



Beyond Accuracy: Rethinking Data Quality as a Strategic Pillar in ERP Implementation

Rushabh Mehta

Financial Analyst, Hammerton, Inc., USA

ABSTRACT

In recent years, a significant number of manufacturing enterprises globally have adopted Enterprise Resource Planning (ERP) systems as a strategic step toward digital transformation, leveraging advancements in cloud-based technologies. ERP systems, characterized by their comprehensive database structures, support advanced capabilities such as Artificial Intelligence (AI), Big Data analytics, Machine Learning (ML), and process automation. Given their integrative potential, these systems effectively consolidate essential business functions, including Sales, Accounting, Manufacturing, Human Resources, and overall management.

Data quality emerges as a critical factor and one of the foundational pillars for the successful implementation of ERP systems. The relevance of high-quality data in ERP deployments is underscored by its direct influence on operational efficiency, departmental integration, and informed decision-making at executive levels. Poor data quality during ERP implementation can result in significant adverse effects, disrupting interdepartmental coordination, and leading to flawed strategic decisions.

This review addresses key data quality issues commonly encountered during the data migration phase, transitioning from legacy systems to modern ERP infrastructures. It highlights prominent data quality challenges, including data inconsistencies, duplication, incompleteness, and misalignment across disparate data sources. Additionally, the paper explores various methodologies and best practices for enhancing data quality, such as rigorous data cleansing, robust governance frameworks, and systematic validation procedures during migration.

Furthermore, this study emphasizes the criticality of maintaining data integrity throughout ERP implementation phases and identifies effective ERP project management practices as vital to ensuring successful system deployment. Insights drawn from recent literature and empirical case studies illustrate the strategies employed to mitigate data quality risks, ensuring the realization of anticipated ERP system benefits.

KEY WORDS: Data Quality, Data Integrity, Quality Control, ERP, Quality Assurance.

1. INTRODUCTION

-Enterprise Resource Planning (ERP) systems are software packages composed of several modules, such as human resources, sales, finance and production, providing cross- organization integration of data through embedded business processes. Many employees in different divisions like accounting and sales use the same information for their various needs. It is not practical forcing employees to maintain separate databases and spread sheets that must be manually merged to generate reports in a real-time business process. So, ERP software Offer some degree of synchronized reporting and automation. Also give a chance to organization members to pull reports from one

system. Having a portal or Dashboard which gives a chance to understand the performance of the business environment is the main feature of the ERP systems. It Help to get quick overview of current business situation. Companies need Data Quality to help make better decisions. Data Quality is generally defined as accuracy, completeness, timeliness or accessibility. Data Quality issues can take place by both conceptual and operational problems.

-Data quality can be succinctly defined as the measure of suitability or fitness of a specific dataset when applied to achieve a particular purpose or objective. It is widely recognized as a multidimensional and multifaceted concept, encompassing various attributes and characteristics that determine its usefulness. Within databases or information systems, data itself does not inherently possess quality or intrinsic value; rather, it holds potential value that is actualized only through meaningful usage or application. Thus, the concept of data quality is context-dependent, becoming fully realized and measurable when data supports informed decision-making, analytical processes, or operational activities.

- The traditional approach to data quality has predominantly focused on isolated, departmental-level practices, emphasizing data accuracy within discrete units such as Sales, Accounting, or Manufacturing. However, such an isolated approach is insufficient for enhancing decision-making at the strategic or enterprise level. Maintaining high data quality solely within individual departments does not inherently support comprehensive strategic insight or facilitate a robust cross-functional collaboration. Due to inherent limitations in data integration and consistency, such fragmented practices often lead to compromised accuracy, ultimately hindering informed and cohesive decision-making across the organization. Therefore, contemporary research advocates for a holistic and integrative data quality framework, emphasizing enterprise-wide governance, comprehensive data integration strategies, and fostering a cross-departmental culture of shared data stewardship and accountability.

