December 2002

DETERMINANTS OF DATA QUALITY MANAGEMENT ADOPTION AND IMPLEMENTATION

Klara Nelson
The University of Tampa

Follow this and additional works at: http://aisel.aisnet.org/amcis2002

DETERMINANTS OF DATA QUALITY MANAGEMENT ADOPTION AND IMPLEMENTATION

Klara Nelson
The University of Tampa
krnelson@ut.edu

Abstract

While high quality data is paramount to many organizations in the information age, many companies assign low priority to investments in data quality. The present study seeks to answer the question of why some firms are more prepared than others to ensure high quality data and identifies major variables that affect the adoption and implementation of data quality management (DQM) initiatives. An understanding of these factors represents an important step towards addressing the pervasive data quality problem and developing effective DQM strategies.

Introduction

The ability of databases to function as a tool for managerial decision-making is greatly impacted by the quality of its data. The importance of data quality management is increasing with the current trend in many industries toward data warehousing and data mining, the significant growth in the usage of automated decisions and processes on the basis of electronic data, and the increased sharing of information between organizations (PriceWaterhouseCoopers 2001).

Prior research has identified a number of data quality dimensions including accuracy, completeness, and consistency (Redman 1992). Based on the definition of high quality data as data that is fit for use by data consumers, Wang and Strong (1996) developed a framework which classifies data quality (DQ) dimensions into four categories: *intrinsic DQ* (accuracy, objectivity, believability, and reputation), *contextual DQ* (value-added, relevancy, timeliness, completeness, and appropriate amount of data), *representational DQ* (interpretability, ease of understanding, representational consistency, and concise representation), and *accessibility DQ* (accessibility, access security).

Because the quality of data cannot be assessed independently of those who use it, data quality management initiatives need to go beyond the traditional intrinsic focus on accuracy and consider the broader and more relative definition of data quality as "fitness for use" (Strong et al. 1997; Ballou and Kumar Tayi 1998). Unfortunately, data management is often "being addressed at the wrong level, in the wrong place, and in the wrong way across too many corporate organizations" (PriceWaterhouseCoopers, 2001, p. 16). Many companies have no plan for managing data quality (Eckerson 2002). Often, data quality problems remain unknown or are ignored (Khalil and Harcar 1999; Stackpole 2001). Companies that measure data quality often narrowly focus on the accuracy dimension instead of addressing some of the other dimensions and attributes outlined above.

Prior research has shown that poor data quality appears to be the norm rather than the exception (Redman 1995). A study of the criminal records system found up to 74.3% of the records exhibiting significant quality problems (Laudon 1986). Others report error rates of 0.5% - 30% when measured at the field level (Redman 1998). Many companies report significant problems as the result of defective data (PriceWaterhouseCoopers 2001), which can impact the operational, tactical and strategic levels (Redman 1998). At the operational level, the cost of poor data has been estimated to be as high as 10 to 25% of total revenues (English 1999). From a tactical perspective, poor data quality compromises decision-making. A recent analysis of the space shuttle Challenger and the USS Vincennes/Iranian Airbus disasters revealed multiple incidents of poor data quality and subsequent faulty decisions that had disastrous consequences (Fisher and Kingma 2001). At the strategic level, poor data quality has the potential for putting companies at a competitive disadvantage by making it more difficult to execute strategies in areas such as data warehousing, customer relationship management (CRM), and e-business (Eckerson 2002).

Given the strategic value of data in the information age, why are companies not doing more to ensure its quality? Empirical research that examines determinants of DQM adoptions and the degree of implementation is sparse. Prior research has developed prescriptions and methodologies for data management from a total quality management perspective that identify *what should be done* to ensure high quality data (e.g., Wang 1998; Dvir and Evans 1996). The research does however not investigate *what is actually done* or *what motivates* a company to adopt data quality management programs. The main objective of this study is fill this void and (1) develop a conceptual framework for identifying the major variables that affect the adoption and implementation of DQM programs and (2) explain why some firms are more prepared than others to invest in data quality.

Adoption and Implementation of Data Quality Management Initiatives

Based on a review of the organizational and technological innovation, total quality management, and IT management literatures, three main classes of variables are considered important in determining the adoption and implementation of data quality management programs: organizational characteristics, characteristics of the IT environment, and the external environment.

