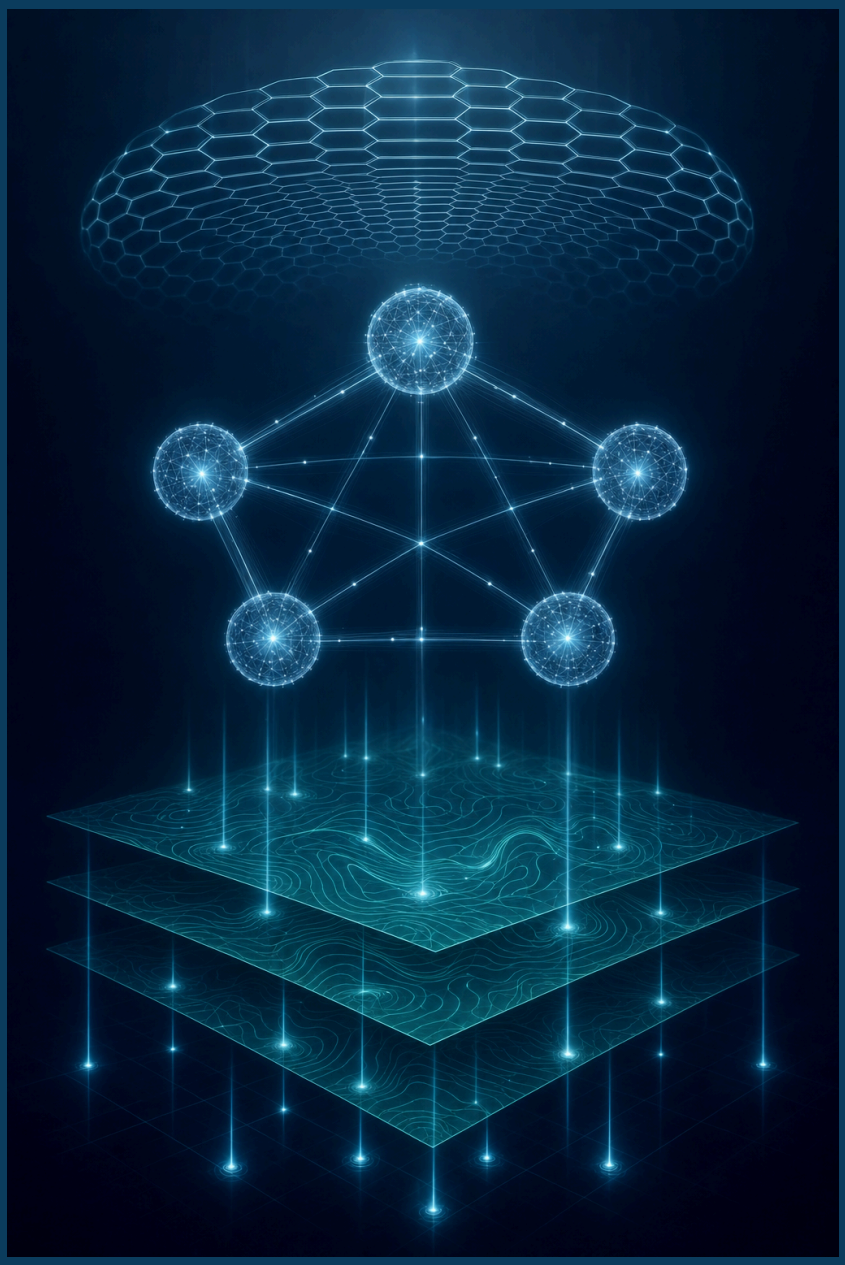

Governing the Agents Your Employees Are Already Running

A Compliance Officer's Implementation Guide for Singapore's IMDA Agentic AI Framework

Authors Bev Waller & Erik Jones
Published June 3, 2026
Publisher Jacobian Engineering · jacobianengineering.com

Governing the Agents Your Employees Are Already Running

A Compliance Officer's Implementation Guide — Fitting Singapore's IMDA Agentic AI Framework into your existing SOC 2, HIPAA, HITRUST, and ISO 27001 programme

