

TDWI Snapshot Series

2022 State of Data Quality

By James Kobielus

A survey of more than 1,000 executives and business users reveals key factors that help today's companies succeed with data quality

Sponsored by:

 ataccama

 RESEARCH

2022 State of Data Quality

A survey of more than 1,000 executives and business users reveals key factors that help today's companies succeed with data quality

By James Kobielus

© 2022 by TDWI, a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or part are prohibited except by written permission. Email requests or feedback to info@tdwi.org.

Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies. Inclusion of a vendor, product, or service in TDWI research does not constitute an endorsement by TDWI or its management. Sponsorship of a publication should not be construed as an endorsement of the sponsor organization or validation of its claims.

This report is based on independent research and represents TDWI's findings; reader experience may differ. The information contained in this report was obtained from sources believed to be reliable at the time of publication. Features and specifications can and do change frequently; readers are encouraged to visit vendor websites for updated information. TDWI shall not be liable for any omissions or errors in the information in this report.

Table of Contents

Executive Summary	2
Research Methodology	2
What Is Data Quality Management?	3
What Is the State of Enterprise Data Quality Management?	4
What Principal Challenges Do Organizations Encounter in Managing Data Quality?	5
What's Required to Establish and Maintain a Mature Enterprise Data Quality Management Practice?	8
What Are the Key Success Factors in Mature Data Quality Management?	14
Recommendations	16
About Our Sponsor	19
About TDWI Research	19
About the Author	19

Executive Summary

Enterprises rely on robust data quality management (DQM) practices to address challenges and seize strategic and operational advantages in today's business environment. TDWI recently conducted a survey of DQM professionals, and three key findings emerged:

- **DQM success factors span deployment, infrastructure, and modernization.** Successful DQM depends on standardization around consistent enterprise DQM practices, a deep stack of DQM functional infrastructure, and DQM modernizations such as AI-automated DQM workflows and reliance on shared data catalogs in support of all functions.
- **Automated DQM is still not universally adopted.** Large enterprises have been more likely than small and midsize businesses to have adopted automation-intensive AI-powered DQM, and hence to have already achieved high DQM maturity.
- **Almost seven in ten enterprises—69% of those who responded to TDWI's practitioner survey—have begun their DQM maturity journeys but have not yet achieved high maturity.** Many enterprises are midway through their DQM maturity journeys, while a substantial number still haven't begun, either because the matter is under consideration or they have no specific plans to take the first step.

Research Methodology

For this study, TDWI conducted an online survey of business technology practitioners and other stakeholders to:

- Assess the state of DQM in the enterprise
- Determine the extent to which organizations have adopted mature DQM practices
- Identify the key features of mature practices that make the greatest contribution to successful DQM

The survey, which was live from late Q1 to early Q2 2022, netted 1,061 responses. Respondents fit the following profile:

Responsible for enterprise IT. Most respondents are corporate IT professionals, including those specializing in business intelligence (BI) and data warehousing (DW). These made up 63% of responses, with business sponsors or users constituting 19%. Fewer than one in five respondents are not responsible for business IT, BI, or DW, with independent consultants or systems integrators making up 12%, vendors 4%, and academics 2%.

Experienced in data quality management (DQM). Most respondents have DQM experience. Nearly four-fifths (79%) say they have performed DQM tasks, have developed applications that rely on DQM, and/or use BI, analytics, and other applications

that rely on DQM. One in five (19%) has only observed DQM from a distance. Just 2% have no experience with DQM.

Employed by enterprises both large and small. Almost one-half (49%) of those completing the survey come from large enterprises with over \$1 billion in annual revenue. Respondents from midsize firms—those with annual revenues between \$100 million and \$999 million—made up almost three in ten (30%) respondents; small businesses represented 16% of respondents; 5% did not answer the question.

Operating in diverse vertical markets. Respondents work in a wide range of industries. Only software/internet (17%), financial services (12%), and retail/wholesale/distribution (11%) industries exceeded 10% of respondents.

Located throughout the world. Respondents from North America made up half of the total, representing 50%. Asia represented slightly over one-quarter (27%), Europe around one-seventh (16%), and Australia/New Zealand about 5%.

What Is Data Quality Management?

Data is the foundation of modern business. Enterprises are ingesting data from a growing range of platforms, including both cloud and on-premises sources.

By adopting mature data quality management best practices, enterprises are assuring users that their data—a precious business asset—remains trustworthy, relevant, and actionable. Keeping data high quality requires enforcement of robust DQM policies and procedures.

DQM—often known as *stewardship*—is the set of procedures, policies, and processes an enterprise uses to maintain trustworthy data in a data warehouse as a system of record, golden record, master record, or single version of the truth. First, however, the data must be cleansed using a structured workflow involving profiling, matching, merging, correcting, and augmenting source data records. DQM workflows must also ensure that the data's format, content, handling, and management are kept in compliance with all relevant standards, regulations, and requirements.

