2007

Data Quality Information and Decision Making: A Healthcare Case Study

Karolyn Kerr
Central Technical Advisory Services, Wellington, Karolyn_kerr@centraltas.co.nz

Tony Norris
Massey University, Auckland, t.norris@massey.ac.nz

Rosemary Stockdale
Massey University, Auckland, r.j.stockdale@massey.ac.nz

Data Quality Information and Decision Making: A Healthcare Case Study

Karolyn Kerr
Information and Analysis Team
Central Technical Advisory Services
Wellington, New Zealand
Email: Karolyn_kerr@centraltas.co.nz

Tony Norris
Institute of Information and Mathematical Sciences
Massey University, Auckland, New Zealand
Email: t.norris@massey.ac.nz

Rosemary Stockdale
Institute of Information and Mathematical Sciences
Massey University
Auckland, New Zealand
Email: r.j.stockdale@massey.ac.nz

Abstract

Defining data quality and realising the need for information that is free of defects and that possesses the right qualities for the task at hand remains a difficult issue. This is particularly so in the healthcare sector where the need for effective decision making is high. This case study addresses the development of a data quality evaluation framework for the NZ health sector. It discusses a data quality strategy that underpins the application of the framework and defines a vision for data quality management in the health sector. It discusses how the framework and strategy combine to increase intelligence density. A significant outcome from the case identified the difficulty of getting data users and managers at all levels to understand the imperative of data quality and accept responsibility for its improvement and maintenance. Recommendations for further research are made.

Keywords

Data quality, decision-making, healthcare, intelligence density

Introduction

It has been estimated that up to 5% of data found in organisations are of poor quality (Redman, 2001) and that the average perceived cost of poor data quality is as high as 10% of an organisation's revenues (Malcom, 1998). In the healthcare sector, lack of data quality has far-reaching effects. Planning and delivery of services rely heavily on data from clinical, administrative and management sources. For example, evidence-based practice (Strauss et al, 2005) requires access to extensive research data, collated and presented in a way that a clinician can use at the time of diagnosis or in other decision-making situations. The higher the quality of the data, the better will be the patient outcomes. Similarly, quality data, particularly with regard to timeliness and accuracy, are needed for administrative purposes such as hospital bed-rostering and for planning services to ensure that they are cost-effective. These different but interlocking data requirements and decisions ensure that health care organisations and their relationships are inherently complex and demanding (Gendron & D'Onofrio, 2001).

Data quality is inextricably linked to the use of Information Systems and the health sector is increasingly an information-driven service (Hovenga, Kidd, & Cesnik, 1996). Information held in databases and other electronic repositories and delivered in a reliable and timely manner is critical to the health and well-being of patients, the wider population, and to the management of health care organisations (Long & Seko, 2002). Raising the level of data quality within an organisation contributes to improving the quality of decision-making enabling the reduction of *uncertainty* and the production of more timely and accurate decision outcomes.

The features of the data/decision-making relationship are illustrated in this paper by a case study from the health sector, although the issues found and the effective solutions are common to many domains. The research uses findings from the case study to inform the development and implementation of a national data quality improvement strategy for the New Zealand (NZ) health sector, that includes data quality assessment. A strategic imperative of the research is to develop a holistic, 'whole of health sector' way of viewing data quality with sufficient flexibility for organisations to implement local innovations through locally developed strategies and data quality improvement programmes. The intention is to make the strategic management of data quality, and the prevention of persistent errors, everyday, 'institutionalised' activities.

This paper first discusses the meaning of data quality before assessing the literature on problems and existing information regarding data quality. An overview of a data quality framework is supported by discussion on an underpinning data quality strategy that contributes to the development of an effective tool for improving data quality and embedding practice within the routine operation of an organisation.

Defining Data Quality

There is on-going confusion between the terms 'data' and 'information' and this distinction needs clarification if we are to address the quality issue successfully. Tayi and Ballou (1998) define data as 'the raw material for the information age'. A datum is a fact; a value assigned to a variable (Saba & McCormick, 2001), a single observational point that characterises a relationship (Shortliffe & Barnett, 2000). Data support managerial and professional work and are critical to all decisions at all levels of an enterprise (Fuller & Redman, 1994; Tayi & Ballou, 1998).

