LTIMindtree

Point of View

Strategies for Achieving AI-Ready Data Management in 2025

Introduction

Artificial Intelligence (AI) is rapidly becoming a key driver of innovation and efficiency across industries. From improving decision-making to automating tasks and unlocking new growth opportunities, businesses are investing heavily in AI to stay competitive. However, while the focus often falls on the technology itself, one critical enabler often goes under the radar—data.

This paper will focus on a foundational element of any successful AI initiative: AI-ready data management. Without high-quality, well-governed, and easily accessible data, even the most advanced AI systems will fall short, delivering limited value, biased outcomes, or incorrect results.

We will begin by breaking down what makes data "AI-ready." What are the essential capabilities your data environment must have? And just as important - what's holding it back? By addressing both, we help you build a clear roadmap for a smarter, more seamless AI-driven ecosystem.

To bring a practical lens, we approach this from two key perspectives: insights from analysts that reflect industry trends, and real-world input from our customers and implementation teams who've seen what works (and what doesn't) on the ground. We'll also explore the role of strong data governance in minimizing bias and hallucination in AI and highlight solutions and technologies that can enable more intelligent data management at scale.

Challenges and Opportunities

The following picture emerged from a Pulse Survey - Data Readiness for the AI Revolution, conducted by HBR[1]; 91% of survey respondents agree that an organization must have a reliable data foundation to successfully adopt AI [1].

When we examine the key data management practices needed to make data AI-ready, a few stand out: data governance, data lifecycle management, data trust, Master Data Management (MDM), and data quality.

What data management practices does your organization need to improve in order to more successfully adopt AI? Select all that apply.

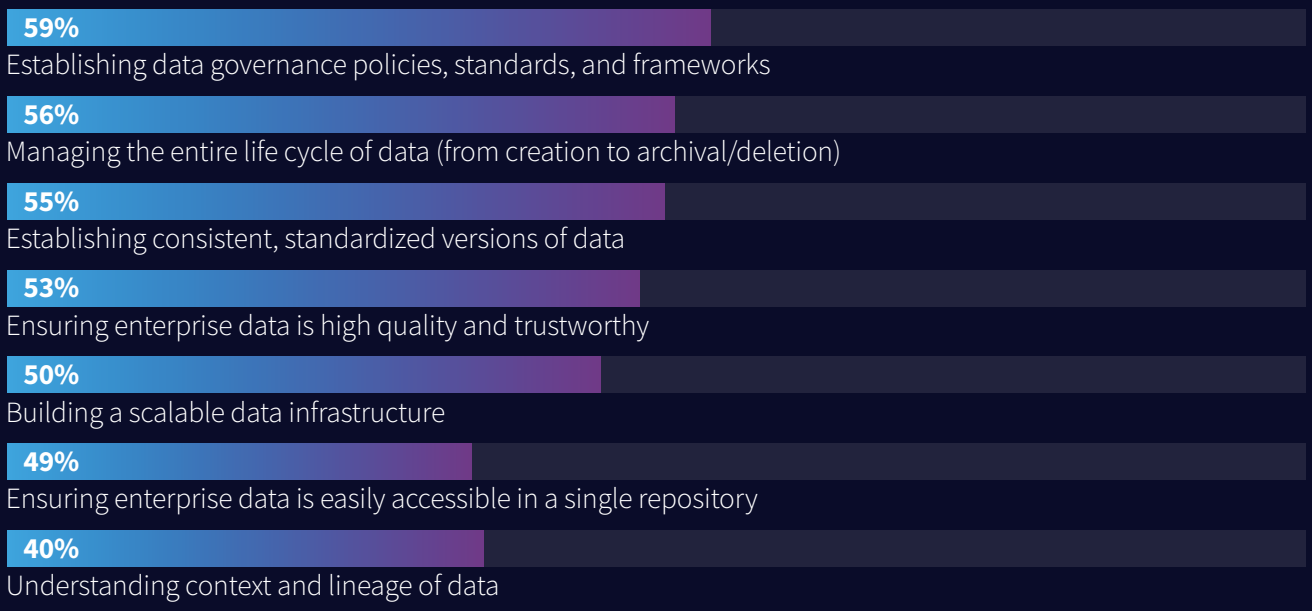

Fig 1: What data management practices does your organization need to improve to adopt AI successfully?
Source: Harvard Business Review Analytic Services survey, January 2024

When asked about the main challenges impeding AI adoption, most answers indicate that inefficient, inadequate, and non-performant data management practices are the root cause.

