ACTIAN™
a division of HCLSoftware

EBOOK

Is Your Data Truly Ready for AI Agents

Trusted AI outcomes start with data reliability, not models. See how data observability makes AI scalable.

Table of Contents

2 Executive Summary

3 When AI Runs on Bad Data, Every Output is at Risk

3 What “AI-Ready Data“ Really Means in Practice

4 5 Steps to Build a Practical, Observability-First Roadmap to AI-Ready Data

- 4 Start with a high-value, high-risk use case
- 4 Make data reliability continuous, not reactive
- 4 Automate triage and resolution to reduce manual work
- 4 Cover the data format for your AI systems' needs
- 4 Put data-health checks inside the AI workflow with MCP

5 How Actian Data Observability Supports AI at Scale

5 Make Data Reliability Non-Negotiable

5 About Actian

Executive Summary

Most organizations already have AI in motion. Chatbots, copilots, recommendation engines, and early-stage AI agents are becoming part of everyday workflows. Although enthusiasm is real, many teams struggle to move AI from pilots into reliable, trusted production use cases.

According to Gartner, through 2026, organizations will abandon 60% of AI projects that are not supported by AI-ready data. This is not because of model choice or algorithm quality, but because the data feeding those systems is unreliable.

Unlike human analysts, AI systems do not question missing values, silent schema changes, or delayed data. They assume inputs are valid and act accordingly. If the data is incomplete, outdated, biased, or poorly monitored, the outputs will reflect those flaws—often at scale and with greater confidence than traditional systems.

This guide focuses on a simple truth: **AI readiness starts with data observability.**

You’ll learn what happens when AI runs on bad data, what AI-ready data looks like in practice, why AI agents increase the need for real-time data reliability, and how data observability helps organizations detect, explain, and prevent data issues before they impact analytics or AI systems. You’ll also see how poor results are often directly tied to gaps in data observability.

So what does failure look like in the real world?

When AI Runs on Bad Data, Every Output is at Risk

If AI relies on bad data, the failure modes are predictable and costly:

Wrong answers at scale	A customer-facing AI assistant references outdated pricing across thousands of interactions before the issue is detected. Human teams must step in to correct errors, eroding trust in both the AI and the organization.
Hidden bias and skew	Incomplete or unbalanced training data leads to biased lending or risk decisions, often without obvious warning signs until regulators or customers flag the issue.
Operational failures	Forecasting, optimization, and planning models trained on inconsistent data misallocate inventory, misjudge demand, or misprice services, directly impacting revenue and customer experience.
Compliance and audit risk	When AI outputs cannot be traced back to validated, documented data sources, audits become difficult and regulatory exposure increases.

In many cases, the models are working exactly as designed. What fails is the assumption that upstream data remains correct, complete, and up to date.

What "AI-Ready Data" Really Means in Practice

AI readiness is often framed as a tooling or infrastructure problem. In reality, it is a data reliability problem.

AI-ready data is data that can be safely used for decisions or automation because its freshness, completeness, and behavior are continuously validated.

In practice, that means data with the following characteristics:

Reliable and current	Data is accurate, complete, and up to date. Missing values, duplicates, schema changes, and volume shifts are detected before they impact models.
Continuously observable	Teams can see how data behaves over time, including anomalies, drift, and pipeline breaks, without relying on periodic manual checks.
Explainable when data changes	When issues occur, teams can quickly understand what changed, where it changed, and why it matters.
Accessible without guesswork	Both humans and AI systems can verify whether data is fit for use before relying on it.

Without these characteristics, teams often get stuck in prototype mode—able to build AI demos, but unable to scale them into reliable production. That limitation becomes dangerous when systems move from informing decisions to making them. What is the common mistake? Treating AI readiness as a one-time certification instead of a continuous validation necessity.

ACTIAN™ a division of HCLSoftware

5 Steps to Build a Practical, Observability-First Roadmap to AI-Ready Data

Organizations do not need to update everything at once. AI readiness improves fastest when data observability is applied deliberately.

1. Start with a high-value, high-risk use case

AI high-risk use cases include underwriting decisions, eligibility determination, compliance reporting, and automated customer actions.

For these use cases:

- Identify the critical datasets
- Define minimum expectations for freshness and completeness
- Document what actions depend on AI outputs

These steps create a clear target for your observability efforts.

2. Make data reliability continuous, not reactive

Traditional data quality relies on dashboards and weekly checks, which often surface issues only after AI outputs degrade or customers are impacted. AI requires continuous validation as data lands.

