Data platform assessment: a 30-point checklist

Use it when you commission an assessment, from us or anyone else. A good one ends in a document you can take to your CFO: what you have, what is at risk, what it costs and what to do next, with prices attached.

Architecture (6)

- ☐ 1 A current-state diagram of platforms, environments and how data moves between them.
- ☐ 2 Identity and access: how people and services authenticate, and who has admin rights.
- ☐ 3 Network design: private connectivity, egress paths, exposure to the internet.
- ☐ 4 Storage layers and table formats (raw, cleaned, curated; Delta, Iceberg, proprietary).
- ☐ 5 Environments: is there a real dev/test/prod separation, and is it defined as code?
- ☐ 6 Single points of failure, including people.

Pipelines and code (6)

- ☐ 7 An inventory of pipelines with owners, schedules and runtimes.
- ☐ 8 A sample review of critical pipelines for reliability and readability.
- ☐ 9 Testing: which pipelines have automated data tests, and which have none.
- ☐ 10 Deployment: how code reaches production (CI/CD or copy-paste).
- ☐ 11 Failure handling: how failures are detected and who is alerted.
- ☐ 12 Documentation and runbooks for the platform's operators.

Cost (5)

- ☐ 13 A cost baseline attributed to teams, pipelines and queries, not just the monthly invoice.
- ☐ 14 Idle and oversized compute.
- ☐ 15 Full refreshes that could be incremental.
- ☐ 16 Unused tables, jobs and environments.
- ☐ 17 A ranked list of savings with estimated value.

Governance and security (6)

- ☐ 18 Critical data elements and their named owners.
- ☐ 19 Data quality controls on the tables that feed reports, regulators and models.
- ☐ 20 Lineage for at least the most important reported figures.
- ☐ 21 Masking and access policies for sensitive data (PII, PHI, card data).
- ☐ 22 Audit evidence: could you show an auditor who accessed what, and when?
- ☐ 23 Regulatory obligations mapped to controls (e.g. SOX, HIPAA, DORA, PCI DSS; only those that apply).

AI readiness (4)

- [ ] 24 The AI and GenAI use cases the business has funded or is planning.
- [ ] 25 Whether the data each use case needs is governed, current and retrievable.
- [ ] 26 Whether permissions will survive into retrieval (see our note on RAG and permissions).
- [ ] 27 How answer quality will be measured.

The deliverable (3)

- [ ] 28 Risks ranked by business impact, not a flat list of findings.
- [ ] 29 A target architecture and, if relevant, a platform recommendation with reasoning.
- [ ] 30 A sequenced roadmap with effort and price estimates for each initiative.

Questions to ask before you sign

- [ ] Who will do the work, and will they be on the delivery team afterwards?
- [ ] What access do you need? (Read-only is enough for an assessment.)
- [ ] What's the fixed fee, and what does it depend on?
- [ ] Can we see a sample report?

From metricoak.com/insights/data-platform-assessment-checklist. Metric Oak runs a three-week, fixed-fee Data Platform Assessment that ends in a priced roadmap: metricoak.com/services/data-platform-assessment