Journal of International Information Management

Volume 9 | Issue 1 Article 4

2000

Managing data quality in the health care industry: Some critical issues

Binshan Lin
Louisiana State University

Houn-Gee Chan
National Chung Cheng University

Managing data quality in the health care industry: Some critical issues

Binshan Lin
Louisiana State University - Shreveport
Houn-Gee Chen
National Chung Cheng University

ABSTRACT

Data quality analysis in health care is a large and ongoing problem. Because the scope of health care information systems is extremely broad and because most such systems have either a direct or indirect impact on the provision of health care, their quality is a topic of critical importance for information management professionals. Of vital concern is the data flow from the point of patient service delivery to an end point of decision support and analysis of the data derived from the encounter. The historical use of these data, their current uses and how industry dynamics have created a need for improved data quality provides the context for a discussion of health care data quality. We address the definition of health care data, assessing data quality, and the need for effective data quality management. Next, the paper reviews some popular methods employed in the health care industry. We then propose an integrated framework to improve data quality in health care, and finally summarize the conclusions of this study.

INTRODUCTION

Data quality analysis in health care is a large and ongoing problem (Horner, 2000). Even with the increased emphasis on quality of coding by the requirements of the contracting arrangements, it remains clear that poor quality data are still inhibiting use of the data set, both for clinical and general management (Cleray, et al., 1994). Problems with data quality in health care were identified as early as 1977 (Institute of Medicine, 1977), and these were emphasized by the introduction of the prospective payment system (Hsia, et al., 1988). Past studies found that the accuracy or completeness of data, or both, ranged from 60% to 90% with certain coding being totally inexplicable (Whates, et al., 1982; Sunderland, 1985).

Today there is a strong evidence that data stored in health care information systems have a significant number of errors (Redman, 1996). Errors in data can cost a company millions of dollars, alienate customers, and make implementing new strategies difficult or impossible (Redman, 1995). Broad and intensive efforts at every level are needed to stem the medical errors that cause as many as 98,000 deaths each year, according to a new report from the Institute of Medicine (IOM, 1999).

Many health care managers are unaware of the quality of data they use and perhaps assume that information technology ensures that data are perfect. Providers of health care utilize coded clinical data to measure outcomes, formulate advantageous payer contract negotiations, deliver cost-efficient treatment, and generate reimbursement. Third party payers use the same data to organize, operate, market, and secure profitable business. Although poor quality appears to be the norm, rather than the exception, they have largely ignored the issue of data quality in health care. Maintaining the quality of the data that is used in a health care organization is becoming an increasingly high priority for health care management.

Because the scope of health care information systems is extremely broad and because most such systems have either a direct or indirect impact on the provision of health care, their quality is a topic of critical importance for information management professionals. Of Vital concern is the data flow from the point of patient service delivery to an end point of decision support and analysis of the data derived from the encounter. The historical use of these data, their current uses and how industry dynamics have created a need for improved data quality provides the context for a discussion of health care date quality. We address the definition of health care data, assessing data quality, and the need for effective data quality management. Next, the paper reviews some popular methods employed in the health care industry. We then propose an integrated framework to improving data quality in health care, and finally summarize the conclusions of this study.

HEALTH CARE DATA DEFINED

Health care data is defined as that information used to provide, manage, pay and/or report on the services used across the entire health care system. Unlike other industries, limited transactions have been standardized, but the basic data quality lessons learned from other industries can be applied. The origin of health care data is an encounter between a patient and a health care provider. This provider may be a physician, hospital, radiologist, physical therapist, lab technician or a physiologist. From this encounter, the provider will record:

- the service rendered (an office visit, a lab test, room and board for an inpatient stay),
- the conditions of the service (the diagnosis, date, place of service),
- patient information (sex, age, patient history, insurance information, height, weight, etc.)
- clinical information (result of tests, prognoses, consultation notes).

