

The AI-Ready Data Checklist

Score your data foundation before you bet an AI initiative on it. Check every box that's true today.

Data quality

- Every key dataset has a named owner who signs off on quality.
- Data quality SLAs exist — and someone gets paged when they break.
- Nulls, formats, and edge cases are documented, not tribal knowledge.
- Golden records exist for core entities (customer, product, order).
- Quality verdicts are machine-readable, not buried in a wiki.

Access & governance

- Permissions are enforced at the row/attribute level, in the data layer.
- PII is classified and handling rules are machine-enforced.
- Consent flags travel with the data, not in a separate spreadsheet.
- Every data read by an automated consumer leaves an audit trail.
- Retention and deletion policies execute automatically.

Semantics

- Business definitions exist for every core metric — written down once.
- Each metric is defined in exactly one place, consumed everywhere.
- Schemas are versioned; breaking changes get a migration path.
- Lineage is traversable: any number traces back to its source.
- Codes and enumerations are documented — no mystery values.

Freshness & reliability

- Freshness SLAs are defined per dataset — with guarantees, not vibes.
- Pipelines are monitored with alerts, not discovered broken by users.
- You can backfill and replay — history isn't a one-way door.
- Schema changes are versioned and announced before they land.
- Key metrics reconcile against the system of record on a schedule.

Score: 16–20 — AI-ready. 11–15 — close, fix the gaps first. Under 11 — foundation work before any AI bet.