Competing with Quality Information

Alan Matsumura and Nadia Shouraboura Diamond Technology Partners

Abstract: Quality of information dictates quality of decisions. This paper presents what quality information means to organizations, and how organizations can improve information quality. In this paper, quality information refers to accurate and accessible information pertinent for effective operation and intelligent decision-making. Accessibility and accuracy together, therefore, form the key dimensions of quality information. To improve information quality, organizations must analyze root-causes of poor information quality. Case examples are presented to explain how organizations improve their information quality to compete effectively in the market place.

1. Introduction

Do you know your customers? Can you distinguish Angus from non-Angus beef? Does your inventory system automatically produce "dummy" order information? Does your organization's information facilitate or hinder its competitive capacity? How good is your organization's information quality (IQ)? Answers to these questions reveal an organization's information readiness for competing in the changing market.

With autonomous, heterogeneous information systems being inter-connected by information super-highways, information from different sources with disparate qualities is becoming increasingly available to decision-makers. To effectively achieve business process engineering, client server computing, data warehousing, internet communication, electronic commerce and virtual corporation, a critical issue that must be addressed is information quality (IQ).

Before discussing information quality, one must first discuss what information is, and how it differs from data. In the literature, information and data are often used synonymously. In practice, managers differentiate information from data intuitively, and typically describe information as data that have been processed. McGee & Prusak [6] articulate this intuition:

"information is not just data collected; rather, it is data collected, organized, ordered, and imbued with meaning and context. Information must *inform*, while data has no such mandate. Information must be bounded, while data can be limitless. In order for data to become *useful* to a decision maker as information, it must be presented in such a way that he or she can relate to it and act upon it."

McGee & Prusak's view coincides with the research findings from MIT's Total Data Quality Management (TDQM) Research Program [7, 10]. It should be noted, however, that data can be viewed as information, and information can be viewed as data. This paper, therefore, will use the term *information* interchangeably with *data*.

Poor information quality can have a severe impact on the overall effectiveness of an organization. This concern has received much attention from both private and public sectors in recent years. In the private sector, numerous articles have been written on issues related to information quality. The Wall Street Journal, in noting research at MIT, reported that,

"Huge databases can be mined to find sales prospects among existing customers, analyzed to unearth costly corporate habits; manipulated to divine future trends. Just one problem: Those huge databases may be full of junk." ("Databases Are Plagued by Reign of Error," The Wall Street Journal, May 26, 1992, page B6).

A recent *Datamation* cover-story also urges organizations to clean up data before warehousing it ("Don't Warehouse Dirty Data," Datamation, October 15, 1995, p. 42). In the academic circle, various issues related to information quality are being addressed. For example, investigators at MIT focus their research on how information quality can be defined, measured, analyzed and improved [3, 8, 9]. In short, information quality has become increasingly critical as firms position themselves to compete with information. Without high quality information that executives can trust in making their decisions, information stored in client/server systems and data warehouses will not be useful, and in some cases disastrous. For firms in the 21st century to compete, survive and achieve total customer satisfaction, it is imperative that problems that may lead to poor-quality information be anticipated and resolved proactively.

2. Information Quality

2.1 What is Information Quality?

Information quality covers the multi-dimensional quality of information that is fit for use. Information quality observed without the context of its business use, therefore, is meaningless. Fifteen dimensions of information quality are defined against the conventional data quality perspective focusing

on accuracy alone [10]. Based on the fifteen dimensions, we devise the key overarching pillars of quality information, which we call the key dimensions of quality information.

Relating quality information to its key dimensions operationalizes quality information in terms of its key quality characteristics so that specific diagnostic tools and improvement solutions can be discovered and applied. Furthermore, capturing the key dimensions can redirect organizational attention to information-related activities and their locus that has a direct bearing on the root-causes of poor information quality.

Accessibility and accuracy are two key dimensions of information quality. Accessible information refers to information that is comprehensible and accessible for information consumers. For example, information that is stored in a database yet difficult to retrieve (not timely, or has technical or procedural barriers), interpret, or integrate, and thus not usable for decision-making has low accessibility. Conversely, information that is usable for business decision-making (collected and stored so as to be comprehensible, and easily integrated and aggregated) for managers refers to highly accessible information.

