PORTFOLIO PROJECT 3 OF 3

Third-Party AI Vendor Risk Due Diligence Toolkit

A Reusable Questionnaire, Scoring Rubric, and Worked Example

Prepared by J Damien Scott, SRMP GRCP IAIP

801-392-9999 | jdamienscott@pm.me | jdamienscottllc.com

July 2026

Portfolio note. Sections 2 and 3 of this toolkit are reusable and vendor-agnostic. Section 4 applies them to a fictional vendor, Cascade AI Concierge, invented to demonstrate the toolkit in use. No real vendor, contract, or customer relationship is described.

Table of Contents

1. Purpose and When to Use This Toolkit

Most organizations already run vendor risk assessments. What changes when a vendor embeds a third-party large language model (LLM) is not the existence of a review process but its content: a vendor questionnaire built for traditional software, focused on uptime, data encryption, and access control, will miss the questions that matter for an AI-enabled product, such as which foundation model the vendor uses, whether customer data trains that model, and what happens when the vendor's upstream model provider changes its behavior without notice.

This toolkit extends, rather than replaces, standard third-party risk management practice. Use it whenever procurement or a business unit proposes a SaaS vendor whose product description mentions generative AI, an AI assistant, or an embedded LLM, and pair it with whatever general vendor security questionnaire the organization already runs (an SIG-lite or equivalent). Section 2 is the AI-specific supplement; Section 3 is the scoring method; Section 4 is a complete worked example.

2. Due Diligence Questionnaire

The questionnaire is organized into six domains. Each question includes a short note on why it matters, so a reviewer without deep AI background can still ask it credibly and recognize a weak answer.

2.1 Data Handling and Privacy

#	Question	Why It Matters
D1	Does customer data submitted to the AI feature train or fine-tune any model, including the vendor's own or an upstream provider's?	If yes, the organization needs a contractual guarantee and an opt-out; if the vendor cannot answer clearly, treat that as a finding, not a pass.
D2	What is the data retention period for prompts and outputs, and can the organization enforce a shorter one?	Retention beyond business need is itself a disclosure risk regardless of the AI feature specifically.
D3	Does the vendor process the organization's data outside jurisdictions the organization has approved?	Cross-border processing may trigger separate regulatory obligations independent of the AI question.

2.2 Model Provenance and Supply Chain

#	Question	Why It Matters
M1	Which foundation model or models power the feature, and is the vendor a reseller of a third-party model provider?	Determines who is actually accountable for model behavior, and whether a second tier of due diligence on the upstream provider is needed.

#	Question	Why It Matters
M2	How does the vendor notify customers of a material model change (version upgrade, provider switch, deprecation)?	An undisclosed model change is functionally equivalent to an unreviewed software update; the organization needs advance notice to re-test.
M3	What subprocessors, including model providers and hosting infrastructure, does the vendor disclose?	Mirrors standard subprocessor disclosure practice, extended to the AI supply chain specifically (OWASP LLM03 Supply Chain).

2.3 Security Controls and Testing

#	Question	Why It Matters
S1	Does the vendor test for prompt injection and excessive agency before release, and can it describe the method?	A vendor that cannot describe its testing approach at all is a stronger signal than a vendor whose testing is merely informal.
S2	Are AI feature actions that change customer data (refunds, account changes) gated behind explicit confirmation?	Directly checks for the excessive-agency pattern named in OWASP LLM06.
S3	What rate and cost controls exist to prevent runaway usage from a single account?	Checks for the unbounded-consumption pattern named in OWASP LLM10, and protects the organization from unexpected pass-through cost.

2.4 Certifications and Compliance

#	Question	Why It Matters
C1	Does the vendor hold SOC 2 Type II, ISO/IEC 27001, or ISO/IEC 42001 certification, and can it provide the current report?	ISO/IEC 42001 specifically evidences an AI management system, not just general information security (International Organization for Standardization, 2023).
C2	Has the vendor assessed its product against the EU AI Act, if the organization operates in the European Union?	Confirms the vendor has done its own risk classification rather than leaving it to the customer to discover.

2.5 Incident History and Notification

#	Question	Why It Matters
I1	Has the vendor had a security incident involving the AI feature specifically, and how was it disclosed?	A vendor with no incident history may simply lack detection, not lack incidents; ask how they would know.
I2	What is the contractual notification window for a future incident involving the AI feature?	Should match or beat the organization's own regulatory notification obligations, not the vendor's general incident clause alone.

