AiOpsOne

Agentic AI Security Scoping Matrix

Right-size AI-agent security to autonomy. Classify each agent by scope, then apply the control depth that scope demands across six dimensions — with the APRA CPS 234 evidence lens.

VERSION 1.0 · JUNE 2026 · AIOPSONE.COM · YOUTUBE @AIOPSONE

How to use this. More autonomy = bigger blast radius when an agent is manipulated. Place each agent in a scope (1–4), then read down the column for the control depth required in every dimension. Re-classify whenever you add a tool or remove an approval step.

SCOPE 1
No Agency
Read-only; recommends. Human executes everything.

SCOPE 2
Prescribed
Proposes changes; every action needs human approval.

SCOPE 3
Supervised
Acts autonomously within set boundaries.

SCOPE 4
Full Agency
Self-initiating; runs continuously, minimal oversight.

— AUTONOMY & REQUIRED CONTROL DEPTH INCREASE →

Six security dimensions × four scopes

DIMENSION	SCOPE 1	SCOPE 2	SCOPE 3	SCOPE 4
Identity Context	User auth	+ agent identity	Scoped to the acting user (AgentCore)	Per-action authz + JIT, no standing access
Data, Memory & State	Stateless	Read-only context	Protected, access-controlled memory	Memory integrity + poisoning detection
Audit & Logging	API call logs	+ decision log	Reasoning-chain + tool-call logs	Full behavioural trace, retained
Agent & LLM Controls	Input/output filtering	Bedrock Guardrails	+ sandboxing	Containment + kill switch
Agency Perimeters	Read-only scope	Mandatory approval gate	IAM/policy boundary (outside the prompt)	Dynamic, enforced constraints
Orchestration	Single tool	Reviewed tool set	Authorization per tool call	Agent-to-agent trust controls

THE CPS 234 LENS

A **Scope-3 or Scope-4 agent is a non-human identity with standing permission to act** — in control terms, the same risk as a human with standing admin (a CPS 234 ¶21 / Essential Eight "restrict admin" finding), except it takes instructions from untrusted text. For an APRA-regulated entity: classify the agent on this matrix *first*, apply the column's controls, and keep the classification + the six-dimension evidence (IAM scoping, memory protection, reasoning-chain logs, guardrail config, policy boundary, tool authorization) as part of your control evidence.

Jayprakash Bilgaye · AWS Certified Security – Specialty

CPS 234 → AWS controls + more resources → aiopsone.com/resources