THE COMPUTERTECH
(An International Peer Review Journal)
VOLUME 12; ISSUE 1 (JAN-JUNE); (2026)
WEBSITE: THE COMPUTERTECH

The Socio-Organizational Dimension of Data Quality: People, Processes, and Culture

Sivananda Reddy Sattiraju[1*]
[1]Trine University, USA

Abstract

Data quality is commonly understood as a technical challenge requiring measurement, validation, and correction of data errors. However, this perspective overlooks the fundamental socio-organizational dimensions that shape data quality outcomes. This review examines the role of people, processes, and organizational culture in determining data quality, drawing on literature from information systems, public administration, and organizational studies. We analyze how human factors—including training, skills, motivation, and behavior—influence data quality at the point of collection and throughout the data lifecycle. We examine the importance of process design, with particular attention to how data quality requirements can be embedded into workflows rather than addressed through post-hoc correction. We analyze the role of organizational culture in shaping attitudes toward data quality and supporting or undermining governance efforts. The review synthesizes evidence from diverse sectors, including healthcare, public administration, and small and medium enterprises, identifying common patterns and sector-specific challenges. We propose a socio-organizational framework for data quality management that complements technical approaches by addressing the human and organizational factors that ultimately determine data quality outcomes.

Keywords: Data Quality; Socio-Organizational Factors; Organizational Culture; Process Design; Human Factors; Data Governance; Training; Public Administration

Introduction

The dominant discourse on data quality has focused primarily on technical dimensions: accuracy, completeness, consistency, timeliness, and other measurable attributes of data. This technical perspective has generated valuable tools and methodologies for assessing and improving data quality. However, it has also obscured the fundamental socio-organizational dimensions that shape data quality outcomes. Data do not exist in a vacuum; they are created, maintained, and used by people within organizations and social contexts. The quality of data reflects not only the characteristics of the data themselves but also the capabilities, practices, and culture of the people and organizations that produce and use them.

This review examines the socio-organizational dimension of data quality, drawing on literature from information systems, public administration, and organizational studies. We analyze how human factors—including training, skills, motivation, and behavior—influence data quality at the point of collection and throughout the data lifecycle. We examine the importance of process design, with particular attention to how data quality requirements can be embedded into workflows rather than addressed through post-hoc correction. We analyze the role of organizational culture in shaping attitudes toward data quality and supporting or undermining governance efforts.

The evidence demonstrates that technical solutions alone are insufficient for ensuring data quality. Organizations that invest in sophisticated data quality tools but neglect the human and organizational dimensions continue to struggle with data quality problems. Conversely, organizations that attend to socio-organizational factors—providing adequate training, designing processes that support quality, and fostering cultures that value data—achieve better outcomes even with modest technical investments.

The remainder of this article is organized as follows. Section 2 examines human factors in data quality, including training, skills, and motivation. Section 3 analyzes the role of process design. Section 4 examines organizational culture and its impact on data quality. Section 5 synthesizes evidence from diverse sectors. Section 6 proposes a socio-organizational framework for data quality management. Section 7 addresses implications for practice and policy. Section 8 identifies future research directions. Section 9 concludes with recommendations.

Human Factors in Data Quality

People are central to data quality outcomes, as data are created, maintained, and used by human actors. The capabilities, knowledge, and behavior of these actors significantly influence data quality. Understanding these human factors is essential for developing effective approaches to data quality management.

Training is a critical determinant of data quality. Staff who are adequately trained in data entry, data management, and quality assurance produce higher quality data than those who are not. Training should cover not only technical procedures but also the importance of data quality and the consequences of poor quality. Effective training is ongoing, with refresher courses and updates as systems and requirements change. It should be practical, with hands-on practice and examples relevant to the work context.

Skills and expertise also influence data quality. Data quality management requires specialized skills, including data profiling, cleaning, validation, and governance. Organizations that lack these skills struggle to maintain data quality. Building capacity requires investment in recruitment, training, and development of data quality expertise. For many organizations, this represents a significant challenge, particularly small and medium enterprises with limited resources.

Motivation and attitudes toward data quality are equally important. Staff who understand the importance of data quality and are motivated to produce high-quality data are more likely to do so than those who view data entry as a low-priority task. Motivation can be enhanced through recognition, incentives, and clear communication of the importance of data quality. However, motivation must be sustained over time, requiring ongoing attention to organizational culture and leadership.

Behavioral factors also influence data quality. Staff may cut corners, make errors, or fail to follow procedures due to time pressure, fatigue, or competing priorities. Behavioral interventions, such as feedback, reminders, and prompts, can support quality. However, these interventions must be designed with an understanding of the work context and the factors that influence behavior.

