AI Risk Gradient for SMEs and SaaS-First Organisations

Recognising Your True Tier

Supplemental Guide to the AI Infrastructure Security Playbook

---

Aligned to ETSI EN 304 223 / UK Code of Practice for AI Cyber Security

V1.0 — February 2026

Risk-tiered · Capability-based · Exposure-driven

Contents

Purpose and Audience

This supplemental guide extends the AI Infrastructure Security Playbook for organisations that primarily consume AI through SaaS platforms rather than building or hosting AI infrastructure. It addresses a critical blind spot: SMEs and SaaS-first organisations are climbing the AI risk gradient involuntarily, enabling capabilities that demand Advanced-tier controls while believing they remain at Foundational level.

The AI Infrastructure Security Playbook defines an eight-tier risk gradient (L1–L8) spanning four assurance zones: Foundational, Standard, Advanced, and High Assurance. That gradient is universal — it applies regardless of whether the AI runs on infrastructure you own or inside a vendor's tenancy. What changes is the implementation context: instead of configuring VPCs and IAM policies on cloud infrastructure, the controls become tenant configuration, governance policies, and vendor due diligence.

This guide is for CISOs, IT managers, compliance officers, and managed security providers serving organisations that rely on enterprise SaaS platforms — Microsoft 365, Google Workspace, Atlassian Cloud, Salesforce, and similar — and are adopting AI capabilities embedded within those platforms.

The Involuntary Climb

Most SMEs do not consciously decide which risk tier they occupy. AI adoption in SaaS-first organisations typically follows three phases, each of which silently escalates the organisation's position on the risk gradient.

Phase 1 — AI Consumption (L1–L2: Foundational)

Employees use vendor-embedded AI features through existing interfaces: CoPilot in Microsoft 365, Gemini in Google Workspace, Atlassian Intelligence, Einstein in Salesforce. They are consumers of AI. They do not configure models, select endpoints, or manage pipelines. The vendor controls the entire AI stack.

The risk at this phase is data exposure through amplification. AI does not create new data access — it surfaces what existing permissions already allow, but faster, more comprehensively, and through natural language queries that bypass the friction that previously protected poorly permissioned content. A user who would never have navigated twelve levels deep in SharePoint to find board papers can now ask CoPilot a question and receive them in seconds.

Controls required: Tenancy hardening. This is the prerequisite for switching on any AI feature safely. The AI is only as secure as the foundation it reads from.

Phase 2 — AI Orchestration (L3–L4: Standard)

Someone in the organisation builds a Power Automate flow with AI Builder, publishes a Copilot Studio agent, configures a Salesforce Einstein Bot, or wires an Atlassian Automation rule with AI actions. They have become orchestrators of AI. They are chaining tools, granting permissions to automations, and creating persistent processes that act on their behalf — often unattended, outside business hours, using their identity.

The risk at this phase is action exposure. The automation doesn't just read data — it does things. It can send emails, update CRM records, write to SharePoint, post Teams messages, and call external APIs via HTTP connectors. The blast radius is everything the user's identity can access, and the automation persists beyond their session. Most SMEs have never configured Power Platform environment policies, connector restrictions, or data loss prevention rules for low-code platforms.

Controls required: Governed automation. Environment isolation, connector DLP, approval workflows, connection lifecycle management.

Phase 3 — AI-Assisted Development (L5–L6: Advanced)

An employee uses Cursor, Replit, Bolt, Lovable, or similar AI coding tools to build a customer-facing application. It may handle PII, connect to production databases, authenticate users, and integrate with company APIs. It gets deployed on a free-tier hosting platform. Nobody in the organisation reviewed the code, scanned the dependencies, assessed the architecture, or knows it exists.

The risk at this phase is deployment exposure. This is shadow IT at a fundamentally different scale — not just using unauthorised SaaS, but building unauthorised software. Hardcoded API keys, insecure authentication, unpatched dependencies, abandoned databases, no logging, no

monitoring, no incident response. The code works, but the security posture is zero. When the person who built it leaves the organisation, nobody can maintain, patch, or decommission it.

