Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2003 Proceedings

Americas Conference on Information Systems (AMCIS)

December 2003

Evaluating Completeness of an Information Product

Yu Cai
Boston University

Mostapha Ziad
Suffolk University

Evaluating Completeness of an Information Product

Yu Cai
Boston University
ycai@bu.edu

Ganesan Shankaranarayanan
Boston University
gshankar@bu.edu

Mostapha Ziad
Suffolk University
mziad@acad.suffolk.edu

Abstract

Research in data quality has identified several dimensions for evaluating data quality including accuracy, timeliness, completeness, and relevance just to name a few. While dimensions such as accuracy and timeliness have received considerable attention from researchers, completeness has not received explicit consideration. In this paper we propose a theoretical method for evaluating completeness as a data quality dimension. This method is built on top of the information product map (IPMAP), a representation scheme for modeling the manufacture of an information product (IP). Moreover, the methods presented in literature for evaluating data quality (focusing on accuracy and timeliness) take an objective view of the evaluation and do not permit the decision-maker to inject his/her own subjective interpretations. We adopt the view that evaluating data quality is not completely objective and requires some subjective interpretation. We take this view because in the dynamic decision environments of today it is not sufficient to inform the decision-maker about the quality of the data. It is also necessary to permit the decision-maker to evaluate the quality by him/herself. This is especially important in dynamic decision environments in which decision-makers deal with very large volumes and widely distributed data sources and with a wide variety and frequency of decision-tasks. The Internet and associated technologies help create such dynamic decision environments within the e-business and m-business concepts that they support. The framework for completeness proposed here is designed to support data quality management in such decision environments.

Keywords: Information quality (IQ), data quality (DQ), completeness, measurement development, IQ/DQ frameworks, IPMAP

Introduction

The quality of decisions made is heavily dependent on the quality of data[1] used in the decision-making process. Poor quality data can have severe impacts on organizations including customer dissatisfaction, increased cost, difficulty to execute strategy, reduced job satisfaction and especially affecting effective decision-making (Redman 1995, 1998; Wang 1998). Despite understanding this, the huge volume of data produced, transferred, processed and stored every day does not meet the required quality standards. The Internet combined with mobile technologies and wireless devices have increased the volume of data handled as well as the frequency of decision tasks while permitting decision-makers to readily access information that is widely distributed. In such environments it is important to assure decision-makers of the quality of information they use in the decision making process.

Research has illustrated that data quality may be evaluated along several different dimensions (DeLone and McLean 1992; Wang and Strong 1996; Redman 1996). Fox et al. (1994) discussed four important dimensions of data quality: accuracy, completeness,

[1]In this paper we use the terms "data quality" and "information quality" interchangeably.

consistency, and currentness. Miller (1996) identified ten dimensions of information quality including accuracy, timeliness, completeness, and relevancy. Based on a two-stage survey, Wang and Strong (1996) analyzed and identified the multiple attributes of data quality from data consumers' perspective. They grouped the attributes into four categories: intrinsic, contextual, representational and accessibility (See Table 1).

Table 1. Data Quality Categories and Dimensions Model of Wang and Strong (1996)

Category	Intrinsic Data Quality	Accessibility Data Quality	Contextual Data Quality	Representational Data Quality
Dimensions	Accuracy Objectivity Believability Reputation	Accessibility Security Ease of operations	Timeliness Completeness Relevancy Value added Amount of data	Interpretability Ease of understanding Concise representation Consistent representation

Wand and Wang (1996) identified completeness as being one of the most important and often-cited quality dimension based on data quality literature. It is identified and analyzed along with other important data quality dimensions-- accuracy and timeliness (Bailey and Pearson 1983; King and Epstein 1983; Ballou and Pazer 1985; Srinivasan 1985; Mahmood 1987; Miller and Doyle 1987). Of these, accuracy and timeliness have received explicit attention (Ballou and Pazer 1995; Ballou et al. 1998). However, completeness has not received much attention. In the decision making process, the completeness of information plays a crucial role. Previous research in information economics theoretically indicates that more complete the information, the higher its normative value (the value calculated through economic modeling), given the assumption of "rational behavior" (McGuire and Radner 1986). In the organizational decision making processes, it is safe to assume most decision makers are rational, or at least, try to be rational. Moreover, the level of completeness of information can also impact the decision makers' judgmental confidences (Alba and Hutchinson 1987; Sanbonmatsu et al. 1992). Ahituv et al. (1998) showed complete information improves the performance (i.e. the quality of decision) in a dynamic environment.

