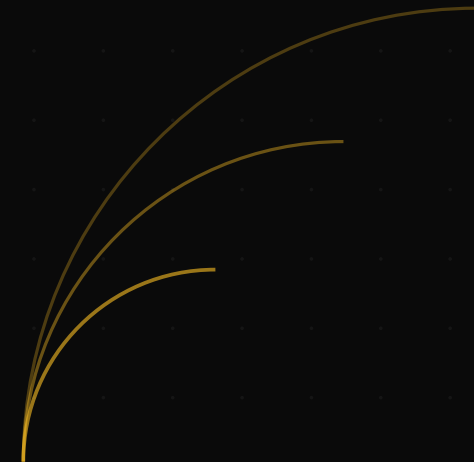


TECHNOLOGY, SECURITY, AND PROCUREMENT LEADERS

How to evaluate an AI or LLM vendor

AI-specific vendor due diligence: model deprecation, training-data rights, a quality SLA distinct from uptime, agentic spend control, exit, and the EU AI Act.



AI vendor evaluation is where buyer-side asymmetry is widest right now, because generic procurement frameworks miss the risks that are specific to models.

WHAT'S INSIDE

6 steps in this guide

- 01 Model deprecation and behaviour change
- 02 Training-data rights and indemnification
- 03 A quality SLA distinct from uptime
- 04 Agentic spend ceilings
- 05 Exit: weights, embeddings, and data
- 06 EU AI Act and ISO 42001

01

Model deprecation and behaviour change

Unlike traditional software, the thing you bought can change without a release you control - a model is deprecated, retrained, or silently adjusted, and your outputs shift. Require contractual notice of model deprecation and material behaviour change, with a defined migration window and a tested fallback.

02

Training-data rights and indemnification

Two questions decide your exposure: will our inputs and outputs be used to train or improve your models, and do you indemnify us against claims that the model's training data infringes third-party rights? Get explicit, written answers. Use of customer data for training without authorisation is both a contractual and a GDPR issue.

03

A quality SLA distinct from uptime

An AI system can be 100% available and still wrong. Uptime SLAs don't capture output quality. Where it matters, negotiate acceptance thresholds on the metrics that fit your use case (precision, recall, latency, refusal rate) and a remedy when quality regresses.

04

Agentic spend ceilings

Usage-based and agentic AI can run up cost fast. Require spend ceilings with auto-pause, per-tenant or per-workflow budgets, and alerting - the controls that prevent a runaway loop from becoming a runaway invoice.

05**Exit: weights, embeddings,
and data**

On exit, what comes back? If you fine-tuned a model, do you get the weights or just lose them? Are embeddings and vector stores portable? Pin down data and artefact return for AI exactly as you would for any platform - the lock-in is subtler but just as real.

06**EU AI Act and ISO 42001**

If you operate in or sell into the EU, the AI Act changes your obligations by risk tier. Confirm whether the system is high-risk, whether the vendor has completed any required conformity assessment, and whether they'll provide technical documentation and notice of changes that affect compliance. The EU's model contractual clauses for AI procurement are a useful baseline even outside the public sector.

PUT THIS TO WORK

Walk into every evaluation with the vendor's own playbook.

- A vendor-proof scope, issued before any proposal
- Gap analysis scoring every proposal against your scope
- A risk-weighted interrogation kit for the vendor meeting
- Scope-drift detection across proposal versions

Start free at benchsideai.com

Your first project is free. No card required.