What data challenges, if any, is your organization experiencing in its adoption of AI? Select all that apply.

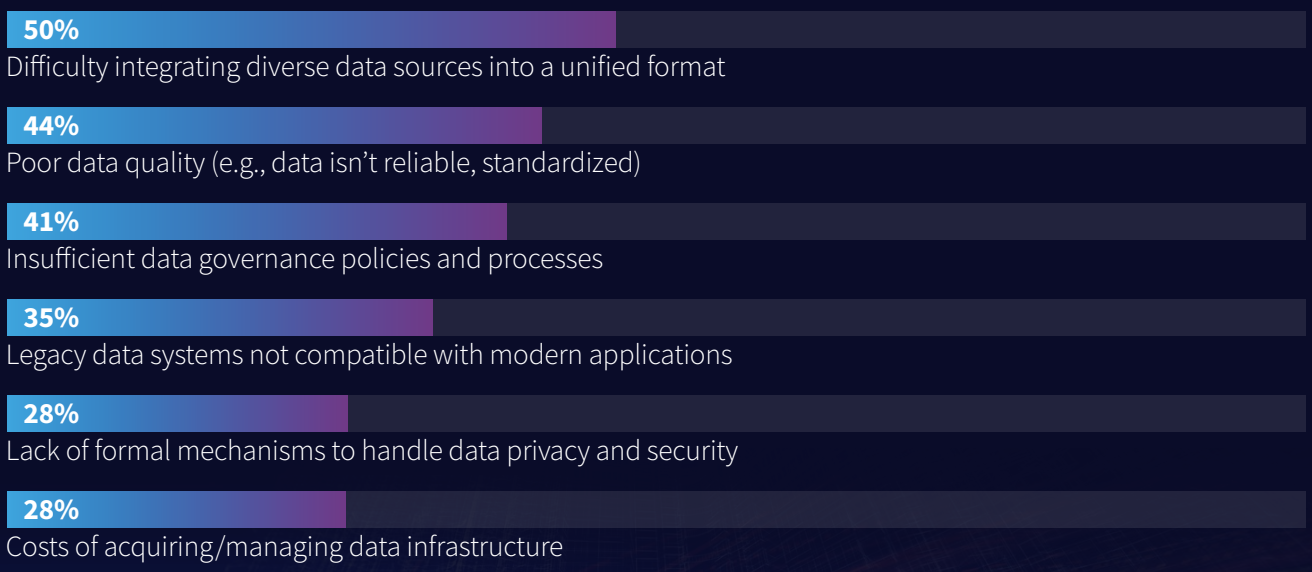

Fig 2: What data challenges is your organization experiencing in adopting AI?
Source: Harvard Business Review Analytic Services survey, January 2024

It is evident that a dependable data foundation with a comprehensive set of capabilities is essential to advance the AI aspirations of organizations.

It was no surprise to us that the analysts' findings aligned closely with several themes we've observed in our customer conversations. Most of them, on their journey to adopt AI and build intelligent systems, have shared their challenges - roadblocks that slow them down and put them behind their competitors. What we heard:

American convenience chain

Our ecosystem isn't set up to seamlessly syndicate product information to modern channels like Facebook and Instagram. To stay competitive, we need a more agile and connected approach that ensures our AI-driven analytics and recommendation systems get the right data for seamless functioning.

Multinational consumer group corporation

Our data catalog is overly complex, making it equally challenging for Business Users and Data Scientists to navigate and use effectively. As a result, the system isn't truly democratized, limiting its accessibility and impact. We need a more user-friendly, persona-based approach to ensure broader adoption and real business value.

Multinational automobile conglomerate

Our MDM is designed mainly for analytical use cases, but we need a more powerful operational MDM that supports real-time processing and streaming interfaces. A more agile and responsive setup will help drive seamless operations and better downstream analytics.

Based on the above assessment, several key areas have been identified that organizations must focus on for a robust AI ecosystem. Prioritizing these elements will create a strong foundation for scalable, efficient, and future-ready AI integration. A point to note - this perspective is explicitly focused on the core areas of data management. That means we haven't touched on things like infrastructure readiness or cybersecurity. Those are important too, but they fall outside the scope of this discussion.

The need for robust data management systems and processes for successful AI adoption cannot be emphasized enough. The true strength of AI lies in its capacity to interpret vast amounts of data created and modified through various sources and channels. If data management processes are not adequately prepared to handle the influx of data, the AI models built on this data will inevitably lack effectiveness. The next generation of AI systems – AI Agents need to be enabled for autonomous operation, which is only possible through access to good data that will ensure intelligent decisions and outcomes.