The perceived quality of the data is influenced by the decision-task. The same data may be viewed with two or more different quality lenses according to the decision-task it is used for. For example, a faculty member trying to place orders for course textbook may find it sufficient to base the decision on an approximate figure of the number of students enrolled. The same faculty member is unlikely to rely on this figure as a representation of his/her class-size when requesting a room (based on seating capacity) for class meetings. Therefore, the decision-maker must not just be informed about the quality of the data he/she is using but also be permitted to evaluate the quality by himself/herself. In this paper we adopt the view that evaluating data quality has a subjective component associated and the decision-maker must be permitted to evaluate the quality of the data by including subjective interpretations.

We propose a method for evaluating completeness as a data quality dimension. This method is based on the information product approach, specifically, the information product map (IPMAP) proposed in (Shankaranarayanan et al. 2000, 2003). The IPMAP lends itself to evaluate completeness of a product as a whole and not just at the data element level. By viewing information as product, the final quality of information product is influenced by and can be traced back to the quality of "raw" input data elements. In developing the framework for completeness, we use the input data elements as anchoring points. We also identify the need for subjectively evaluating completeness. We propose a comprehensive framework for evaluating completeness of an information product by examining both the objective and subjective views of quality. In developing a product and service performance model for information quality (PSI/IQ), Kahn et al. (2002) argued that one important aspect of quality is conformance to specification. Following their logic, without an appropriate measurement, it is impossible to know the level of conformance to specification. Therefore, we believe our work towards developing a measurement for completeness is an important step to push data quality research further.

The remainder of this paper is organized as follows. Section 2 presents an overview of the relevant literature. We summarize the IP approach, the information manufacturing system (IMS) upon which our framework for completeness is built. In section 3, we propose a measurement framework that accounts for both objective and subjective measurement of completeness. We use a supply chain example to illustrate the application of the measurement framework. The conclusion and the future research directions are presented in section 4.

Relevant Literature

Ballou and Pazer (1985) defined completeness based on the concepts of "no missing value" and "all values for a certain variable are recorded". Miller (1996) pointed out that complete information for one person may be incomplete for another and different tasks may require different levels of detail. However, this research neither defined completeness, nor provided a method for evaluating it. Redman (1996) regarded completeness as a data quality dimension pertaining to data values: "Completeness refers to the degree to which values are present in a data collection". His definition focuses on the completeness of attribute values in a database. In his opinion, for an individual datum, the completeness can only have two situations-- either a value is assigned or not. Wang (1998) used an example of a client account database to define completeness as a quality measurement: "the percentage of non-existent accounts or the number of accounts with missing values in the industry-code field". In an empirical research based on the data from air force simulation, Ahituv et al. (1998) examined the influence of different attributes of information on decision makers who faced dynamic decision environment (how to utilize airborne and other defensive resources to defeat a massive aerial attacks). Their analysis on completeness was a comparison of complete information vs. incomplete information scenarios; no theoretical assessment for completeness was introduced. In this paper, we propose a theoretical approach for computing the completeness of an information product.

Pipino et al. (2002) identified three pervasive functional forms (simple ratio, min or max operation, and weighted average) for the objective data quality assessments. When addressing completeness, they identified three types of completeness: schema, column, and population completeness. They further argued that each of them could be measured by a simple ratio function (i.e. the ratio of the number of incomplete items to the total number of items and subtracting from 1). Their perspectives, as with other previous research (Ballou and Pazer 1985; Redman 1996; etc.) on completeness, focus at the data level instead of the information product (IP) level. Data quality researchers have viewed data as asset and resource (Redman 1995, Levitin and Redman 1998). The IP approach adopted in this paper, views the information as a product instead of a "by product" of information systems (Wang et al. 1998). In the real world, the dynamic decision environments require swift decision-making. In most cases, decision makers do not pay close attention to the large volume of relevant data in full detail. To accomplish the mission, highly comprehensive and customized reports are generalized to facilitate the decision makers' job. Therefore, the assessment of quality at the data level is insufficient. A method capable of providing the data quality measurement at the IP level has more practical significance. In this paper we address the completeness of an IP instead of at the data level.

