RESOURCE · PROCUREMENT · RES-001

AI Vendor Security Questionnaire

VERSION

v1.0

TRACKED AT

agentmodeai.com/holding/RES-001/

SOURCE

agentmodeai.com/resources/

LICENCE

CC BY 4.0

Agent Mode AI publishes editorial work for senior IT leaders. Every claim including the premise of this template is tracked on a public ledger. Verdict updates and corrections at agentmodeai.com/holding/.

Version 1.0 · Published 4 May 2026 · 90-day review cycle

Forty-seven questions across seven sections, each tied to an AI-specific failure mode the standard cloud-procurement questionnaires (CAIQ v4, SIG) do not surface. Use as an addendum to existing procurement frameworks, not a replacement.

Source explainer: https://agentmodeai.com/resources/ai-vendor-security-questionnaire/ Holding-up tracking: https://agentmodeai.com/holding/RES-001/

How to use

- Send to the vendor BEFORE the proof-of-concept, not after.
- Score each section binary: answered with evidence, or unanswered.
- Plausible prose without documentation/audit log sample/contract reference scores as unanswered.
- The completed pack becomes an artifact in your vendor risk register.

Section 1 — Model lineage and provenance

#	Question	Evidence required
1	Name the foundation model(s) underpinning this product, including version numbers.	Model card URL
2	Disclose any fine-tuning, distillation, or post-training applied to the foundation model.	Dataset description
3	State whether the deployed model can change without customer notification, and the SLA window for notification when it does.	Written SLA commitment
4	Provide the model card or equivalent technical documentation conforming to NIST AI RMF Map function 1.1.	Document
5	Confirm whether model outputs are deterministic given identical inputs and parameters, and if not, state the temperature, top-p, and seed defaults.	Configuration export

Section 2 — Training data and inference data handling

#	Question	Evidence required
1	State whether your model was trained or fine-tuned on data that includes personal information about EU, UK, or California residents. If yes, provide the lawful basis under GDPR Article 6 or CPRA equivalent.	Legal basis statement
2	Confirm in writing that customer prompts, completions, and tool-use traces from this contract are NOT used to train, fine-tune, evaluate, or improve any model — yours or a third party's.	Contract clause reference
3	State the data retention window for prompts, completions, embeddings, and intermediate agent state. Provide deletion-on-request SLA.	Written SLA
4	Disclose every geographic region in which inference data is processed.	Region list
5	Disclose every third-party sub-processor that receives prompt or completion data.	Sub-processor list URL
6	State whether prompt data crosses any jurisdictional boundary that would trigger GDPR Article 44–49 transfer requirements.	Transfer mechanism (SCC/adequacy)
7	Confirm whether model providers (Anthropic, OpenAI, etc.) receive customer prompt data under your enterprise contract with them, and the retention applied at that layer.	Upstream contract terms
8	Provide the technical mechanism that enforces non-training of customer data — contract language alone is insufficient.	API parameter / endpoint configuration

Section 3 — Identity, authentication, and authorisation

#	Question	Evidence required
1	Describe the identity model used by agents this product deploys. Name the protocol (OAuth 2.0 client-credentials, mTLS, SPIFFE, etc.).	Architecture document
2	Confirm that agent identities are distinct from any human user identity and cannot be impersonated through credential reuse.	Identity isolation diagram
3	Provide the access-scoping mechanism — how a procurement-team agent is prevented from calling finance APIs.	RBAC/ABAC policy export
4	State the agent-credential rotation SLA and the procedure for emergency credential revocation.	Written runbook
5	Disclose every external system this product's agents will call as part of standard operation (databases, SaaS APIs, MCP servers, internal tools).	Inventory
6	Confirm whether agent-to-agent authentication uses signed tokens or relies on network-position trust.	Architecture document