In contrast, information is useful data that have been processed in such a way as to increase the knowledge of the person who uses the data (McFadden, Hoffer, & Prescott, 1999). English (1999a) builds on the idea of information being data in context, with knowledge being information in context, where you know the significance of the information. Translating information into knowledge requires personal experience and reflection. Knowledge itself may be processed to generate decisions and new knowledge (Saba & McCormick, 2001) including the results of formal studies and also commonsense facts, assumptions, heuristics (strategic rules of thumb), and models – any of which may reflect the experience or biases of people who interpret the initial data (Shortliffe & Barnett, 2000).Thus, whilst the transformation of data into information is normally an explicit, repeatable and easily conveyed procedure, the further translation of information into knowledge often involves tacit processes that are much more difficult to capture and explain to others (Davidson & Voss, 2002).

High-quality data and derived information are needed to create institutional knowledge (stored information) plus reasoning processes that help an organisation extract the maximum benefit from the resources. This approach, now accepted under the term 'knowledge management' (Davenport, 1998; Davidson & Voss, 2002), draws together the tangible and intangible elements of data and shares them amongst all workers.

Data quality is now emerging as a discipline, with specific research programmes underway within universities, the most significant being that of the Sloan School of Management Information Quality Programme at the Massachusetts Institute of Technology (MIT). The field encompasses the well established Quality Discipline, drawing on the work of Deming (1982), with the adaptation of the *plan, do, check, act, cycle'* of Crosby (1980), through the notion that 'quality is free' because of the cost of doing things wrong (Juran & Godfrey, 1999). Further exemplars include the utilisation of Six Sigma and Total Quality Management, adapted to Total Data Quality Management (TDQM) and the management of information as a product (Wang, Lee, Pipino, & Strong, 1998).

Klein and Rossin (1999) note there is no *single* definition of data quality accepted by researchers and those working in the discipline. Data quality takes a consumer-focused view (consumers being people or groups who have experience in using organisational data to make business decisions) that quality data are 'data that are fit for use' (Loshin, 2001; Redman, 2001; Wang, Strong, & Guarascio, 1996). Data quality is 'contextual'; the user defines what is good data quality for each proposed use of the data, within its context of use (Pringle, Wilson, & Grol, 2002; Strong, Lee, & Wang, 1997). Therefore:

> *Data are of high quality if they are fit for their intended uses in operations, decision-making, and planning. Data are fit for use if they are free of defects and possess desired features.*
>
> *(*Redman, 2001, p. 241)

The Origins of Data Quality Problems

It is well recognised that no data are completely accurate. The real concern with data quality is to ensure not that the data are perfect, but that they are accurate enough, timely enough, and consistent enough for the organisation to make appropriate and reliable decisions.

This research includes an assessment of data quality problems commonly found in the health sector but the insights are applicable across many other sectors. In this study, data quality problems were often identified only when the data were used to provide information in reports. Those analysing the data for information are often the first to identify problems that declare the data as not 'fit for use' for their analytical purpose. The authors found that trust was lost and reports either rejected or not used when reported data were perceived as inaccurate. Late data were also noted as a problem as data collection staff had little understanding of the impact of not supplying the data on time and the impact on reporting requirements and the management of services.

The study also notes that the absence or incomplete availability of data also affected decision-making. For example, a process of devolving services from the Ministry of Health to health care provider organisations ran into problems because it was not supported with access to historical information on service provision. Equally difficult was the need to make decisions using out-of-date data, where no current data were available.

Other workers have identified additional data quality factors that influence choices. For example, health care has always been a 'mobile' profession and Shankaranarayan, Ziad & Wang (2003) have noted the impact of distributed data collection and application through new technologies such as wireless and the Internet. Such environments empower decision makers to make 'on-the-spot' decisions through ready access to possibly large volumes of data. This requires data quality management to be efficient and to inform the decision maker on the quality of data. Increasingly, however, distributed data sources reduce the likelihood of knowing the source and therefore the quality of the data. Hence, a step–by-step process to data management known as information product mapping is required to provide an audit trail tracking the quality of the data throughout their life-cycle. Adding this layer of knowledge about data origin and transformation can enhance perceptions of believability and reliability of the data.

Another issue is the context of data collection (Dravis, 2004); that is the intended purpose for which data are captured and the policies and procedures that govern acquisition, storage and use. The quality requirements, such as the demands for accuracy and currency, may well differ from one context to another, or change over time (Lee, 2004). Problems can arise when practitioners are not informed of the context and they make incorrect assumptions. This is particularly important in health care where data are collected from multiple disparate sources (Strong et al., 1997).

