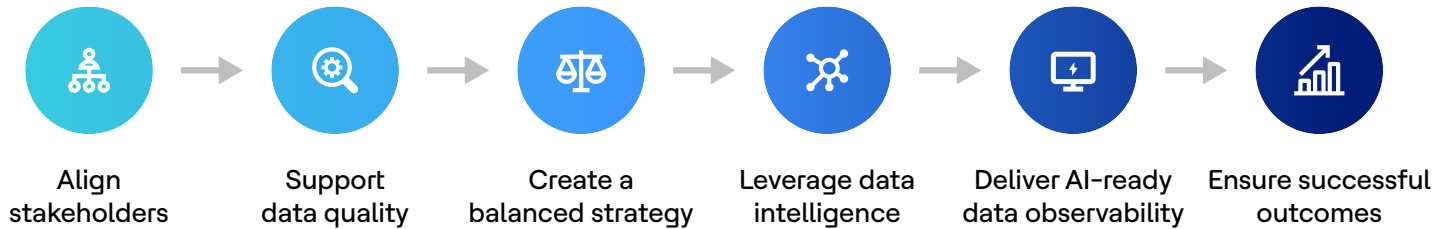


# The Ultimate Checklist for AI Data Readiness

Follow these steps to ensure your data is  
AI-ready before making the first prompt.



## The Path to AI Success



Before you can deliver value with AI, you must have confidence in your data. Quality, trusted, AI-ready data is essential to success. Yet there's a disconnect between AI ambitions and data quality.

According to Gartner®<sup>1</sup>:

- 63% of organizations either do not have or are unsure if they have the right data management practices for AI.
- Through 2026, organizations will abandon 60% of AI projects unsupported by AI-ready data.
- Organizations that fail to realize the vast differences between AI-ready data requirements and traditional data management will endanger the success of their AI efforts.

This checklist ensures your organization is truly prepared to optimize the full power of AI. It can also help you avoid data roadblocks that slow adoption, increase costs, or compromise outcomes.

## 6 Steps to Ensure Trusted, AI-Ready Data

Ensure data quality, readiness, and confidence at every point along your AI journey. These practical steps will enable you to deliver trusted, AI-ready data for your use cases.



### 1. Align stakeholders with a data readiness strategy

Ensure every team understands their role in preparing data for AI. Ask the right questions to drive the right conversations, then take action to turn talk into progress.

#### Business Users

- ☐ What are my AI use cases?
- ☐ What data is needed—and not needed—to achieve them?
- ☐ What's needed to ensure data readiness?

*Define specific goals, such as hyper-personalizing customer experiences or improving demand forecasting, and identify the necessary datasets. Eliminate non-critical data early to reduce noise and complexity.*

<sup>1</sup> Gartner, "Lack of AI-Ready Data Puts AI Projects at Risk," February 2025.

### Data Stewards

- ☐ How will we verify data quality and compliance?
- ☐ How is data anonymized to ensure we're compliant with privacy and security regulations?
- ☐ What involvement is needed by IT developers and data engineers?

*Work closely with IT and engineers to test, monitor, and document quality checks.*

*Apply anonymization techniques and enforce role-based access for sensitive data.*

### IT Developers / Data Engineers

- ☐ How will we automate secure, real-time data integration while enforcing rules and standards?

*Build and connect scalable pipelines that validate data at ingestion, apply governance policies, and maintain performance service level agreements (SLAs).*

### Business Analysts

- ☐ How will we prepare trusted data for analytics and share the data with business users?

*Standardize and automate data preparation workflows and optimize user-friendly tools to deliver accurate data to where business users and others need it.*

### Data Scientists

- ☐ How will we establish our models and define the right training data?

*Collaborate with domain experts to set model parameters and ensure the appropriate data quality.*



### Business benefit

Involving business users, data scientists, engineers, stewards, and analysts in collaborative planning aligns stakeholders. This ensures data meets technical, compliance, and business needs to support use cases and accelerate time to value.



## 2. Define and support data quality standards

Ask critical questions to strengthen your foundation for AI-ready data. Take action to mitigate risk and ensure trusted data from the start.

- ☐ How much data volume do we need to manage?  
*Determine capacity needs early to avoid scale and performance issues later.*
- ☐ How will we use this data?  
*Map data for each AI use case to eliminate redundancy and focus resources.*
- ☐ How will we clean and normalize the data?  
*Apply consistent formatting, deduplication, and standardization rules to improve data usability across systems and teams.*
- ☐ How will we optimize machine learning for use cases?  
*Align capabilities with specific goals for each use case to improve results and train models.*
- ☐ How will we ensure accuracy and quality for training data, including data sources outside our company?  
*Validate third-party data and establish rules to maintain trust and reduce data hallucinations from both internal and external sources.*
- ☐ How will we prevent poor quality data from being injected upstream?  
*Implement upstream quality checks to identify and stop issues before they impact downstream use cases.*
- ☐ What risks does poor data quality pose?  
*Quantify potential impacts of poor quality data, such as inaccurate predictions or compliance failures, to build the business case for proactive quality management.*



### Business benefit

High-quality data leads to better AI performance, reducing data hallucinations and rework while building trust and confidence in the data.





### 3. Manage data quality as a lifecycle, not a one-time task

Ensuring long-term success requires thinking beyond initial deployment. Asking the right questions helps teams future-proof AI initiatives by building in flexibility, data quality, and accountability.