The data quality issue is inherent in the information systems management (DeLone and McLean 1992). Errors may come from data input, data update and other processes (Morey 1982). Several approaches such as data tracking and statistical process control (Redman 1996), data cleansing (Hernandez and Stolfo 1998) and IP approach proposed by Wang et al. (1998) have been developed based on the different perspectives for examining data quality. Information product is the deliverable that corresponds to the specific requirements of the end users/consumers. Examples of IPs include invoices, business reports, and prescriptions. The IP approach provides a foundation for developing systematic frameworks for tracing, evaluating and managing data quality in information systems. It takes both data element and data processing into consideration. Compared with other approaches, the IP approach offers some distinct advantages. First, IP approach views information as product. As a result, we can adopt total quality management (TQM) techniques that have been successfully applied to manage quality in physical product manufacture to implement total data quality management. Furthermore, as the IP approach is akin to physical product manufacture, it is easy for the business managers to understand it. Second, the spirit behind IP approach is highly customer oriented. The goal of the data quality management is meeting the information product user's requirements. Moreover, the IP approach considers the quality of an information component not just in an isolated fashion but also in the larger context of the "manufacture" of an IP. The IP approach integrates the concept of an information supply chain, as it is possible to trace/visualize the flow of data across business units and departmental boundaries and through the manufacturing processes that create an IP. In today's increasingly networked environment, the ability to trace multiple upstream sources of information products is a desirable feature.

Ballou et al. (1998) proposed a modeling method called the information manufacturing system (IMS). They defined information manufacturing system as "information systems that produce predefined information products". We believe their work is very valuable as a starting point for data quality evaluation. But their framework largely adopts an objective view for evaluating data quality. Considering the complex reality of decision-making, we believe that the objective view alone is insufficient. In this paper, we present the drawbacks of the objective views and propose a more balanced framework by taking subjective issues into account, focusing on completeness. Furthermore, the IMS does not permit explicit representation of certain "processes" or "events" in IP manufacturing such as the crossing-over of organizational or system boundaries that are typical in the networked environments of today. The proposed framework for evaluating completeness is hence based on the IPMAP that offers a more explicit and comprehensive representation of the manufacture of IPs.

Evaluating Completeness

The framework for evaluating completeness proposed in this paper is based on the IP approach for managing data quality. Completeness, as quality dimension, requires understanding *when* and *how* data elements are included / processed to create the final information product. On the other hand, to evaluate the completeness of an information product, it is necessary to examine all the constituent data elements of this product starting with the source of each element. The IP approach facilitates this view of completeness and allows us to comprehensively evaluate completeness.