Information on Data Quality

Data quality information derived from metadata (data about data) that can be included with the data themselves to specify the level of quality (Fisher, Chengalur-Smith, & Ballou, 2003). Metadata are defined as:

> *all the characteristics that need to be known about data to build databases and applications and to support knowledge workers and information producers. (English, 1999b).*

Research has shown that including information about the quality of data can impact decision-making (Chengalur-Smith, Ballou, & Pazer, 1999) by enabling decision makers to utilise data more efficiently and effectively (Even, Shankaranarayanan & Watts, 2006). For example, decision makers need to have access to sufficient data quality information to gauge the reliability of the data (Shankaranarayan, et al., 2003). The presentation of the data quality information also can also have an impact (Chengalur-Smith, Ballou, & Pazer, 1999; Ballou & Tayi, 1999). Where simple decision-making is required, Chengalur-Smith et al., found that the participants in their study were prepared to use complex data quality information to assist their choices whereas when faced with complex decisions they resorted to simple metadata. The authors propose that this difference is due to 'information overload' where complex metadata simply compound the complexity of decision-making and make it too difficult.

Decisions made in a health care environment are often complex whether due to the multidisciplinary nature of clinical care or the limited resources available for planning and delivering services. However, what is a complex decision to one user may not be complex to another and the use of data quality information is found to increase as experience levels progress from novice to professional (Fisher et al., 2003). Data quality information also assists those who need to make decisions under time pressure, such as in crisis situations (Fisher et al., 2003).

The Case Study

This section discusses the case study and the development of a data quality evaluation framework for the NZ health sector. It first addresses the research design before discussing a data quality strategy that underpins the application of the framework and defines a vision for data quality management in the health sector. Finally, this section discusses how the framework and strategy combine to increase intelligence density.

Research Design

A qualitative research methodology was used to study learning and change resulting from data quality initiatives within the social setting of the New Zealand Ministry of Health. An interpretive, participatory, action research approach (Mumford, 2001) enabled the principal researcher to observe and examine an iterative cycle of learning and institutional change within the case environment. Action research supports the importance of people within the research context; people are a particularly important element in achieving data quality improvement in organisations (Lee et al., 2004). The principal researcher and the study participants took part in an educational and change initiation process through two cycles of action research over a period of two years.

Data were collected at workshops, focus groups, semi-structured meetings and interviews and via questionnaires over the two year period. All study participants were employed by the NZ Ministry of Health and were initially drawn from the Clinical Analysis Team and the Data Quality Team. Participants were later drawn from across the Ministry to elicit data from specific role perspectives. For example, managers of Health Information Systems were interviewed regarding their perceptions of change and the CIOs of District Health Boards regarding development of strategy.

Grounded theory (Pauleen and Yoong, 2004) was used for the analysis of data through inductive coding and constant comparison in the analysis phase of the action research iterative cycle.

Data Quality Framework

At its most basic, a data quality framework is a tool for the assessment of data quality within an organisation (Wang et al., 1996). Such frameworks seek to assess areas where poor quality processes or inefficiencies may reduce the performance of an organisation A framework can go beyond the individual elements of data quality assessment, and become embedded into the processes of the organisation.

The current research developed a data quality evaluation framework for the New Zealand health sector to provide data users with extensive data quality information and a consistent assessment tool across the national health information databases and registries. The framework was modelled on the findings of the Canadian Institute for Health Information (CIHI) (Canadian Institute for Health Information, 2003), an organisation that has pioneered work on health care information quality. The development and evaluation of the framework (Kerr, 2006) included an investigation of:

- the applicability of the data quality dimensions, characteristics, and criteria for the collection being assessed;
- the language used in the framework;
- the language and examples provided in the user manual;
- the length of time required to complete the assessment using the framework;
- the value to users of the information provided by the framework;
- the table of contents for the Data Quality Documentation Folder.

Figure 1 outlines the structure of the framework, adapted for New Zealand's requirements, from the CIHI model. Sixty-nine primitive criteria for data quality are aggregated into 24 characteristics which are further grouped into six high-level dimensions. Each criterion requires an assessment process against a derived metric and provides guidance to users on what data quality information to collect and document to understand and encourage data quality improvements.

