

AI Ready Data Checklist

From AI-ready baseline to decision-ready data products

Use this checklist to audit your data and identify the gap between readable and decision-ready.

1. AI Ready Data Baseline

Technical readiness — data AI can read

- ☐ Data is cleaned and formatted consistently
No malformed values, consistent types, standard encoding.
- ☐ Structure is documented and stable
Schemas, dictionaries, and change processes are known.
- ☐ Data is catalogued and discoverable
Teams can find relevant datasets without tribal knowledge.
- ☐ Accessible via API, warehouse, or vector store
Models and applications can reach it programmatically.
- ☐ Quality checks run at pipeline time
Validation, profiling, and anomaly detection are automated.
- ☐ Security basics are in place
Encryption, access lists, and credential management covered.

2. Decision-Ready Upgrade

Business trust — data AI can act on

- ☐ Every dataset has a named business owner
Accountability sits with the domain that understands it.
- ☐ Business semantics are explicit
Terms like 'active customer' are defined and versioned.
- ☐ Freshness, quality, and access policies are bound to the data
Not just documented — enforced at the point of use.
- ☐ Lineage is traceable end to end
From source system to the AI decision and back.
- ☐ Governance is active, not manual
Policy is applied automatically when data is consumed.
- ☐ Data is packaged as a reusable product
One product serves multiple AI systems and use cases.

3. Readiness Questions

Ask these before launching an AI initiative

- ☐ Can a non-engineer explain what this data means?
- ☐ Can we trace any AI output back to its source data?
- ☐ Do we know who is accountable for freshness and quality?
- ☐ Is access policy enforced automatically at consumption?
- ☐ Can this data product be reused for the next AI use case?
- ☐ Would we bet a business decision on output from this data today?

4. 30-Day Action Plan

Close the gap from baseline to decision-ready

- ☐ **Pick one decision AI must support**
Start with the business outcome, not the dataset.
- ☐ **Map the data products it needs**
Identify owners, sources, consumers, and contracts.
- ☐ **Document semantics and policies**
Turn implicit knowledge into explicit, versioned rules.
- ☐ **Enforce policy at the point of use**
Bind rules to the product so AI cannot bypass them.
- ☐ **Ship one reusable product**
Validate it with a copilot, agent, or analytic before scaling.