Research in healthcare provides instructive examples. Studies of Ethiopian public health facilities found that inadequate training, weak governance, and insufficient infrastructure compromised the timeliness, completeness, and accuracy of care, directly affecting patient outcomes. Similarly, qualitative research in multiple healthcare contexts has documented how staff behaviors and attitudes contribute to data quality problems, with implications for patient safety and clinical decision-making.

Process Design and Data Quality

Process design profoundly influences data quality. When data quality requirements are embedded into business processes, quality is achieved more efficiently and effectively than when quality is addressed through post-hoc correction. Understanding the relationship between process design and data quality is essential for sustainable data quality management.

Embedding quality into processes means designing workflows that support quality from the start. This includes clear procedures for data entry, validation checks at the point of entry, and feedback mechanisms that identify and correct errors promptly. Embedded quality is more efficient than post-hoc correction, as it prevents errors from entering the system rather than detecting and correcting them afterward.

Information quality fragments represent one approach to embedding quality into processes. These modular components integrate checks for accessibility, completeness, accuracy, and consistency into business process models. By incorporating quality checks directly into process models, organizations can ensure that quality is addressed systematically and consistently. Validation by domain experts demonstrates the feasibility of this approach, with potential for reducing modeling time and increasing confidence in data quality.

Process design must also account for the specific characteristics of data quality dimensions. Accuracy requires validation against authoritative sources; completeness requires specification of required fields; consistency requires cross-referencing across systems; timeliness requires monitoring of update frequency and latency. Each dimension requires different process interventions, and these must be integrated into overall process design.

The relationship between process design and data quality is particularly evident in public administration. Studies of Ukrainian e-government reveal that staff skills, process design, and management practices are crucial for ensuring high-quality information in government. Similarly, research in criminal justice has shown that poor information quality undermines workflows by introducing inefficiencies and errors, highlighting the need to align process design with information system requirements from the outset.

Organizational Culture and Data Quality

Organizational culture shapes attitudes toward data quality and the effectiveness of governance efforts. Culture refers to the shared values, beliefs, and norms that characterize an organization. A culture that values data quality supports effective governance; a culture that does not undermines even well-designed technical systems.

Leadership commitment is a key cultural factor. When leaders demonstrate commitment to data quality, through resource allocation, personal attention, and accountability, the organization follows. Leadership commitment signals that data quality is a strategic priority, not an afterthought. Conversely, when leaders neglect data quality, staff receive a signal that quality is not important.

Accountability structures also reflect culture. Organizations with clear accountability for data quality—with defined roles, responsibilities, and consequences—are more likely to achieve quality than those without. Accountability must be supported by authority and resources; staff cannot be held accountable for quality without the means to achieve it.

Values and norms shape day-to-day behavior. When quality is valued and norms support quality, staff are more likely to follow procedures, correct errors, and report problems. When quality is not valued, staff may cut corners, hide errors, and prioritize other concerns.

Culture is also reflected in the approach to errors. Organizations that treat errors as learning opportunities—analyzing causes and implementing improvements—achieve better quality than those that punish errors without addressing underlying causes. A just culture, which balances accountability with learning, supports continuous improvement in data quality.

Research in public administration demonstrates that technology alone does not ensure data quality. Persistent shortcomings in government, including insufficient staff training, weak governance structures, and organizational silos, undermine data reliability. Successful e-government initiatives combine clear governance, adequate training, process redesign, and technological solutions.

Evidence from Diverse Sectors

Evidence from diverse sectors supports the importance of socio-organizational factors in data quality. In healthcare, the stakes are highest, with data quality affecting patient safety, clinical decision-making, and public health monitoring. Research consistently demonstrates that inadequate training, weak governance, and insufficient infrastructure compromise data quality. Conversely, organizations that implement comprehensive governance frameworks and invest in staff training achieve better outcomes.

In public administration, research shows that information technology alone does not ensure data quality. Organizational factors, including staff skills, governance, and culture, significantly influence data reliability. E-government initiatives that address organizational factors achieve better outcomes than those that focus solely on technical solutions.

In small and medium enterprises, socio-organizational factors are particularly important due to limited resources and technical capabilities. SMEs often lack the specialized expertise and infrastructure of larger organizations, making them more dependent on staff capabilities and organizational practices. Research on UK SMEs found that deficits in technical skills and resources limit the adoption of robust practices. However, SMEs that prioritize data quality and invest in staff training achieve competitive advantages.