Another key consideration for AI-driven systems is staying mindful of data sensitivity and privacy, ensuring such data is not intentionally or accidentally exposed. At the same time, AI models need access to high-quality data to avoid bias and hallucination.

Key capabilities for ensuring AI strategy adherence

In this section, we dive deeper into specific data management areas that need to be improved and practices that need to be adhered to for building competent AI systems.

For each area, we've highlighted four key aspects: how it enhances AI readiness, the latest next-gen capabilities, features that Informatica's IDMC (Intelligent Data Management Cloud) platform enables, and the role of LTIMindtree accelerators in driving transformation.

(We have chosen Informatica's IDMC platform as an industry benchmark and a leading cloud native platform powered by CLAIRE, which brings in AI and Gen AI capabilities to Informatica's robust data management ecosystem.)

Data sourcing and ingestion

By integrating diverse data sources into a unified format, organizations ensure consistency and accessibility across the board. This seamless integration enhances the effectiveness of AI models.

Next-gen capabilities

- Unstructured data collection and parsing
- AI-powered schema mapping and transformation
- Real-time and streaming ingestion
- Integrated data quality and governance

IDMC components

- CDI (Cloud Data Integration)- Allows building and running complex integrations with pre-built functions and templates.

LTIMindtree enablers and accelerators

- Agentic AI framework for data quality

Data quality

Ensuring high-quality data involves stringent data quality checks, regular cleansing, validation, and standardization processes to eliminate errors and discrepancies. This enhances the accuracy of AI models and builds trust in the insights generated.

Next-gen capabilities

- AI-driven data observability
- Self-healing data pipeline
- Real-time data validation

IDMC components

- CDQ (Cloud Data Quality)- Verify, standardize, and improve the quality of your business data.

LTIMindtree enablers and accelerators

- Gen AI-powered DQ for data profiling and rule recommendation.

Data governance and lineage

Establishing a comprehensive data governance framework ensures that data is managed, protected, and utilized responsibly. This includes defining data ownership, access controls, and compliance policies to safeguard sensitive information and ensure regulatory adherence.

Next-gen capabilities

- Automated lineage tracking
- Dynamic policy enforcement
- Metadata-driven governance
- AI-powered compliance monitoring

IDMC components

- CDGC (Cloud Data Governance and Catalog) - Manage, analyze, and interpret extracted metadata, policies, and workflows.

LTIMindtree enablers and accelerators

- Agentic AI Framework for Metadata Management
- Gen AI-powered lineage tracking and visualization

Data marketplace

Data marketplace provides a centralized platform for data discovery, sharing, and monetization. Organizations can enhance their AI models and drive innovation through data collaboration by facilitating easy access to diverse datasets.

Next-gen capabilities

- AI-driven data discovery
- Interoperability and standardization

IDMC components

- CDMP (Cloud Data Marketplace)- Publish data collections for business purposes and access data collections published by others within their organization.

LTIMindtree enablers and accelerators

- Data marketplace solution comprising a persona-specific user experience and data recommendation

Entity resolution and a single source of truth

A single source of truth ensures that all data is consistent, accurate, and accessible from a centralized location. By maintaining a unified data repository, organizations can eliminate data silos, reduce discrepancies, and enhance the reliability of AI models.

Next-gen capabilities

- AI-powered match/merge
- Decentralized and federated models
- Active Metadata Management
- Automated data quality checks

IDMC components

- MDM SaaS - Manage the master data domains like Customer, Product, and Supplier

LTIMindtree enablers and accelerators

- Match Merge Framework based on LLMs

How can LTIMindtree help?

LTIMindtree's robust Trust Framework is a blueprint for implementing an AI-ready data management platform. It is a four-stage approach that takes an organization from assessment to implementation in its data management journey. A range of enablers and accelerators facilitate a seamless transition to AI-enabled systems and decision-making. Our approach identifies and involves the people, processes, and technologies needed for a successful data management journey.

We start with the Assessment phase, in which we aim to understand the organization's data management journey's goals, objectives, and priorities. We suggest a phased approach to building its capabilities and how to tackle each challenge. This is followed by the Design phase, which prioritizes data management capabilities and identifies scope and use cases for a Pilot phase. In the Pilot phase, we implement the use cases and demonstrate the business benefits and technical improvements. We set the foundation for these capabilities to be enhanced and improved in the Expand phase.