Some standardization exists in the way data is captured. Most of this standardization has come about in the last 10 years because of the government's Medicare program. For a provider to be reimbursed by Medicare through a fiscal intermediary (i.e., the payor), the provider must submit a "claim" for services rendered. Historically, the provider submitted a nonstandard claim, which minimally included the patient name, age, sex, a unique employer/group identifier, the services provided and the data they were delivered. This information was used almost exclusively by the payer to reimburse the provider.

At times, providers report services improperly for reimbursement at a rate more than appropriate for the services provided. Studies have shown that less than five percent of providers actually submit fraudulent claims, but they account for most overpayments. As concern of fraud and abuse escalate from this apparent lack of control, standards began to develop. Certain data elements were submitted based on standard code definitions for the procedure or service performed, and the diagnosis. The service codes are maintained by the American Medical Association for physician services while diagnosis codes are maintained by the government's Health Care Financing Administration.

Once implemented, these coding rules become the basis for payment to the provider while concurrently being captured as "reimbursement" data by the payor. The complexity of correct coding creates two additional problems.

- There are approximately 8,000 procedure codes and 16,000 diagnosis codes. A physician's office staff salary is typically a minimum wage and these individuals may not have had adequate training on coding rules and methodologies.
- These methodologies change annually, but the training and materials may be viewed as an unnecessary expense. While standardization has evolved further, data actually captured are recorded and resubmitted by the provider for reimbursement, but may still not be representative of the encounter since the clinical aspects are recorded on paper and stored by the provider.

ASSESSING DATA QUALITY

The quality paradigm is difficult to describe because of its "amorphous" nature. Therefore, different authors tend to emphasize different aspects (Fox & Frakes, 1997). When the quality paradigm was formed emphasis was given only to an inspection to achieve quality - conformance to a standard or a specification. Rapid changes in the last years have lead to new definitions of quality. One of the most important is the IEEE standard definition (IEEE, 1998) in which the quality is defined as the totality of features and characteristics of a product or service that bears on its ability to satisfy given needs.

Redman (1996) defines the data quality in its broadest sense. This implies to data that are relevant to their intended uses and are of sufficient detail and quality with a high degree of accuracy and completeness, consistent with other sources and presented in appropriate ways. Tayi and Ballou (1998) have defined the term "data quality" as "fitness for use" which implies that the concept of data quality is relative. Data appropriate for one use may not possess sufficient quality for another use. A related problem with multiple users of data is also that of semantics. The data designer and/or gatherer as well as initial user may fully agree with the same definitions regarding the meaning of the various data items, but probably this will not be a view of the other users. Such problems are becoming increasingly critical as organizations implement data warehouses. To the same time the conceptual view on data is becoming more and more important as a possible solution for the mentioned problems.

Orr (1998) introduces a kind of measurement view on this term. It is defined as a measure of the agreement between the data views presented by an information system and the same data in the real world. Of course no serious information system has data quality of 100%, but tries to ensure that the data is accurate enough, timely and consistent for the enterprise to survive and make reasonable decisions. Actually the real problem with data quality is change. Data in any information systems are static, but in the real world they are changing. This is one reason more to have a conceptual view.

In some circumstances the term "information" refers to both data and information (Strong, 1997). Data usually refer to information at their early stages of processing and information to the product at a later stage. Rather than switching between the terms the information is used to refer to data or informaiton values at any point in the process. But still we must bear in our minds that different informaiton definitions depend upon different points of view. For example:

- The information management point of view -- Information is processed data (Redman, 1996).
- The information theory point of view -- Information is the non-redundant part of a message (Redman, 1996).
- The informaiton technology for management point of view -- Informaiton is data that have been organized so that they have a meaning to the user (Turban, 1996). However once a point of view is fixed, no conflict should arise and once again it is important to recognize that the prerequisite for the information quality is a data quality.

Data quality in nealth care can be classified into two dimensions: operational data and metadata. Operational data is the atomic data or summary information stored in data sets. End users inquire the operational data to answer the health care questions. On the other hand, metadata is the data about the operational data. Usually metadata is stored in a data dictionary or a data repository. End users inquire metadata to determine what type of information are available.

