Friday, July 10, 2026

ET Edge | CISOFORUM

Security's new blind spot: The rise of machine-speed decisions without machine-speed oversight

Posted on July 3, 2026 by CISO Forum Bureau

AI agents and automation are making decisions faster than enterprises can monitor or explain them. This growing oversight gap is creating new challenges around accountability, governance, and cyber resilience.

The workforce shift is not coming. It is here.

The honest version of the AI story is simpler than the industry wants to admit. Enterprises are not deploying AI agents to help their teams. They are deploying them to absorb work their teams currently do. The language is "augmentation." The spreadsheet says headcount reduction.

And the numbers are moving fast. Gartner projects that 15 percent of day-to-day work decisions will be made autonomously by agentic AI by 2028, up from zero in 2024. Forty percent of business applications are expected to feature autonomous agents by the end of this year.

These are not chatbots answering FAQs. Agents are triaging support tickets, matching invoices, routing approvals, executing security playbooks, screening candidates, and managing infrastructure. Each of those is a decision. Each decision has consequences. And each consequence will eventually need to be explained to someone: a regulator, an auditor, a board, or a court.

The economic case for agents is settled. The question that has not been answered is a different one entirely: when those explanations are demanded, will the enterprise actually be able to provide them?

Shomiron Das Gupta
Founder & CEO
Bloo.io

The governance gap is not a policy problem. It is an infrastructure problem.

Most organizations treat agent governance as a matter of writing the right policies. Define the boundaries. Document the escalation paths. Assign accountability. Put it in a framework. Review it annually.

That approach worked when decisions were made by people and recorded in systems designed for human workflows. It does not work when decisions are made by autonomous software executing across multiple systems at machine speed, often without a human in the loop at all.

The data is already showing the strain. A 2026 report from Gravitee surveying over 900 executives found that 81 percent of organizations are past the planning phase on AI agent deployment, but only 14.4 percent have full security approval. Eighty-eight percent confirmed or suspected security incidents tied to their agents. Only 22 percent treat agents as independent identities. The rest still rely on shared API keys.

A separate study from Kiteworks found that 63 percent of enterprises cannot enforce purpose limitations on their AI agents, 60 percent cannot quickly terminate a misbehaving one, and 33 percent lack evidence-quality audit trails. Sixty-one percent have fragmented logs spread across disconnected systems.

These are not policy failures. These are infrastructure failures. You cannot audit what you did not record. You cannot explain a decision you cannot reconstruct. And right now, most enterprises cannot reconstruct what their agents did last week, let alone six months ago.

Gartner predicts that over 40 percent of agentic AI projects will be canceled by the end of 2027. The three reasons they cite are escalating costs, unclear business value, and inadequate risk controls. That third reason is the one that should concern security leaders most, because it is the one that turns an operational problem into a regulatory one.

The real problem: agents are not deterministic, and that changes everything

Here is the part of the conversation the industry has not caught up with yet.

For decades, enterprise audit has relied on one fundamental assumption: that systems are deterministic. If you know the inputs and the code, you can reconstruct the output. The logic is fixed. The path is repeatable. If something went wrong, you read the code, trace the logic, and arrive at the same answer the system did. In that world, the code itself is the audit trail.

AI agents break that assumption completely.

An agent does not follow a fixed path. It reasons. It weighs context. It synthesizes information from multiple sources. And critically, it may arrive at a different answer each time, even when given the same inputs. The decision is not a function of static logic. It is a function of model state, the specific data available at that moment, and a reasoning process that is inherently variable.

This has a structural consequence that most governance frameworks have not absorbed: you cannot audit an agent by reading its code. You cannot reproduce its decision by re-running it. The model may have been updated. The data may have changed. The context window that existed during that reasoning chain is gone.

In a deterministic world, the code is the record. In a non-deterministic world, the telemetry of the actual execution is the only record. If you did not capture the inputs the agent consumed, the reasoning path it followed, and the action it took at the moment it happened, that evidence is gone. Permanently. There is no re-run. There is no replay.

This is what makes agent observability fundamentally different from traditional monitoring. Monitoring tells you what a system is doing right now. What agents require is memory: a complete, immutable record of what they did, what data they used to decide, and what outcome they produced. Not a summary. Not a sample. Not a dashboard metric. The full record, retained long enough to satisfy the regulator, the insurer, or the board when they come asking.

And here is where the gap becomes structural. Most enterprise telemetry architectures retain 30 to 90 days of hot, searchable data. IBM's 2025 Cost of a Data Breach report found that the average time to identify a breach is 181 days. Mandiant's M-Trends 2026 report found that median dwell time for

espionage intrusions was 122 days. By the time most enterprises discover a problem, the telemetry covering the first half of the story has already been overwritten.

That was already a problem with human-operated systems. With autonomous agents making thousands of decisions per day across finance, operations, IT, and security, it becomes an existential governance risk. Every decision that was not captured at full fidelity is a decision that can never be explained.

Regulation is converging on exactly this point

Regulators have noticed the gap, and they are moving.

In February 2026, NIST launched a dedicated initiative to develop standards for autonomous AI agents, covering agent identity, action logging, and containment boundaries. The EU AI Act imposes fines up to 35 million euros or 7 percent of global turnover for prohibited practices, and requires full documentation of decision processes for high-risk systems. SEC Regulation S-K Item 106 requires public companies to describe board oversight of cybersecurity risks. DORA mandates continuous operational resilience testing across financial services.

A governance framework published in June 2026 goes further: it mandates that audit logs capture trigger events, inputs, actions, timestamps, and responsible owners for each autonomous action.

The pattern is clear. Regulators are not asking whether you deployed agents. They are asking whether you can explain what your agents did. That is not a document you can write after the fact. It is a telemetry infrastructure you either have in place or you do not.

The moat is not the agent. The moat is the memory.

Every major platform vendor is building agents. The agent layer will be commoditized. Where differentiation will emerge is in the layer underneath: the infrastructure that captures what every agent did, across every domain, at full fidelity, in a format that both humans and machines can reason over.

That layer does not exist in most enterprises today. SIEMs were not designed for it. Observability platforms were not designed for it. Cloud-native logging was not designed for it. What is needed is a telemetry substrate that treats every autonomous action as a first-class event, captured with full context, retained for as long as the enterprise might need to explain it, and structured so that audit tools, compliance systems, and yes, other agents, can reconstruct the causal chain.

Think of it this way. The enterprise data warehouse became critical infrastructure not because it made decisions, but because it held the truth about what the business did. The next equivalent is the telemetry layer that holds the truth about what the machines did.

The organizations that build this layer first will be the ones that can scale agent deployment safely, satisfy regulators, and maintain accountability over autonomous operations. The ones that skip it will join the 40 percent that Gartner predicts will cancel their projects.

The agent is the capability. The memory is the control. Enterprises investing in the first without the second are building speed without the ability to explain where they went.

Sources:

- Gartner, “Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027,” June 2025
- Gravitee, “State of AI Agent Security 2026,” February 2026
- Kiteworks, “Data Security and Compliance Risk: 2026 Forecast Report,” May 2026
- IBM, “Cost of a Data Breach Report 2025”
- Mandiant (Google), “M-Trends 2026 Report”
- NIST, Autonomous AI Agent Standards Initiative, February 2026
- Microsoft, “2026 Work Trend Index,” May 2026
- Deloitte, AI Governance Maturity Research, 2025

–Authored by Shomiron Das Gupta, Founder & CEO, Bloo.io

Author

CISO Forum Bureau

View all posts

Posted in View Points

Tagged AI, security

Related Posts