

As AI agents take on more work, are humans really in control?

Today's top story, from writer Judith Aquino:

Have you heard the story about an AI agent deleting a critical production database? Jeremy Crane, Founder and CEO of car rental platform PocketOS sure has, and he lamented it in a widely read post on X.

It's a mistake that raises an increasingly urgent question as agents take on more responsibility: are workers meaningfully steering these tools—or just signing off on choices AI has already made?

Until now, digital systems and tools “were built around humans,” Crane told IBM Think. “Now, we have computers making decisions in split seconds, and our DevOps and our security protocols aren't prepared for that world.”

55% of organizations are actively developing or deploying an agentic AI operating model, according to a 2026 IBM Institute for Business Value (IBV) report, “The blueprint for agentic operations.” And 60% plan to adopt next-generation delivery structures where AI agents coordinate integrated workflows across finance, supply chain, HR, procurement, physical operations and customer service.

Meanwhile, as expectations increase for humans to apply critical thinking and judgment to AI-generated outputs, organizations are reaching a crossroads: whether to invest in systems and governance that enable real human judgment, or risk reducing the process into a rubber-stamping exercise.

Crane's deleted database is a casualty of that risk, says Phaedra Boinodiris, Global Leader for Trustworthy AI at IBM Consulting. “This is not a story about a rogue AI,” she told IBM Think over email. “This is a story about liability laundering at infrastructure scale.”

Liability laundering, Boinodiris explains in an article for IBM Think, is when a human is in charge of overseeing an AI's work, but “the conditions for genuine oversight—the architectural transparency, the training, the authority to halt, the feedback mechanism—were not established.”

As AI's capabilities and reach expand, having a human in the loop is more critical than ever. But the value that humans add when working with AI agents is meaningless without operational infrastructure that allows them to apply judgment, underpinned by governance, transparency, guardrails and escalation paths to supervise agentic systems at machine speed.

As Maryam Ashoori, VP of Product and Engineering for IBM watsonx.governance, said in an interview with IBM Think, “Organizations need clear structures that keep humans accountable for the decisions that matter most. The question isn't whether humans are in the loop; it's whether they're in control.”