82% of U.S. Companies Have Seen AI Agents ‘Go Rogue’ in the Last 12 Months, Gravitee Reveals

As Agentic AI implementation ramps up, companies are still racing to deploy agents even as early rollouts expose errors, data leaks, and security risks

NEW YORK, NY, UNITED STATES, November 18, 2025 /EINPresswire.com/ -- Gravitee, a leading provider of API management and agentic AI solutions, has released new research that shows signs of both confidence and chaos in the agentic AI era. This new study shows 82% of U.S.-based companies using AI agents have already seen those agents act in an unexpected way, such as making incorrect decisions, exposing data, or triggering security breaches. Despite these rogue agents, according to the same survey, 60% of organizations still plan to launch more than 15 additional agents by the end of 2026. The findings underscore the growing tension between innovation and control as businesses race to scale agentic AI responsibly.

Despite the risks, companies continue to move forward with their agentic AI plans. Gravitee’s findings reveal an interesting reality for enterprise leaders. While a vast majority have already witnessed rogue behavior from their AI agents, most leaders are not deterred by these issues and may view them more as growing pains, not red flags. Every company surveyed says they plan to launch at least five agents by the end of the 2026. They see agentic AI as too powerful to pause, citing gains in automation, customer service, and data-driven decision-making. This comes on the heels of previously released research by Gravitee that approximately 1.2 million AI bots will enter the U.S. workforce by 2026, replacing roles at nearly two-thirds of firms. The race to scale is exposing a deeper problem: innovation is moving faster than governance, leaving organizations struggling to balance security with pace.

For most companies, “rogue” agent behavior is an operational reality. Gravitee’s research found that AI agents may have already made incorrect business decisions, exposed confidential data, and even triggered security incidents. In some cases, agents may have acted on outdated or incomplete information; in others, they may have escalated tasks or shared data across systems without human oversight. These missteps underscore the risks of deploying autonomous systems without guardrails. As organizations accelerate their AI adoption, the line between autonomy and anarchy is growing dangerously thin.

To regain control, many organizations are now fortifying the connective tissue of their AI ecosystems. Gravitee’s research shows that only 42% of companies have already enlisted third-

party API management to bring governance to what's become known as "agent sprawl." As AI protocols emerge, such as Google's Agent-to-Agent (A2A) protocol, which governs how agents discover, authenticate, and interact with one another, the need for visibility and security between those connections has never been greater. In the agentic AI world, knowing how agents talk is just as critical as what they say.

"Agentic AI isn't the problem - unchecked autonomy is," said Rory Blundell, CEO, Gravitee. "Right now, companies are giving their AI agents keys to the enterprise without knowing which doors they're opening. The rapid increase of the number of agents employed at companies across the world means these agents are now talking to each other, making decisions, and sharing data faster than humans can track. Without an API-level control layer, that's a recipe for a cybersecurity disaster. The companies that rise to the top of the next wave of AI won't be the ones moving fastest, they'll be the ones building the strongest guardrails."

The findings point to a new inflection point for AI adoption. As agentic AI becomes the backbone of digital operations, organizations must decide whether they'll build for control or speed. Governance can no longer be an afterthought patched on after deployment, it has to be architected from the first line of code.

For more information on Gravitee and how it can help enable organizations to apply policy, monitoring, and governance across those A2A interactions, please visit gravitee.io.

To support enterprise leaders navigating this transition, Gravitee hosted the inaugural A2A Summit on November 6, 2025, in New York, in partnership with The Linux Foundation. The event explored the future of agent-to-agent (A2A) orchestration and autonomous enterprise systems, bringing together technology leaders from Google, Microsoft, Gartner, and others to provide actionable insights to help organizations tackle agent sprawl and unlock the full potential of AI-driven decision-making.

-ENDS-

About Gravitee
Gravitee.io , with a valuation of over $300m, is the open-source leader in Agentic API & Event Management. The Gravitee platform empowers enterprises to design, secure, and govern APIs, event streams, and AI-driven interactions across hybrid, multi-cloud, and edge environments. With a federated, agent-ready approach and native support for real-time traffic and autonomous agents via the Gravitee Agent Mesh, Gravitee enables secure, scalable, and intelligent connectivity in an increasingly complex ecosystem.

About the survey
This research was conducted in September 2025 by Opinion Matters and is based on a survey of

500 technology leaders of companies with 250 or more employees: 250 based in North America.

Nicholas Bennett
Gravitee
nicholas.bennett@graviteesource.com
Visit us on social media:
LinkedIn
