CDAO CHICAGO 2026

Who Authorized the Agent?

Governing AI that acts on someone's behalf

The question behind every stalled agent pilot: who granted the authority to act?

Karthik Mahalingam

IEEE Chicago • August 12, 2026 • 2:30 pm

The authorization gap

Agent risk is not wrong answers. It is wrong actions.

Nothing stood between decision and irreversible action.

Pilots stall at sign-off — not at the model

What the numbers actually say

>40%

PROJECTS CANCELED
by end-2027

$ cost · ? unclear value · ! inadequate controls

Gartner forecast · Jun 2025

17%

DEPLOYED TODAY

60%+ expect deployment within two years

Gartner CIO Survey · 2026

<15%

PILOTS REACH PRODUCTION

while 78% of enterprises are already piloting

650 tech leaders · Mar 2026

MODEL CAPABILITY ISN'T ON THE LIST. No one signs off on what they can't scope, trace, or revoke.

A new kind of actor

The legal and technical shift underneath every stalled sign-off

A tool, used **BY** a person

- The human remains the actor
- Permissions attach to the person
- Every action maps to a user session
- Accountability is settled law and settled IT

VS

An agent, acting **AS** a person

- Research Agent
- Data Agent
- Policy Agent
- Ops Agent

- Exercises delegated authority
- Acts while no one is watching
- Can spawn and task other agents
- Accountability diffuses across five parties

21.9% Treat agents as independent identity-bearing entities

45.6% Still run agents on shared API keys

25.5% Can create or task other agents

Identity. Permissioning. Lineage. This isn't legal's problem. It's ours.

Singapore wrote the first agent-specific playbook

IMDA Model AI Governance Framework for Agentic AI

1

Assess & bound risks upfront

Score impact × likelihood before anything ships.

2

Make humans meaningfully accountable

Named owners. Approval at significant checkpoints.

3

Implement technical controls

Least privilege. Unique identity. Sandboxing.

4

Enable end-user responsibility

Users know when the agent acts and how to stop it.

EU — binding law, generic to high-risk AI

US — patchwork, evolving

Singapore — voluntary, agent-specific, operational

v1.0 Jan 2026 • v1.5 May 2026

60+ organizations • 10+ case studies

Inside the framework: three moves

How IMDA turns principle into practice

01

SCORE THE RISK

RISK = IMPACT × LIKELIHOOD

- Impact: sensitivity · read/write · reversibility
- Likelihood: autonomy · complexity · third parties
- v1.5: complexity itself raises the score

02

BOUND THE AGENT

THE CEILING RULE

- Least privilege — only what the task needs
- Unique identity — no shared credentials
- Never more authority than the human delegator

