



The Gen AI Data Readiness Checklist

Findings from the Actian Gen AI Data Readiness Study

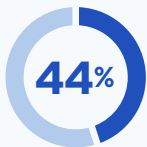
A Second Look at Gen AI Data Readiness

Many organizations are prioritizing the deployment of Gen AI for a number of mission-critical use cases, but they may not be as ready as they think. A lack of proper data preparation could lead to costly and time-consuming setbacks. But when approached correctly, proper data preparation can help accelerate and future-proof your Gen AI deployment.

In order to determine organizational readiness to adopt Gen AI, Actian surveyed 550 professionals—70% of which were Director or higher—from 6 countries and 7 industries about their deployment experiences. Their answers indicated that they believe they’re prepared for Gen AI—but how ready are they really?

The stakes are high

Respondents are deploying Gen AI for mission-critical use cases



Implementing today



Just starting



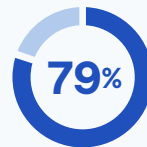
Planning/considering



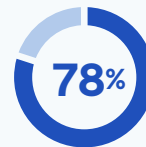
Not planning to implement

People think their data is ready

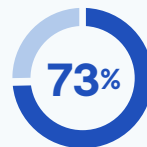
Rate themselves “above average” and “outstanding” in data readiness



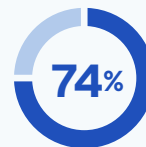
Data quality & cleanliness



Data preparation



Data management and staffing

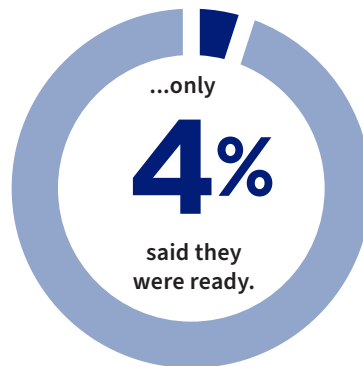


Data governance

Source: Actian’s Gen AI Data Readiness Study, 2024.

But, how ready is your data, really?

When Gartner asked people in charge of AI data readiness...



The Gen AI Data Readiness Checklist

1 Have the right conversations with the right stakeholders

Your Business Users

- What are my use cases?
- What data do I need (and not need) to achieve them?
- What do I need to do to get my data ready for Gen AI?

Your Data Scientists

- How will we establish and create our data model?

Your IT Developers/Data Engineers

- How will we automate real-time data integration / enforce rules & standards?

Your Data Stewards

- How will we verify our data set for quality and compliance, and what involvement do we need from our IT developers/data engineers?
- How will we anonymize data and ensure compliance with privacy and security regulations?

Your Business Analysts

- How will we prepare and share the data for consumption by the business users?

2 Define data quality

- How much data volume do we need to manage?
- How will we use this data?
- How will we clean and normalize the data?
- How will we optimize ML specifically for this use case?
- How will we ensure the accuracy and quality of the training data, including data sources outside our company?
- How will we prevent poor quality data from being injected upstream?
- What kinds of risks does poor data quality pose?

3 Plan for data quality lifecycle management

- How are my use cases and data sets going to shift over time?
- How will we build flexibility and agility into this process?
- How will we extend data quality to existing data flows/pipelines?
- How will we conduct readiness tests and debug our workflows before we entrust some of our most important business processes to them?
- How much time and money would we lose if we skip or shortcut one of the above steps?
- How will we track data quality over time?

4 Create a balanced strategy

Consider the pros and cons of either moving too quickly or too slowly on Gen AI adoption

- What are the risks of moving too quickly? How much would it cost and how long would it take to re-do our Gen AI data workflows later if we don't get it right the first time?
- What are the risks of moving too slowly? How long would it take to catch up to a competitor who tackles their Gen AI data readiness challenge before we do?
- How much damage would result if we automate decisions and functions based on hallucinations?
- Where do we feel stuck? And how could we leverage a proven data approach (i.e., data readiness for Cloud) to get us un-stuck?

Determine the risk of faulty data and how to avoid generating wrong answers that sound real in Gen AI

- How do we make our data ready in a way that prevents hallucinations?
- How do we define data accuracy?

Explore the need to account for emerging regulations and standards, and to find a partner who can help

- How do we build flexibility into our approach?
- Which vendors have the best track record in flexible, trusted data readiness?
- How can that vendor help us find the fastest path to real value from Gen AI?



www.actian.com



ACTIAN[™]
a division of HCLSoftware

Actian makes data easy. We deliver cloud, hybrid, and on-premises data solutions that simplify how people connect, manage, and analyze data. We transform business by enabling customers to make confident, data-driven decisions that accelerate their organization's growth.