Four attributes of data quality in operational data -- *timeliness,accuracy, consistency, and completeness* -- porposed in the literature (Ballou, 1985; Wang, 1996; Bryant, 1998) are the initial intent of the study.

- *Timeliness* means that the age of data must be appropriate for the task at hand. Usually the end user will require data to be available or updated monthly, weekly, daily or even real-time. If data delivery meets user's satisfaction, the data is timely.
- *Accuracy* is the degree to which data stored in an operational database truly reflects the meaning of the object. It means that data must be correct, reliable, and certified free of error. For example, the patient name stored in medical records is spelled correctly and the patient address is up-to-date.
- *Consistency* means the form and data should be consistently defined in order to be shared by different objects. For example, if the ob-gyn department code is defined by a 4-digit code. all the objects or tables that have the ob-gyn department code as its attributes have to

be defined in the same way. Implementing the data consistency will allow the data to be integrated across the different applications and different platforms.

- *Completeness* means that all the possible data values are included in the proper domain. It implies that data must be of sufficient breadth, depth, and scope for the task at hand. For example, a department table should contain all the possible department codes. All the line items should be captured for a payment record so that calculation of the billing amount will be correct.

In the context of health care data, several assessing issues became significant. First, as the data was submitted by several different clients that covered a three year time span, timeliness of the data could not be measured adequately from the perspective of the currency of data. Timeliness is viewed from the perspective of the data supplier, and refers to the use of the most timely standard for coding methodologies. Coding methodologies such as CPT codes and ICD-9 codes, change on an annual basis. If claims are submitted with deleted codes, they should be considered an error on the timeliness dimension. While this is a significant component of data quality, we exclude this dimension in this study. In a follow-up study, we plan to automate checking the data with the coding reference standard to better determine the value of timeliness.

Second, we can use adherence to standards as another dimension of quality. Several industries have not only tried to establish standard formats for their transactions, but have also established standard reference data that allows the codification of information. Healthcare has attempted to do this for many years and is a continuing effort. Since 1983, HCFA has proposed standards around coding on the HCFA 1500 and UB-92 data collection forms. These standards include the HCPC Levels I, II, and III which codifies the services performed during an encounter. It includes the CPT procedural coding (Level (< HCPC's Level II which codifies supplies and non-physician services, and Level III which is variable as it codifies local standards at a state level for Medicaid or Medicare billing.

Third, it is likely that interdependencies could exist among attributes. For example, laboratory data that are generated or arrive too late cause an unacceptable amount of incompleteness and timeliness. Similar relationships exist between timeliness and accuracy, and completeness and accuracy (Kon et al., 1993). Such interdependencies may make it difficult to define a minimal and orthogonal set of data quality attributes in health care.

To trim costs and maximize productivity and value, health care entities are, for the first time, turning to their data and decision support tools to validate their cost and quality initiatives. For health plans to effectively monitor activity within their networks, they must have an accurate picture of encounters that take place between provider and patient. Ideally, the correct data to analyze the effectiveness and quality of these interactions would be the original clinical data collected by the provider at the point of service. Unfortunately, the clinical data are not currently available because of a lack of automation and standard methodologies.

Health plan analytical and reporting demands have dramatically increased during the past three years. The traditional focus on production and financial reporting gave way in the 1990s to enterprise-wide access by business users needing information on medical management, enhanced account reporting, marketing, etc. Because of these users' demand for data, health plans have begun to realize the need to install warehouses of historical data to address these business needs. As users began accessing data, the enthusiasm to provide information quickly turned to frustration. Many times the numbers pulled did not match other reports. Data were found to be incomplete and inconclusive, and there was generally a lack of understanding by analysts as to anomalies and inconsistencies in these data.

THE NEED FOR EFFECTIVE HEALTH CARE DATA QUALITY MANAGEMENT

The culmination of these data quality problems points to the fact that there is no true data quality ownership. Operations areas are charged with decreasing the turnaround time from receipt of data to the date a claim is paid. They are also audited on financial accuracy. Information Systems departments maintain a legacy system, and have projects backlogged to bring the system up to the current pricing and flexible contractual requirements.