Data observability enables teams to:

- Monitor pipelines and datasets in real time
- Detect anomalies proactively
- Identify data drift before it degrades AI performance

3. Automate triage and resolution to reduce manual work

Before agents were used, teams chased alerts across tools and logs. With Data Reliability Agents, issues are automatically validated, explained, and routed. Data Reliability Agents manage the full lifecycle of data quality issues:

- Validation Agent checks data as it lands
- Incident Diagnosis Agent explains root causes in plain language
- Lineage Agent shows upstream and downstream impact
- Insight Agent surfaces trends and risks
- Orchestration and Routing Agents help coordinate resolution
- Help Agent answers questions about data health through natural language

Instead of chasing alerts, teams get clear explanations and guided resolution.

4. Cover the data format for your AI systems' needs

Most observability tools focus on structured data (relational database tables, data warehouse/lakehouse tables, spreadsheets, time-series data). Real AI pipelines rely on PDFs, XML files, contracts, invoices, claims, logs, and partner feeds that are rarely monitored with the same rigor as tables.

It is imperative to extend data reliability to PDF and XML data formats, which are common in regulatory filings, invoices, claims, logs, and partner feeds.

This gives teams visibility into semi-structured and document-based pipelines that are typically out of scope for traditional monitoring tools.

5. Put data-health checks inside the AI workflow with MCP

AI systems need a safe way to verify data health before acting. A Model Context Protocol (MCP) server easily exposes data quality status, incidents, alerts, monitors, and validation results directly to AI assistants and agentic workflows.

It allows AI systems to check data health before taking action, rather than relying on assumptions or outdated signals.

Now here's what it looks like when you operationalize observability with Actian.

ACTIAN™
a division of HCLSoftware

How Actian Data Observability Supports AI at Scale

Actian Data Observability focuses on making data reliability automatic and usable, with an approach designed for teams running AI in regulated, high-stakes environments where trust, explainability, and speed matter (Figure 1).

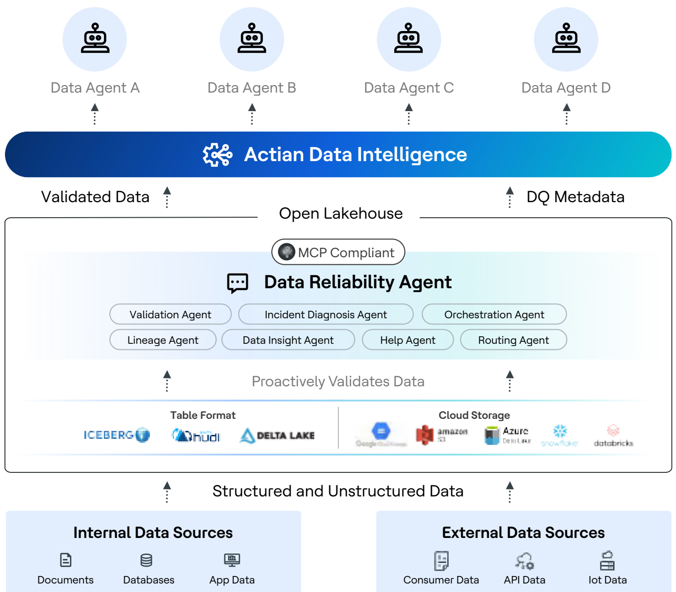

Figure 1. Actian Data Intelligence with Data Reliability Agents

The Actian solution provides:

- Continuous validation for data coming from Iceberg, Delta Lake, and Hudi
- Data Reliability Agents that detect, explain, and resolve issues quickly
- Support for structured, semi-structured, and document-based data
- MCP connectivity that brings observability context directly into AI tools

Together, these capabilities reduce risk, speed resolution, and allow teams to scale AI without adding complexity or headcount.

Make Data Reliability Non-Negotiable

Most AI failures are not model failures. They are data reliability failures that surface too late.

If AI outputs are inconsistent, hard to explain, or not trusted, the right question is not, "Which model should we use?" It is, "Can we verify that the data is reliable right now?"

Actian Data Observability helps you answer that question continuously, before data reaches executives, customers, or autonomous systems.

If your AI system cannot verify data health at the moment of action, it is not ready for autonomous decisions. With Actian Data Observability, your teams can move faster with confidence—turning data health into a built-in control that keeps AI decisions trusted, explainable, and ready for scale.

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 and AI division of **HCLSoftware**, at **actian.com**.

Tel +1 512 231 6000 Fax +1.512.231.6010 710 Hesters Crossing Road, Suite 250, Round Rock, TX 78681 **Actian.com**

ACTIAN™ a division of HCLSoftware