Accurate information refers to information that is correct, free of errors, and relevant for information consumers. For example, information collected with incorrect definitions, units, and values or not relevant for the purpose of tasks, and thus not useful for decision-making has low accuracy. Conversely, information that is useful for business decision-making for managers refers to highly accurate information

2.2 Quality Information for Quality Decisions

We use the following example to illustrate quality information in a real organizational setting. Comco's (pseudonym) experiences exemplify the definition, power and value of managing quality information for quality decisions.

Comco: At Comco, a major commodity company, multiple databases in different functional areas were serving different business purposes. Historically, business rules, expectations, and understanding of operations of its business in different departments were different. The production department was only interested in how the cows are cut out. The rules of cutting were based mainly on the weight of a cow. Different breeds of the same weight cows were cut in the same way. For example, an 800 LB Angus (a type of breed: has fat layers in between muscles, tastes good, higher premium value) cow and an 800 LB

non-Angus cow are cut in the same way. The production department focused on cutting 800 LB and 1400 LB animals differently. All 800 LB animals were treated equally. What mattered to the production is how heavy the animals were. To the sales department, however, recognizing 20-30 % premium for Angus cows was considered important. Thus, the sales department needed to track and record different breeds. To the sales department, not all cows were equal.

Consequently, the Production Department tracked weight and how they work on it; while, the Sales Department tracked how they discriminate price based on different breeds and kinds. The information from the production system was accurate. The information from the orders management system (sales) was accurate. The result was a collision. Both systems contained accurate information which was not accessible and usable for making business decisions at a company-wide ("global") level. The challenge for Comco was how to interpret and utilize high volume of totally "accurate" and not-accessible information for its critical business decision-making. It was nearly impossible to evaluate Comco's profit and cost meaningfully let alone comprehending and planning future business. One Comco's senior manager summed up their problems: "The things we produce are not the things we sell." Obviously, Comco needed to recognize the overarching view at a company-wide level to resolve the mismatching and disjointed information problems.

Comco exemplifies improving company-wide (global) level accessibility of locally accurate information. It exemplifies that accurate information can be inaccessible at a company-wide level. It also exemplifies that integrating disparate information harvests strategically valuable and highly accessible information for a company as a whole. The data warehousing projects utilized multi-dimensional data modeling techniques to integrate Comco's production and sales information. The capability for aggregating and slicing information also achieved to create an overarching value across functional areas. Most importantly, quality information helped to create and use quality decision-making mechanisms for Comco.

Valuable lessons can be learned from making historical and accumulated business information accessible to information consumers. Sales or accounting data for a month may not give managers a clear trajectory of their business. Comparing this month with last year's accounting can give you some insights. Accounting data is originally prepared for reporting and tax purposes. Companies need to be alerted before accounting systems show a minus after the fact. Accurate and accessible leading indicators for future and past are critical information for quality decision-making. Information stored in databases must be accessible and represented in a form that managers can easily manipulate and comprehend, so as to be useful for their tasks.

Comco's experience illustrates value of accessible information. The company was able to aggregate data and represent it in the way that managers can value and use information for their

decision-making. Pieces of business history was integrated and compiled to be viewed in a summarized fashion, or "sliced" and "diced" in different manners to raise and answer various business questions including "what if" questions. Managers were also involved in deciding "what if" questions needed for future business. Now in Comco, information is being collected by distributed systems and autonomous departments, yet this information is being integrated and aggregated to be viewed and used across, for example, departments, time, and business purposes. With improved accessibility of their information, Comco's information readiness for making quality decisions is improved.

3. A Framework for Quality Information

We have highlighted the importance of information quality for competing in the marketplace. In this section, we present the unique benefits of viewing business from the IQ lens.

3.1 The Power and Value of Information Quality Lens

Why do companies need to think of their business in terms of information quality? Business process re-engineering (BPR) and systems infrastructure approach are two popular approaches widely practiced in industry today. They play important roles in defining and meeting corporate information requirements.

Many firms, however, continue to suffer from poor quality information in their effort to compete more effectively in the market place. Well-articulated BPR plans and sophisticated information infrastructure can be merely blue prints or crude tools for firms if the quality of information supported by the information infrastructure for supporting business functions go un-checked. When un-checked, the long-term costs are severe.

BPR assumes that information is perfect and the issue of systems to support BPR can be easily resolved. As a result, it failed developing the necessary quality information to support the re-engineered business process. On the other hand, a systems infrastructure perspective focuses on hardware selection and software development. It often overlooks that the goal is to provide users with highly accessible and

accurate information for business processes. More importantly, no methods exist to bridge the disconnect between the two popular perspectives.