This failure to manage data quality is ironic because the data are the only true asset held by these organizations. Some have accepted the notion that bad data is better than no data. While some visionaries attempt to tackle the issues affecting data quality, this daunting task is complicated by the fact that data quality issues are many and mitigating resources available to address them are few. Ensuring data quality does not neatly fall into one area of an organization -- it must be addressed by the entire organization and senior management commitment is critical to successful implementation.

Most data quality initiatives are born where the need is the greatest, usually in analytical or decision support departments. These departments establish their own quality needs, use the data to support database accuracy levels and pass reports to operations areas. Through these relationships, operations is influenced to consider tis data requirements and change or upgrade its level of data capture. This activity has proven to be ineffective since many analytical requirements directly conflict with operations processing requirements. Without controls in place at the point of entry, little will change to ensure more accurate, consistent data entry in the future.

MANAGING DATA QUALITY METHODS IN HEALTH CARE

A problem with existing data quality methods in health care is that although there has been much recent debate on the subject, very few of the methods proposed or implemented go beyond basic frameworks which require further development. As yet, there is no single universal accepted methodology that is supported by statute, standard, or professional association. Although several methods exist, many are characterized by a very definite purpose and scope which makes their universal adoption difficult. Some popular methods to manage health care data quality are briefly reviewed below.

Cleansing Front-End Data

Front-end data cleansing can be performed on almost any platform, but it must be "programmable" by the user. An interface must exist for the user to input a variety of business rules and controls tailored to unique business needs. These rules may define whether a field i complete or dictate the conditions under which a particular claim is entered by setting up logic that defines the relationship between fields. The interface must also access clinical rules that determine the appropriate use of service and diagnosis codes. Some statistical algorithms may also be used to detect whether the charge is beyond payment scope for a service indicated or is an incorrect submission. The number and complexity of cleansing rules will vary from plan to plan. A system that has more than adequate controls and the ability to incorporate contract specifications in its adjudication system will have fewer rules than one that is not as flexible or sophisticated.

Many health care organizations are beginning to integrate dynamic cleansing front-end data and standardization routines with their front-end Web applications to verify patient addresses, spelling, and completeness of health care information. E-commerce sites in health care are also implementing new processes to handle marginal cases in a more customer-friendly way.

Sampling and Auditing Front-End Data

The second function for a data quality tool is periodic auditing of a sample that has passed through the front end. A sophisticated methodology is required to test for front-end cleansing performance, which, in turn, identifies data anomalies and opportunities to add to the rules in the front end. Additionally, these audits serve as a quality improvement process by feeding specific information on errors being made back to processors, supervisors and trainers, describing which areas are most prone to errors.

Claim auditing is usually a laborious manual process that is typically supported by a system-generated, paper audit trail. User selection of the sample is often manual and, at best, these data are captured in a spreadsheet for manipulation and analysis. With these limitations, the audit method is incapable of producing the most meaningful result.

A sampling tool must be developed to allow the user to establish a random selection of claim records by category (analysts, group, provider type, etc.). This selection would then populate an interface for the auditor to use in verifying the claim submission's completeness and accuracy. This interface supports the identification of errors and categorizes by the types of errors observed. Collected in a database, this information can then be used for reporting in a variety of ways. As an example, accuracy can be assessed at any time in the month by the processor's supervisor, providing the processor with knowledge of his/her current quality level. This allows that individual to address substandard performance early in the month by focusing on corrective action to increase quality performance. Valuable information is flowed to trainers, identifying areas where most processors are having difficulty. Targeted retraining and/or a restructuring of the training process to achieve greater efficiencies may result. As quality issues are discovered, auditors will report any new business rules identified from the audit for incorporation into the front end.