Executive Summary

On 20 May 2026, Singapore's IMDA published the first government-issued governance framework written specifically for agentic AI. Your employees are already running agents — Microsoft Copilot, Cursor, Cline, Salesforce Einstein agents, ServiceNow Now Assist, Notion AI, Glean, AWS Bedrock Agents — and policy has not caught up. This guide is for compliance officers, CISOs, and AI risk officers who now have to fit the framework into a programme that already runs SOC 2, HIPAA, HITRUST, and ISO 27001. The governance overlay precedes the new technical controls.

The Shadow-AI Problem Compliance Has Inherited

Walk a floor at any growing SaaS or healthcare-technology company in 2026 and you will find agents your IT team did not sanction. Not because employees are reckless — because the agents work. A finance analyst is using Microsoft Copilot for a board memo. An engineer is letting Cursor or Cline write a script that touches production secrets. A sales rep is using Salesforce Einstein agents to triage opportunities. Support leans on ServiceNow Now Assist, marketing has plugged Notion AI into the research workspace, and a platform engineer has stood up an AWS Bedrock Agent for billing reconciliation. Glean quietly searches across three CRMs and a data room.

None of those uses was malicious. Each one cleared a desk. The cumulative result is that the compliance officer has inherited a fleet of decision-makers and action-takers that do not appear in any existing asset inventory, do not map to a documented control set, and do not have a named human owner. This is shadow AI, and it is the operating reality compliance has to govern around.

The 2026 compliance question — already arriving in customer security questionnaires and surfacing in SOC 2 walkthroughs — is simply: "What is your agentic AI governance posture?" Teams that can answer with an inventory, a policy, a risk-register entry, and a named owner per agent will close enterprise deals faster and walk into examinations with less anxiety.

Singapore's IMDA Framework as a Structural Backbone

Singapore's IMDA framework matters because it is the first government-issued governance document written for AI systems that plan, call tools, and act over multiple steps. The EU AI

Act's conformity guidance references it as a touchstone for high-risk and general-purpose AI, and national AI strategies in the UK, Canada, and Japan are aligning to it. Voluntary guidance is doing the work of de facto regulation in 2026.

For a compliance officer, the four dimensions translate cleanly. **Dimension 1 — Assess and bound risks upfront** is risk analysis under a different name. **Dimension 2 — Make humans meaningfully accountable** is the control-environment work your SOC 2 examiners already expect. **Dimension 3 — Implement technical controls and processes** is access management, monitoring, change management, and logging applied to a new class of principal. **Dimension 4 — Enable end-user responsibility** is policy, training, and notice-and-disclosure work your privacy and awareness programmes already run.

The engineering-side deep walkthrough lives with our AI services division at TrustEdge: see the Singapore IMDA Agentic AI Regulations Hub. The remainder of this guide stays in the compliance lane.

Mapping the Four Dimensions to Your Existing Frameworks

Most of what IMDA expects already maps to mature SOC 2, HIPAA, HITRUST, NIST AI RMF, and ISO 27001:2022 programmes. The work of the next two quarters is largely a crosswalk and a scoping exercise — bringing agents into evidence categories you already produce.

Dimension 1: Risk Assessment Upfront

The artifact is a risk-register entry treating each agentic system as a discrete asset. **SOC 2 CC3.1–CC3.4** already require risk identification — extend scope to capture agent autonomy tier, data scope, and tool-call surface. **HIPAA §164.308(a)(1) Risk Analysis** is the home for any agent touching ePHI; if Microsoft Copilot has tenant access to a SharePoint that carries PHI, the agent belongs in the Risk Analysis. **NIST AI RMF — Map function** is this work. **ISO 27001:2022 Annex A.5.7 (threat intelligence) and A.5.9 (asset inventory)** extend to agentic systems. **HITRUST CSF 03.b** matches. Beyond your current evidence library, expect a documented autonomy-tier classification per agent and a residual-risk justification for any Act-with-Approval or Autonomous agent.