Organizational Characteristics

Major attributes of the organizational environment include the customer focus of the organization, its general quality orientation, and the degree of management support for investments in data quality.

Customer Focus of the Organization
Customer focus has been identified as a critical TQM implementation construct, which must be reflected in the overall planning and execution of quality efforts (Ahire et al. 1996). This is particularly important in the information age where an increasing number of organizations are implementing CRM strategies that are built on providing high-quality products and services on the basis of personal knowledge of customers. Successful implementation of such relationship marketing requires a high degree of data quality (Khalil and Harcar 1999).

Quality Orientation of the Organization
The way the entire organization perceives and deals with quality initiatives is expected to greatly influence the adoption and success of a DQM program (Stylianou and Kumar 2000). IT units in quality-oriented organizations are likely to be under pressure to improve systems quality, an important component of which is data quality. The quality orientation of the host organization was found to be significant in explaining the swiftness and intensity of TQM adoption in systems development (Ravichandran 2000).

Management Support
The commitment of top management to the diffusion of innovations throughout an organization has been well documented. Top-level commitment and support has been shown to be a critical success factor in the implementation of any kind of project including successful quality management implementations (Ahire et al. 1996). More recently, IS management support for quality was found to be significant for both the swiftness and intensity of TQM adoption in systems development (Ravichandran 2000). Executive perceptions of the *relative advantage* of high quality data, of *data as a strategic resource,* and of *existing data quality* are likely to play a role in management's support for DQM initiatives. Relative advantage, defined as the degree to which an innovation is perceived as being better than the idea it supersedes, indicates the benefits and costs resulting from adoption of an innovation (Rogers 1995). Top management support for DQM initiatives is more likely when a business case can be made. In addition, senior managers who do not realize the value of their data as strategic assets do not appear to place importance on and invest in data quality (PriceWaterhouseCoopers 2002). Finally, faulty perceptions that overestimate data quality and underestimate the costs of errors may be the reason for lack of data quality programs (Eckerson 2002).

IT Environment

Three major characteristics of the IT environment that influence the adoption and implementation of DQM include the overall IT strategy, IT management sophistication, and resources available to the IT unit.

IT Strategy

A firm's overall technology strategy relative to major competitors can be classified according to its degree of innovativeness into four major categories (Sambamurthy and Zmud 1992). *Innovators* are on the leading edge of IT and are willing to incur considerable technological risk. *Early imitators* track the efforts of innovators and are willing to incur moderate amount of technological risk. *Mid-cycle imitators* wait for an information technology to be proven before investing in it while *late entrants* only invest in IT when technology becomes a strategic necessity. Greater innovativeness in IT is likely to be related to a tendency to adopt an innovation such as DQM (Thong and Yap 1995).

IT Management Sophistication

The overall level of IT management capability has also been referred to as IT management sophistication (Karimi et al. 2000). Dimensions of IT management sophistication include: IT planning, IT control, IT organization, and IT integration. Firms with greater IT management sophistication exhibit a higher degree of alignment of IS plans with the firms' business plans and technology integration for better exploitation of IT within the firm. IT departments with greater IT management sophistication are more likely to realize the business value of DQM and support its adoption (Karimi et al. 2000).

IT Unit Resources

A resource-based view of implementation suggests that greater financial resources and expertise increase the likelihood of IS implementation success (Thong 2001). Smaller IT units may not have the resources necessary to initiate and sustain data quality management initiatives. This includes the presence of a separate quality function within the IS unit which was found to facilitate the adoption of TQM in systems development (Ravichandran 2000).

External Environment

Information Intensity of Industry

Information intensity refers to the degree to which information is present in the product or service of business (Thong and Yap 1995). Some industries are considered more information intense than others. The health care industry for example includes many types of organizations that are data driven and in which the quality of data is critical (Gendron and D'Onofrio 2000).

Conceptual Model

The above arguments are synthesized in the conceptual model for the organizational adoption and implementation of DQM programs as depicted in Figure 1. The following propositions can be posited based on the discussion above.

Organizational Environment

Proposition 1:   There is a positive relationship between the quality orientation of the organization and the adoption of DQM.