In customer relationship management, socio-organizational factors have been shown to influence adoption and effectiveness. Poor validation erodes user trust and compromises adoption, particularly in SMEs facing frequent inconsistencies. Successful implementations combine standardized data entry, automated validation, continuous monitoring, and user feedback integration.

A Socio-Organizational Framework for Data Quality Management

We propose a socio-organizational framework for data quality management that complements technical approaches by addressing the human and organizational factors that ultimately determine data quality outcomes. The framework is organized around four pillars: people, processes, culture, and governance.

The people pillar addresses the human factors that influence data quality. This includes training and skill development, motivation and engagement, behavior and attitudes, and recruitment and retention. Organizations should invest in building capabilities, fostering motivation, and supporting positive behaviors. Training should be practical, ongoing, and tailored to specific roles and responsibilities. Recognition and incentives should reward quality and support continuous improvement.

The processes pillar addresses the design and implementation of workflows that support data quality. This includes embedding quality into processes, developing clear procedures, implementing validation checks, and establishing feedback mechanisms. Organizations should design processes that make it easy to produce high-quality data and difficult to produce low-quality data. Continuous improvement should be built into process design, with regular review and updates.

The culture pillar addresses the values, beliefs, and norms that shape attitudes toward data quality. This includes leadership commitment, accountability structures, quality values and norms, and approaches to error and learning. Organizations should foster cultures that value data quality, support accountability, encourage learning from errors, and promote continuous improvement.

The governance pillar addresses the structures and processes through which data quality is managed. This includes roles and responsibilities, policies and standards, monitoring and enforcement, and continuous improvement. Organizations should establish clear governance structures that provide direction, oversight, and support for data quality management.

Implications for Practice and Policy

For practitioners, the framework highlights the importance of addressing socio-organizational factors in data quality management. Technical solutions are necessary but not sufficient; organizations must also invest in training, process design, culture, and governance. Data quality management should be integrated into overall organizational management, with leadership commitment, resource allocation, and accountability structures.

For policymakers, the framework underscores the importance of addressing organizational and human factors in public sector data quality initiatives. Technical solutions alone will not improve data quality; attention must also be given to staff training, process design, governance, and culture. Policy should require and support comprehensive approaches to data quality that address both technical and socio-organizational dimensions.

Future Research Directions

Several research directions are needed to advance understanding of socio-organizational factors in data quality. Research on the effectiveness of training and development is needed to identify best practices for building data quality capabilities. Research on the relationship between organizational culture and data quality outcomes is needed to understand how culture shapes quality and how culture can be changed. Research on the effectiveness of different governance structures is needed to identify the most effective approaches for different contexts. Research on sectoral differences is needed to understand how socio-organizational factors vary across sectors and how approaches should be tailored accordingly.

Conclusion

This review demonstrates that data quality is fundamentally a socio-organizational as well as a technical challenge. The quality of data reflects not only the characteristics of the data themselves but also the capabilities, practices, and culture of the people and organizations that produce and use them. Training, skills, and motivation influence quality at the point of collection; process design shapes quality throughout the lifecycle; culture and governance support or undermine quality efforts. Technical solutions alone are insufficient for ensuring data quality. Organizations that invest in sophisticated tools but neglect socio-organizational factors continue to struggle. Conversely, organizations that attend to people, processes, culture, and governance achieve better outcomes even with modest technical investments. A socio-organizational framework for data quality management, organized around people, processes, culture, and governance, provides guidance for addressing the human and organizational factors that ultimately determine data quality. As data become increasingly central to organizational operations, decision-making, and scientific research, the importance of socio-organizational factors will only intensify. Organizations that attend to these factors will be better positioned to realize the benefits of their data and avoid the costs of poor quality. The challenge is not simply technical but fundamentally organizational and human, requiring sustained attention to the capabilities, practices, and cultures that shape data quality outcomes.

References

[1] Batini, C., & Scannapieco, M. (2016). Data and Information Quality: Dimensions, Principles and Techniques. Springer.

[2] Aluri, Y. S. (2025). Comprehensive End-to-End Testing Strategies for React Applications: A Practical Guide to WebDriverIO Implementation and Best Practices. *Journal of Computer Science and Technology Studies*, *7*(12), 237-243.

[3] Redman, T. C. (2008). Data Driven: Profiting from Your Most Important Business Asset. Harvard Business Press.

[4] Mallempati, A. (2023). Smart Data, Smart Decisions: The Future of MDM & Governance. *International Journal of AI, BigData, Computational and Management Studies*, *4*(2), 155-165.