Cleansing Back-End Database Data

Back-end data cleansing of the analytical database against predefined business rules is the third function of a data quality tool. Concurrent with installing a front-end cleanser, an analysis on prior years' data is also needed to bring these legacy data to the same standard as the current data being entered. With an increasingly consolidated marketplace, data from these once-separate entities require an enormous amount of mapping from one data set to another. There are also definitional differences that must be agreed upon. The cleanser must also produce and map fields so they are more meaningful to the analyst, and use standard terms and definitions for categorizing fields rather than system-generated mnemonics. The cleanser may also group the data into categories by product, benefit or episode of care. The process supports a more accurate and timely pre-analysis of certain aspects of the data, which, in turn, may offer opportunities to enhance the front-end edits prior to full data quality implementation.

Other Data Entry Points

Accounting for the three previously mentioned functions of a data quality tool would provide an effective framework to ensure accuracy in claims data. The next step is to look at all other significant points of data entry into the system and propose a similar process. Entry points include provider, enrollment/eligibility, utilization information and authorization data. These areas also contribute to data quality problems because the data are constantly changing and the maintenance requirements are often unsupported with adequate resources. While these data directly affect financial transactions, often the controls were never built into entry screens and system audits.

Database Reuse

Database reuse is a process of using existing database conceptual models or parts of them rather than building them from scratch. Typically, reuse involves abstraction, selection, specialization and integration of reusable parts, although different techniques (when they will be defined) may emphasize or de-emphasize some of them.

The primary motivation to reuse database components (conceptual models or parts of them) is to reduce the time and efforts required when building a conceptual model, actually a database. Because the quality of software systems is enhanced by reusing quality software artifacts, which also reduces the time and efforts required to maintain software, (similarly) reusable database components can, first of all, influence logical and physical database design and (not at last) also the database maintenance.

AN INTEGRATED FRAMEWORK

In consideration of the limited choice of data quality methods it was decided that an integrated framework should be developed that might be of use in health care organizations. The methodology is therefore presented in its entirety, identifying each and every prerequisite for the improvement of data quality in health care. Some organizations may find that they already possess the knowledge to satisfy some of these steps. For example, they may already have a mission statement with clearly identified data quality objectives; if this is the case they will be able to skip the relevant steps.

There are four major stages in the framework: Promote, Identify, Evaluation, and Account (P-I-E-A).

Promote Stage

The major purpose of this promote stage is to promote support and cooperation for the data quality. There are three steps:

1. Promote the benefits of the data quality. First, ideally the organization should hold a training or seminars which explain the importance of data quality and why the organization needs one, training on data and information, their roles in the enterprise and how they may be improved. Second, adopt new philosophy. The organization can no longer live with currently accepted levels of information (data) quality. Third, it is crucial to understand that it is a very difficult task to ensure data quality, since data is usually entered into the health care information systems by end users typing into free-form text fields. These entries are prone to spelling errors and inconsistencies such as abbreviations for names, titles, companies, and addresses.
2. Foster cooperation throughout the organization. This can be achieved by circulating a passport letter or newsletter signed by the chief executive that succinctly reiterates the issues addressed (Hamilton 1987). Application, functional and business domain ensure a free flow of high data quality across the organizational boundaries.
3. Carry out a preliminary survey of the organization to make initial assessment of the level of awareness and value of data quality throughout the organization by a simple informal walk-around.

Once the promote stage has been completed there will exist, at the very least, greater understanding of the importance and purpose of the data quality and, hopefully, greater cooperation and support for the data quality assessment process.

Identify Stage

This stage begins with a top-down strategic analysis of the organization which builds up a rich picture of the organization's mission, environment, structure, and culture. Towards the latter part of this the organization's information resources and information flows are identified. These includes the following four steps:

1. Identify and define the organization's mission.
2. Identify and define the management and cultural environment.
3. Identify the health care information flows.
4. Identify the organization's information resources.

Once this stage has been completed the organization will be able to create a structure in top management that recognizes the importance of data and information and their relationships to the rest of the business. Business domain supports this recognition and the top management can always find a support for understanding data in existing health care information process.