Proposition 2:   There is a positive relationship between the extent of management support for quality and the adoption of DQM.

Proposition 3:   There is a positive relationship between the customer focus of the organization and the adoption of DQM.

Proposition 4:   Management support is a function of executive perceptions of relative advantage, data as a strategic resource, and existing data quality.

    (a) Perceptions of relative advantage are positively related to management support.

    (b) Perceptions of data as a strategy resource are positively related to management support.

    (c) Perceptions of existing data quality are negatively related to management support.

IT Environment

Proposition 5:   There is a positive relationship between the extent to which a company is considered an IT innovator and the adoption of DQM.
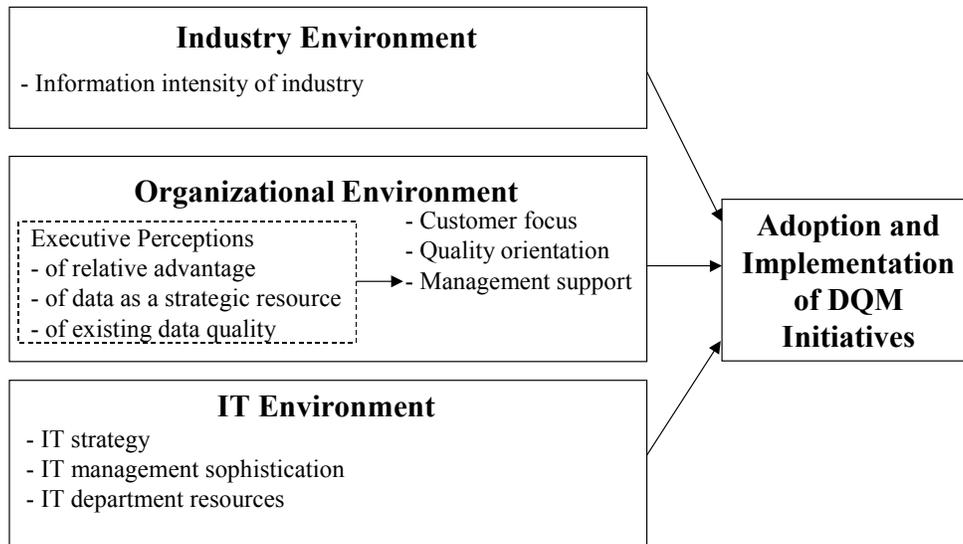
Figure 1. A Conceptual Framework of DQM Adoption and Implementation

Proposition 6:   There is a positive relationship between the IT management sophistication of the IT department and the adoption of DQM.

Proposition 7:   There is a positive relationship between the resources available to the IT unit and the adoption of DQM.

External Environment

Proposition 8:   There is a positive relationship between an organization's industry information intensity and the adoption of DQM.

Discussion

This paper tried to fill a void in data quality research and developed a model describing the relationship between external, organizational, and IT-department characteristics and the adoption and implementation of data quality management initiatives. One contribution of this study would be its extension of prior research on organizational and IT innovations, IT management, and total quality management to data quality management. The results of this study are also expected to provide useful insights into the current status of DQM in organizations. Finally, this study is expected to represent an important step toward understanding and addressing the seemingly pervasive data quality problem that exists in organizations today.

The next stage of this research consists of validating the model, operationalizing the variables, and testing the above propositions. Given the exploratory nature of this study, the general research paradigm for use in this study is based on Churchill (1979) as adapted by Little and Gibson (1999). The first step is to achieve conceptual clarity on what is to be measured (Bagozzi 1982) based on the literature review and interviews with data quality managers to elicit data about their understanding of data quality management. Interviews may result in modifications to the conceptual model. The next step then involves developing a valid and reliable instrument to measure (1) the factors that affect the adoption of DQM initiatives and (2) data quality initiatives and their implementation. Where possible, existing scales will be used. After pre-testing the preliminary instrument with a panel of experts for feedback on clarity, wording, appropriateness, and face and content validity, a large-scale survey of data quality and IS managers is planned using the revised questionnaire.

References

Ahire, S. L., Golhar, D. Y., and Waller, M. A. "Development and Validation of TQM Implementation Constructs," *Decision Sciences* (27:1), Winter 1996, pp. 23-55.