Section 4 — Audit, observability, and explainability

#	Question	Evidence required
1	Provide a sample of the audit log produced for one agent invocation, including prompt, completion, tool calls, and timing.	Log sample (JSON Lines)
2	State the audit log retention window and the export format (JSON Lines, OpenTelemetry, CEF).	Written commitment
3	Confirm whether logs include the model version that produced each completion and the parameters used.	Log schema
4	State whether tool-use calls (functions, MCP tools, retrieval queries) are logged with arguments and results.	Log schema
5	Provide the SLA for customer-initiated log export, including format and maximum delay.	Written SLA
6	Disclose whether the vendor retains logs after customer deletion, and the basis for any retention.	Retention policy
7	State whether the vendor produces explainability artifacts (decision rationale, retrieval attribution) on request, and the SLA.	Sample artifact
8	Confirm whether logs are immutable post-write and the cryptographic basis for that immutability.	Architecture document

Section 5 — Incident response and kill-switch

#	Question	Evidence required
1	Define the kill-switch SLA — maximum elapsed time from customer notification to all agents halted.	Written SLA
2	Identify by role every party who can invoke the kill-switch. Confirm at least one customer-side party has authority.	Runbook
3	Describe the rollback mechanism — how an agent decision (database write, email sent, payment authorised) is reversed when discovered to be in error.	Architecture document
4	Provide the incident-notification SLA for security incidents, model malfunctions, and unintended autonomous actions.	Written SLA
5	Confirm whether the product supports rate-limiting and budget-capping at the agent level, with customer-configurable thresholds.	Configuration export
6	Describe the procedure for preserving forensic state when an incident is detected.	Runbook

Section 6 — EU AI Act, GDPR, and regulatory posture

#	Question	Evidence required
1	State whether the vendor classifies this product as a high-risk AI system under EU AI Act Annex III. If yes, provide the risk classification rationale.	Written classification
2	Provide the EU AI Act conformity assessment artifacts — Annex IV technical documentation, declaration of conformity, CE marking where applicable.	Documents
3	Confirm whether the vendor will supply the deployer documentation a customer needs to satisfy Article 26 obligations.	Sample documentation
4	Provide the GDPR Data Protection Impact Assessment (DPIA) the vendor has conducted, and confirm cooperation with customer DPIAs under Article 35.	DPIA document
5	Disclose whether the product or any deployment configuration triggers German BetrVG §87(1) point 6 co-determination, Dutch WOR Article 27 consent, French CSE consultation, or equivalent works-council requirements in EU jurisdictions.	Jurisdictional analysis
6	Provide the named data protection officer (DPO) under GDPR Article 37, with contact information.	DPO contact card
7	State the vendor's position on AI training-data copyright and the indemnification provided to customers if model outputs reproduce copyrighted material.	Contract clause
8	Disclose any pending regulatory enforcement action, supervisory authority investigation, or class-action litigation involving the product or model.	Written disclosure

Section 7 — Contract, indemnification, and exit

#	Question	Evidence required
1	Provide the standard MSA + AI-specific addendum + DPA. Identify every clause changed in the last 18 months and the rationale.	Documents + redline summary
2	State the indemnification position on hallucinated outputs that cause customer harm — financial loss, reputational damage, regulatory penalty.	Contract clause
3	Confirm whether the vendor indemnifies the customer against IP infringement claims arising from model outputs.	Contract clause
4	Provide the data-portability mechanism on contract termination — export format, retention window for customer data, deletion certification.	Written procedure
5	State the model-deprecation notice period and the supported migration path when a foundation model the product depends on is sunset.	Written SLA
6	Disclose whether the vendor reserves the right to suspend service for usage patterns it deems problematic, and the appeal mechanism.	Contract clause

Scoring rubric

For each section, calculate `answered_with_evidence / total`. Section is **passing** at ≥80%. Vendor is **deployment-ready** when 6 of 7 sections pass. Vendor scoring 5/7 or below is approved for low-risk use cases only, with the gap analysis as the basis for the next vendor review.

Versioning

- v1.0 — 4 May 2026 — initial publication.
- Next review: 2 August 2026 (post-EU AI Act enforcement window opens).

Frameworks referenced

- Cloud Security Alliance CAIQ v4 (October 2024)
- Shared Assessments SIG (2025 release)
- NIST AI Risk Management Framework AI RMF 1.0
- ISO/IEC 42001:2023 (AI management systems)
- EU AI Act Regulation 2024/1689, Annexes III and IV
- GDPR Articles 6, 22, 26, 35, 37, 44–49
- ENISA cybersecurity guidance for AI

Licence