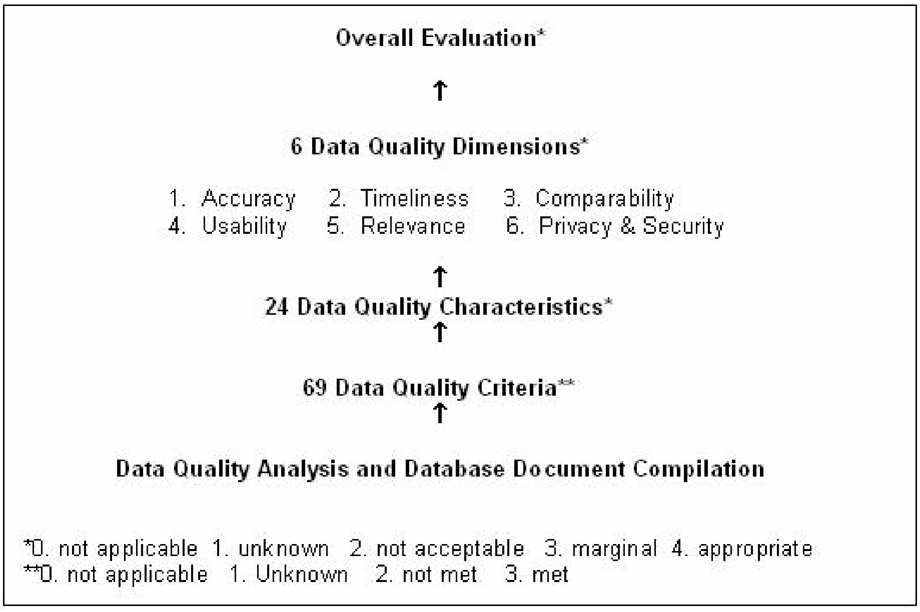

Figure 1: The CIHI Data Quality Framework Model as Applied in the New Zealand Framework

The six information quality dimensions chosen for the framework are as follows with accuracy being the most extensive:

1. Accuracy
2. Relevancy
3. Timeliness
4. Comparability
5. Usability
6. Security and privacy

A review of the primitive criteria in the context of a given data collection helps to prioritise the dimensions that are the practical constructs for defining metrics. This prioritisation (Ballou & Pazer, 1995) ensures that all relevant dimensions are taken into account whilst allowing any trade-offs needed, for example, by the user's inability to collect sufficiently accurate data within a restricted time frame (timeliness). The summary information gained from assessments of all collections can also be collated to form a prioritised list of data quality improvement initiatives across the organisation. Ongoing assessment using the framework provides information on the success of data quality improvement initiatives.

This research found that, for the development and application of a data quality evaluation framework, it is important to:

- define the underpinning data quality criteria carefully involving all stakeholders to ensure common understanding and direction;
- consider the critical quality dimensions that reflect how the organisation uses data and how data flow throughout the business processes;
- document business processes identifying data sources and their reliability;
- appreciate that the framework requires practical supporting tools, e.g. documentation, to make it effective;
- customise the language of user guidelines or manuals with regard to the level and experience of the intended users;
- be aware of the importance of both education and training at all necessary stages and levels – training is essential to affect the culture change that must accompany the realisation of the importance of data quality;
- be aware that application of the framework is an iterative, on-going process – the required outcomes cannot be achieved in a single-pass.

Data Quality Strategy

Whilst a data quality framework models the data environment and identifies the quality characteristics, an underpinning quality strategy is broader in scope. It establishes the business purpose and context and applies the framework to define key functions such as data acquisition and conversion, and database design, creation and maintenance. A data quality strategy is not an Information Technology strategy, nor an Information Systems strategy. Although such strategies may provide insight and tools to assist in a data quality improvement strategy, these improvements cannot be attained merely through information technology, the problem is one of processes and people. Hence, the research project developed a strategy, outlining the vision of the New Zealand health sector for data quality management and the work required to realise this vision. The strategy applies the theory of appreciative enquiry (Fry, 2002) to encourage change and the utilisation of existing organisational knowledge. Simple rules, such as the total data quality management (TDQM) process (Lee et al., 2004) and data quality dimensions guided the change, leaving room for innovation.

Some approaches to data quality management target specific errors within a collection or an entire collection but often do not devise solutions to prevent systemic problems. In contrast, TDQM focuses on two key aspects of data management: the data flow processes that constitute the organisation's business, and the recognition that information is a product, rather than a by-product of these processes (Wang, 1998). Regarding the processes themselves, TDQM seeks to ensure that none of them change the initial meaning of the data leading to systematic errors and repeated information quality problems. Systematic process errors can be prevented by several means, some of which will depend upon the nature of the business unit and its data.