2.6 Contractual and Governance Terms

#	Question	Why It Matters
G1	Does the vendor accept audit rights specific to the AI feature, not just general information security audit rights?	General audit clauses often predate the AI feature and may not cover it explicitly.
G2	Who is contractually liable if the AI feature produces an output that causes customer harm?	Liability allocation for AI-generated output is still an evolving area and should be addressed explicitly, not assumed.

3. Scoring Rubric

Score each of the six domains from 1 (unacceptable: no meaningful control or answer) to 5 (strong: documented, tested, and contractually backed). Weight domains by consequence rather than treating all six equally; the weights below reflect that a data-handling failure or a security-control gap is typically more consequential than a governance-documentation gap, though a reviewer should adjust weights to the organization's own risk appetite.

Domain	Weight	Score (1–5)	Weighted Score
Data Handling and Privacy	25%	—	—
Model Provenance and Supply Chain	20%	—	—
Security Controls and Testing	25%	—	—
Certifications and Compliance	10%	—	—
Incident History and Notification	10%	—	—
Contractual and Governance Terms	10%	—	—

Risk tier: Green (weighted total 4.0–5.0, approve under standard delegation), Yellow (2.5–3.9, approve with named remediation conditions and a follow-up review date), Red (below 2.5, escalate to full governance committee review before any contract signature). This three-tier structure mirrors the triage approach standard in vendor and contract risk workflows generally, not an AI-specific invention.

4. Worked Example: Cascade AI Concierge (Fictional Vendor)

Cascade AI Concierge is a fictional SaaS customer-engagement platform that Meridian Grid Energy Services (introduced in Portfolio Projects 1 and 2) is considering for a live-chat overflow queue. Cascade embeds a third-party foundation model behind its own orchestration layer. The completed assessment below shows the toolkit in use.

Domain	Finding	Score
Data Handling and Privacy	Cascade confirmed customer data does not train the underlying model and provided a 30-day default retention window, adjustable on request. Data is processed only in-region.	4

Domain	Finding	Score
Model Provenance and Supply Chain	Cascade discloses its upstream model provider and offers 30 days' notice of a material model version change, but could not commit to advance notice of a full provider switch.	3
Security Controls and Testing	Cascade described a documented prompt-injection test suite run pre-release and confirmed all account-changing actions require explicit user confirmation. Rate and cost caps are configurable per customer.	4
Certifications and Compliance	Cascade holds SOC 2 Type II; ISO/IEC 42001 certification is in progress but not yet issued.	3
Incident History and Notification	No disclosed AI-specific incidents to date; contractual notification window is 72 hours, which meets Meridian's internal standard.	4
Contractual and Governance Terms	Standard audit-rights clause does not yet name the AI feature specifically; Cascade agreed to add specific language during redline.	3

Weighted total: (4 × 0.25) + (3 × 0.20) + (4 × 0.25) + (3 × 0.10) + (4 × 0.10) + (3 × 0.10) = 3.7. This places Cascade in the Yellow tier: approve with named remediation conditions, specifically, AI-feature-specific audit-rights language added at contract redline and a commitment to notify Meridian of any full model-provider switch, with a follow-up review scheduled for the ISO/IEC 42001 certification once issued.

5. Recommended Contract Clauses

The following clause topics translate the questionnaire's findings into contract language procurement or legal can negotiate directly. Exact language should be reviewed by counsel; these are the substantive points to cover, not a final draft.

- Data processing addendum (DPA) that explicitly names the AI feature and confirms customer data is not used for model training absent separate written consent.
- Model-change notification clause requiring advance written notice of any material model version change or upstream provider switch, with a defined notice period (Cascade example: 30 days).
- Audit rights extended explicitly to the AI feature, including the right to review test results for prompt-injection and excessive-agency controls, not only general information-security audit rights.
- Incident notification clause with a window that matches or beats the customer organization's own regulatory notification obligations.
- Subprocessor disclosure and update clause covering the model provider and any AI-specific infrastructure vendor, with advance notice of any change.

6. Presenting This Project in an Interview

This project demonstrates the ability to extend an existing GRC discipline, vendor and third-party risk management, into a new technology area without discarding the underlying method. In an interview, walk through the Cascade example's weighted score and explain why the Yellow tier, not a flat approve or reject, is the more defensible outcome. That judgment, tying a numeric score to a specific remediation condition, is the skill a governance or risk-engineering role is actually hiring for.

References

International Organization for Standardization. (2023). ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system.

OWASP Foundation. (2025). OWASP Top 10 for LLM Applications 2025. https://genai.owasp.org/llm-top-10/