Modern DQM tools automate profiling, monitoring, parsing, standardizing, matching, merging, correcting, cleansing, and enhancing data for delivery into enterprise data warehouses and other downstream repositories. The tools enable creating and revising data quality rules. They support workflow-based monitoring and corrective actions, both automated and manual, in response to quality issues. Stewardship workflows enable nontechnical users to identify, quarantine, assign, escalate, resolve, and monitor data quality issues. They provide visualizations and other analytics that give stewards and other stakeholders insights into data quality and assist in correcting any issues discovered.

What are your organization’s business objectives, goals, and/or metrics for data quality? (Select all that apply.)

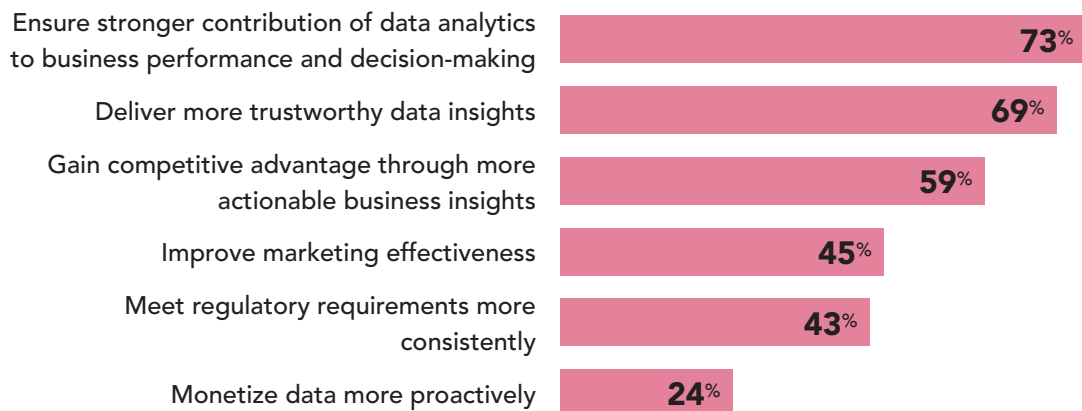


Figure 1. Enterprise DQM business objectives.

What Is the State of Enterprise Data Quality Management?

The state of enterprise data quality varies widely by organization. We asked respondents to break out the background, challenges, progress, and success of their enterprises’ DQM initiatives.

Organizations see data quality as a key pillar of business success (see Figure 1). When asked what their organization’s business objectives, goals, and/or metrics for data quality are, more than half of respondents say they include ensuring a stronger contribution of data analytics to business performance and decision-making (73%), delivering more trustworthy data insights (69%), and gaining competitive advantage through more actionable business insights (59%).

Organizations take data quality seriously. When asked about the priority their organization gives to roles, processes, tools, and platforms for ensuring high-quality data, nearly all respondents (97%, not shown) say their organization considers it an extremely or somewhat important priority. Only 2% say these elements are not important.

Organizations vary widely in impetus—internal versus external mandates—for data quality management (see Figure 2). We asked to what extent data quality initiatives were a response to external regulatory and legal mandates at respondents’ organizations. The responses were nearly split in half: 47% say their organizations’ data quality initiatives are either exclusively or mostly a response to internal strategic, business, and operational imperatives, and 53% say their DQ initiatives are either exclusively or frequently a response to external mandates.

To what extent are your organization's data quality initiatives a response to external regulatory and legal mandates?

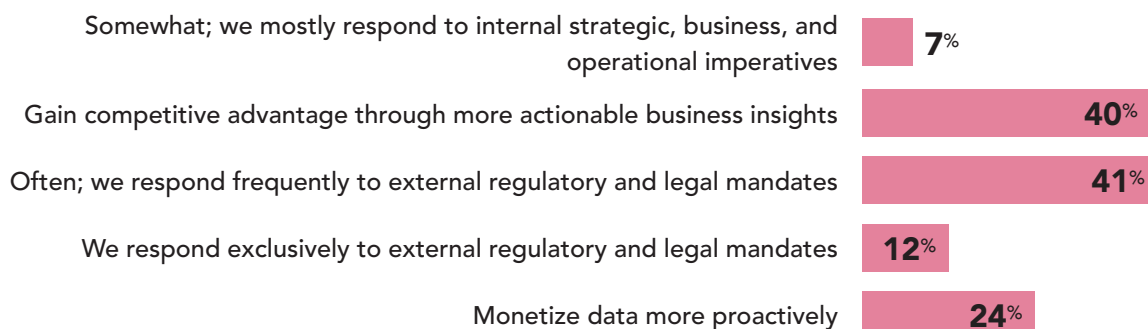


Figure 2. The drivers for enterprise DQM.