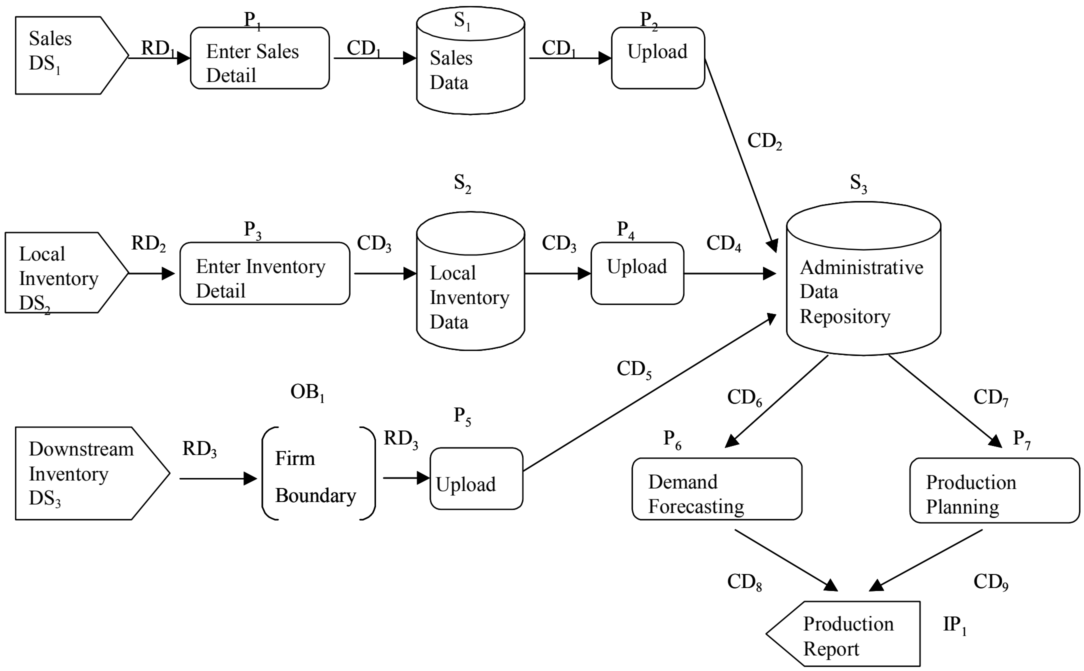

Figure 1. An IPMAP Example in Supply Chain Management

The information product map (IPMAP) (Shankaranarayanan et al. 2000) permits the systematic representation of the manufacturing processes that create an information product. The IPMAP is a graphic model similar to Data Flow Diagram (DFD). It's an extension of the IMS for a more explicit and comprehensive representation of the information product manufacturing process. The IPMAP permits representing the flow of data across organizational and information system boundaries. The IPMAP allows the decision-maker to visualize not only the widespread distribution of data and other resources but also the flow of data elements and the sequence by which these data elements are processed to create the required IPs. Therefore, the IPMAP helps decision makers understand the IP. Moreover, it also provides a powerful vehicle to evaluate IP quality. The constructs in an IPMAP include the source block, the processing block, the data quality check block, the organizational boundary block, the information system boundary block, and the sink block. Complete details on the constructs in an IPMAP are available in (Shankaranarayanan et al. 2000). The constructs and the layout of IPMAP are illustrated using an example in figure 1. This IPMAP is for a requirements planning report in a supply chain. This example is also used to illustrate the framework for completeness in the later part of this section.

Current methods for evaluating data quality dimensions do so in an objective manner. Methods that objectively evaluate data quality are important and desirable. Each data element that goes into creating an information product may be input /maintained /managed by several individuals /roles /departments. If we permit users (data creators, custodians) to subjectively assign values to quality dimensions of data elements that each is responsible for (for personal gains) without any objective criteria, this would bias the quality evaluation and render the process useless. However, objectively evaluating data quality in dynamic decision environments is difficult and may often be misleading and useless. Consider the completeness dimension as an example. An inventory report (say inventory levels of products in multiple distribution centers) is an IP and may be used by several decision-makers. Let us further say that this report has 10 constituent data elements of which 5 are missing. An objective evaluation of this report would deem it 50% complete. However, for a specific task, the decision maker may only need a part of this report (say 5 of the 10 data elements). If the 5 elements available are the ones he/she needs the report, in his/her perspective, it is a 100% complete. In this case, the objective evaluation may be considered as rather misleading. Alternatively, the same report may have 8 of the 10 data elements available making it 80% complete. If the decision task requires the two data elements that are missing from this report, then the decision-maker is correct in gauging the report (though 80% complete) as being useless and its evaluated completeness value even more so. Therefore, in many circumstances, an isolated "objective" figure without context is not very meaningful or even misleading. Since the data quality of a data element is dependent on the decision-task it is used for, decision-makers must have the flexibility to evaluate data quality from the viewpoint of a decision-task. Introducing subjective considerations into data quality measurement is a necessary step towards understanding the implications for data quality.

A classic definition of quality is fitness for use, or the extent to which a product successfully serves the purposes of customers (Juran et al. 1974). This implies the data quality is a concept that depends on the various purposes that the data (IP) are intended for. In a recent paper, Kahn et al. (2002) defined completeness as "the extent to which data is not missing and is of sufficient breadth and depth for the task at hand." The later part of the definition, "task at hand" identifies the need for subjective consideration. Pipino et al. (2002) briefly addressed the issue of subjective data quality assessments, but they did not provide a detailed analysis. Contrary to their argument that the subjective and objective assessments are two independent processes and must be considered separately, we attempt to incorporate both objective and subjective considerations into the measurement framework for completeness.

In this paper, we define the completeness of an information product as follows. An information product is complete if it includes all the data elements needed by the decision-maker for the decision-task. In other words, an information product may have missing values for some data elements but still be perceived as complete by the decision-maker. Although completeness is a highly contextual dimension of data quality, an objective measurement framework is still very helpful for facilitating the common understanding of completeness. We use the following notations to define the completeness measurement adopting a purely objective approach.

