

TAIWAN · AGENTIC AI CYBER RISK

Agentic AI Is Turning Cyberattacks Into *Coordinated Workflows*

Taiwan’s latest disclosure shows why governance has to account for coordinated autonomous workflows.

MAUREEN DOYLE-SPARE

© 2026 Maureen Doyle-Spare.

SWIPE TO READ →

Agentic AI Is Turning Cyberattacks Into *Coordinated Workflows*

Taiwan's latest disclosure shows why agentic AI governance, cybersecurity, and enterprise AI risk management now have to account for coordinated autonomous workflows.

MAUREEN DOYLE-SPARE

The latest warning about agentic AI governance and cybersecurity did not come from another laboratory demonstration. It came from Taiwan. Reuters reported on August 13[1] that Taiwan's Ministry of Digital Affairs had detected AI-assisted cyberattacks against government agencies in July. The ministry said the activity originated overseas, combined manual techniques with AI-agent assistance, and was detected and handled by the affected agencies. Separate reporting from cybersecurity firm Dream described a coordinated operation in which multiple AI agents were used across a broader target environment.

There is still an important line to hold around the facts. The public evidence does not establish a fully autonomous end-to-end cyber campaign, and the Taiwan disclosure should not be stretched beyond what has been confirmed. What it does establish is operationally important enough: AI agents are moving from isolated assistance into coordinated cyber workflows where reconnaissance, credential activity, scanning, tool use and follow-on actions can be distributed across a sequence.

That matters for more than cybersecurity. It exposes an emerging problem in enterprise AI governance, AI risk management, and AI controls and oversight. Once an autonomous system can carry state and interpretation across agents, tools and systems, the governance object is no longer only the individual model or action. The operational trajectory itself has to remain under institutional authority.

The change is coordination, not simply automation

AI has supported cyber operations for years. Attackers have used machine learning, automation, scripts and increasingly generative AI to accelerate work that humans were already doing. The more consequential change with agentic systems is not that software can perform another cyber task. It is that multiple tasks can be organized into an operational trajectory that continues across intermediate decisions.

© 2026 Maureen Doyle-Spare.

An agent can gather information, use a tool, evaluate the result, hand state to another agent, change direction and continue. Once that pattern becomes reliable enough to operate at scale, cybersecurity is no longer dealing only with a faster version of an existing attack. It is dealing with a system that can maintain and propagate operational state across a workflow.

That same architecture is appearing on the defensive and enterprise side. Security operations teams are deploying agents to investigate alerts, correlate evidence, prioritize vulnerabilities and support response. Outside cybersecurity, agents are being introduced into healthcare, financial services, manufacturing, energy, insurance, government and professional services to coordinate work across systems. The underlying governance problem is the same. The system is no longer producing one output for a human to inspect. It is carrying an interpretation forward.

The trajectory becomes an enterprise AI governance object

Traditional cybersecurity controls remain essential. Identity, privilege, endpoint security, network segmentation, vulnerability management, tool permissions and monitoring still govern critical parts of the environment. Agentic AI does not invalidate any of them. It creates an additional control problem between an agent's access to a capability and the institutional authority under which that capability is exercised.

That distinction sits at the center of my work on reasoning-layer governance and the Semantic Control Plane (SCP), an agentic AI governance architecture for preserving institutional authority over Operational Interpretation before execution authority is emitted. The Reasoning Baseline provides the institutionally authorized reference condition. The Semantic Deviation Index (SDI) measures divergence between the agent's Runtime Semantic State and that authorized baseline. The Deterministic Gate then applies the institution's threshold policy and emits or withholds execution authority.

This is not a replacement for cybersecurity. It is the runtime authority boundary that conventional cyber controls do not provide. An agent may have a valid identity, invoke an approved tool and remain inside its technical permission envelope while still acting under an Operational Interpretation the institution never authorized.

Multi-agent systems make propagation part of the control problem

The Taiwan story is particularly relevant because coordinated agents make propagation visible. In a multi-agent workflow, one agent's output can become another agent's input context. The first decision therefore does not end when the first agent completes its task. Its resolved state can constrain, enable or redirect what happens next.