Whether internally or externally driven, the good news is that organizations are largely successful in managing data quality. When asked about their success, slightly less than one-half of respondents (46%) report that their organizations have been nearly completely or quite successful at managing data quality (see Figure 3). Close to one-third (30%) have been somewhat successful. Only 4% say they almost always face problems with data quality, and 20% often face such problems.

How successful has your organization been at managing data quality?

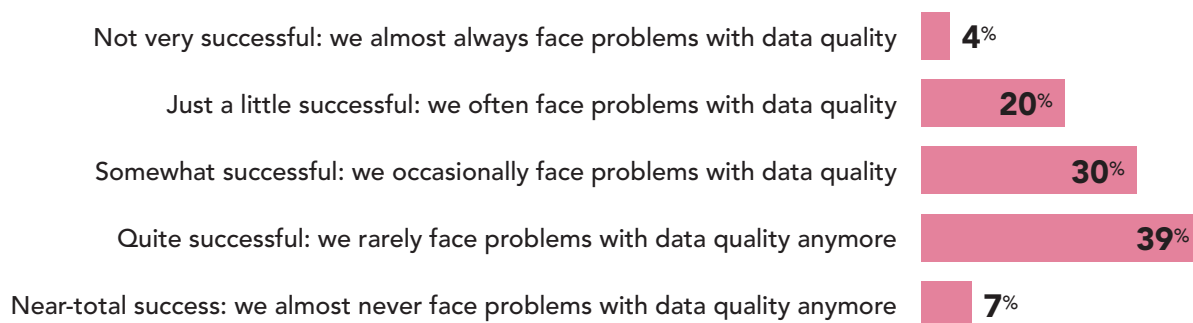


Figure 3. Enterprise success managing data quality.

What Principal Challenges Do Organizations Encounter in Managing Data Quality?

Business culture issues loom large in organizations' data quality management initiatives. Even respondents who are largely successful at DQM are reporting challenges in managing data quality. This is not surprising given the diversity of data types and the volume of data.

Enterprises face numerous DQM challenges. These include both technical issues associated with the complexities of the requisite DQM infrastructure, tools, and applications as well as cultural and management issues that may impede an enterprise on its journey to DQM maturity.

Maintaining a sustainable DQM operation requires robust technical infrastructure. The principal enterprise challenges in this regard include:

- Scaling DQM to the entire enterprise
- Mitigating DQM vulnerabilities associated with self-service data analytics applications
- Applying DQM processes to new sources of semistructured and unstructured data
- Enforcing consistent DQM across hybrid clouds, multiplatform computing, and other complex enterprise data environments
- Migrating DQM processes, platforms, applications, and workloads to modern cloud-native platforms

DQM's cultural and management challenges are just as multifaceted. They include:

- Convincing employees to adhere to DQM policies
- Training employees in DQM compliance
- Creating clear and usable DQM policies
- Keeping the DQM bureaucracy lean and agile
- Receiving consistent executive support for DQM
- Staffing the DQM board effectively
- Policing employee behavior relative to DQM
- Coordinating multiple DQM boards and programs

We asked "What challenges has your organization encountered managing data quality?" Respondents state that the principal DQM challenges span technical, cultural, and management issues (see Figure 4).

Technical. The top challenge reported is "identifying sensitive data via quality metadata, cataloging, and lineage" (44% of responses). This is clearly a technical matter with management and cultural overtones. It requires investment in DQM infrastructure and tools that combine data profiling, metadata management, data catalogs, and data lineage analysis.

Cultural. Close behind in importance is "creating quality policies that are clear and useful." Reported by 43% of respondents, this comes down to ensuring that the policy guidance offered to data stakeholders, as well as what's implemented in DQM tools, makes sense and is workable in everyday operations while minimizing the need for

specialized DQM training or excessive technical support assisting users in applying DQM policies.

Management. Near the top of the list, indicated in 40% of responses, is “keeping data quality bureaucracy lean and agile,” followed closely (at 38%) by “receiving strong and consistent executive support.” What’s clear from these results is that enterprise DQM professionals want their organizations to commit top-to-bottom to this practice and be able to scale it without greatly expanding the requisite headcount.

What challenges has your organization encountered managing data quality? (Select all that apply.)



Figure 4. Enterprise DQM challenges.

What's Required to Establish and Maintain a Mature Enterprise Data Quality Management Practice?

The state of enterprise data quality is improving thanks to a steady flow of investments in human resources, intelligent platforms and tools, and innovative practices.

These efforts help organizations achieve greater maturity in their DQM initiatives. Enterprises begin their DQM maturation journey when they begin to plan a DQM team, process, and platform for at least one business unit, functional group, or data domain. At the very least, starting the DQM journey requires assigning data stewardship responsibilities to business and/or technical personnel who will be equipped from the start with self-service tools for data quality reporting, analysis, and dashboarding.