Evaluation Stage

The major purpose of this stage is to analyze and evaluate the organization's information resources and to formulate action plans to improve data quality. There are four steps:

1. Evaluate the information resources, based on their strategic importance, utility, and associated problems in order to identify appropriate management strategies for each information resource (Buchanan & Gibb, 11998).
2. Assess the risks of poor data quality. In fact, as data quality degrades, the legal and financial risks the organization faces increase exponentially. There are some events or situations which will cause the poor data quality. For example, the unexpected computer downtime will prevent the users from receiving information.
3. Formulate action plans. Checkland and Schole (1990) suggest a soft systems methodology that provides a practical step-by-step method to deal with complex, unstructured, or poorly defined problematic situations.
4. Evaluate performance, review and repeat. Be sure we reach high quality components for building health care information systems models.

Account Stage

This stage is to cost the data quality in order to assign costs to the data quality and associated management strategies and action plans. The costing valuing of data quality is recognized as being a problematic area (Badenoch, et al., 1994). Orna (1990) and Burk and Horton (1988) emphasize the need to liaise with the organization's accountants to ensure that there is consistency and comparability within the business. Three approaches can be employed as follows:

1. Output based specification (OBS): OBS is a quality performance measurement system that also provides a mechanism to link payment to quality performance by identifying the minimum quality standards and quality indicators for each data resources other than the costs.
2. Activity based costing (ABC): ABC identifies the costs for information resources by measuring the causal relationship between activity cost and data quality efforts (Turney, 1996).
3. Glazier's model: Glazier's model (Glazier, 1993) is a novel approach to the measurement of information assets in order to identify opportunities to improve revenue streams, reduce process costs, and focus on customer demand as the most tangible evidence of delivered value.

CONCLUSION

Concerns about data quality among IS professionals remain great, especially with regard to consistency and the application of guidelines for users in the health care industry (Bryant, 1998). Within a multidimensional concept of quality care (Lin & Schneider, 1992), some dimensions may be poorly addressed by the database approach. However, it is in the standardization of the data set that the real value of health care databases lie.

Over the years many approaches to improve information quality have been developed and employed in various situations. It is important to recognize that most of these approaches are only vaguely aware that the prerequisite for information quality is data quality, despite the fact that in information technology aggressive steps to improve just the data quality are being taken. In the last few years information technology has been spectacularly successful in automating many operations and making data available to more people. The advances of information technology also have had an impact on poor data quality. But unfortunately just as it is natural to assume that computerized data are correct, so it is natural to blame the information technology when the data are incorrect. These problems can grow out of proportions especially in the data warehouse environments as well as on the Internet.

The framework described is based on an analysis of existing approaches and practical experience derived from the development of data quality management in the health care industry. Pure technological approaches in the form of technology driven solutions are necessary, but not sufficient, to improve sustained data quality improvements (Abate, et al., 1988). To ensure long-term data quality improvements, research efforts should also include data driven solutions directed at information processes and systems to ensure both compliance and efficiency. The other potential benefits of the framework are:

- It provides a complete step-by-step pragmatic solution to data quality.
- It provides a management tool kit that can be tailored to individual requirements.
- It provides new approaches to costing data quality

However, there are also several potential barriers to successful implementation. For instance:

- Synthesis between stages may not always be clear and unambiguous due to the multi-dimensional nature of the data quality management.
- The scale of the data quality efforts and associated resources requirements may make it impractical for organizations.

As discussed above, the framework is intended to be wide-ranging and of general applicability in health care, but it is recognized that health care organizations may need to make compromises, may desire to use a sub-set of the steps, or may need to enhance or tailor to their specific requirements for data quality.

REFERENCES

Abate, M. L., Diegert, K. V., & Allen, H. W. (1998). A hierarchical approach to improve data quality. *Data Quality Journal, 4*(1), 1-10.

Badenoch, D., Redi, C., Burton, P., Gibb, F., & Oppenheim, C. (1994) The value of information. In *The Value and Impact of Information,* M. Feeney and M. Grieves (Eds.), pp. 9-78. East Grinstead, UK: Bowker-Saurm.

Bryant, G. (1998). Aspects of data quality in the new millennium. *Topics in Health Information Management, 18*(4), 81-88.