-Major objective of the review is to make aware that Data Quality and Risk management is the strategic pillar for the ERP system implementation. Data Quality issues can have huge impact on financial performance and legal consequences. Quality data is imperative for success of Accounting Information system, unfortunately most of the accounting system has a significant amount of inaccurate data. Inaccurate data lie at the root of many of the most important issues of the day – from causing dissatisfaction among the end users and customers and it makes it difficult for the top management to make the decision. Data-related problems cause supplier relationships to deteriorate, reduce internal productivity and substandard customer service.

- To empirically assess the quality of data in the ERP system with respect to data quality dimensions.
- To critically evaluate whether adopting data quality and risk management procedure could lead to reduction in the cost of operation.
- To examine the extent to which application of data quality and risk management tools has led to significant improvements in organizational performance

2. Data Quality and ERP Risks.

Data Quality within the Data and risk management plays a significant role in success and failure of the ERP system implementation. Data have no actual quality and value, they only have potential value that is realized only when someone uses that data to make any decisions.

• 2.1 Security Risk

-Security issues might occur due to a lack of data quality. Companies must prioritize online security for their application servers. ERP systems are integrated with online apps. When one item is overlooked, the security risks of these web applications increase, lowering the software's quality. Cross Site Scripting (XSS) and SQL injection are

common threats to web applications that lack sufficient protection. SQL injections occur when an intruder alters an existing SQL request to publish secret data, crush vital values, or perform harmful orders in the database. The program gathers data from Internet users and utilizes it to generate a SQL request.

Cross Site Scripting (XSS) is an attack that takes advantage of vulnerability in a website that does not verify user input. XSS uses numerous approaches to insert and execute scripts written in languages like JavaScript. This type of attack aims to collect identifiable information from cookies or deceive users into providing it to the attacker. ERP systems rely heavily on web apps for client transactions, making security critical. So it should consider all stages of the development cycle. However, during the application's design process, it must be given the highest priority.

• 2.2 Dirty Data

-ERP relies heavily on transaction data for daily operations and decision-making. This data can be input both electronically and manually. Following that, information was categorized, controlled, and extracted for decision-making purposes. Entered data may be utilized to help with constructing, shipping, and invoicing items, while extracted data can be used to assess manufacturing and sales force performance in the near term. Long-term data is used to make business choices, including determining operational efficiency and safeguarding data integrity. Dirty data refers to data that has been used by an organization for a long time and has dissimilar structures, such as spelling discrepancies, multiple account numbers, address variations, incomplete or missing data, a lack of legacy data standards, actual data values differing from meta-labels, and the use of free-form fields. Cleaning up different data repositories in corporations can help solve these issues.

Improper order processing, defective items, or problems in packing, paperwork, or invoicing can all lead to inaccurate statistics. Dirty data may lead to consumer dissatisfaction, loss of shareholder trust, increased labor and material expenses, and wasted time rectifying inaccuracies. If we do not check in the system to avoid human mistake, we will end up with filthy data.

"Bad data can put a company at a competitive disadvantage" some computer-literate criminals simulating accidents and taking use of filthy data in corporate databases A research indicated that filthy data led to the loss of \$56 million by many insurance providers to a single fraud ring.

• 2.3 Data Integrity

-Data integrity means that following data entering, it is systematically updated or revised by 'Experts' to remove inaccuracies. "Duplicate data or data that is incomplete or extraneous". Data integrity necessitates knowledge and dirty data management. A collection of data has integrity if the data is logically consistent and correct. To ensure data integrity, updates must be reflected in all storage places and consistent across mediums. Data integrity needs users to comprehend its significance in the context of the company. Data integrity demands a systematic method for processing, storage, sharing, modification, and reporting. Data errors can lead to financial losses, consumer dissatisfaction, and hinder the implementation of new initiatives. Businesses who process debits and credits or dispose of surplus inventory realize the importance of data accuracy. Each credit or debit is anticipated to cost the corporation \$75 million in clerical work for analysis, production, and distribution of the document. Otherwise, it will be expensive to test these ERP software products.