Enterprises advance to intermediate DQM maturity when they execute on these initial plans at the business-unit level. Operationalizing DQM teams, processes, and platforms usually involves building on existing ETL tools by investing in data governance, master/reference data management, data compliance, metadata management, and other enabling infrastructure and tools.

Enterprises attain high maturity in DQM when their practices have achieved:

Sustainably high commitment and success. In a mature DQM practice, the organization has top-to-bottom commitment to data quality (rating it “extremely important”) and has achieved consistent success (rated as “near-total success” or “quite successful”) in keeping data correct, consistent, and trustworthy for all intended uses.

Standardized deployment. In a mature DQM practice, an enterprise has deployed a standard, sustainable DQM team, process, and platform. This practice may span all lines of business, functional groups, and application domains, or it may comprise decentralized DQM practices in various business units consistent with a common enterprise-wide set of DQM policies, procedures, and structures.

Deep infrastructure. In a mature DQM practice, an enterprise has implemented a comprehensive range of functional DQM infrastructure. This may include standard software for data quality management (such as data profiling, data preparation, or data cleansing and transformation tools) and may also include more advanced tools, such as a data catalog or automated tools (more about that later in this report) for monitoring data quality.

| In a mature DQM practice, an enterprise has deployed a standard, sustainable DQM team, process, and platform.

Intelligent modernization. In a mature DQM practice, an enterprise has modernized its infrastructure and processes by investing in greater AI/ML-driven automation, team-based stewardship, sophisticated policy management, and closed-loop issue remediation. These intelligent modernizations build on the self-service and real-time

DQM capabilities found in earlier stages in the maturity journey. Their effectiveness depends on implementation of a data catalog within an enterprise-wide DQM team, process, and platform.

Figure 5 lists the criteria TDWI used to determine the stage of the DQM maturity journey in which each respondent fell.

Criteria for Data Quality Management Maturity Tiers

HIGH	
Commitment and Success	Have indicated that their DQM practice is extremely important and has achieved near-total success or is quite successful
Deployment	Have deployed an enterprise-wide data quality management team, process, and platform
Infrastructure	Intermediate infrastructure profile plus add-on: data catalog
Modernization	Intermediate modernization profile plus add-ons: using AI to automate more data quality management processes; centralized data quality policy and rules library; collaborative data quality stewardship, management, and governance with a shared data catalog as a hub; and human-in-the-loop data quality exception handling augmented by AI-driven contextual recommendations

BEYOND INTERMEDIATE	
Commitment and Success	Have indicated that their DQM practice is less than extremely important and that their DQM practice has been less than quite successful
Deployment	Same as intermediate profile
Infrastructure	Same as high profile
Modernization	Same as high profile, including at least one and possibly all add-ons

(continued on next page)

INTERMEDIATE

Commitment and Success	Have indicated that their DQM practice is less than extremely important and that their DQM practice has been less than quite successful
Deployment	Have deployed data quality teams, processes, and platforms that are dedicated to specific business units, application domains, data sets, and/or regions
Infrastructure	Same as beginner profile plus add-ons: data governance, master/reference data management, data compliance, data privacy, data security, data modeling, and metadata management
Modernization	Same as beginner profile plus add-on: real-time data quality monitoring

BEYOND BEGINNER

Commitment and Success	Have indicated that their DQM practice is less than extremely important and that their DQM practice has been less than quite successful
Deployment	Same as beginner profile
Infrastructure	Same as intermediate profile plus at least one but not all add-ons
Modernization	Same as intermediate profile

BEGINNER

Commitment and Success	Have indicated that their DQM practice is less than extremely important and that their DQM practice has been less than quite successful
Deployment	Implementing data quality team, processes, and platforms at the enterprise and/or business-unit level, but don't currently have a deployed data quality management practice anywhere in the organization
Infrastructure	ETL/data integration/data engineering, data profiling, data preparation/transformation, and data cleansing/augmentation
Modernization	Self-service data quality reporting, analysis, and dashboarding

Figure 5. DQM maturity level criteria.

From the survey responses, 16% of respondents fit the criteria of high DQM maturity and slightly over half are in the intermediate stage. Fewer (about 15%) are beginners (see Figure 6). These are described in more detail in the next section.

Enterprise data quality management maturity

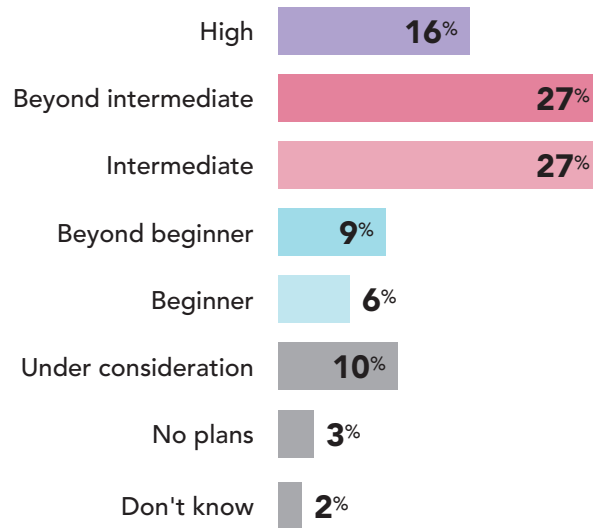


Figure 6. Enterprise progress in their DQM maturity journey.