Buchanan, S. & Gibb, F. (1998). The information audit: An integrated strategic approach. *International Journal of Information Management, 18*(1), 29-47.

Burk, C. F. & Horton, F. W. (1988). *InforMap: A complete guide to discovering corporate information resources.* Englewood Cliffs: Prentice-Hall.

Checkland, P. & SCholes, J. (1990). *Soft systems in action.* Chichester: John Wiley and Sons.

Cleary, R., Beard, R., Coles, J., Devlin, B., Hopkins, A., Sehumacher, D., & Wickings, I. (1994) . Comparative hospital databases: Value for management and quality. *Quality in Health Care, 3*(3), 3-10.

Fox, C. & Grakes, W. (1997). Elements of the quality paradigm. *Communications of the ACM, 40*(6), 26.

Glazier, R. (1993). Measuring the value of information. *IBM Systems Journal, 32*(1), 23-35.

Hamilton, S. (1987). *Communication audit handbook: Helping organizations communicate.* London: Pitman.

Hsia, D. C., Krushat, W. M., Faga, A. B., Tebbutt, J. A., & Kusserow, R. P. (1988). Accuracy of diagnostic coding for medicate patients under the prospective-payment system. *New England Journal of Medicien, 318*, 352-355.

Horner, P. R. (2000). Health services research and medical informatics: A new frontier. *OR/MS Today, 27*(1), 24-27.

IEEE (1989). *IEEE standards - third edition.* New York: IEEE.

Institute of Medicine (1977). *Reliability of hospital discharge abstracts.* Washington, DC: National Academy of Sciences (IOM-77-01).

IOM. (1999). To err is human: Building a safer health system. *www.nas.edu.*

Kon, H., Lee, J., & Wang, R. (1993). *A process view of data quality.* Working Paper TDQM-93-01, MIT TDQM Research Program, E53-320, Cambridge, MA 02139.

Kueger, C. W. (1992). Software resue. *ACM Computer Surveys, 24*(2), 131-184.

Lin, B. & Schneider, H. (1992). A framework for measuring quality in health care. *International Journal of Health Care Quality Assurance, 5*(6), 25-31.

Orna, E. (1990) *Practical information policies: How to manage information flows in organizations,* Gower, Aldershop.

Orr, K. (1998). Data quality and systems theory. *Communications of the ACM, 41*(2), 66-71.

Redman, T. C. (1996). *Data quality for the information age.* Boston, London: Artech House.

Redman, T. C. (1995). IMprove data quality for competitive advantage. *Sloan Management Review, 36*(2), 99-107.

Reifer, D. J. (1997). *Practical software reuse.* Toronto: John Wiley & Sons.

Rine, D. C. (1997). Supporting reuse with object technology. *IEEE Computer, 30*(10), 43-45.

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). 10 potholes in the road to information quality. *IEEE Computer, 30*(10), 38-46.

Sunderland, R. (1982). Inaccurate coding corrupts medical information. *Arch Dis Child, 60,* 593-594.

Tayi, G. K. & Ballou, W. (1998). Examining data quality. *Communications of the ACM, 41*(2), 54-57.

Turban, E., McLean, E., & Wetherbe, J. (1996). *Information technology for management.* New York: John Wiley & Sons.

Turney, P. B. B. (1996). *Activity based costing: The performance breakthrough,* London: Kogan Page.

Whates, P. D., Birgzalis, A.R., & Irvine, M. (1982). Accuracy of hospital activity analysis codes. *BMJ, 284,* 1857-1858.

Welzer, T. (1995). Model of the database system with eplicit built-in expert knowledge. Doctor of Science Thesis.

Welzer, T. (1997). Database reuse a new way to the conceptual database design. *Short Notes Proceedings of the 22nd Euromicro Conference*, pp. 81-88. New York: IEEE Computer Society Press.

Welzer, T. & Rozmasn, I. (1998). Information quality by meta model. *Proceedings Software Quality Management VI,* pp. 81-88. London: Springer.