Controls required: Controlled creation. Application registry, mandatory security scanning, deployment governance, supply chain verification, ownership lifecycle, decommissioning process.

Recognising Your True Tier

The following diagnostic helps organisations identify where they actually sit on the risk gradient. If any condition in a higher tier is true, the organisation has reached that tier — regardless of whether the corresponding controls are in place.

Risk Tier	You are here if…	But you need…
L1–L2 Foundational	CoPilot, Gemini, or Einstein features are enabled. Employees use AI assistants through vendor UIs. No custom automations or integrations.	MFA enforced everywhere. Conditional access policies. SharePoint/Drive permissions audited for least privilege. Sensitivity labels on confidential content. DLP policies for sensitive content types. Encryption at rest verified.
L3–L4 Standard	Anyone has built a Power Automate flow, Copilot Studio agent, Salesforce Flow, or Atlassian Automation rule that uses AI connectors. Any automation sends data between applications or calls external APIs.	Power Platform environment strategy (default locked, managed for approved flows). Connector DLP policies blocking HTTP/SQL/custom connectors in unmanaged environments. Approval workflow for automations touching data outside originating app. Connection inventory and orphan process. Audit logging for all automation activity.
L5–L6 Advanced	Anyone has used AI coding tools (Cursor, Replit, Bolt, Lovable, ChatGPT) to build an application that handles company data, connects to company APIs, or serves customers. Any AI-generated code is deployed on any platform.	Citizen-built application registry with ownership. Mandatory security scanning before deployment. Acceptable use policy for AI-assisted development. Dependency and supply chain audit. Deployment gates for anything touching PII or production systems. Decommissioning and lifecycle management.
L7–L8 High Assurance	The organisation operates in a regulated sector (financial services, healthcare, legal) AND any of the L5–L6 conditions are true. AI systems make or influence decisions with regulatory, financial, or clinical consequences.	All L5–L6 controls plus: formal change management for all AI deployments, regulatory compliance mapping, model output validation, third-party audit trail, data sovereignty verification. Consider engaging the full AI Infrastructure Security Playbook.

Key principle: Your risk tier is determined by the highest-risk activity occurring anywhere in the organisation, not by the activity you intended to authorise. If one employee has built a Power Automate flow with an HTTP connector, the organisation is at L3–L4 — even if the CISO has never heard of Power Platform. Tiers are cumulative: controls from lower tiers remain mandatory at higher tiers.

L1–L2 Foundational: Tenancy Hardening

These controls are prerequisites for safely enabling any AI feature in a SaaS environment. They should be in place regardless of AI adoption because they address fundamental tenancy security — but AI makes their absence critically dangerous.

Identity and Access [P4, P5]

- Enforce MFA for all users with no exceptions. Conditional access policies requiring compliant devices for sensitive applications.
- Audit and remediate SharePoint, OneDrive, Google Drive, and Confluence permissions to enforce least privilege. Remove inherited permissions that grant organisation-wide read access to sensitive libraries.
- Implement Just-in-Time access for administrative roles. Remove standing global admin assignments.
- Configure session timeout policies appropriate to data sensitivity.
- Disable legacy authentication protocols and basic authentication where still enabled. In SME tenancies, legacy auth is common and AI amplification of compromised credentials makes it critically dangerous.

Data Classification and Protection [P5, P8]

- Deploy sensitivity labels (Microsoft Purview, Google DLP, or equivalent) for confidential, internal, and public content. AI assistants honour sensitivity labels — but only if they exist.
- Configure DLP policies for sensitive content types: financial data, PII, health records, legal privileged material. Block or warn on external sharing of labelled content.
- Review and restrict anonymous link sharing across SharePoint and Drive repositories. Anonymous links bypass all classification and DLP controls and represent a major AI amplification vector.
- Verify encryption at rest is active across all repositories. Microsoft 365 enables this by default but verify configuration has not been modified. Google Workspace Client-Side Encryption for high-sensitivity content.