3. Strategic Importance of Data Quality in ERP

• 3.1 High-Impact Use Cases: SCM, Finance, CRM

-ERP system brings everyone together and makes sure that it helps Sales, Finance, Accounting, Manufacturing to make better decisions making. While Implementing the ERP system, one of the most important factors plays a major

role is training of everyone because everyone will start using this system like a legacy system and it will create a lot of data quality issues because one wrong information in one module will carry forward by different modules and it will create a chain of misinformation within the ERP system. One of the other high impact things in terms of data quality is that it will be very difficult for the top management, managers and other teams to do any sort of analysis like Revenue forecasting, sales forecasting, demand forecasting, product margins, cash flow analysis, inventory valuation, accounting etc. Data Quality has one of the highest impacts is Customer Satisfaction and managing relationships with the customer.

- **3.2 Risk of Ignoring Data Strategy**

- Training and structuring of the data tables from the legacy system is one of the most important for data strategy and ignoring these will cost the company a fortune because it will create a lot of dirty data and there will be no data integrity which will have high impact is understanding the customer needs and it will follow the domino effect as it will be affect the supply chain and inventory team, Finance and accounting team as well. There are various potential risks if the company does not have data strategy.

1. Poor Data Quality and Inaccurate Insights
2. Missed opportunities for data-driven Insights
3. Increased Vulnerability to Security Breaches
4. Operational and Technical Inefficiencies
5. Compliance Issues and Legal Consequences
6. Loss of Competitive Advantages.

Data Strategy is crucial for organizations to effectively manage and utilize their data assets, mitigate risks, and achieve their business goals. By addressing these risks proactively, organizations can ensure that their data becomes a valuable asset that drives innovation, improves decision-making, and enhances their competitive position in the market.

- **3.3 Cost and Consequences of Low-Quality Data**

- Low Quality data incurs significant financial costs for businesses, impacting productivity, customer relationships and operational efficiencies. Organization can go over-budget because of the low quality data. It can also cost millions of dollars annually due to inaccurate data, leading to waster resources, poor decision making and compliance issues. There are 2 most important Costly Consequences of Low Quality data.

1. Financial Costs:

a. Low Productivity: Low Quality data takes a significant amount of time in fixing errors to make it even more reliable information for employee to start their analysis.

b. Ineffective Decision Making: Low quality of data can lead to misguided strategic decisions, missed revenue opportunities and wasted resources.

c. Operational Inefficiencies: Inaccurate information within the system can disrupt workflows which can cause delay in their day-to-day deadlines and reduces their productivity, it can also cause delay in order fulfillment, inventory management.

d. Compliance Issues: Failure to meet data quality standards can result in fines and legal repercussions.

e. Reputational damage: Data breaches or customer service issues stemming from poor data management can harm a company's reputation.

2. Consequences for Business:

- a. Reduced customer trust and loyalty:** Customers who experience errors due to bad data, such as misdirected communications or incorrect bills, may lose trust and loyalty, leading to decreased sales.
- b. Missed Revenue Opportunities:** Inaccurate data can prevent businesses to understand the market and build strategy, it will affect the competitive advantages and it will slow in capitalizing on new opportunities.

• 3.4 The Data Quality–User Trust–Adoption Cycle

- The relationship between data quality, user trust and adoption form a reinforcing cycle that is critical for successful data driven decision making and the effective use of analytics and AI systems. Working of the Cycle is as follows:

a. Data Quality: High-quality data—defined as reliability, completeness, consistency, timeliness, and validity—is the core of any reliable analytics or AI system. Organizations must build comprehensive data quality frameworks and regularly monitor, profile, and enhance their data assets to preserve this foundation.

b. User Trust: When people meet dependable, accurate, and relevant data, their trust in it grows. Users' trust is damaged when they encounter data inaccuracies, inconsistencies, or a lack of transparency regarding data lineage and quality processes. Trusted data is vital for consumers to believe in the outcomes of analytics tools, dashboards, or AI applications.

c. Adoption: High user trust leads to increased use of data tools and analytics platforms. Users who trust the data are more likely to integrate analytics into their everyday processes, make data-driven choices, and push for more data efforts inside the business.

d. Reinforcement: Increased usage leads to increased input and involvement, which may be used to improve data quality procedures and governance. As adoption rises, firms are encouraged to invest more in data quality, which completes the cycle.