The IQ perspective bridges the disconnect: (1) IQ lens offers a perspective to view information from both the systems and user's perspectives, and thus, bridges the missing link between BPR and systems integration, C/S computing, and data warehousing, and (2) IQ lens provides a rich set of operationalized concepts and vocabulary (hence tools and methods) to easily uncover and solve critical IQ problems that are essential from both the systems and the users perspectives.

To compete and to thrive in the 21st century, organizations need to manage the newly emerging value-critical asset that can serve as the core of the firm's competitive weapon. Quality information is the core of a firm's competence besides physical, financial, and human assets. Quality information is essential for gaining fact-based knowledge about supply chain, distribution channel, customers, products and services, and the market. With quality information, organizations can make strategically intelligent decisions so as to effectively manage internally and to compete externally in the market. Firms, therefore, must pro-actively address quality of their information.

3.2 A Framework for IQ

What is information quality? Although there is an increasing body of literature addressing issues related to information quality, there is no well-formulated definition of information quality that is easy to adopt in real organizations.

3.2.1 A High-Level IQ Framework

Broadly, we can view Quality Information as accurate and accessible information, as presented in the previous section. Figure 1 depicts a high-level framework. Information must be accurate in terms of correctness of its value and meaning, and also in the context of relevancy for the task at hand. Information must be accessible to users in terms of comprehensibility and accessibility.

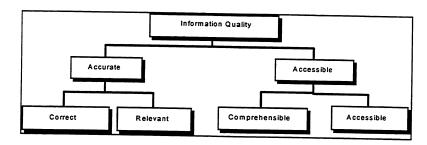


Figure 1: A High-Level IQ Framework

3.2.2 Information Quality Matrix

The two overarching dimensions of information quality, accuracy and accessibility represent the level and characteristics of quality information that both technology and business users' camps of a firm need. When information is highly accurate, its accessibility needs to be examined. Conversely, when information is highly accessible, its accuracy needs to be examined as well. Understanding both dimensions is highly useful for understanding and prioritizing information quality projects. For example, stock market traders need highly accurate and highly accessible financial market information at hand. Information for junk mail usually has low accuracy. We can easily assume that when the junk mail's information is highly accurate and accessible, it would not be called junk mail. It would be called intelligent data-marketing mail, instead. This is an example of how quality of information changes not only the quality of operation but also the strategy and characteristics of business. The matrix for quality information covers the key dimensions and their levels (Figure 2).

It is important to note that the dimensions of information quality need to be addressed and evaluated with the following in mind: objectives, boundary, task contexts, and time.

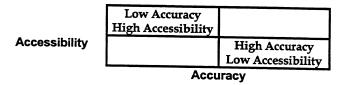
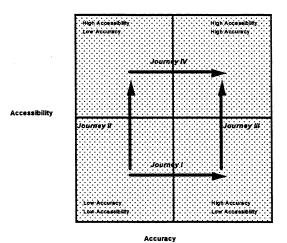


Figure 2: Quality Information Matrix

4. Improving Information Quality

Quality information encompasses highly accurate and highly accessible information that is fit for use. To achieve quality information, organizations need to participate in a transformational "journey": the journey from the diagnosed status to a targeted status. The concept of journey toward highly accurate and accessible information is a direct extension of the recommendations of the quality management perspective [2]. Because of the nature of information, the process of achieving both highly accurate and highly accessible information has to be optimized and strategized with priorities [1]. The four journeys in Figure 3 represent the priorities and transformation stages of IQ. Selected cases of successful projects highlight these journeys, ranging from low - high accuracy to low - high accessibility.



Legend: Journey I: Selco; Journey II: Datco; Journey III: Comco; Journey IV: Finco

Figure 3: Information Quality Journey

4.1 Journey I: Accurate information from "dummy" data

This journey involves an improvement focused on accuracy of information (Figure 4). The improvement covers accurate representation of reality, accurate collection of reality, accurate storage of reality. The remedial projects include data scrubbing for stored data in database and process streamlining for accurate collection procedures. The priority is given to establishing streamlined and accurate information collection procedures.

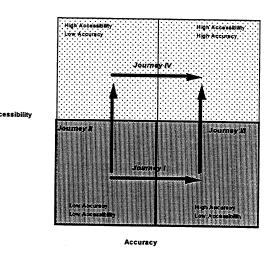


Figure 4: Information Quality Journey I

Selco's Journey: Selco streamlined "dummy" entry data.