Our survey asked respondents to break out the components of their enterprises' DQM maturation journeys. Here are the key findings.

Progress on the DQM maturity journey is mixed. Organizations vary widely in their progress on the DQM journey.

- **Not yet begun to consider DQM.** About one in seven respondents hasn't yet begun a DQM maturity journey. A small group of respondents (3%) have no plans for managing data quality programs. One in ten (10%) has it under consideration; 2% don't know if their organizations have taken the first step or are even considering such a journey.
- **Have begun but not yet deployed their DQM practices.** Enterprises that are just beginning their DQM maturity journeys constitute about one in seven (15%) respondents. These organizations are implementing a DQM team, process, and platform at the enterprise or business-unit level but haven't yet deployed them.
- **Stand midway or are well along toward high maturity in their DQM journey.** Two in five respondents (40%, not shown) have deployed and are managing a data quality team, process, and platform dedicated to specific business units, application domains, data sets, and/or regions. Slightly less than one-third (30%) have deployed and are managing an enterprise-wide DQ management team, process, and platform.

DQM is almost evenly split between manual and automated. Automation is a hallmark of high maturity in enterprise DQM. Organizations range widely on the maturity spectrum in their adoption of automation in managing data quality. When asked what percentage of their organization’s DQM functions are manual versus automated, few respondents were entirely at one end or the other of this spectrum. One-half (50%, not shown) say their DQM functions are entirely manual, largely manual, or more manual than automated, and a similar percentage (48%) say theirs are entirely automated, largely automated, or more automated than manual.

Enterprises are adopting a deep stack of DQM functional infrastructure and tools. Organizations mature in DQM when they add specialized infrastructure and tools for this purpose to supplement or extend their existing data integration and engineering platforms. Many of the respondents predominantly use existing data preparation, data engineering, and ETL tools to manage data quality (see Figure 7).

What platforms, tools, and other technical infrastructure does your organization use to manage data quality? (Select all that apply.)

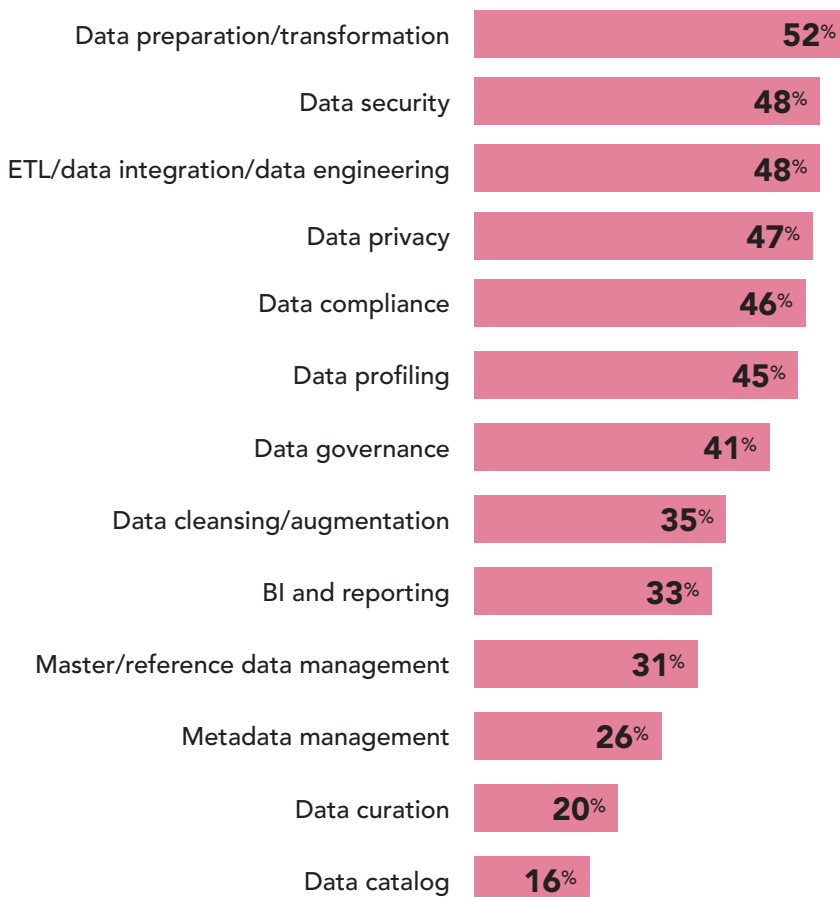


Figure 7. Implementation of a deep stack of DQM functional capabilities.