The Data Quality–User Trust–Adoption Cycle is a self-reinforcing cycle in which good data quality fosters user trust, which then encourages adoption of data tools and analytics. As use grows, companies receive more feedback and motivation to enhance data quality, forming a virtuous loop that drives effective data-driven change.

3.4.1. Importance of the Cycle:

-Without Data Quality: Users become distrustful, resulting in low confidence and adoption of analytics or AI solutions. This leads to missing business opportunities and misused technological investments.

-With Data Quality: Trust is built, which motivates adoption. Employees make smarter decisions, and corporations see actual results from data projects.

4. Characterizing Data Quality in ERP System – Dimensions of Data Quality

- There are various parameters through which data quality metrics can be assessed. Data Quality Dimensions is very much a qualitative and subjective matter because it depends on the understanding and familiarity of the ERP system through which the parameters are decided if the data quality is accurate, reliable and trustworthy.

	SELECTED LITERATURE							
Dimensions	Loshin [4]	Pipino [29]	Kerr et al [18]	Berti-Equille [11]	Batini & Scannapieco [1]	Chapman [2]	Olson [6]	Fisher & Kingma [15]
Accuracy	✓	✓	✓	✓	✓	✓	✓	✓
Completeness	✓	✓		✓	✓	✓	✓	✓
Consistency	✓	✓		✓	✓	✓	✓	✓
Timeliness	✓	✓	✓	✓	✓	✓	✓	✓
Correctness		✓						
Currency	✓	✓			✓	✓		
Comparability			✓					
Usability			✓					
Relevance			✓				✓	✓
Understood							✓	
Trusted							✓	
Volatility					✓			
Interpretability					✓			
Accessibility		✓			✓	✓		
Reliability					✓			
Readability					✓			
Uniqueness	✓							
Flexibility						✓		
Transparency						✓		
Lineage						✓		
Security			✓			✓		

Table 1: Summary of Literature Review Identifying Data Quality Dimensions Mostly Used

(Reference Link - <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7898777e8a16ca22aef6c58e8ffae5f1820047c5>)

The table 1 compares data quality dimensions discussed across selected academic literature. It reveals that while core dimensions like accuracy, completeness, consistency, and timeliness are widely acknowledged across studies, other dimensions such as usability, relevance, security, and interpretability receive less consistent attention.

- Accuracy, completeness, and timeliness are the most frequently cited dimensions.
- Broader aspects like accessibility, trust, uniqueness, and reliability are discussed by only a few authors.
- The diversity in dimensions indicates a lack of standardization in defining and measuring data quality across ERP and information systems research.

The table highlights the multi-faceted nature of data quality and supports the need for comprehensive frameworks when managing ERP data.

5. Data Quality Management

- Data Quality Management in a during the ERP implementation phase ensures the accuracy, completeness, consistency, and reliability of the data used across various business functions. It includes implementation of various processes, preventing data quality issues, and improving the efficiency of ERP system and organizational overall operations.

Parameters for Data Quality Management in ERP system:

1. Identifying Data Quality issues which are as follows:

- a. Data Audits
- b. Data Profiling
- c. User Feedback

2.Improving Data Quality Issues:

- a. Data Cleansing
- b. Data Validation
- c. Data Migration

3.Preventing Future Data Quality Issues:

- a. Data Governance
- b. Master Data Management
- c. Data Quality Monitoring

There are many Data Quality software and ERP system built-in data quality management features and functionalities which can be used to maintaining the data quality. Data Quality matters in the ERP system because it helps top management make better decision making, It also helps in reducing errors, delays, and rework which is enhancing the efficiency and productivity of every single individual within the company. Maintaining data quality can increase customer satisfaction, have competitive advantages and have better risk management techniques.