03

INSERT THE HUMAN

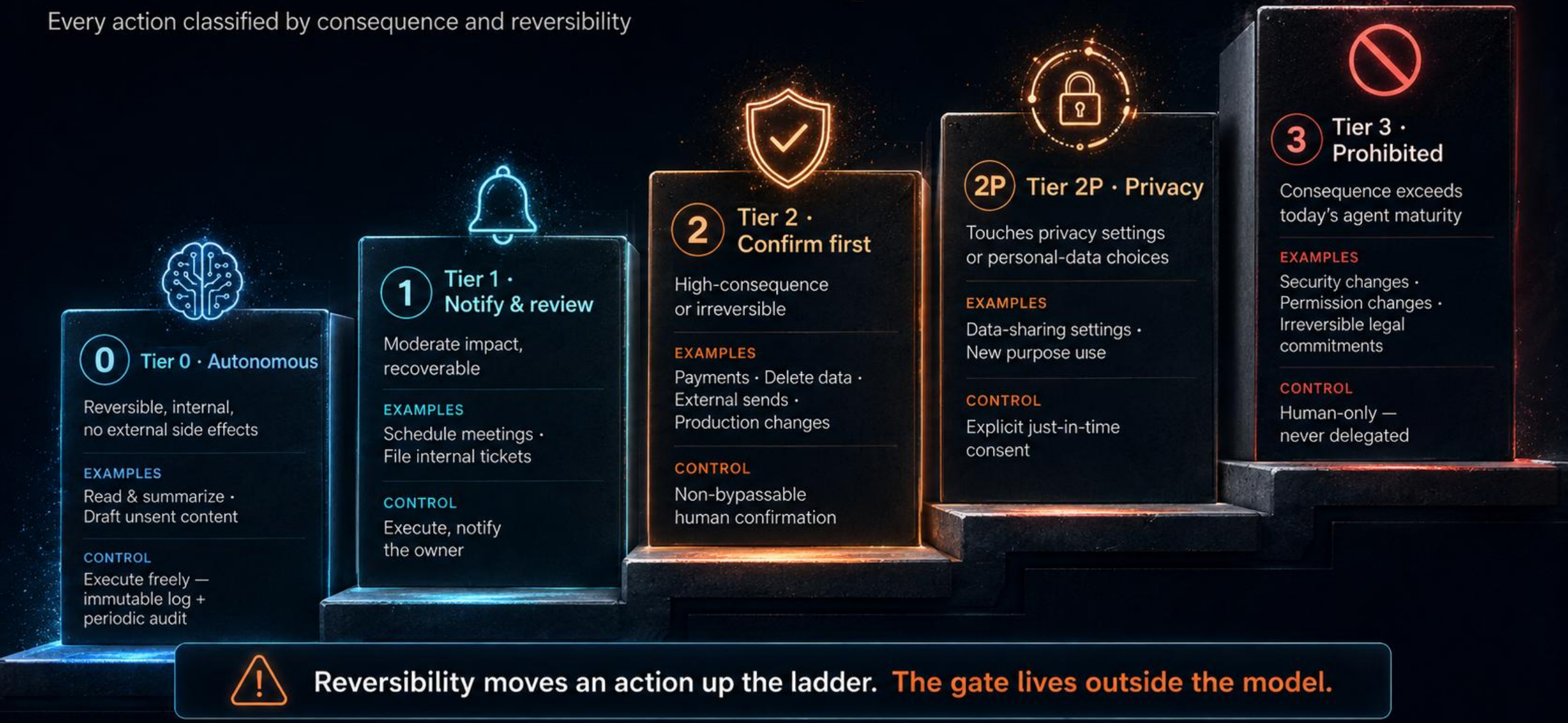
DETERMINISTIC CHECKPOINTS

- High-stakes · irreversible · outlier actions
- Delete · send · pay · user-set thresholds
- Track overrides to prove oversight is real

They never asked how good the model was. They asked how bad the action could be.

Govern the action, not the agent

Every action classified by consequence and reversibility



Reversibility moves an action up the ladder. The gate lives outside the model.

The four primitives of delegated authority

If the architecture cannot express all four, the agent is not ready for production

Shared service accounts break all four at once.

If you can’t demonstrate all four for an agent, it doesn’t ship.

The audit trail is what regulators ask for first

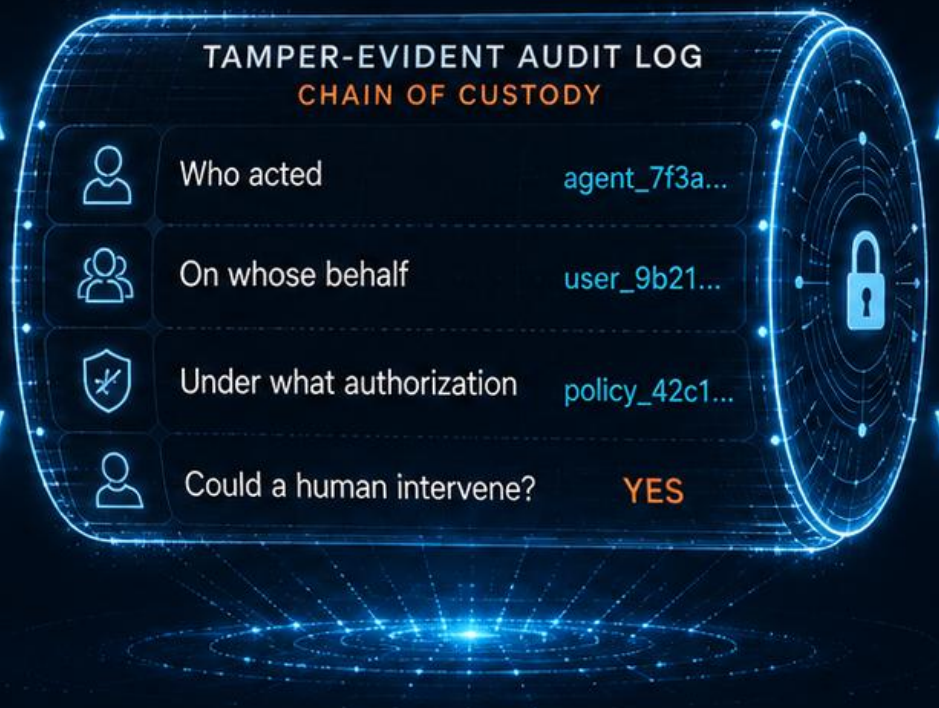
Different regimes, one convergent artifact

EU AI Act

Event logging · Human oversight · Retention

United States

Human review rights · auditability



Singapore IMDA v1.5

Verifiable identity + authorization trail

Payment rails

Agent identity + signed consumer consent

One artifact satisfies them all: **a tamper-evident log.**

Monday morning

No law required — start with the structure that already exists

1

This week

Inventory every agent and agent-like automation. Named owner + unique identity for each. No shared credentials.

2

30 days

Tier every action by impact × reversibility. Deterministic confirmation for financial and irreversible actions. Just-in-time consent for privacy actions.

3

60–90 days

Stand up Consent, Scope, Revocation, Audit — and make ‘show me all four’ the production gate.

4

This quarter

Shape the audit trail to Article 12 readiness: actor, principal, authorization, tools, timestamps, human decisions. Tamper-evident. 6+ month retention.

5

Ongoing

Watch override rates and approval times. A checkpoint no one rejects isn’t oversight.

The teams that reach production don’t have better models.
They have answered: who authorized the agent?

Karthik Mahalingam • IEEE Chicago • linkedin.com/in/karthikprivsec