The resulting Data Quality Strategy provides the Ministry of Health and the sector with guidelines on how to develop and implement TDQM throughout all levels of the health sector. The strategy uses the dimensions found in the framework to set data quality standards for designing, developing, and maintaining the national health data collections and data management throughout New Zealand. Roles and responsibilities are clearly defined, along with data ownership. A series of projects provides the required development for business as usual initiatives that institutionalise data quality into every day practice and make use of existing sector knowledge through the development and dissemination of best practice guidelines.

Prevention is now a large part of the Ministry's work through a proactive approach that includes:

- regular minimum data quality initiatives for all collections with specific initiatives for single collections only when identified as a specific requirement;
- preventing poor data quality in new collections through a 'data quality plan' initiated in the early phases of developing a new national data collection. Data quality is embedded in the system and processes of the new collection prior to 'go live'.
- continuing what works, as the organisation has learnt a considerable amount about the operational requirements of maintaining data quality levels;
- the endorsement of standards for the collection, management, storage and use of specific data elements by the national Health Information Standards Organisation. This way all stakeholders know and agree to all steps in the data flow process;
- regular 'business as usual' processes that review repeated data quality errors from suppliers and feedback information on issues to suppliers, with support provided for improvement.

This research found that for a data quality improvement strategy, it is important to:

- derive and impose standards that facilitate data and information transfer whilst preserving quality;
- re-engineer the business processes to deliver the quality data needed for efficient service planning and the effective practice of integrated patient care;
- identify and disseminate best practice to reduce the development time needed to improve data quality;
- ensure data quality levels are not unnecessarily rigorous to maintain user ownership and workloads at reasonable levels;
- define user accountabilities for data quality and the mechanisms to enforce them;
- seek to embed the search for data quality in normal working practices and recognise its achievement in appropriate ways such as accreditation.

Whilst some of the terminology and examples used here are specific to a public health sector, environment many of the principles are generic and the relationship between data quality and decision-making and the approaches to improvement translate to other business sectors.

Data Quality and Intelligence Density

The data quality framework together with the implementation strategy and the attendant policies and procedures comprise an effective tool for improving data quality and embedding the thinking and practice of improvement in the routine operation of an organisation. In the case study described here these techniques have been used to alert an organisation to the importance of data quality and to set it on the road to better decision-making. As the organisation ascends further along the learning curve, it will automate, integrate, and then optimise its processes to establish best data quality and handling practice throughout the whole administration.

These successive phases require the progressive introduction of information systems that access, store, and make available the high-quality data and metadata needed to provide information to decision makers The effective use of these decision support systems leads to the interaction between people and systems that characterises knowledge management (Davidson & Voss, 2002). However, if these systems are built employing artificial intelligence (AI) techniques such as fuzzy logic, genetic algorithms, neural networks, etc (Dhar & Stein, 1997) they can become adaptive, self-aware, and agile thereby transforming the organisation into what is known as an *intelligent enterprise* (Delic & Dayal, 2003).

Intelligent enterprises can create and monitor dynamic models of themselves and their environments and then self-adjust to reflect changes in these models and their key parameter values. In this way, organisations increase their *intelligence density* defined by Delic and Daval as the amount of useful decision support information that a decision maker gets from using organisational tools and methods. Low information density indicates information that has poor relevance and utility with a propensity for information overload. In contrast, high information density suggests high utility in which the elements that facilitate good decision-making are easy to locate and use.

The data quality framework and strategy that are the topics of this paper can combine to increase intelligence density in several ways. Firstly, by design, they reduce inaccuracies, inconsistencies, and other data flaws that decrease the value and utility of decision support information. In addition, the co-availability with data of the metadata that define context minimise errors of interpretation (Olson, 2003) whilst in the self-aware, intelligent enterprise they improve the accuracy of the AI algorithms that control the automated adaptation processes.

Metadata context also increases relevance and therefore data utility to individual practitioners; an important feature in a multidisciplinary environment such as healthcare where the same data can be used for many different purposes. Where context is historical, metadata will moreover increase data longevity so that decision makers can employ data from historical collections even when the processes and procedures used to capture the data have changed over time.

The aggregation of low-level data quality criteria and characteristics into (the six in this study) high-level quality dimensions is a further contributor to the avoidance of information overload.

Conclusions and Further Work

The increasingly information intensive nature of business demands a proactive and strategic approach to data quality to ensure that the right information is available to the right person at the right time in the right format whilst guaranteeing the privacy rights of individuals to have their personal data (especially health data) protected and used in an ethical way.