Monitoring and Audit [P8, P12]

- Enable unified audit logging. Microsoft 365 audit logs, Google Workspace audit logs, Atlassian audit log, Salesforce event monitoring.
- Configure log retention to meet regulatory and organisational requirements (minimum 90 days, 365 days recommended).
- If using a managed SOC or MSSP: verify AI-related telemetry is included in monitoring scope (see Section 8).
- Automation execution logs must be retained for forensic analysis in the event of suspected data exfiltration or misuse. Align retention periods with incident response requirements.

Vendor Due Diligence [P3, P7]

- Review vendor AI data handling policies: does the vendor use your data to train models? Microsoft, Google, and Salesforce have published commitments — verify them for your specific licence tier.
- Confirm data residency: where is AI processing performed? Does it match your regulatory requirements?
- Map shared responsibility boundaries: understand exactly what the vendor secures versus what you must configure.

L3–L4 Standard: Governed Automation

Once anyone in the organisation has created an AI-powered automation, these controls become mandatory. The gap between “using CoPilot” and “building Power Automate flows” is where most SMEs lose control — because the platform makes it effortless to cross that boundary with no governance in place.

Environment Strategy [P2, P4]

- Configure Power Platform environment isolation: a default environment restricted to standard connectors only, and a managed environment for approved automations requiring premium or custom connectors.
- Equivalent controls for other platforms: Salesforce sandbox strategy, Atlassian project-level permission schemes for automation rules.
- Block self-service environment creation. All new environments require IT approval.

Connector and Integration DLP [P2, P5, P6]

- Create Power Platform Data Loss Prevention policies that classify connectors into Business, Non-Business, and Blocked categories.
- Block HTTP, SQL, Custom Connector, and Azure Functions connectors in unmanaged environments. These are the connectors that allow data egress to arbitrary external endpoints.
- Restrict cross-tenant sharing connectors unless explicitly approved.
- Equivalent controls: Salesforce Connected Apps policies, Google Workspace Marketplace app whitelisting, Atlassian Forge app permissions.

Approval and Change Management [P3, P9]

- Require approval for any automation that: moves data between applications, calls external APIs, sends communications on behalf of users, or accesses content labelled confidential or above.
- Maintain an automation inventory: what flows exist, who owns them, what connectors they use, what data they access, when they were last reviewed.
- Implement a review cycle (quarterly minimum) for all active automations.

Connection and Identity Lifecycle [P4, P5, P13]

- Audit connection objects (stored credentials used by automations). Identify connections using individual user credentials versus service accounts.
- Implement an offboarding process that identifies and reassigns or disables automations owned by departing staff. Orphaned flows running under deleted-user credentials are a critical risk.
- Service account governance: dedicated service accounts for production automations with least-privilege permissions, MFA exemption only via conditional access policy with compensating controls.

- Automations must not operate using identities with privilege escalation capability or broad administrative scope. This is a global invariant: no automation should be able to elevate its own permissions at runtime.

L5–L6 Advanced: Controlled Creation

When employees use AI coding tools to build applications, the organisation has entered shadow development. Traditional shadow IT policies assumed employees might sign up for unauthorised SaaS — they did not anticipate employees building software. The controls at this tier address a fundamentally different threat model.

Acceptable Use Policy for AI-Assisted Development [P3]

- Establish a clear policy covering: which AI coding tools are permitted (if any), what types of applications may be built, what data categories may be processed, and what deployment platforms are acceptable.
- Distinguish between exploration (prototyping on synthetic data) and production (handling real data or serving real users). Exploration may be encouraged; production requires governance.
- Communicate the policy as enablement, not prohibition: “Here’s how to build things safely” rather than “Don’t build things.”

Application Registry and Ownership [P3, P9, P13]

- Maintain a registry of all citizen-built applications: name, owner, purpose, data categories processed, hosting platform, dependencies, last review date.
- Assign a technical steward (may be the MSSP) responsible for periodic review.
- Define lifecycle expectations: every registered application must have a review date, an owner, and a decommissioning plan.