Dimension 2: Human Accountability

The artifact is a RACI for agent oversight — named owners, defined responsibilities, unambiguous escalation criteria. **SOC 2 CC1.1–CC1.5** cover control environment and accountability — absorb agent owners as an extension of the system-owner concept. **NIST AI RMF — Govern function** matches: GOVERN-1.x organisational governance, GOVERN-2.x accountability structures. **ISO 27001:2022 Annex A.5.2 and A.5.4** extend to agent-action approvers. **HIPAA §164.308(a)(2) Assigned Security Responsibility** has to grow to treat agents as workforce-equivalent principals — a non-trivial shift your auditor will ask about. **HITRUST CSF 01.a and 02.a** match. Expect to show a named human owner per agent, an escalation SOP, and evidence the reviewer is actually engaging rather than rubber-stamping.

Dimension 3: Technical Controls and Processes

The artifact is a control matrix connecting each agentic system to access, monitoring, change-management, and logging controls. **SOC 2 CC6.1–CC6.8 (logical access) and CC7.1–CC7.5 (system operations)** already require continuous monitoring and anomaly detection — Datadog, PagerDuty, and AWS GuardDuty signal flows extend naturally to agent activity, provided agents have distinct identities. **NIST AI RMF — Measure and Manage functions** are the operational counterparts. **ISO 27001:2022 Annex A.8 — particularly A.8.2 (privileged access), A.8.15 (logging), A.8.16 (monitoring), and A.8.32 (change management)** apply with minimal interpretation. **HIPAA §164.312(a)(1), §164.312(b), and §164.312(c)(1)** cover the regulated-data side. **HITRUST CSF 09.aa, 09.ab, 09.j, and 10.k** round it out. Expect agent-distinct identities in IAM, multi-agent emergent-behaviour testing (Garak, PyRIT, Lakera Guard, NeMo Guardrails fit here), and a change-management trail capturing prompt updates and tool grants as configuration changes.

Dimension 4: End-User Responsibility

The artifact is an AI Use Policy addressing agent-specific behaviour, a training programme reaching the people interacting with agents, and notice-and-disclosure logs. **SOC 2 CC2.1–CC2.3** cover internal policy communication and external disclosure. **NIST AI RMF GOVERN-4.x and MANAGE-4.x** address awareness, training, and post-deployment communication. **ISO 27001:2022 Annex A.6.3 (awareness and training) and A.5.10 (acceptable use)** match directly. **HIPAA §164.308(a)(5) Workforce Training** is the home for agent-specific training records. **HITRUST CSF 02.e** matches. Expect an AI Use Policy that names agent-specific clauses (not just generative-AI clauses) and an automation-bias measurement programme.

The genuinely new asks — agent identity at the IAM layer, multi-agent emergent-behaviour testing, automation-bias monitoring, and tradecraft retention — are the controls our 90-day compliance checklist walks through with concrete first-move guidance.

ISO/IEC 42001 — The AI Management System Standard You Should Know About

ISO/IEC 42001 is the international standard for AI Management Systems, published in December 2023. Where NIST AI RMF is a voluntary framework you map to, ISO/IEC 42001 is a certifiable management system you operate against — the same relationship ISO 27001 has to NIST CSF. It defines requirements for establishing, implementing, maintaining, and continually improving an AI management system, and it is being positioned as the natural management-system home for the controls IMDA, NIST AI RMF, and the EU AI Act all expect.

Why this matters for compliance teams in 2026: certification will become a vendor-questionnaire ask within twelve to eighteen months, and the enterprise customers most worth selling to will treat ISO 42001 the way they treat ISO 27001 today. We will cover ISO 42001 implementation in depth in a separate whitepaper — this section is your awareness primer, and it is enough to start the internal conversation about whether your team pursues certification in the next planning cycle.

A Policy Template You Can Adapt

Most existing AI Use Policies were written for ChatGPT-style assistants — chat, draft, summarise — and are silent on agents that take action. Below are six agent-specific clauses to add. Each one is presented with sample language your legal and compliance team can adapt verbatim.