Our survey found that:

- Close to one-half report using data preparation/transformation (52%), ETL (48%), and data profiling (45%) tools for DQM
- Specialized data governance (41%), data cleansing/augmentation (35%), and master data management (31%) tools are used in around one-third of instances
- Metadata management (26%) and data curation (20%) tools are each used in around one-quarter of respondents and data catalogs are used by one in seven (16%)

Organizations are implementing a wide range of technology modernization tools and techniques to improve their management of data quality. Respondents told us their organizations have adopted the following:

- **Collaborative data quality stewardship** refers to an online team environment that provides the workflow, sharing, and other collaboration features necessary for business and technical professionals to work together on data quality management functions. It has been implemented by 58% of respondents (figure not shown).
- **Self-service data quality reporting, analysis, and dashboarding** refers to DQM tools that allow business and technical personnel to configure, consume, analyze, and share data quality information unassisted. It provides a task-focused user experience through low-code visual tools that can be customized to the specific functions of each user and stakeholder. It has been implemented by 44% of respondents.
- **AI-driven DQM automation** refers to the use of embedded machine learning models to handle one or more DQM functions reliably, repeatedly, and accurately without needing direct human oversight and assistance. It has been implemented by 40% of respondents, with large enterprises more likely to have implemented AI-driven automated DQM (50% of respondents with revenues greater than \$1 billion report having implemented it) than midsize businesses (36% of respondents with revenues between \$100 million and \$999 million) and small businesses (23% of respondents with revenues less than \$100 million).
- **Human-in-the-loop DQ exception handling** refers to a DQM automation capability under which human data stewards are alerted to data quality issues and guided individually or in a structured workflow to resolve these issues, aided by AI-driven contextual recommendations presented to them in their respective self-service data quality reporting, analysis, and dashboarding environments. It has been implemented by 29% of respondents.
- **Centralized data quality policy and rules library** refers to a shared resource that enables teams of DQM professionals to author, revise, introspect, and supervise the business rules embedded in—and enforced by—one or more DQM functional components. It has been implemented by 39% of respondents.

- **DQ metrics** refers to a dashboard that enables business and/or technical professionals to track how well the deployed DQM environment has been delivering on its intended business outcomes, such as correction of inaccurate data and speed at resolving inconsistencies among two or more data sources. It has been implemented by 37% of respondents.

What Are the Key Success Factors in Mature Data Quality Management?

TDWI grouped respondents who self-identified as successful, somewhat successful, or not successful into three groups and compared them against each other to explore success factors. When we analyzed the extent to which the success of an enterprise DQM practice is correlated with these factors, the findings were clear.

DQM automation delivers results. Automation is one of the paramount success factors for DQM. Companies with successful DQM initiatives rely more on automated processes (70% of these organizations run DQM in an automation-intensive manner) than organizations with somewhat successful initiatives (41% of these are automation-intensive) or unsuccessful ones (19% of them are automation-intensive). This correlation of AI-driven automation in DQM to success in managing data quality proved statistically significant (see Figure 8.)

Even though they are using AI-driven DQM automation now, successful enterprise DQM practices are also more likely to be planning to expand their use of it in the next 12 months than those who are only somewhat successful or unsuccessful. In other words, they are continually looking to do better.

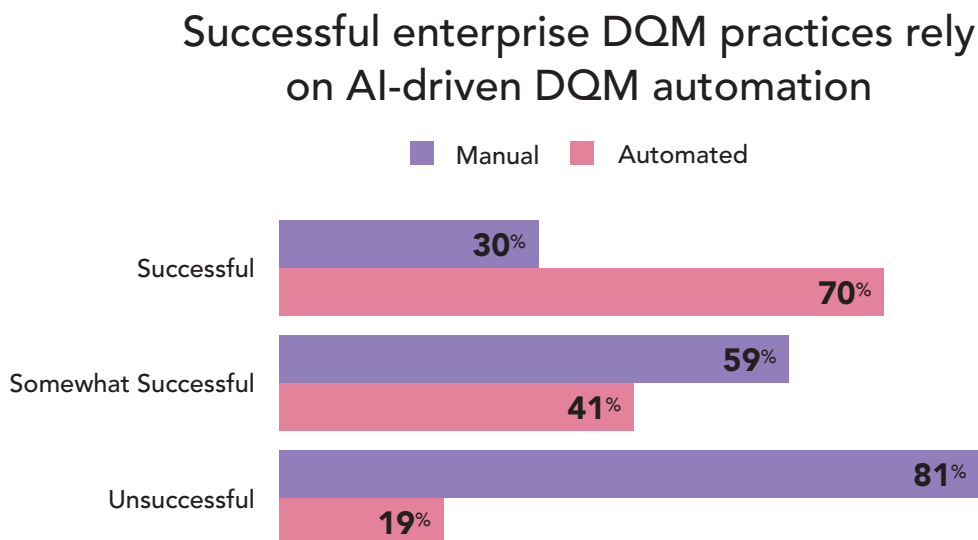


Figure 8. The percentages of manual versus automated functions in enterprises with unsuccessful, somewhat successful, and successful data quality management practices.