6. Factors Affecting Data Quality

Efforts to improve data quality often lack a strategic viewpoint, hindering their efficacy. Maintaining data quality requires excellent staff management, organizational elements, and technical systems. The categorization identifies technical, organizational, and environmental elements that impact data quality. These variables can have major ramifications at the technological, organizational, and legal levels. These are summarized below.

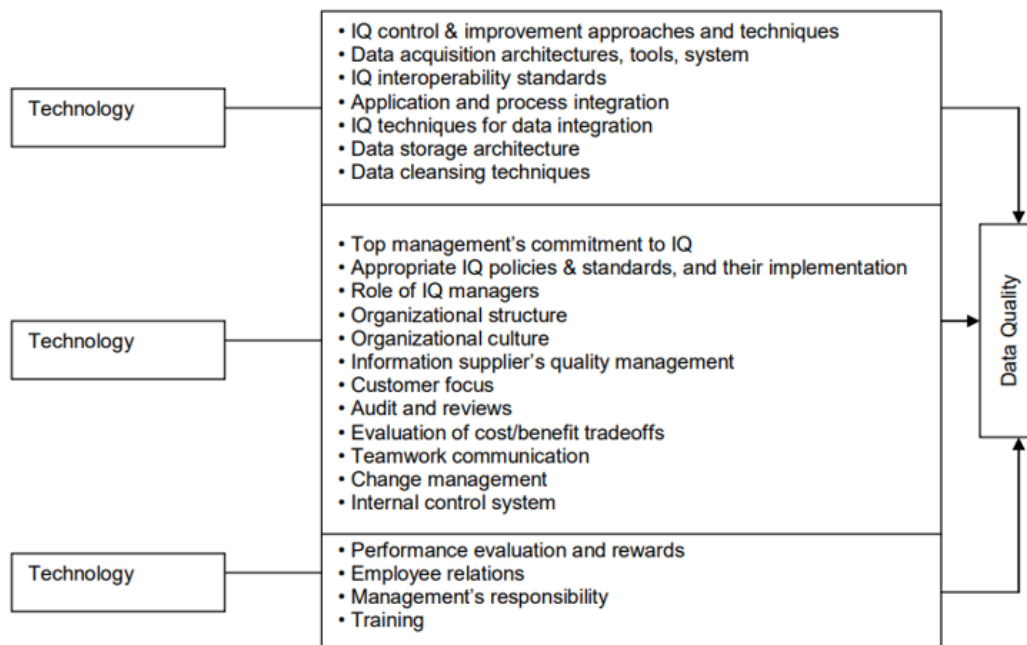


Image 1: Summary of Factors Influencing Data Quality

(Reference Link: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7898777e8a16ca22aef6c58e8ffae5f1820047c5>)

The table 2 illustrates a three-layer framework for achieving high data quality (DQ) in ERP systems. It highlights that DQ is influenced by:

1. **Technical Factors** – such as data integration, cleansing techniques, storage architecture, and process interoperability.
2. **Organizational Factors** – including management commitment, policies, data governance roles, audits, and internal controls.
3. **Human Factors** – like employee training, performance evaluation, and accountability.

Together, these dimensions emphasize that maintaining ERP data quality requires a balanced approach across technology, governance, and people.

7. Future Directions

• 7.1 Predictive & Explainable AI for DQ in ERP

- Predictive and explainable AI tools for Data Quality issues can be used for the data quality management in ERP systems. Predictive AI can be used in proactively detecting and correcting the data quality issues in real time, it also automates the process of data cleansing, forecasts potential problems even before they occur. Explainable AI (XAI) makes AI-driven decisions transparent, helping the employee and auditors within the companies determine the automation process is correctly applied or not and with that it can also develop the trust it requires within the ERP system, which can also be vital for the compliance adoption. Future prospects include AI-Powered data fabrics that unify data quality across several sources, context-aware data quality rules that adapt to business requirements, and the use of generative AI to generate synthetic training data. However, obstacles remain in expanding governance, closing talent shortages, and assuring ethical AI use.