Agent inventory and registration

> *All AI agents — including embedded agent-mode features in licensed software — must be registered in the AI Agent Inventory before being granted access to company data. Registration captures the agent name, vendor, owner, business sponsor, autonomy tier, data scope, tools the agent may call, and date of next scheduled review.*

Autonomy tier classification

> *Each agent is classified into one tier. **Assist**: produces suggestions a human reads but never acts on. **Suggest**: proposes an action a human performs. **Act-with-Approval**: executes after explicit per-action human approval logged in an auditable trail. **Autonomous**: acts without per-action approval inside a documented and bounded scope with periodic human review. Tier assignments are reviewed at least annually and on any material capability change.*

Permitted data scopes by tier

> *Assist and Suggest agents may access any data the requesting user is authorised to access. Act-with-Approval agents are limited to non-regulated data unless additional authorisation from the Data Protection Officer and Compliance Lead is on file. Autonomous agents are limited to non-sensitive operational data and prohibited from acting on PHI, PII, payment data, or other regulated classes without a documented exception approved by the AI Risk Committee.*

Human-in-the-loop requirements by tier

> *Act-with-Approval agents require explicit per-action human approval recorded in an auditable log capturing approver identity, timestamp, action approved, and time spent on the decision. Autonomous agents are subject to monthly sampling review of at least five percent of agent actions, with exceptions investigated as security events.*

Prohibited tool combinations

> *Agents may not chain external-tool calls — outbound network calls, file system writes, or third-party API requests — with read access to PHI, PII, or other regulated data classes without explicit written authorisation from the Compliance Lead.*

Vendor onboarding gates

> *Any vendor whose product includes agentic AI capability — including agent-mode features embedded in licensed SaaS — must complete the AI Vendor Questionnaire and be reviewed against this policy before Procurement issues a signed order.*

An Agent Inventory and Risk Register Template

Below is a worked example for a hypothetical mid-market SaaS company at roughly 150 employees, with the agent mix we see most often in 2026. Treat it as a starting structure to fork into your own risk register.

Agent	Tier	Data Scope	Residual Risk	Control Set	Owner	Cadence
Microsoft Copilot for M365	Assist	Internal	Low	M365 tenant scope + DLP + sensitivity labels	IT Ops	Quarterly
Salesforce Einstein agents	Act-with-Approval	Confidential	Medium	Permission sets + approval gate + audit	Sales Ops	Quarterly
ServiceNow Now Assist	Suggest	Internal	Low	RBAC + change-management audit	IT Ops	Quarterly
Cursor / Cline	Suggest	Confidential (source code)	Medium	Repo permissions + secret scanning + mandatory PR review	Engineering	Monthly
Glean (enterprise search)	Assist	Confidential	Medium	Permission inheritance + access review + query audit	Security	Quarterly
Browser copilot (Arc, Comet)	Assist	Internal	Low	Browser policy + no enterprise SSO context + allowlist	IT Ops	Quarterly
Internal RAG with agentic retrieval	Act-with-Approval	Confidential	High	ACL inheritance + query audit + Lakera Guard	Security + Data Eng	Monthly
Finance reconciliation agent (month-end)	Autonomous	Regulated (financial)	High	Bounded transaction range + dual-write audit + SOX mapping	Finance	Quarterly

Forking instructions: copy this table into your risk register, replace each row with an agent your team is running, recalibrate residual risk to your data sensitivity and customer commitments, and replace the owner column with a named individual. The control-set column should evolve toward specific IDs linked to your SOC 2 evidence library or HITRUST control statements.

The Vendor Questionnaire You Will Be Asked (and Asked to Ask)

Two uses for this list. First, these are the questions your customers will start including in vendor security questionnaires this year. Second, these are the questions your team should be asking of your own agentic vendors — Microsoft, Salesforce, ServiceNow, Glean, Notion, Datadog's AI-assisted alerting, and any SaaS product that has shipped an agent-mode feature in its last release. Run the list both directions.