Organizations are placing a high priority on AI-driven automation of all data management processes, with an emphasis on data compliance and privacy. About one-half of respondents report that their top priorities include data management automation (52%) and/or AI/machine learning (50%) (figure not shown). They would like to use AI to automate a wide range of DQM functions, with priorities encompassing automation of DQM monitoring, cleansing and standardization, anomaly detection, and on-demand evaluation.

Data catalogs play a central role in DQM plans. Another key infrastructure element for DQM is a data catalog. When asked to discuss how they're addressing the success factors in their enterprises' DQM journeys, respondents are also placing a high priority on deployment of centralized data catalogs. Close to half of respondents (45%, not shown) say their organizations are implementing a central data catalog in the next 12 months. Respondents who reported that they have been successful in managing data quality were more likely to be planning to implement a data catalog for DQM over the coming 12 months.

The shift toward standardized DQM deployment is a clear trend. Standardization of how DQM is operationalized is a trend revealed in the survey. Given the wide range of companies responding to the survey, it's no surprise that they have disparate organizational models in mind for their DQM practices. Slightly more than two in five (43%) respondents either have an enterprise-wide DQ team or plan to create one; more than one-half (51%) either already deploy business-unit-level DQ teams or plan to do so (figure not shown). However, there wasn't a statistically significant correlation between having implemented enterprise-wide standardization of DQM practice and having achieved success in management of data quality.

Close to half of respondents say their organizations are implementing a central data catalog in the next 12 months.

Other DQM modernizations. In addition to AI-driven DQM automation and plans for central data catalogs, respondents who say they are successful at DQM are also more likely to be implementing closed-loop data quality remediation (which accelerates responses to DQM issues through real-time alerting and response mechanisms within the DQM stewardship workflow) and DQM observability and monitoring (which supports continuous measurement of DQM metrics such as data completeness, consistency, and accuracy, as well as unified views of data logged on systems throughout the DQM pipeline).

Recommendations

Enterprises rely on mature DQM practices to keep their data assets continually trustworthy, relevant, and consumable for a wide range of uses.

Data management professionals play an important role in moving their organizations up the maturity curve toward a sustainable, standardized, scalable, and efficient DQM practice. With the broader business stakes in mind, DQM professionals should heed these strategic recommendations:

Assess your enterprise's requirements for comprehensive DQM. When identifying your enterprise's DQM requirements, remember that this practice focuses on ensuring that business data remains trustworthy, relevant, and consumable for all business use cases. Though it might be managed with some of the same tools, platforms, and personnel, DQM is distinct from data security and data protection.

Assess your enterprise's DQM maturity level. Enterprises must assess the gap between evolving needs and current capabilities in DQM. Data professionals need to assess the extent to which their current DQM practices ensure the ready availability of correct, current, clean, and actionable data for all business analytics and other uses. They should also consider the extent to which these practices enable organizations to respond fully to compliance mandates related to such issues as privacy protection and anti-bias.

To determine the maturity of their current DQM practices, enterprises should conduct an audit to determine the incidence of rogue data sets, untagged data assets, and any other quality and actionability issues associated with analytics sandboxes, data lakes, and self-service data practices.

Though it might be managed with some of the same tools, platforms, and personnel, DQM is distinct from data security and data protection.

Establish a sustainable enterprise-wide DQM practice. Before evaluating commercial solutions, DQM champions' first order of business should be to kick-start their enterprises on their journey to maturity. Ideally, an enterprise should have steadfast DQM sponsorship at the C-level and a budget to fund these ongoing activities, as well as any necessary investments in enabling tools and platforms. First and foremost, they should advocate for creating an enterprise-wide coordinating council and operational team as well as the necessary processes, platforms, and tools for managing these functions.

To avoid creating a rigid governance bureaucracy, the processes of the DQM coordinating council and operational team should be kept lean and agile. It may be wise to bring in professional services and consulting firms to standardize, establish, and promote your organization's DQM practices among stakeholders.

Instill a holistic culture for enterprise-wide DQM. Enterprises should build a culture in which all technical and business roles place a high priority on maintaining strong DQM. At the very least, this will involve orchestrating an ongoing campaign to convince employees to adhere to DQM policies. To ensure that DQM policies are put into practice, this guidance should be clear and usable, and most of its ongoing execution and monitoring should be automated with highly usable, self-service tools.