Let $\boldsymbol{Dj}$ define the completeness of the j^{th} raw data element.

$$\boldsymbol{Dj} = 0, \text{ if the value is missing} \qquad (1)$$
$$\boldsymbol{Dj} = 1, \text{ if the value is present}$$

Let $\boldsymbol{E_i}$ define the completeness of the i^{th} component of the IP

$$E_i = \frac{\sum_{J=1}^{m} D_j}{m} \qquad (2)$$

m = the number of raw data points needed to get E_i

An IP is typically composed of both raw data elements as well as component data elements. A raw data element is a data element that is obtained from a source and is used directly without undergoing any processing that changes the data element in any way. A component data element is one that is created by processing one or more data elements. Even if a raw data element goes through an inspection process (where it may just be examined and nothing changed, in this paper we treat this as a component data element to differentiate it from a raw data element that does not undergo any kind of processing.

The completeness of a raw data element i, D_i, is easily specified by whether its value is available or is missing, as defined in (1). If a component data element is made up entirely of just raw data elements, its completeness value may be computed using equation (2) given above. We refer to a component data that is composed entirely of raw data elements as an IP component.

However, there may exist component data elements that are in turn composed of one or more different component data elements and raw data elements. The final information product, IP, is an example of one such component data element. Such component data elements may also be outputs of one or more intermediate stages in an IPMAP. For example, if you are only interested in the average life expectation of males in the United State, it is a final IP. However, if you want a report on the average life expectation of the population in the United States that consists of both males and females, the information on males is only one component of the report, which it's the final IP in this case. Equation 3 given below can be used to evaluate the completeness of an IP or any intermediate component data.

Let C define the completeness measure of an IP.

$$C = \sum_{i=1}^{n} w_i E_i \qquad (3)$$

N = The number of IP components that make up the final IP

$\boldsymbol{W_i}$: The weight of E_i in Information Product C; ($\sum w_i = 1$)

The weight "W_i" makes it possible to incorporate the decision maker's subjective consideration. "W_i" may be subjectively assigned a value between 0 and 1 to specify how relevant is the completeness of a specific IP component in calculating the overall completeness of the IP. Decision-makers may assign and/or change the weights assigned to evaluate the completeness of the IP. (For clarification, the purpose of the proposed measurement is to provide an indicator on the completeness of information products to decision-makers. It's not a statistical test or method).

The following illustrative example demonstrates how this framework to evaluate completeness may be used. The IPMAP corresponding to the information product (called the Production Report for convenience) is shown in figure 1. Assume a company Widget Inc. makes several widgets, one of which is widget *M*. Widget *M* is purchased from Widget Inc. by five customers (named X_1 through X_5). Each customer places one order for widget *M* every month. Widget Inc. has an ERP system to update inventory data daily. Two out of five customers agree to share their inventory data with Widget Inc. in exchange for price discounts and better service. They provide their inventory levels of widget *M* on a monthly basis.

Widget Inc. needs to forecast its demand for the purpose of planning its production. In order to obtain a superior forecast, Widget Inc. believes that the historical sales data is not sufficient. The fact that the quantity of *M* ordered by customer X_1 was high in the previous month does not necessarily imply a large order this month. This is because the inventory level at customer X_1 also plays an important role in determining customer X_1's order this month. Ideally, Widget Inc. would like to have the entire inventory level data from all of its five customers.

Widget Inc. uses the previous months' data on order volumes and inventory levels to estimate their future orders. Even though it has the order volumes corresponding to all of its five customers, it can only access the inventory level data corresponding to two customers. So the data on order volume is 100% complete, but the input inventory level data from customers is only 40% complete. In order to get demand forecasting based on incomplete data, Widget Inc. uses the average of the inventory levels available from the two customers as an estimate of the average inventory level of all five customers. The estimated average inventory level data is also 40% complete (By using equation 2).

By combining the order volume data (completeness = 1.0) and estimated average inventory levels data at the customers (completeness = 0.4), Widget Inc. could arrive at the demand forecast for product *M*. Using equation 3 (C= $\sum_{i=1}^{n} w_i E_i$), the completeness of the demand forecast can be evaluated as 0.5*1.0+0.5*0.4 = 0.7 (assuming the sales data and estimated average inventory levels are equally weighted in demand forecasting). Widget Inc. uses this demand forecast data along with the on-hand inventory levels of product *M* for arriving at its production plan. Therefore, the completeness of the production plan is 0.5 * 0.7 +0.5 * 1.0 = 0.85 (also assuming that W_i for both IP components are equal here).