Selco is one of the largest computer retailers in the U.S.. For the last several years the volume of sold computer parts doubled and the company experienced a continuous growth. Its profit margin, however, decreased and overall performance suffered. Selco's management with the help of a consulting firm conducted a detailed business analysis. They have discovered that the long payment collection process and a large volume of inventory considerably affected the profit margin. As for the collection process, the ordering systems in place had a manual, not-friendly interface for entering price. It generated numerous numbers of pricing errors. This pricing error was directly reflected in the invoice and the customers delayed payments due to the billing errors. As for the large volume of inventory, the sales personnel had no other way to reserve a certain amount of products but to place a "dummy" order in the system. The large volume of dummy orders were processed as real orders by the inventory ordering system and as a result extra volume of inventory was ordered automatically. This systematically inaccurate data was the basis for many important business decisions. Selco's management realized that inaccurate pricing data and dummy orders created the IQ problems for Selco. The company made a decision to improve the accuracy of Selco's data by modifying the existing order entry process and improving the quality of the software accordingly.

Selco's journey demonstrates the value of basic accurate information. Without accurate price and inventory information, business decisions and software tools are systematically in error for solving business problems.

4.2. Journey II: Accessible Information from faceless data

This journey covers improving accessibility of information. Priority was given to storage and utilization processes (Figure 5).

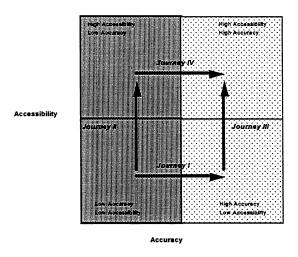


Figure 5: Information Quality Journey II

Datco's Journey: Redefined data-marketing for getting to know customers better

Datco is a large direct marketing company. It supports marketing databases for over eighty companies in the U.S. Up to recently, Datco's clients were mainly interested in clean, up-to-date mailing lists and Datco was one of the market leaders for cleaning, consolidating and updating name and address lists and providing a sufficiently accurate data.

Recently, database-marketing became more important for Datco's clients. Companies were trying to cut down marketing budget by selecting target-markets for marketing-activities. This new demand required more sophisticated analysis of the results of data-base marketing. Simply mailing out marketing materials with address and names was not sufficient for satisfying the clients and the their customers. To satisfy growing customers demands, the companies needed convenient access to their data to learn more about each individual customer, to evaluate customer value, and to understand how to distinguish between valuable customers and non-valuable ones.

Datco received a large number of requests for providing clients with higher accessibility to their marketing data. Some clients inquired if Datco was interested in extending their services to providing a detailed marketing analysis of the data. At that time Datco was not capable of satisfying their growing clients' demands. The company did not have experience and techniques for redesigning client's databases.

Datco's management evaluated the situation and realized the importance of providing highly accessible information and made a decision to invest in designing a new approach. As a result of this effort a new tool was developed. This tool was incorporated in a set of existing tools and allowed Datco to keep the necessary level of details and be able to easily summarize data on demand. Datco established a special training program for Datco's employees to learn this new tool and within several months it was capable of successfully migrating client's databases to a new environment.

By exploiting the multi-dimensional quality of information Datco achieved a considerable competitive advantage. It was capable of providing a higher value service to its clients.

Datco's journey demonstrates the power and value of providing accessible information. It was particularly valuable for Datco since its clients are engaged in the business that depends on satisfaction by

continuing customers. Without highly accessible information about detailed customer information, marketing can not be targeted effectively. Without highly accessible information about the effectiveness of various marketing methods and tools, the management cannot evaluate their efforts for customer satisfaction. In short, Datco redefined data-marketing as delivering quality information about customers instead of delivering accurate names and addresses.

4.3. Journey III: Accessible Information from Heterogeneous Data

This journey involves an improvement focused on accessibility of information (from low accessibility to high accessibility). The improvement covers accessible rearrangement of locally accurate information (Figure 6). The remedial project included the data warehouses for easily accessible use of data from different departments. The priority is given to establishing accessible storage and retrieval techniques for information users tasks particularly for cross-departmental analyses.

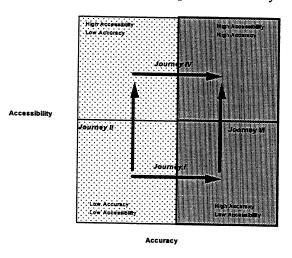


Figure 6: Information Quality Journey III

<u>Comco's Journey</u>: Differentiating Angus with highly accessible information by integrating locally accurate and heterogeneous information

Comco's journey covers integration of data for making quality business decisions and understanding business situations across traditionally autonomous functional areas. Integration of databases create valuable information not only for understanding its business in a coherent and integrated way but also for future business decisions. This case was introduced in the earlier section.