Security Scanning and Supply Chain [P7, P9]

- Require security scanning before any citizen-built application handles company data or serves external users. This can be as simple as running a SAST tool or having the MSSP review the deployment.
- AI-generated code routinely includes outdated or vulnerable dependencies. Dependency scanning is non-negotiable.
- Check for hardcoded credentials, API keys, and connection strings — the most common vulnerability in AI-generated code.

Deployment Governance [P2, P4]

- If it handles PII, connects to production systems, or serves external users: it goes through change management, regardless of who built it or how.
- Restrict deployment targets: free-tier hosting platforms with no security controls, no logging, and no SLA are not acceptable for production workloads.
- Require authentication and access controls on any externally-facing application.
- All externally accessible citizen-built applications must implement request logging and error logging sufficient to support incident investigation. Without this, disaster recovery and incident response are impossible.

Decommissioning [P13, P6]

- When the owner leaves: automated alert to the technical steward, 30-day window to reassign or decommission, automatic suspension after 30 days with no new owner.
- Decommissioning procedure: revoke API keys and credentials, delete or archive data, remove deployment, update registry.
- Before decommissioning, assess whether retained data must be archived to meet regulatory retention obligations. Premature deletion may constitute a compliance breach.
- Key material destruction: revoke any service account credentials, API tokens, or OAuth grants associated with the application.

ETSI EN 304 223 Provision Mapping

The following table maps each ETSI EN 304 223 provision to the SaaS-context controls described in this guide. Provision numbering aligns to the AI Infrastructure Security Playbook cross-cloud alignment table.

Ref	Provision	SaaS Context	Applicable Tier
P2	Network Security	Environment isolation, connector DLP, deployment target restrictions	L3–L4 (connector policies), L5–L6 (deployment governance)
P3	Risk Management	Tier recognition diagnostic, vendor due diligence, acceptable use policy	All tiers
P4	Access Control	Conditional access, MFA, environment RBAC, deployment permissions	L1–L2 (MFA/CA), L3–L4 (environment RBAC), L5–L6 (deployment gates)
P5	Data Management	Sensitivity labels, DLP policies, permissions audit, information barriers	L1–L2 (classification), L3–L4 (connector DLP, connection governance)
P6	Crypto & Key Mgmt	Encryption at rest verification, credential management, key destruction on decommission	L1–L2 (encryption verification), L5–L6 (credential/key lifecycle)
P7	Supply Chain	Vendor AI data policies, dependency scanning for AI-generated code	L1–L2 (vendor diligence), L5–L6 (dependency audit)
P8	Monitoring & Logging	Unified audit logs, AI telemetry in SOC scope, automation activity monitoring	L1–L2 (audit logging), L3–L4 (automation monitoring), L5–L6 (application logging)
P9	Integrity Verification	Automation approval workflows, security scanning for citizen-built applications	L3–L4 (approval gates), L5–L6 (scanning)
P11	Incident Management	Managed SOC / MSSP with AI-aware detection rules and response playbooks, automation kill procedures	All tiers (SOC engagement scales with tier)
P12	Compliance & Audit	Log retention, automation inventory review, regulatory mapping	All tiers
P13	Decommissioning	Orphaned flow remediation, application lifecycle, offboarding procedures, credential revocation	L3–L4 (flow lifecycle), L5–L6 (application decommissioning)

Questions for Your Managed SOC / MSSP

Most SMEs do not operate their own security operations centre. They rely on a managed SOC or MSSP — which means security monitoring is only as good as the scope of the engagement. The following questions help organisations assess whether their MSSP is equipped to support AI-era risks.

CoPilot and Embedded AI Monitoring

- Are Microsoft 365 CoPilot audit logs (or equivalent) included in your monitoring scope?
- Do your detection rules cover unusual AI query patterns — for example, bulk data summarisation requests, queries targeting sensitivity-labelled content, or AI-assisted search across mailboxes?
- Can you alert on CoPilot access to content that the user has never previously accessed manually?