Adopt DQM platforms that use ML-driven process automation. Enterprises should automate DQM functions while augmenting the productivity of technical and business personnel who administer these functions. The most agile automation involves embedded machine learning models that proactively sense and remediate data quality deficiencies and other issues before they become showstoppers.

This can enable more consistent, continuous application of DQM policies across all data types, in all use cases, throughout the enterprise. It can also drive human-in-the-loop data quality exception handling through contextual recommendations. ML-driven DQM automation can reduce human intervention and thereby accelerate responses and boost productivity of data stewards and other operational personnel.

Implement a collaborative data stewardship and policy management environment. DQM is a workflow that involves many practitioners and stakeholders. Be sure to choose a core DQM solution platform that leverages a shared data catalog, a centralized policy, and a common rules library. The environment should support collaborative workflows among business and technical roles involved in management of policies surrounding data quality.

The environment should also incorporate embedded machine learning models that trigger closed-loop collaborative review and remediation of data quality issues and other indicators of anomalous governance events. Furthermore, it should support self-service access to DQM and policy management tasks through low-code visual tools that are customizable to the specific functions of each stakeholder.

Choose a core DQM solution platform that leverages a shared data catalog, a centralized policy, and a common rules library.

Provide DQM observability metrics to all stakeholders. DQM should be a continuous business process. Make sure your chosen solutions leverage a shared data catalog and data lineage tools to support continuous monitoring, measurement, and sending of threshold-based alerts on all data governance metrics. This can allow an enterprise to set up a continuous feedback loop to improve data trustworthiness, compliance, and other metrics.

Solutions should offer built-in monitoring and auditing capabilities to simplify and manage this process. Monitoring should encompass the entire DataOps pipeline, data

platforms, and data apps, and should extend to monitoring metrics associated with assets sourced from data marketplaces and other channels.

Stay ahead of innovation. As TDWI sees it, a highly mature enterprise DQM practice must continue to evolve in keeping with innovations in platforms, processes, and teaming models for this as an operationalized business function. Over time, the following innovations will need to be absorbed into enterprise DQM practices everywhere:

- Shared data catalogs will evolve to support search, curation, and monitoring of a wider range of data, metadata, ML models, and other artifacts relevant to DQM use cases. Data catalogs will continue to play an important role in automating and scaling DQM functions through the metadata they store. They will also be important enablers for emerging data management practices, such as the data fabric and data mesh.
- AI-automated DQM will evolve to address unstructured and other new source data, incorporate more sophisticated AI for new purposes, and drive more contextual, adaptive guidance in real time for stakeholders.
- Collaborative DQM environments will evolve to address more flexible team models, more sophisticated workflows, and a wider range of data stewardship and policy authoring and administration capabilities.
- Self-service DQM tools will be tailored to a wider range of task, outcome, role, and personalization profiles.

Recognize that DQM is an ongoing journey. In the final analysis, DQM maturity is not a “one and done” proposition. In fact, a highly mature and successful DQM practice is more of a moving target than an end state. Organizations that attain high DQM maturity in the present day may find themselves out-of-date in a few years’ time unless they make continued investments aligned with evolving business and technical trends.

About Our Sponsor



Ataccama enables enterprise data democratization with a unified platform for automated data quality, MDM, and metadata management across cloud and hybrid environments. We enable business and data teams to collaborate on creating high-quality, reusable data products and massively scale data-driven innovation while maintaining data accuracy, control, and governance.

About TDWI Research

TDWI Research provides industry-leading research and advice for data and analytics professionals worldwide. TDWI Research focuses on modern data management, analytics, and data science approaches and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of business and technical challenges surrounding the deployment and use of data and analytics. TDWI Research offers in-depth research reports, commentary, assessments, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

About the Author



James Kobielus is senior director of research for data management at TDWI. He is a veteran industry analyst, consultant, author, speaker, and blogger in analytics and

data management. He focuses on advanced analytics, artificial intelligence, and cloud computing. Kobielus has held positions at Futurum Research, SiliconANGLE Wikibon, Forrester Research, Current Analysis, and the Burton Group and also served as senior program director, product marketing for big data analytics, for IBM, where he was both a subject matter expert and a strategist on thought leadership and content marketing programs targeted at the data science community. You can reach him by email (jkobielus@tdwi.org) on Twitter ([@jameskobielus](https://twitter.com/jameskobielus)) and on LinkedIn (<https://www.linkedin.com/in/jameskobielus/>).

Featured Ataccama Partners

Ataccama would like to thank its partners for their ongoing support in creating a world where enterprise data is easier to manage.



TDWI Research provides research and advice for data professionals worldwide. TDWI Research focuses exclusively on data management and analytics issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of data management and analytics solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.



A Division of 1105 Media
6300 Canoga Avenue, Suite 1150
Woodland Hills, CA 91367

E info@tdwi.org

tdwi.org