Similarly, as the final report is created using both the demand forecast and the production plan, the completeness of the final report can be evaluated as 0.5 * 0.7 + 0.5 * 0.85 = 0.775 (again assuming that the values assigned to W_i is the same for both).

If a decision-maker in Widget Inc. believes that the on-hand inventory data is more important (or has a greater impact) than the demand forecast data in creating the production plan (because the inventory in stock data are more reliable than the "forecasted" data), he/she can assign weights to the two IP components in a manner consistent with his/her beliefs. For example, W_i for the on-hand inventory data may be assigned 0.6 and that for demand forecast 0.4. The completeness of the production plan will now be 0.6 * 1.0 + 0.4 * 0.7= 0.88. The change here will impact the completeness of the final IP report. The new completeness for the final report will now be 0.5 * 0.7 + 0.5 * 0.88 = 0.79 (assuming both input components are viewed as being equally important / relevant).

Conclusions and Directions for Further Research

Data quality problem can have severe impacts on organizations. Completeness is among the most often-cited dimensions of data quality. While dimensions such as accuracy and timeliness have received considerable attention from researchers, completeness has not received explicit consideration. In this paper we propose a theoretical method for evaluating completeness as a data quality dimension not only at the data level but more importantly at the IP level. This method is built on top of the IPMAP, a representation scheme for modeling the manufacture of an IP. Moreover, the methods presented in literature for evaluating data quality (focusing on accuracy and timeliness) take an objective view of the evaluation and do not permit the decision-maker to inject his/her own subjective interpretations. We argue that the quality of the data is dependent on the decision-task and the same data may be viewed with two or more different quality lenses according to the decision-task it is used for. In the dynamic decision environments of today it is not sufficient to inform the decision-maker about the quality of the data. It is also necessary to permit the decision-maker to evaluate the quality by him/herself. This is especially important in dynamic decision environments in which decision-makers deal with very large volumes and widely distributed data sources and with a wide variety and frequency of decision-tasks. Therefore, we adopt the view that evaluating data quality is not completely objective and requires some subjective interpretation. We believe our work towards developing a comprehensive measurement framework for evaluating completeness of an information product by examining both the objective and subjective views is an important step to push data quality research further.

Although it's important to incorporate decision-maker's subjective concerns into data quality evaluation, subjectively assigning weight to define quality has an inherent drawback. Users could assign "false" weights for personal gains. An incentive system must be implemented that accounts for such spurious weight assignments. Such incentive scheme(s) must be in place to effectively implement and manage data quality in a manner that is beneficial to decision-makers and organizations.

Decision-making is a very complex process. The framework we developed based on IP approach helps inform the completeness measure of an IP to decision makers. Further research can be done to understand the decision-makers' common behavior and individual differences in using completeness measurement in various decision tasks.

We foresee the possible uses of the completeness measure in the following ways: 1) Decisions can be adjusted based on the completeness of an IP. For instance, in the above example in section 3, if the completeness measure of the production plan is below 0.90, Widget Inc. will increase the inventory level by 5%; if the completeness measure is below 0.75, it will increase the inventory level further by 10% to offset the higher uncertainty; 2) It can help make the decision more accurate and efficient. In the same Widget Inc. setting, based on the analysis of past data, manager found with above 0.85 completeness of production report, 1000 units of inventories are good enough. So the manager could make better inventory management decisions by knowing the completeness measure.

Understanding how to measure the data quality dimensions is important, but it's not sufficient to solve the pervasive and costly data quality problem. The IPMAP upon which our framework for completeness is built has the capability to trace the flow of data/IP components in the IP creation process, as well as the widespread distribution of data and other resources even across system/organizational boundaries. By using the graph theory to analyze the IPMAP graphic diagram, it's quite possible that we can not only evaluate data quality of an IP, but more importantly, trace back and find out where and how the error happened. Further research in this direction has great promise in facilitating solving the data quality problem from its origin.

Currently we are in the process of developing a prototype system using the IPMAP for managing data quality. This system will incorporate the completeness framework proposed and will permit decision makers to evaluate data quality based on the decision task. We further propose that the system be offered as a web service.

References

Ahituv, N., Igbaria, M. and Sella, A. "The Effects of Time Pressure and Completeness of Information on Decision Making", *Journal of Management Information Systems*, (15:2), 1998, pp. 153-172.

Alba, J.W., and Hutchinson, J.W. "Dimensions of Consumer Expertise". *Journal of Consumer Research*. (13:4), 1987, pp. 411-454.