[5] Lee, Y. W., Pipino, L. L., Funk, J. D., & Wang, R. Y. (2006). Journey to Data Quality. MIT Press.

[6] Jaladi, D. S., & Mallempati, A. (2024). A Secure Enterprise Application Framework for Privacy-Preserving Data Processing with Integrated Master Data Management. *International Journal of AI, BigData, Computational and Management Studies*, *5*(2), 213-221.

[7] Loshin, D. (2011). The Practitioner's Guide to Data Quality Improvement. Morgan Kaufmann.

[8] Olson, J. E. (2003). Data Quality: The Accuracy Dimension. Morgan Kaufmann.

[9] Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, *12*(4), 5–33.

[10] Yuvaraj, N., & Kumar, M. S. (2025). Agentic AI for Zero-Touch Customer Experience: A Governance-Constrained Framework for Autonomous Service Systems. *American International Journal of Computer Science and Technology*, *7*(2), 100-111.

[11] Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. Communications of the ACM, *45*(4), 211–218.

[12] Kumar, M. S. (2024). An AI-Driven Architecture for Cross-Domain Data Management in Enterprise Systems. *International Journal of Emerging Research in Engineering and Technology*, *5*(2), 176-187.

[13] Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. Communications of the ACM, *40*(5), 103–110.

[14] Ehrlinger, L., & Wöß, W. (2022). A survey of data quality measurement and monitoring tools. Frontiers in Big Data, *5*, 850611.

[15] Kumar, M. S. (2023). A Scalable Architecture for Automated Data Classification and Sensitive Information Discovery Using Artificial Intelligence. *International Journal of Emerging Research in Engineering and Technology*, *4*(2), 158-169.

[16] Haug, A., & Zachariassen, F. (2016). Data quality in CRM systems. Industrial Management & Data Systems, *116*(9), 1896–1914.

[17] Miller, T. (2023). A hierarchical model for data quality terminology. Data & Knowledge Engineering, *143*, 102087.

[18] Mallempati, A. (2024). Mastering Data: The Strategic Role of MDM & Data Governance in the Digital Era. *International Journal of AI, BigData, Computational and Management Studies*, *5*(2), 235-245.

[19] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, *3*, 160018.

[20] Kumar, M. S., & Yuvaraj, N. (2024). Predictive Customer Experience Orchestration Using Governed Data Pipelines and Intelligent Service Signals. *International Journal of Emerging Trends in Computer Science and Information Technology*, *5*(1), 206-215.

[21] Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lehnert, K., Nosek, B., ... & Wyborn, L. (2019). Make scientific data FAIR. Nature, *570*(7759), 27–29.

[22] Aluri, Y. S. (2024). Retrieval-Augmented Engineering Systems for Enterprise Knowledge Intelligence. *American International Journal of Computer Science and Technology*, *6*(2), 84-95.

[23] Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: Fair enough? European Journal of Human Genetics, *26*(7), 931–936.

[24] Aluri, Y. S. (2025). Frontend Performance Optimization of Large-Scale E-commerce Landing Pages: A Comprehensive Analysis. *International Journal of Computational and Experimental Science and Engineering*, *11*(4).

[25] European Commission. (2020). Turn FAIR data into reality. European Commission.

[26] GO FAIR. (2020). The FAIR Data principles. GO FAIR International Support and Coordination Office.

[27] Lamprecht, A. L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., ... & Gonzalez-Beltran, A. (2020). Towards FAIR principles for research software. Data Science, *3*(1), 37–59.

[28] Aluri, Y. S. (2025). Agentic AI Frameworks for Autonomous Enterprise Software Development Workflows. *International Journal of AI, BigData, Computational and Management Studies*, *6*(1), 217-226.

[29] Sansone, S. A., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A. L., & Thurston, M. (2019). FAIRsharing as a community approach to standards, repositories and policies. Nature Biotechnology, *37*(4), 358–367.

[30] Kumar, M. S., & Yuvaraj, N. (2024). Predictive Customer Experience Orchestration Using Governed Data Pipelines and Intelligent Service Signals. *International Journal of Emerging Trends in Computer Science and Information Technology*, *5*(1), 206-215.

[31] Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., ... & Schultes, E. (2020). FAIR principles: Interpretations and implementation considerations. Data Intelligence, *2*(1-2), 10–29.

[32] Jaladi, D. S., & Mallempati, A. (2025). A Real-Time Enterprise Application Architecture for High-Volume Data Processing with Integrated Master Data Management. *International Journal of Emerging Research in Engineering and Technology*, *6*(1), 126-134

[33] Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. ACM Computing Surveys, *41*(3), 1–52.