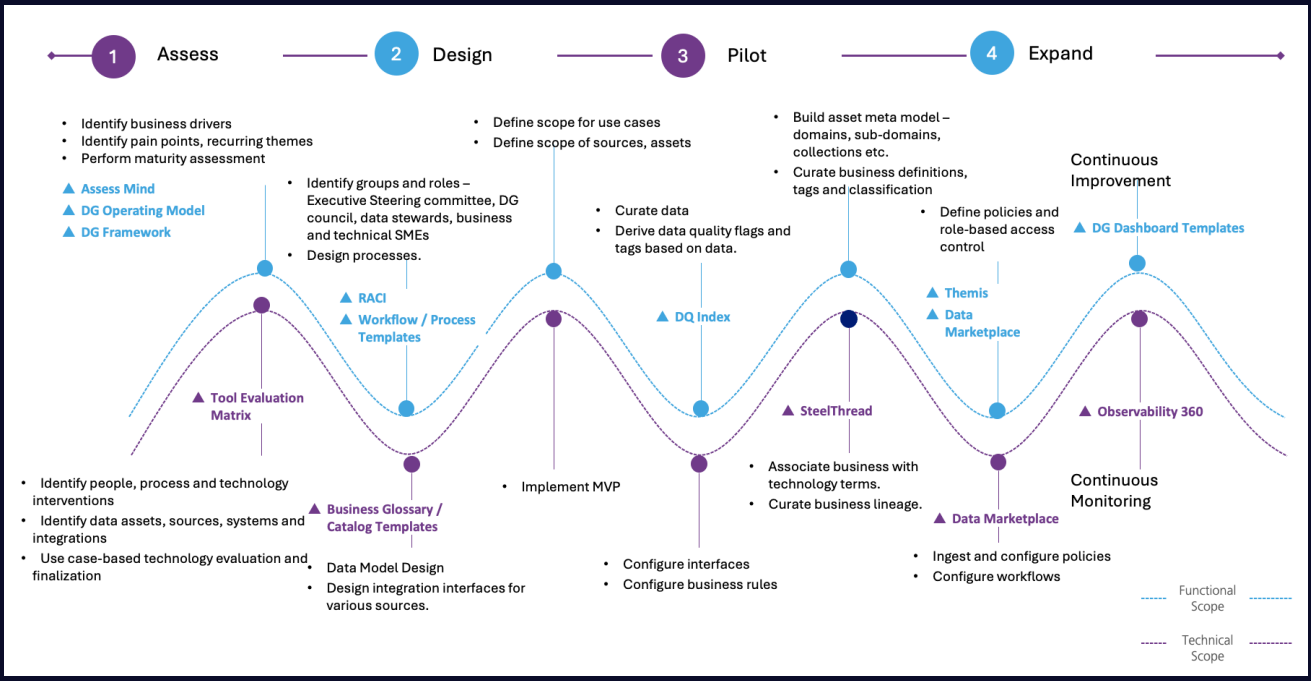

Fig 7: Data governance implementation roadmap

Conclusion

Achieving AI-ready data management requires a holistic approach encompassing data integration, quality, governance, and continuous improvement. Organizations can unlock AI's full potential and drive meaningful insights and innovation by implementing these strategies.

LTIMindtree is driving the future of data and AI architecture by embedding trust into the core of data management and modern data architecture.

By leveraging the advanced capabilities of next-gen platforms like Informatica's IDMC, organizations can ensure their data is well-prepared for AI applications. This comprehensive strategy not only enhances data reliability and accessibility but also drives meaningful insights and innovation, paving the way for successful AI adoption.

References

1. *Data Readiness for the AI Revolution, Harvard Business Review,* May 01, 2024: *https://hbr.org/sponsored/2024/05/data-readiness-for-the-ai-revolution*

LTIMindtree

About the authors

Pooja Joshi

Principal, Data and Analytics Practice, LTIMindtree

Pooja has nearly 20 years of experience helping clients implement large-scale MDM programs. Her specialization includes architecting high-availability MDM applications that can handle large data volumes and advising on next-gen MDM platforms leveraging the latest techniques: NoSQL, cloud-native technologies, AI, ML, and Gen AI.

Prashant Kondi

Associate Principal - Architecture, Data & Analytics Practice, LTIMindtree

Prashant has over 18 years of experience in Informatica Master Data Management (MDM), Data Quality (DQ), and Data Governance (DG) consulting, solution architecture design, MDM and DQ implementations, data migrations, and MDM modernizations.