

Vendor Technologies Supporting Modern Data Quality Architecture Principles

Version 1.0

Copyright © 2026 Grumpy Old IT Guy Publishers
Angelo R. Bobak

Introduction

Modern data quality architectures are no longer dependent on a single cleansing tool or isolated quality process. Instead, leading platforms distribute quality responsibilities throughout the data ecosystem by embedding validation, governance, monitoring, automation, and intelligence directly into data pipelines and analytical environments. The following vendor technologies represent examples of platforms that align with one or more of the key architectural principles of modern data quality management.

Vendor / Technology	Primary Data Quality Capabilities	Architectural Principles Supported
Microsoft Purview	Enterprise data cataloging, metadata management,	Metadata-Driven Governance – Uses metadata, lineage, and classification to

Vendor / Technology	Primary Data Quality Capabilities	Architectural Principles Supported
Microsoft Fabric	lineage tracking, classification, governance workflows, and integration with Microsoft data platforms.	improve visibility and governance. Distributed Responsibility – Connects governance across multiple data environments.
	Unified analytics platform supporting data engineering, Lakehouse architectures, pipelines, real-time analytics, AI integration, and governance capabilities.	Cloud-Native Scalability – Designed for large-scale cloud analytics environments. Continuous Monitoring – Supports monitoring of data pipelines and analytical workloads. AI-Augmented Quality – Integrates AI capabilities throughout the data lifecycle.

Vendor / Technology	Primary Data Quality Capabilities	Architectural Principles Supported
Databricks Lakehouse Platform	Data engineering, machine learning, Delta Lake quality controls, automated monitoring, and AI-driven analytics.	Lakehouse Scalability – Designed for enterprise-scale data processing. AI-Augmented Quality – Uses machine learning and AI capabilities to analyze and improve data. Distributed Responsibility – Quality checks can be embedded directly into pipelines.
	Open-source data quality framework supporting automated validation rules, expectations, profiling, and pipeline integration.	Distributed Responsibility – Quality rules are embedded directly into data workflows. Continuous Monitoring – Automated validation can occur throughout data movement.

Vendor / Technology	Primary Data Quality Capabilities	Architectural Principles Supported
Monte Carlo Data Observability Platform	Data observability, anomaly detection, lineage analysis, incident management, and automated monitoring.	Continuous Monitoring – Focuses on real-time detection of data reliability issues. AI-Augmented Quality – Uses machine learning to detect unusual data behavior.
Collibra Data Intelligence Platform	Data governance, business glossary, stewardship workflows, metadata management, and policy management.	Metadata-Driven Governance – Uses metadata and business context to manage enterprise information. Distributed Responsibility – Enables business and technical collaboration.
Informatica Intelligent Data Management Cloud (IDMC)	Data quality, master data management, metadata intelligence,	AI-Augmented Quality – Applies AI to improve discovery, matching, and automation.

Vendor / Technology	Primary Data Quality Capabilities	Architectural Principles Supported
AWS Glue Data Quality	Automated data profiling, quality rules, validation, and integration with AWS analytics services.	Cloud-Native Scalability – Designed for large-scale cloud data environments. Distributed Responsibility – Quality checks can be incorporated into AWS data pipelines.

Mapping Technologies to the Five Modern Data Quality Principles

No single technology can deliver enterprise data quality on its own. Instead, organizations achieve lasting success by aligning the right tools with the fundamental principles of profiling, cleansing, standardization, governance, and continuous monitoring.

This section illustrates how modern data management technologies support each of these principles, enabling organizations to build a scalable, integrated, and sustainable data quality architecture.

Architectural Principle	Representative Technologies	How the Principle Is Implemented
Distributed Responsibility	Microsoft Fabric, Databricks, Great Expectations, Informatica, Talend	Data quality rules are embedded within ingestion pipelines, transformation processes, and analytical workflows instead of relying on a single centralized cleansing application.
Metadata-Driven Governance	Microsoft Purview, Collibra, Informatica, Ataccama	Metadata, lineage, classifications, and business definitions provide the foundation for understanding and governing enterprise data.
AI-Augmented Quality	Databricks, Microsoft Fabric, Informatica, Ataccama, Monte Carlo	Artificial intelligence assists with anomaly detection, duplicate identification, rule recommendations, and predictive quality scoring.

Architectural Principle	Representative Technologies	How the Principle Is Implemented
Continuous Monitoring	Monte Carlo, Microsoft Fabric, Databricks, Ataccama	Automated monitoring detects changes in data behavior, pipeline failures, quality degradation, and emerging issues.
Cloud-Native Scalability	Microsoft Fabric, Azure, AWS Glue, Databricks, Snowflake ecosystem	Data quality capabilities operate across large-scale cloud data platforms, lakehouses, and distributed processing environments.

Important Consideration for Organizations

No single vendor platform completely solves enterprise data quality. Successful implementations typically combine multiple capabilities, including governance, metadata management, automated validation, observability, stewardship workflows, and AI-assisted analysis. The selected technology should support the organization's data architecture, regulatory requirements, business processes, and maturity level.

The fundamental lesson remains consistent: **modern data quality is not a standalone application—it is an enterprise capability embedded throughout the entire data lifecycle.**

© 2026 Angelo R. Bobak
All rights reserved.

First Edition: February 2026

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written permission of the publisher.

Grumpy Old IT Guy Publishers
Westchester County, NY, USA
www.grumpyolditguy.com