

BUYER'S FRAMEWORK

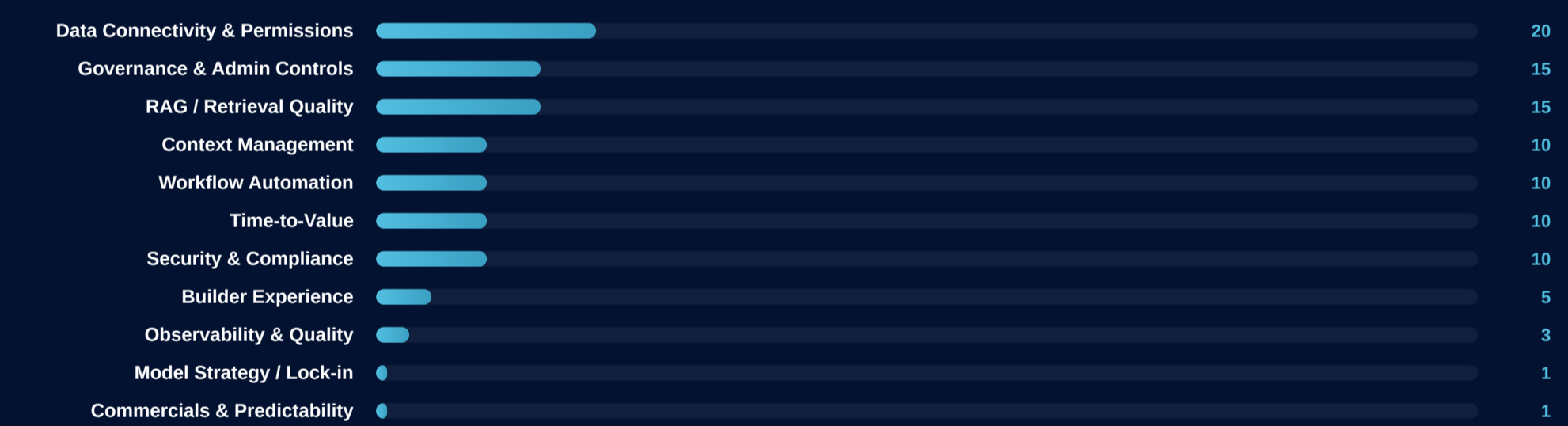
How to Evaluate an Enterprise AI Platform in Under 2 Weeks

A structured scorecard to run an apples-to-apples comparison — so your team can make a confident, defensible AI platform decision faster.

A — THE SCORECARD · WEIGHT EACH CATEGORY (TOTAL = 100) · SCORE 1-5 · MULTIPLY

<p>Data Connectivity & Permissions (20 pts)</p> <p>Native connectors to your core systems with minimal admin overhead; least-privilege access by group, user, and agent.</p> <p>QUICK TEST List your top 5 systems. Confirm connectors, setup steps, and required permissions.</p>	<p>Governance & Admin Controls (15 pts)</p> <p>Centralized policy, usage analytics, and audit logs enforced at the group, user, and agent level.</p> <p>QUICK TEST Walk through admin console: policies, logs, dashboards, and data exports.</p>	<p>RAG / Retrieval Quality at Scale (15 pts)</p> <p>Intelligent pre-processing and indexing; fast, accurate retrieval across hundreds of documents with low hallucination rates.</p> <p>QUICK TEST Load a representative corpus. Run 20 queries. Check relevance and latency.</p>
<p>Context Management (10 pts)</p> <p>Reusable context layers at company → team → project → user levels. Consistent outputs, minimal prompt sprawl.</p> <p>QUICK TEST Run the same task across 3 teams. Differences should reflect context/permissions only.</p>	<p>Workflow Automation & Orchestration (10 pts)</p> <p>Multi-step Flows with human-in-the-loop approvals, retries, and deterministic steps where precision matters.</p> <p>QUICK TEST Build a 5-step flow. Run it 10 times. Inspect failures and retries.</p>	<p>Time-to-Value (≤ 2 Weeks) (10 pts)</p> <p>2–3 real workflows live with measurable time saved and adoption across roles within 14 days.</p> <p>QUICK TEST Pick 2 workflows. Measure time to first usable version + active users by day 14.</p>
<p>Security & Compliance (10 pts)</p> <p>SOC 2 Type II (+ HIPAA as needed), SSO, configurable data retention, and procurement-ready documentation.</p> <p>QUICK TEST Request security package. Verify SSO, retention policies, and subprocessor list.</p>	<p>Builder Experience (5 pts)</p> <p>Non-technical team members can create, iterate, and publish Agents safely using templates, sharing, and versioning.</p> <p>QUICK TEST Time a business user building and publishing a safe agent from a one-page spec.</p>	<p>Observability & Output Quality (3 pts)</p> <p>Versioning, eval loops, feedback capture, failure analysis, and cost/latency visibility across all runs.</p> <p>QUICK TEST Review last 20 runs: errors, sources, cost, latency, and root-cause tooling.</p>
<p>Model Strategy & Portability (1 pt)</p> <p>Multi-model support across OpenAI, Anthropic, Google, and more — easy switching without rebuilding workflows.</p> <p>QUICK TEST Swap models on the same workflow. Compare quality and cost. Confirm portability.</p>	<p>Commercials & Predictability (1 pt)</p> <p>Transparent pricing, predictable usage-based model, clear overage rules, and scalable spend guardrails.</p> <p>QUICK TEST Model base and peak scenarios. Confirm unit economics and spend caps.</p>	

B — RECOMMENDED WEIGHTS (DEFAULT, TOTAL = 100)



C — QUICK EVALUATION PILOT (SINGLE SPRINT)

<p>≥ 80%</p> <p>Accuracy Threshold</p> <p>Outputs accepted with light edits — no heavy rework required by your team.</p>	<p>≤ 10s</p> <p>Latency on Typical Queries</p> <p>Responses fast enough for real-world use — not a bottleneck to your workflow.</p>	<p>≥ 10</p> <p>Active Pilot Users by Week 2</p> <p>Real adoption with repeat usage — not a demo that lives in a slide deck.</p>
--	---	---

START WITH THESE TWO WORKFLOW TYPES

<p>WORKFLOW 1</p> <p>Sales / BD — Account brief generation & personalized outreach</p>	<p>WORKFLOW 2</p> <p>Ops / Support — Policy Q&A with cited sources from internal docs</p>	<p>MUST DEMONSTRATE</p> <p>Governance in action: audit trail + role-based access control (RBAC)</p>
---	--	--

ARTIFACTS TO REQUEST FROM ANY VENDOR

REQUEST LIST

- Admin & Policy Demo
- Security Package
- Connector List
- Permission Model Walkthrough
- Pricing — 3 Usage Scenarios
- Audit Log Access
- RBAC Evidence
- SOC 2 Report
- Subprocessor List

SEE HOW ELVEX SCORES

Ready to run this evaluation?

elvex is purpose-built to score 5/5 across every category that matters most. Let's walk through it together.

elvex.com
SOC 2 TYPE II · HIPAA COMPLIANT