Bagozzi, R. P., and Phillips, L. W. "Representing and Testing Organizational Theories: A Holistic Construal," *Administrative Science Quarterly* (27:3), 1982, pp. 459-489.

Ballou, D. P., and Kumar Tayi, G. "Enhancing Data Quality in Data Warehouse Environments", *Communications of the ACM* (42:1), 1998, pp. 73-78.

Churchill, G. A. "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16, pp. 64-73.

Dvir, R., and Evans, S. "A TQM Approach to the Improvement of Information Quality," *Proceedings of the 1996 Conference on Information Quality*, MIT, downloaded from web.mit.edu/tdqm/papers/other/evans.html.

Eckerson, W. W. "Data Quality and the Bottom Line: Achieving Business Success through the Commitment to High Quality Data", The TDWI Report Series, 2002.

English, L. *Improving Data Warehouse and Business Information Quality*, John Wiley and Sons, New York, 1999.

Fisher, C. W., and Kingma, B. R. "Criticality of Data Quality as Exemplified in Two Disasters", *Information and Management*, 39, 2001, pp. 109-116.

Gendron, M. S., and D'Onofrio, M. J. "An Exploratory Study Investigating Data Quality in the Healthcare Industry: What are the Implications for Data Warehousing?", *Proceedings of the 2000 Americas Conference on Information Systems,* August 10-13, 2000, Long Beach, California, pp. 126-128.

Karimi, J., Bhattacherjee, A., Gupta, Y.P., and Somers, T. M. "The Effects of MIS Steering Committees on Information Technology Management Sophistication," *Journal of Management Information Systems* (17:2), Fall 2000, pp. 207-203.

Khalil, O. E. M., and Harcar, T. D. "Relationship Marketing and Data Quality Management", *SAM Advanced Management Journal*, Spring 1999, pp. 26-33.

Laudon, K. C. "Data Quality and Due Process in Large Interorganizational Record Systems," *Communications of the ACM* (29:1), 1986, pp. 4-11.

Little, R. G., and Gibson, M. L. "Identification of Factors Affecting the Implementation of Data Warehousing," *Proceedings of the 32nd Hawaii International Conference on System Sciences 1999*, pp. 1-10.

PriceWaterhouseCoopers *Global Data Management Survey 2001*.

Ravichandran, T. "Swiftness and Intensity of Administrative Innovation Adoption: An Empirical Study of TQM in Information Systems," *Decision Sciences* (31:3), Summer 2000, pp. 691-724.

Redman, T. C. *Data Quality: Management and Technology*. New York: Bantam Books, 1992.

Redman, T. C. "Improve Data Quality for Competitive Advantage", *Sloan Management Review*, Winter 1995, pp. 99-107.

Redman, T. C. "The Impact of Poor Data Quality on the Typical Enterprise," *Communications of the ACM* (41:2), February 1998, pp. 79-82.

Rogers, E. M. *Diffusion of Innovations* (4th edition). New York: The Free Press, 1995.

Sambamurthy, V., and Zmud, R. W. *Managing IT for Success: The Empowering Business Partnership*, Financial Executives Research Foundation, Morristown, New Jersey, 1992.

Stackpole, B. "Wash Me: Dirty Data is the Dirty Little Secret that can Jeopardize your CRM Effort," *CIO Magazine*, February 15, 2001, downloaded from www.cio.com.

Strong, D. M., Lee, Y. W., and Wang, R. Y. "Data Quality in Context," *Communications of the ACM* (40:5), May 1997, pp. 103-110.

Stylianou, A. C., and Kumar, R. L. "An Integrative Framework for IS Quality Management," *Communications of the ACM* (43:9), September 2000, pp. 99-104.

Thong, J. Y. L. "Resource Constraints and Information Systems Implementation in Singaporean Small Businesses," *Omega* 29, 2001, pp. 143-156.

Thong, J., and Yap, C.S. "CEO Characteristics, Organizational Characteristics and Information Technology Adoption in Small Business," *Omega* (23:4), 1995, pp. 429-442.

Wang, R. Y. "Total Data Quality Management", *Communications of the ACM* (41:2), February 1998, pp. 58-65.

Wang, R. Y., and Strong, D. M. "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems* (12:4), Spring 1996, pp. 5-34.