is vendor blindness. Many organizations will soon discover that they didn’t directly use AI, but their vendor did. Training libraries, policy generators, compliance tools, translation services, quiz builders, incident investigation aids, and LMS content features may all involve AI somewhere in the workflow. That doesn’t make them unsafe. It does mean the employer needs to ask better questions. Who reviews the content? How is jurisdictional accuracy verified? How often is it updated? Are legal references traceable? Can content be customized? Are records retained? Is there a documented quality process?

None of this means safety leaders should avoid AI altogether. That would be unrealistic and, in some cases, counterproductive. AI can help safety teams draft first versions, simplify complex language, convert procedures into toolbox talks, create scenario questions, build checklists, translate plain-language reminders, and identify gaps for human review. Used properly, it can reduce administrative burden and help safety professionals spend more time on higher-value work.

But AI has to stay in the right role.

It should support competent people, not replace them. It should accelerate drafting, not bypass review. It should help organize knowledge, not become the legal authority. It should make training easier to produce, not easier to fake.

The practical standard should be simple: if the document would matter after an injury, inspection, enforcement action, workers’ compensation claim, lawsuit, grievance, or fatality investigation, it shouldn’t be used without competent human review.

That includes policies, procedures, hazard assessments, inspection forms, training scripts, competency checklists, supervisor guides, incident investigation tools, orientation materials, LMS quizzes, compliance matrices, and worker instructions.

A defensible AI-assisted safety process should include several controls.

Start with an AI use inventory. Safety leaders should know where AI is being used in the safety system, including by vendors, consultants, trainers, supervisors, HR, operations, and communications teams. If AI is quietly generating safety talks, policies, or incident summaries, the organization needs to know.

Classify the risk of the output. A poster reminding workers to hydrate is one level of risk. A lockout procedure, confined space rescue plan, fall protection plan, respiratory protection procedure, or harassment investigation framework is another. High-risk content needs higher scrutiny.

Require jurisdictional identification. Every compliance-related output should be tied to the relevant jurisdiction, sector, and legal framework. If the answer applies differently in Canada and the United States, or across provinces and states, that difference should be visible.

Require competent review. The reviewer should understand the hazard, the applicable law, and the workplace operation. In some cases, that may be an OHS professional. In others, it may include maintenance, engineering, HR, legal, supervisors, the JHSC or safety committee, or an external specialist.

Document the sources checked. The organization doesn’t need a law review memo for every safety talk, but it does need a way to show that critical content was based on more than an AI answer. Legislation, regulator guidance, standards, manufacturer instructions, internal incident history, and workplace-specific procedures should be part of the review trail where appropriate.

Test the training against the work. Ask a simple question before rollout: could a worker use this to do the job safely in our workplace? If the answer is no, the training isn’t ready. For high-risk work, build in demonstrations, supervisor observations, field verification, or scenario-based checks.

Retain version history. If AI helped draft the content, that’s less important than whether the final approved version is controlled. Employers should know which version was assigned, who approved it, when it changed, and who completed it.

Audit the system. Once AI-assisted content is in circulation, it should be reviewed periodically. Laws change. Equipment changes. Work processes change. Hazards change. Training content that was acceptable last year may be weak today.

This is where safety training platforms and content systems have to evolve. The value of a training system isn’t just that it gives employers more content. More content can create more risk if it’s generic, outdated, unreviewed, or disconnected from the work. The real value is controlled content, consistent assignment, practical learning design, defensible records, and the ability to show what was trained, who completed it, when it happened, and how the organization followed up.

That’s the difference between paper safety and defensible safety.

Paper safety says, “We have a policy.”

Defensible safety says, “Here’s the approved policy. Here’s why it applies. Here’s who reviewed it. Here’s when workers were trained. Here’s how supervisors reinforced it. Here’s how we verified understanding. Here’s what changed after the incident, inspection, audit, or legal update.”

AI can help build parts of that system. It can’t replace the system.

The FDA warning letter should be treated as an early signal. Regulators are unlikely to accept “the AI told us” as an explanation for compliance failure. Inspectors in workplace safety won’t either. After a serious incident, nobody will care that the safety talk sounded polished. They’ll care whether it was correct, current, specific, reviewed, understood, and applied.

That’s the standard safety leaders should use now.

The organizations most at risk aren’t the ones experimenting with AI. They’re the ones using AI casually inside compliance workflows without governance. They’re the ones allowing unreviewed outputs to become official documents. They’re the ones confusing speed with diligence, language with accuracy, and completion records with competence.

AI will not eliminate the need for safety expertise. It will expose organizations that don’t have enough safety expertise reviewing what AI produces.

And when that happens, the ownership question won’t be complicated.

If AI writes the safety policy, the employer still owns the mistake.