
BLACKTECHSTARTUP EDUCATION

Buy AI Like Infrastructure, Not Like Magic

A procurement checklist for small companies choosing model APIs and AI vendors: data rights, evaluation, cost, portability, security and failure behavior.

AI vendors can create powerful leverage, but the buying decision has unusual hidden dependencies: model behavior changes, terms change, cost scales with usage, prompts may contain sensitive data, and a product can become tightly coupled to one provider. NIST and FTC guidance make risk and data commitments part of the product decision, not paperwork after deployment.

Write the acceptance test before choosing the vendor

Build a representative evaluation set of real tasks, including edge cases and failure cases. Define minimum quality, latency, cost, citation/grounding behavior, safety requirements and escalation rules. Run the same set against candidate systems. A demo should not outrank your own test data.

If the use case affects money, health, legal rights, employment, safety or sensitive decisions, define where human review is mandatory and what the system is prohibited from doing.

Interrogate data terms

Ask whether prompts, files, outputs and metadata are used to train or improve models; what enterprise controls change that; retention periods; subprocessors; region; deletion; incident notification; and whether administrators can disable risky features. Save the exact terms that applied when you bought the service.

FTC has warned AI companies that privacy and confidentiality promises must match actual data practices. Your company should make the same commitment to customers.

Design portability before dependence

Put provider calls behind your own service layer where practical. Store prompts/templates, evaluation cases and business rules in your system—not only inside a vendor dashboard. Keep your source documents and embeddings strategy portable. Log outputs in a provider-neutral format.

You do not need to support five vendors on day one. You do need to know what it would take to move if price, quality, policy or availability changes.

Track total operating cost

Measure model usage, retrieval/storage, tools, media, retries, human review and engineering overhead. Compare vendors on cost per successful task, not published token price alone. A cheaper model that requires three retries may be more expensive.

Review the vendor quarterly against the same evaluation set and business metrics. AI procurement is ongoing performance management.

The production test

Before calling an AI/data system ready, define the normal case, difficult case, unacceptable failure, cost ceiling, latency ceiling, privacy boundary and human escalation path. Then build tests for each one. A demo proves possibility; a test suite proves repeatability.

Keep a failure log. Every meaningful failure should become a test, a product rule, a data-quality fix or a clearly documented limitation. That is how reliability compounds.

Use a vendor scorecard that includes the exit

Score candidates across task quality, failure behavior, latency, total cost, privacy/data terms, security controls, administration, observability, geographic/industry requirements and portability. Weight the categories before the demo. Otherwise the most impressive interface can dominate a decision that should have been about risk and operating fit.

Run an exit drill on paper. If the provider doubles price, changes a key policy, removes a model, has a prolonged outage or no longer meets privacy requirements, what must change in your application? Identify provider-specific prompts, tools, file stores, embeddings, fine-tunes, agent definitions and evaluation dashboards. You do not need zero lock-in; you need to understand the price of leaving.

Keep the acceptance dataset. Re-run it after material model or product changes and on a regular schedule. Track quality, cost and latency together. An AI vendor relationship is closer to infrastructure management than buying office software because output behavior itself can change even when your own code does not.

Put operating commitments in the commercial review too. Ask what support exists during outages, how material changes are communicated, whether usage data can be exported, what happens at termination, and which promises live in a binding agreement rather than a sales deck. For a critical workflow, a provider with slightly weaker benchmark performance but clearer controls, support and exit terms can be the lower-risk choice.

SOURCE DESK

Research behind this guide

Use the primary sources below to verify current rules, eligibility and program details before acting. Program terms can change.

NIST — AI RMF Generative AI Profile



FTC — AI privacy/confidentiality commitments



Google Cloud — Generative AI evaluation and operations



NIST — AI Resource Center



Important: This article is educational information, not individualized legal, tax, investment or compliance advice. For decisions with legal or financial consequences, use a qualified professional who can review your facts.

SHARE THIS GUIDE ↗

POWERED BY: BLACK TECH STARTUP

blacktechstartup.com