Comco exemplifies improving company-wide *accessibility* of locally accurate information. It exemplifies that integrating disparate information harvests strategically valuable and highly accessible information for a company. The data warehousing projects deliver quality information.

4.4 Journey IV: Accurate Information from World-wide Customers

This journey involves establishing an IQ team for improving accuracy of data. It focused on information collection and storage (Figure 7).

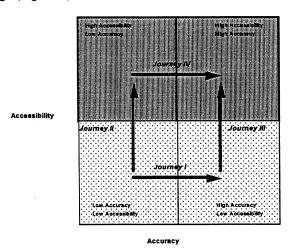


Figure 7: Information Quality Journey IV

Finco is an international bank with branches established in New York, Singapore, London, Tokyo, among others. Every country branch had its own customer file, which made it difficult to estimate global risk limits for international clients. Some clients were assigned with multiple keys since they applied for an account from multiple countries. Finco built a master customer information file, which was fed by the customer files in all branches. Joining all customer files revealed that Finco had a more serious problem with the accuracy of their customer information. Finco's customer database was lack of rigorous definition for its data fields. Different account managers assigned different definition to the same fields, which created inconsistency. For example, the same client was assigned with a different domicile codes depending on account managers and countries. In addition, Finco did not have an established company-wide information policies that could manage information ownership and thus would empower activities that generate quality information. All changes on all accounts were reported to account managers but the account managers did not have appropriate tools for evaluating the quality of account data. Most of account managers were simply collecting piles of reports without any procedures in place for reviewing and solving the problems with their account data. Finco's management recognized the importance of solving information accuracy problem and established a team of data quality for improving accuracy of their static data and for establishing a company-wide procedures for evaluating and improving quality of their information.

Finco's journey highlights establishing an IQ team for improving information accuracy companywide. Finco streamlined the information collection and storage processes so that the established capacity for information accessibility would not suffer from low accuracy of information. A designated IQ team along with established IQ policies can be examples for other companies to follow. In sum, the four types of journeys are presented in Table 1.

Table 1: Improving IQ: The Four Journeys

Table 1. Improving 1Q: The Four Journeys				
Journey Types	Objectives	Example Projects:		
		How to get there		
Journey I:	Improve Accuracy	Accurate collection of information		
Highly accurate information	Accurate representation of reality: process and information value	Accurate storage of information Accurate use of information Process stream-lining Data scrubbing		
Journey II: Highly accessible information	Improve Accessibility Intelligent/quality decision- making based on available information: information integration and aggregation	Accessible storage of information Accessible use of information Analysis streamlining Data warehouse		
Journey III: Highly accessible and highly-accurate information	Improve Accessibility Intelligent/quality decision-making based on highly accurate information: information integration	Accessible storage of information Accurate collection of information Accessible use of information Increase analysis capability Data warehouse, information integration and aggregation		
Journey IV: Highly accurate and highly-accessible information	Improve Accuracy Accurate data collection for accessible data use	Accurate storage and representation of information Accessible use of information New procedures for accurate collection, storage and retrieval Increase analysis capability across multiple departments		

4.5. Implementing Quality Information Projects

The process of implementing quality information improvement projects in organizations involves the following three steps: design, execution and deployment (Table 2); the detailed steps for the design are illustrated below (Table 3) [4]. Initiating an information quality improvement effort can be most effective when introduced immediately after a major data quality issue has been identified. It is best executed as an independent project by a small team of business, technology and data experts (an IQ team).

Table 2: Implementation Steps for Quality Information Projects

Step 1: Design	Step 2: Execution	Step 3: Deployment
 Define and diagnose IQ issues/problems Design cost efficient steps for fundamental and measurable improvements of IQ Design a dynamic process for supporting high quality information 	 Implement data quality improvements Put in place a dynamic process for supporting high quality information 	_ Ensure that implemented improvements have a high impact on decision-making capabilities _ Design the most effective way to use information throughout an organization