1. Do any of your systems include autonomous AI agents that take actions on customer data without per-action human approval?
2. How do you authenticate AI agents accessing customer data, and are agent identities distinct from human and service-account identities?
3. What is your agent identity model — service accounts, OAuth on behalf of users, or a dedicated agent IAM layer?
4. How do you scope, time-bound, and revoke agent permissions, and what is the time-to-revoke if an agent misbehaves?
5. What evaluation framework do you use to detect agent misbehaviour or unintended action chains — Garak, PyRIT, NeMo Guardrails, or equivalent?
6. Do you perform multi-agent emergent-behaviour testing as part of your evaluation cycle, and how often?
7. Have you mapped your AI systems to a named governance framework — NIST AI RMF, Singapore IMDA, ISO/IEC 42001, or EU AI Act?
8. How do you audit human reviewers for automation bias, and what measurements do you retain?
9. What evidence do you retain of agent decision logs and human override actions, and for how long?
10. Are agentic capabilities — and the sub-processors that power them — described in your DPA and sub-processor list?
11. How do you train end-users on agent behaviour, limitations, and the route to report unexpected agent activity?
12. Have you completed an external assessment of your agentic AI controls, and is the report available under NDA?

For deeper technical phrasing, our AI services division at TrustEdge maintains a fuller vendor diligence questionnaire in their engineering implementation guide. Build the answers before you are asked.

A 12-Month Implementation Roadmap

The blog companion laid out a 90-day tactical checklist. This roadmap is the longer horizon, broken out by quarter.

Quarter 1 (first 90 days)

- Stand up the AI Agent Inventory and populate every known agent — standalone tools and embedded agent-mode features alike.
- Update the AI Use Policy with the six clauses in Section 6 and route through your existing policy-approval cadence.
- For HIPAA-covered entities, refresh §164.308(a)(1) Risk Analysis and §164.308(a)(8) Evaluation to bring agentic systems into scope.
- Pre-brief your SOC 2 auditor on the agent-class control set — Datadog, PagerDuty, AWS GuardDuty signal flows extended to agent identities. See our 90-day tactical checklist.

Quarter 2

- Run a vendor questionnaire campaign — issue Section 8 to every agentic vendor in your stack.
- Pilot an IAM-layer agent identity for one production agent; document time-to-revoke.
- Baseline an automation-bias measurement programme — capture reviewer time-on-decision and modify-versus-accept rates.
- Conduct the first internal multi-agent emergent-behaviour test using Garak, PyRIT, or equivalent.

Quarter 3

- Bring agentic systems into HITRUST scope before the next interim — speak with your authorised external assessor about a mid-cycle update.
- Close SOC 2 evidence gaps from the mid-year readiness check, particularly around agent identity and human oversight.
- Run an ISO/IEC 42001 readiness gap analysis.
- Standardise the agent-class control set across business units so finance, engineering, sales ops, and support work against one template.

Quarter 4

- Run the annual AI Use Policy review with lessons learned from the first three quarters baked in.
- Complete the vendor recertification cycle with the agentic questionnaire as a standing section.
- If pursuing certification, run an ISO/IEC 42001 internal audit pilot.
- Deliver board-level reporting on AI governance maturity — agents inventoried, residual-risk distribution, exceptions investigated.

How Jacobian Helps + Where to Go Next

Jacobian Engineering has been helping growing SaaS, healthcare-technology, and regulated mid-market organisations stand up and evidence compliance programmes since 2005. We are an employee-owned, HITRUST Certified Assessor, and our compliance assessment and managed-services teams know what your examiners and your enterprise customers are about to ask for. Our AI Model Risk Management practice pairs with our core SOC 2, HIPAA, and HITRUST programmes.

Book a Free Assessment and we will spend an hour with your team — no slides, no pitch — walking your current programme against the framework and the roadmap above. You can reach us at (415) 644-8208 or through the contact form.

For the engineering side, our AI services division at TrustEdge has the companion material at the IMDA Regulations Hub and the Building Agentic AI Systems Aligned With the IMDA Framework blog. The two sides reinforce each other.