- ☐ How will our use cases and data assets evolve over time?  
*Revisit use cases along with input and output requirements regularly. Build roadmaps that anticipate scale, new data sources, and increasing complexity to keep AI efforts aligned with business needs.*
- ☐ How can we build flexibility and agility into data processes?  
*Use modular architecture with low-code or no-code tools when possible. This allows you to quickly integrate or change components without rebuilding the entire pipeline.*
- ☐ How will we ensure current data flows and pipelines meet quality standards?  
*Extend the data quality framework to existing systems by adding observability checkpoints and automated validations at every pipeline stage.*
- ☐ How will we test readiness and debug workflows before deploying data in business-critical environments?  
*Establish a "safe zone" staging environment for simulation and testing with version control and rollback capabilities. Include human review checkpoints in the validation process.*
- ☐ What are the time, cost, or operational risks if we skip or shortcut these steps?  
*Calculate the potential cost of rework, poor decision making, model hallucinations, or reputational damage from untrustworthy outputs.*
- ☐ How will we continuously monitor and measure data quality?  
*Implement a real-time data quality dashboard with alerts tied to key metrics like accuracy, completeness, and freshness. Review trends monthly and act on any anomalies before they impact outputs.*



#### Business benefit

Automated quality checks with data observability and governance frameworks build flexibility and confidence into AI operations. This delivers sustainable value over time by performing readiness tests and ensuring long-term monitoring across pipelines.



#### 4. Create a balanced strategy that meets your specific needs

Every organization will have different requirements and expectations for using AI. Design a strategy for your specific business needs that considers timelines, resources, capabilities, and skillsets.

- ☐ Have you assessed the risk of moving too fast—or too slow—on AI adoption?  
*Balance urgency with strategy by aligning your AI approach with proven data readiness frameworks. This avoids costly rebuilds later by not jumping in before your data is ready, while ensuring trusted outcomes for a competitive edge.*
- ☐ Can we prevent bad data from creating convincing but wrong answers?  
*Establish accuracy criteria and upstream controls to prevent low-quality data from feeding data models or for other AI use cases.*
- ☐ How will we define data accuracy?  
*Consult business and technical stakeholders to create a shared definition of accuracy for each use case. The definition can cover timeliness, consistency with source systems, or alignment with specific use cases.*
- ☐ Are you preparing for evolving regulations and standards?  
*Choose tools and partners that offer—even prioritize—flexibility, compliance, and scalability. Look for vendors that have a track record in trusted data readiness, can help you find the fastest path to value, and enable your data infrastructure to scale to meet evolving AI needs.*



##### Business benefit

A balanced AI strategy helps avoid costly pitfalls, ensures more accurate and reliable AI outputs, and positions your organization to scale confidently as use cases and regulations evolve. Building trust into every step prepares your infrastructure for current and emerging requirements.



## 5. Leverage data intelligence to support AI

AI success depends not just on data volume, but on data understanding. Data intelligence provides the metadata, lineage, and context needed to ensure AI models are using the right data in the right way.

- ☐ Do we have clear visibility into where our data comes from and how it flows?  
*Implement data lineage tools to trace data movement and transformations across systems. This enables better oversight, faster debugging, and more confident use of data.*
- ☐ Are we capturing and using metadata to enrich AI inputs?  
*Automate metadata usage with your data catalog. Rich metadata provides essential context around source, structure, and semantics, helping AI models generate more relevant, accurate outputs.*
- ☐ Do we understand how data transforms as it moves through systems?  
*Use graph-powered data lineage to visualize how data evolves across your infrastructure.*
- ☐ Are we using business context to inform how AI understands and applies data?  
*Build data contracts and business glossaries to define terms and relationships in a way that aligns with how your organization operates.*
- ☐ How do we make data intelligence accessible to both business and technical users?  
*Implement a user-friendly data intelligence platform with self-service access so users across roles can explore and trust the data powering AI.*



### Business benefit

Deploying a unified data intelligence platform enables discovery, governance, and lineage tracking across teams. The platform gives users the ability to find, understand, and trust data, which is essential for scalable and compliant AI initiatives.



## 6. Ensure modern data observability capabilities are in place

Powering AI with trusted, high-quality data requires visibility into how data flows, transforms, and behaves across the entire pipeline. You need the ability to see and mitigate any quality issues before they move downstream or cause blind spots in your pipeline.

- ☐ Can we monitor data quality across every stage of our pipeline in real time?  
*Deploy observability tools that assess data quality continuously—without relying on sampling. Look for solutions that provide end-to-end visibility and detect issues before they impact AI outcomes.*
- ☐ How quickly can we detect and resolve data issues before they affect outputs?  
*Establish data quality policies supported by automated alerts and remediation workflows. Create SLAs for issue resolution to minimize the risk of AI generating flawed or misleading outputs.*
- ☐ Do our observability tools integrate with our data environment?  
*Choose a solution that monitors and analyzes data without straining source systems. Look for a solution that allows you to maintain performance and scale monitoring across hybrid and multi-cloud environments.*



### Business benefit

A modern observability solution prevents blind spots, protects SLAs, and ensures data flowing into AI is accurate, complete, and compliant. Using real-time, AI-powered observability detects quality issues across your full data lifecycle, before they impact use cases.



# Don't Just Build AI. Build It Using the Right Data

AI is only as powerful as the data that fuels it. Success isn't just about having more data, it's about having the right data that's clean, complete, compliant, and contextual.

By using this checklist, your organization can move beyond hype and hope to take deliberate, confident steps to ensure your AI initiatives are grounded in trusted, business-ready data. With the right strategy, tools, and collaboration across teams, you'll set the foundation for AI success today and also for scalable, sustainable innovation well into the future.

## About Actian

Actian empowers enterprises to confidently manage and govern data at scale. Organizations trust Actian data management and data intelligence solutions to streamline complex data environments and accelerate the delivery of AI-ready data. Designed to be flexible, Actian solutions integrate seamlessly and perform reliably across on-premises, cloud and hybrid environments. Learn more about Actian, the data division of **HCLSoftware**, at [actian.com](https://actian.com).