Table 3: Design Steps for Quality Information Projects

	Table 3: Design Steps for Quanty Information Projects
Establish a working team	 Organize an IQ team of several full-time participants Identify business community leaders and establish a team for defining business rules and providing consultation to the IQ team Identify Information Technology community leaders and establish a team for defining information infrastructure rules and consulting the IQ team
Use IQ lens	 Understand IQ problems by discussing major IQ problems with both business and technology community participants Formulate IQ problems from IQ lens perspective
Define logical data entities in the context of its business use	 Obtain a list of physical data entities being collected by an organization and combine them into logical objects. Map existing information collection processes. Identify individuals responsible for collection of information Map existing business processes which use information Identify individuals/systems who are using information for decision-making Identify when each logical data entity is being created, read, updated and deleted in relation to data collection and usage processes Estimate the volume of information being collected daily
Prioritize dimensions of information quality	 Introduce multi-dimensional quality of information to the team of business experts Conduct sessions to prioritize dimensions of information quality Define which dimensions are most critical for organization. We will denote the set as {D} For every element d of {D} estimate the current level of IQ in dimension d sketch the desired level of IQ and an optimal scenario for using high information quality in dimension d
For every dimension formalize IQ problem and design IQ metric	 For every element d of {D} formalize the existing information quality problem in dimension d formalize the desired state of information quality in dimension d identify critical business processes for collection and usage of information in d Design a metric for measuring IQ in dimension d. The metric can be qualitative Measure current IQ for every d of {D} Research the connection between elements of {D}
problem in multi- dimensional space	2. Assign some weights to every d of {D}3. Identify a utility function U({D})
Design methods for improving IQ	 Design methods for improving information quality for every d of {D} Estimate the impact of every method on other dimensions Using IQ metric estimate the effectiveness of every method
Define optimal solution	 Estimate cost effectiveness of every method Evaluate all methods from System Infrastructure and Business Process perspective Choose an optimal solution Define the execution plan
IQ support	 Define IQ dynamics Design a process for periodic IQ measurement Design a dynamic process for supporting quality information

5. Conclusion

The persuasiveness of IQ problems in organizations and the severe financial and operational costs of such problems have become a critical issue in organizations [7]. Exploiting quality information as a competitive core capacity yields a new organizational capacity for making quality decisions. This newly

developed capacity, in turn, harvests advances in new product, service, and market strategies in the changing environment.

This paper presents definitions and a methodology for diagnosis and improvement of information quality. With the increasing dependence of organizations on the quality of their information, the business process improvement approach or the systems infrastructure approach alone is no longer a viable solution. Organizations must learn to view their business through the IQ lens. They must recognize the potential value and power of collecting, storing, and using quality information and proactively develop solutions. This requires knowledge of the information manufacturing process and understanding why the process performs or fails to perform as it should [5].

With this knowledge, managers and IS professionals can proactively develop long-term solutions to IQ problems in their organizations. In particular, early adopters of quality information perspective should be able to exploit new opportunities from improved accessibility and accuracy of information as demonstrated in this paper.

6. References

- [1] Ballou, D. P. and H. L. Pazer, Designing Information Systems to Optimize the Accuracy-Timeliness Tradeoff. Information Systems Research (ISR), 6(1) 1995, pp. 51-72.
- [2] Deming, E. W., Out of the Crisis. Center for Advanced Engineering Study, MIT, Cambridge, 1986.
- [3] Firth, C. P. and R. Y. Wang, Data Quality Systems: Evaluation and Implementation. Cambridge Market Intelligence Ltd., London, 1996.
- [4] Lee, Y. and R. Y. Wang (1996). Information Space: Where does quality information come from? (No. TDQM-96-01). MIT Sloan School of Management.
- [5] Lee, Y. W. Why "Know Why" Knowledge is Useful for Solving Information Quality Problems in Proceedings of The Second AIS Conference. Phoenix, AZ, 1996.
- [6] McGee, J. and L. Prusak, Managing Information Strategically. The Earnest & Young Information Management Series, John Wiley & Sons, Inc., New York, 1993.
- [7] Strong, D. M., Y. W. Lee and R. Y. Wang, Data Quality in Context. Forthcoming in Communications of the ACM, 1996.
- [8] Wand, Y. and R. Y. Wang, Anchoring Data Quality Dimensions in Ontological Foundations. Forthcoming, Communications of the ACM, November 1996.
- [9] Wang, R. Y., V. C. Storey and C. P. Firth, A Framework for Analysis of Data Quality Research. IEEE Transactions on Knowledge and Data Engineering, 7(4) 1995, pp. 623-640.
- [10] Wang, R. Y. and D. M. Strong, Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems (JMIS), 12(4) 1996, pp. 5-34.