Ballou, D. and Pazer, H. "Modeling Data and Process Quality in Multi-input Multi-output Information Systems," *Management Science*, (31:2), 1985, pp. 150-162.

Ballou, D. and Pazer, H. "Designing Information Systems to Optimize the Accuracy-timeliness Tradeoff", *Information System Research*, (6: 1), 1995, pp. 51-72.

Ballou, D., Wang, R. Y., Pazer, H., and Tayi, G. K. "Modeling Information Manufacturing Systems to Determine Information Product Quality", *Management Science,* (44: 4), 1998, pp. 462-484.

Bailey, J. E. and Pearson, S. W. "Development of a Tool for Measuring and Analyzing Computer User Satisfaction", *Management Science*, (29: 5), 1983, pp. 530-545.

DeLone, W. H., McLean, E. R. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research*, (3: 1), 1992, pp.60-95.

Fox, C., Levitin, A. and Redman, T. "The notion of data and its quality dimensions", *Information Processing & Management*, (30: 1), 1994, pp. 9-20.

Hernandez, M. A. and Stolfo, S. J. "Real World Data is Dirty: Data Cleansing and the Merge/Purge Problem", *Journal of Data Mining and Knowledge Discovery*, (2: 1), 1998, pp. 9-37.

Juran, J. M., Gryna, F. M.J., and Bingham, R.S. Quality Control Handbook (3rd ed.). McGraw-Hill Book Co, New York, NY, 1974.

Kahn, B. K., Strong, D. M. and Wang, R. Y. "Data quality Benchmarks: Product and Service Performance," *Communications of ACM*, (45:4), 2002, pp. 184-193.

King, W. R. and Epstein, B. J. "Assessing Information System Value," *Decision Sciences*, (14:1), 1983, pp. 34-45.

Levitin, A.V. and Redman, T.C. "Data as a Resource: Properties, Implications, and Prescriptions", *Sloan Management Review*, (40:1), 1998, pp.89-101.

Mahmood, M. A. "Systems Development Methods—A Comparative Investigation," *MIS Quarterly*, (11:3), 1987, pp. 293-311.

McGuire, C.B., and Radner, R., eds. "Decision and Organization", 2nd ed. Minneapolis: University of Minnesota Press, 1986.

Miller, H. "The Multiple Dimensions of Information Quality", *Information Systems Managemen*t, (13:2), 1996, pp. 79-82.

Miller, J. and Doyle, B. A. "Measuring Effectiveness of Computer Based Information Systems Financial Services Sector," *MIS Quarterly*, (11:1), 1987, pp. 107-124.

Morey, R. C. "Estimating and Improving the Quality of Information in the MIS", *Communications of ACM*, (25:5), 1982, pp. 337-342.

Pipino, L. L., Lee, Y. W. and Wang, R. Y. "Data Quality Assessment", *Communications of ACM*, (45:4), 2002, pp.211-218.

Redman, T. C. "Improve Data Quality for Competitive Advantage", *Sloan Management Review*, (37:2), 1995, pp. 99-107.

Redman, T.C. (Ed.) Data Quality for the Information Age, Artech House, Boston, MA, 1996

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

Sanbonmatsu, D.M., Kardes, F.R., and Herr, P.M. "The Role of Prior Knowledge and Missing Information in Multiattribute Evaluation", *Organizational Behavior and Human Decision Making Process*, (51), 1992, pp. 76-91.

Shankaranarayanan, G., Wang, R. and Ziad, M. "IPMAP: Representing the Manufacture of an Information Product", *The Proceedings of MIT Data quality Conference (IQ 2000),* Boston, MA, 2000.

Shankaranarayanan, G., Ziad, M.and Wang, R. "Managing Date Quality in Dynamic Decision Environment: An Information Product Approach", Forthcoming in the Journal of Database Management, 2003.

Srinivasan, A. "Alternative Measures of System Effectiveness: Associations and Implications," *MIS Quarterly*, (9:3), 1985, pp. 243-253.

Wand, Y. and Wang, R. Y. "Anchoring Data Quality Dimensions in Ontological Foundations", *Communication of the ACM*, (39:11), 1996, pp. 86-95.

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

Wang, R.Y., Lee, Y. W., Pipino, L. and Strong, D. M. "Manage your Information as a Product", *Sloan Management Review*, (39:4), 1998, pp. 95-105.

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