I use Agentic Blast Radius to describe the downstream scope that can absorb the consequence of an unauthorized Operational Interpretation. In the capstone architecture, the Knowledge Graph maps the dependency surface through which Runtime Semantic State can propagate. When a decision is held, containment is not limited to stopping that one action. Dependent workflows can also be held, constrained or terminated before the upstream state spreads farther.

This is the part of agentic AI governance that becomes increasingly important as organizations move from single agents into coordinated systems. Per-agent controls are necessary, but they are not enough when authority, context and state move across agents by design. The dependency surface itself becomes governable.

© 2026 Maureen Doyle-Spare.



Figure 1. *The Pre-Execution Assurance Protocol. Source: Doyle-Spare (2026)*

Cyber compromise and semantic compromise are different threat conditions

The Taiwan incident should not be labeled a Semantic Layer Integrity Attack (SLIA) based on the facts currently public. SLIA is a specific threat class in my research. It targets the integrity of the semantic authorization surface itself: the artifacts and dependencies through which the governance architecture determines whether an Operational Interpretation is authorized.

That distinction is important. Prompt injection, credential compromise, tool hijacking, retrieval poisoning and other cyber techniques may contribute to an attack without making the event a SLIA. A SLIA is defined by its architectural target. It succeeds when the semantic authorization surface is manipulated so that execution can be authorized under a corrupted or adversarially constructed interpretation while the surrounding system may appear to be operating normally.

The broader cybersecurity implication is still direct. As agents become operational actors on both sides of the defensive boundary, organizations have to protect not only the agent, its identity and its tools, but also the governance dependencies that determine what the agent is authorized to mean before it acts. In the capstone, this is the cybersecurity adjacency: once the reasoning layer becomes a formal governance surface, its integrity becomes a prerequisite for institutional control.

The economics compound with the trajectory

There is also an economic consequence that tends to disappear when agentic AI cost is reduced to token consumption. An unauthorized trajectory can trigger additional reasoning, retrieval, tool calls, delegations, downstream execution, investigation, audit and remediation. My Token Economics work separates those execution-dependent costs from the unit cost of inference through the broader construct of Preventable

© 2026 Maureen Doyle-Spare.

