Platform Pilot Scorecard

Evaluate AI, semantic layer, and BI pilots based on operating readiness, not technical novelty.

Objective: Evaluate new platform capabilities based on operational readiness rather than technical novelty.

How to use it

1. Rate each prompt from 1 to 5.
2. Average the three scores in each section.
3. Divide that average by 5, then multiply by the section weight.
4. Add all section scores for a total out of 100.

Rating scale

1 = No chance
2 = Weak / unclear
3 = Partial fit
4 = Strong fit
5 = Proven and production-ready

Pilot details

Pilot / Feature	Business Owner	Evaluator	Date

1. Strategic Alignment (Weight: 30%)

Criteria	What to ask	Score	Notes
Decision Velocity	Does this feature reduce the time from "Question" to "Action" for a business stakeholder?		
Unit Economics	Is the cost-per-insight predictable and traceable to a budget-owning department?		
Exit Strategy	If the pilot fails or pricing changes, how hard is it to roll back or migrate the logic?		

Category score formula: (Average section rating / 5) x 30 = _____ / 30

2. Operational Governance (Weight: 40%)

Criteria	What to ask	Score	Notes
Ownership Clarity	Is there a named data steward accountable for output accuracy?		
Auditability	Can we see how an AI-generated query reached its conclusion? Think lineage, not hand-wavy magic.		
Security Integration	Does it follow Azure Entra ID / Unity Catalog policies without creating a shadow access model?		

Category score formula: (Average section rating / 5) x 40 = _____ / 40

3. Engineering Efficiency (Weight: 30%)

Criteria	What to ask	Score	Notes
Context Switching Tax	Does this fit the current workflow, including Git and CI/CD, or force a separate interface?		
Maintenance Load	Will this need human-in-the-loop monitoring, or is observability automated?		
Skills Gap	Do we have the internal seniority to manage this, or are we creating vendor dependence?		

Category score formula: (Average section rating / 5) x 30 = _____ / 30

How to interpret the total

0-50 | Reject

High risk of creating technical debt and data chaos.

51-75 | Caution

Pilot in a sandbox with strictly defined, time-bound KPIs. Do not integrate into production.

76-100 | Proceed

Scale with a focus on documenting the operating model alongside the deployment.

Need a no-BS read before you greenlight a pilot?

Gambill Data helps teams pressure-test platforms, architecture, governance, and rollout plans before budget disappears into a very expensive science fair project.

Contact

chris.gambill@gambilldataengineering.com
www.gambilldataengineering.com