• 7.2 Multi-cloud and Distributed ERP Data Governance

- Unified Governance Framework: Organizations will prioritize centralized policies and cross cloud observability tools to make sure that it meets standard data quality and integration of the multi-cloud environments. This also includes all the different permissions and controls, authentication, automated compliance check which can help in maintaining the data governance.
- Shift – Left Data Observability: Quality problems won't spread throughout clouds if real-time monitoring is integrated early in the data lifecycle. This includes automatic validation for ERP-critical data (such as vendor and customer information), defined KPIs, and SLAs with cloud providers.
- Interoperability and Portability: Adapting to data fabrics and open standards will centralize the governance principles and meta data across hybrid ERP implementations, guaranteeing smooth integration. Workflows that are containerized, such order processing, will make cloud migration easier.
- Security and Compliance Automation: Policies for the sensitive ERP modules code frameworks and encryption methods will be automated GDPR/CCPA compliance. Vulnerabilities in distributed architecture will be addressed by ongoing threat detection and risk modeling.

To guarantee data integrity across multi-cloud settings, future ERP systems will rely on AI-augmented governance and interoperable frameworks, lowering operational risks while striking a balance between scalability and regulatory requirements.

8. CONCLUSION

- ERP is the Cloud-based software package that can be adopted in the companies which are manufacturing companies. Data Quality plays a major role in the ERP system. Data Quality is the one of pillars for the successful implementation of the ERP system. Data Quality and data management policies enable enterprises to respond proactively and deliver products and services which can achieve maximum customer satisfaction. Implementation of the data quality centralized policies at all levels ensure that right information is available to the right person at the right time, format, accurate and reliable which enhances their productivity.

REFERENCES

1. NetSuite Inc. (2014). NetSuite products. Retrieved November 22, 2014, from <http://www.netsuite.com/portal/resource/articles/erp/what-is-erp.shtml>
2. Rothlin, M. (2014). An exploratory study of data quality management practices. Retrieved November 20, 2014, from <http://books.google.lk/books?id=wyheQllcyh0Cpg=PA292lpg=M.Rothlin>
3. El-Rayyes, E. K., & Abu-Zaid, I. M. (2012). New model to achieve software quality assurance (SQA) in web applications. **International Journal of Science and Technology, 2*(7)*.
4. Venkitaraman, R. (n.d.). Software quality assurance. Department of Computer Science, The University of Texas, Dallas.
5. Core. (n.d.). Retrieved from <https://core.ac.uk/download/pdf/235049621.pdf>
6. Liepins, G. E. (1989). Sound data are a sound investment. **Quality Progress, 22*(2)*, 61–64.
7. Loshin, D. (2009). **Master data management**. Elsevier Inc.
8. Pipino, L., Wang, R., Kopcsó, D., & Rybolt, W. (2005). Developing a measurement scale for data quality dimensions in information quality. In V. Zwass (Ed.), **Advances in management information systems** (pp. 37–51). M. E. Sharpe Inc.
9. Kerr, K., Norris, T., & Stockdale, S. (2007). Data quality information and decision-making: A healthcare case study. In **Proceedings of the 18th Australasian Conference on Information Systems** (pp. 1017–1026), Toowoomba.
10. Batini, C., & Scannapieco, M. (2006). **Data quality concepts, methodologies, and techniques**. Springer.
11. Olson, J. E. (2003). **Data quality: The accuracy dimension**. Morgan Kaufmann Publishers.
12. Fisher, C. W., & Kingma, B. R. (2001). Criticality of data quality as exemplified in two disasters. **Information and Management, 39*(2)*, 109–116.