Citizen Developer and Automation Monitoring

- Are Power Platform / Salesforce Flow / Atlassian Automation audit logs ingested?
- Do you alert on new connector registrations, particularly HTTP, SQL, and custom connectors?
- Do your playbooks cover scenarios where an automation exfiltrates data through an external API call?
- Can you detect orphaned automations running under disabled or deleted user accounts?

AI-Assisted Data Exfiltration

- Do your DLP monitoring rules account for AI-mediated data movement — for example, CoPilot summarising confidential documents into an email or Teams message that then leaves the organisation?
- Are sensitivity label violation events triggering alerts in your SOC?
- Can you detect bulk AI-assisted summarisation of sensitivity-labelled content? This is a primary amplification-based exfiltration vector.

Vibe-Coded Application Monitoring

- If the organisation maintains a citizen-built application registry, can you incorporate those endpoints into monitoring?
- Do you offer lightweight application security review services for citizen-built deployments?
- Can you detect and alert on credential exposure (API keys, tokens) in public repositories or deployment logs?

Implementation Roadmap

The following phased approach allows SMEs to systematically address AI security without overwhelming limited resources. Each phase builds on the previous one and should be completed before progressing.

Phase 1: Discover (Weeks 1–2)

- Run the tier recognition diagnostic from Section 3 to determine your actual risk tier.
- Inventory existing AI features enabled across all SaaS platforms.
- Inventory Power Automate flows, Copilot Studio agents, and equivalent automations across all platforms.
- Survey staff to identify any AI-assisted development or vibe-coded applications.
- Review current MSSP scope against the questions in Section 8.

Phase 2: Harden (Weeks 3–6)

- Implement L1–L2 Foundational controls: MFA, conditional access, permissions audit, sensitivity labels, DLP policies.
- Verify encryption configuration across all platforms.
- Enable and configure audit logging with appropriate retention.
- Engage MSSP to expand scope for AI telemetry if gaps identified.

Phase 3: Govern (Weeks 7–10)

- If L3–L4 activity identified: implement Power Platform environment strategy, connector DLP, approval workflows.
- Establish automation inventory and review cycle.
- Implement offboarding process for automation and connection ownership transfer.
- Communicate governance policies to citizen developers as enablement guidance.

Phase 4: Control (Weeks 11–16)

- If L5–L6 activity identified: establish acceptable use policy, application registry, security scanning requirements.
- Define deployment governance for citizen-built applications.
- Implement decommissioning process for applications and automations.
- Schedule first quarterly review of all registered automations and applications.
- Schedule annual reassessment of risk tier using the diagnostic in Section 3. Organisations move up the gradient as adoption expands; controls must scale accordingly.

Vendor Security Documentation Signposts

This guide provides the governance framework and decision structure. The following vendor documentation provides the implementation detail for specific platforms. These links were verified at time of publication; check vendor sites for current versions.

Microsoft 365 and CoPilot

- Microsoft CoPilot for Microsoft 365 security documentation and data privacy commitments
- Microsoft Purview sensitivity labels and information protection
- Power Platform environment and connector DLP administration
- Microsoft Defender for Cloud Apps configuration guide
- Microsoft 365 audit log reference

Google Workspace

- Google Workspace Gemini Enterprise data handling and privacy
- Google Workspace Client-Side Encryption administration
- Google Workspace DLP configuration
- Google Workspace audit and investigation tool

Salesforce

- Salesforce Einstein Trust Layer documentation
- Salesforce Shield encryption and event monitoring
- Salesforce Flow and automation governance

Atlassian

- Atlassian Intelligence trust and data handling
- Atlassian Cloud organisation security policies
- Atlassian Access (SSO, SCIM provisioning, audit log)

This supplemental guide is maintained alongside the AI Infrastructure Security Playbook. The risk gradient, ETSI EN 304 223 provision mappings, and cross-cloud control alignment tables in the main playbook provide the full technical reference. This guide translates those controls for the SaaS tenant configuration context that SMEs and SaaS-first organisations operate within.