

Six steps to data reliability

When it comes to unlocking value from your organizational data, data quality has always been a challenge. Even in our AI era, more than half of those responsible for data still don't trust it.¹ The lack of trust has significant consequences—from poor decision-making to increased operational costs. With the acceleration of AI adoption and a rapidly evolving regulatory landscape, manual data quality processes are no longer viable. Successful organizations are turning to machine learning and automation to ensure data quality. Here are six essential steps to ensure data reliability and foster trusted decision-making.



Step 1 Profile the data

Start by discovering all your data sources and classifying them, with a particular focus on identifying sensitive data types.

1

55% of those responsible for data don't trust it²



Step 2 Define data policies and business rules

Define policies and business rules that guide how different data types should be handled.

2

37% of CDAOs say their efforts to improve data quality have been successful³



Step 3 Detect anomalies

Establish a baseline for normal behavior within your data. Use machine learning algorithms to monitor data continuously and detect deviations from this baseline.

3

1/5 Nearly one fifth of a data scientist's time is spent discovering suitable data sets⁴



Step 4 Monitor for impact

Correlate anomalies with unintended changes and other events to identify the root cause and assess the potential impact.

4

67% of executives cite potential errors as the top risk of AI adoption⁵



Step 5 Notify key experts

Provide contextual alerts to relevant stakeholders—including data engineers, analysts, business leaders and compliance officers—to initiate remediation processes.

5

Poor data quality costs organizations \$12.9m annually⁶



Step 6 Optimize continuously

Evolve your data policies, rules, and reports based on insights gained from monitoring and anomaly detection.

6

AI systems are **only as good as the data that informs them**⁷



[Read our ebook](#) to learn how to get your organization to data reliability in the AI era

1. Source: <https://www.idc.com/getdoc.jsp?containerId=US51397423>

2. Source: [Source: IDC PlanScape: Data Quality Management, 2023](#)

3. Source: [Wavestone: Data and AI Leadership Executive Survey 2024](#)

4. Source: [Dataset Discovery and Exploration: A Survey | ACM Computing Surveys](#)

5. [OFO Drive: Flawed data ranked as top AI risk](#)

6. <https://www.idc.com/getdoc.jsp?containerId=US51397423>

7. [Forbes: A data gap continues to inhibit artificial intelligence, 2023](#)