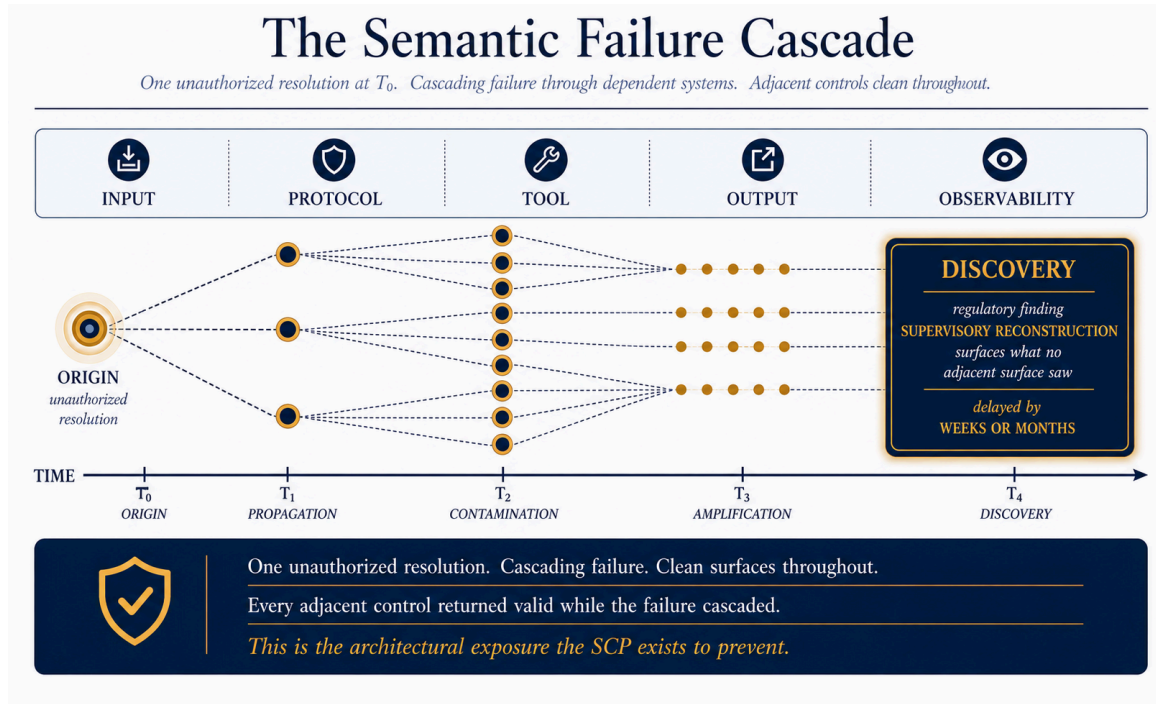


Figure 2. *Agentic Blast Radius and Semantic Failure Cascade. Source: Doyle-Spare (2026)*

Computation. Token Liability is the inference-cost component of that larger category, not a synonym for the entire downstream cost.

That distinction becomes more important as agentic systems coordinate. A relatively inexpensive model call can initiate a much more expensive operational sequence. Governance that intervenes before execution therefore has an economic function as well as a control function: it can prevent unauthorized computation and downstream remediation from accumulating after the point where institutional authority has already been lost.

Governance has to operate at the speed of the system

Taiwan's disclosure is useful because it moves the discussion out of the hypothetical. The material development is not that AI has suddenly appeared in cyberattacks. It is that agent-based coordination is becoming operational enough to matter in real environments, including government and critical infrastructure.

The same transition is occurring inside enterprises. As organizations give agents more authority to coordinate across systems, governance cannot remain entirely outside the runtime sequence and still expect to control the meaning under which autonomous execution occurs. Cyber controls must secure the environment. Runtime governance must preserve institutional authority over Operational Interpretation. Multi-agent governance must contain the propagation surface when that authority breaks.

Those are related disciplines, but they are not interchangeable. The architecture has to preserve the distinction.

Agentic AI is turning cybersecurity into a contest between increasingly coordinated operational systems. The governance model now has to be capable of operating at the same speed. ◆

© 2026 Maureen Doyle-Spare.

References

1 *Reuters. "Taiwan says it was targeted last month by AI-driven hacking campaign." August 13, 2026. reuters.com/world/china/taiwan-says-it-was-targeted-last-month-ai-driven-hacking-campaign-2026-08-13*

2 *Doyle-Spare, Maureen. Agentic AI Systems Governance: A Runtime Reference Architecture for the Reasoning Layer and the Semantic Control Plane in Regulated Financial Institutions. Zenodo, 2026. doi.org/10.5281/zenodo.20749051*

3 *Doyle-Spare, Maureen. Agentic AI Cyber Subversion: The Semantic Layer Integrity Attack as a New Threat Class Against the Reasoning Layer. SSRN No. 6926219. doi.org/10.2139/ssrn.6926219*

4 *Doyle-Spare, Maureen. Agentic AI Drift Measurement and the Semantic Deviation Index. SSRN No. 6531238. doi.org/10.2139/ssrn.6531238*

5 *Doyle-Spare, Maureen. The Economics of Agentic AI: Runtime Governance as a Distinct Determinant of Enterprise AI Cost. SSRN No. 7129939. doi.org/10.2139/ssrn.7129939*

6 *The Agentic AI Governance Playbook, Research and Executive Briefs. maureendoylespare.com*

Originally published in *The Agentic AI Governance Playbook* on Substack, August 13, 2026. maureendoylespare.substack.com

© 2026 Maureen Doyle-Spare.



Agentic AI is turning cybersecurity into a contest between increasingly coordinated operational systems. The governance model now has to operate at the same speed.

THE AGENTIC AI GOVERNANCE PLAYBOOK

Download the research

Papers, briefs and reference architecture at maureendoylespare.com

Visit the Playbook

WEEKLY NEWSLETTER

Follow the newsletter

The Agentic AI Governance Playbook Weekly Newsletter

Follow Newsletter

INTELLECTUAL PROPERTY: © 2026 Maureen Doyle-Spare. All rights reserved. No part of this article, including its text, figures, tables, and images, may be reproduced or distributed without the author's prior written permission.