A data quality framework and strategy will contribute significantly to a Total Data Quality Management programme through a continuous cycle of work that will incrementally improve the quality of data. The addition of contextual metadata will also improve the intelligence density of data so that the integrated programme will facilitate better decision-making throughout the organisation. With the introduction of appropriate information systems and artificial intelligence software, the TDQM programme will hep the organisation become an intelligent enterprise which will self-adapt, self-heal, and self-manage itself evolving like a biological system to meet the demands of its environment and users.

Further research is required in several areas before this vision can be realised. As mentioned, this case study describes the first but important step in an organisation’s awareness of the importance of data quality. One of the object lessons was the difficulty of getting data users and managers at all levels to understand the imperative of data quality and accept responsibility for its improvement and maintenance (Kerr, 2006). We need a data quality strategy with generic but customisable principles for managing this change process and assigning accountability for data quality.

Considerable research is still required to understand the data quality information needs of data users in different organisations and domains. This research needs to take place within the organisational environment to understand fully the operational issues specific to that environment.

Mechanisms are also needed to involve data users in decisions on the granularity of data quality information. Findings from this research show that data consumers require the most detailed information on how quality assessments were made for each data criterion whereas managers needed only summary information (Chengalur-Smith et al., 1999). The structure of the framework should therefore allow for the aggregation of data quality information to minimise information overload.

A further area of research concerns the presentation of data quality information in an appropriate format that facilitates the effective and efficient decision-making. For example, metadata should be linked dynamically to the data it relates to and wherever possible at the interface with the user to ensure visibility and ease of access during data manipulation (Devillers, Bedard & Jeansoulin, 2005).

References

Ballou, D. P., & Pazer, H 1995. Designing information systems to optimise the accuracy-timeliness tradeoff. *Information Systems Research*, vol. 6 no. 1, pp. 51-72.

Ballou, D. P., & Tayi, G. K. 1999. Enhancing data quality in data warehouse environments, *Communications of the ACM*, vol. 42 no. 1, pp. 73-78.

Canadian Institute for Health Information 2003 *Data Quality Framework*. Viewed on 18 Jun 2006 from http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=quality_e

Chengalur-Smith, I., Ballou, D. P., & Pazer, H. 1999. The impact of data quality information on decision making: an exploratory analysis, *IEE Transactions on Knowledge and Data Engineering* vol. 11, pp. 853-864.

Crosby, P. 1980 *Quality is Free*. New York: Penguin Group.

Davenport, T. H. 1998. *Working Knowledge: How Organisations Manage What They Know*. Boston: Harvard University.

Davidson, C. and Voss, P. 2002. *Knowledge Management: An Introduction to Creating Competitive Advantage from Intellectual Capital*. Auckland: Tandem Press.

Delic, K. A; & Dayal, U. 2003. The Rise of the Intelligent Enterprise. *The Virtual Strategist*, Issue 5, 10 pages.

Deming, W. E. 1982. *Out of the Crisis*. Cambridge: Massachusetts Institute of Technology.

Devillers, R., Bedard, Y. & Jeansoulin R. 2005 Multidimensional Management of Geospatial Data Quality Information for it Dynamic Use Within GIS, *Photogravimetric Engineering and Remote Sensing*, vol 71, no. 2, pp. 205-215.

Dhar, V, Stein, R. 1997, *Seven Methods for Transforming Corporate Data into Business Intelligence*, Englewood Cliffs: Prentice-Hall.

Dravis, F. 2004. *Data quality strategy: A step by step approach*. Paper presented at the Ninth International Conference on Information Quality, Boston.

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

English, L. 1999b. *Information quality in the knowledge age*. Viewed online on 21 Aug 06, from http://www.dmreview.com/master.cfm?NavID=198&EdID=1452

Even, A. Shankaranarayanan, G. Watts, S. 2006. *Enhancing Decision Making with Process Metadata: Theoretical Framework, Research Tool, and Exploratory Examination*. Proceedings of the 39th Hawaii International Conference on System Sciences.

Fisher, C., Chengalur-Smith, I., & Ballou, D. P. 2003. The impact of experience and time on the use of data quality information in decision making, *Information Systems Research* vol. 14, pp. 170-188.

Fry, R. 2002. Appreciating your past in order to transform the future. Viewed online 26 Sept 2005 from http://www.clevelandshrm.com/

Fuller, B., & Redman, T. C. 1994. Data Quality lessons learned at Telecom Australia, *QUALCON 94*, pp. 379-386.

Gendron, M. S., & D'Onofrio, M. J. 2001. Data Quality in the Healthcare Industry. *Data Quality Journal*, vol 7, no.1.

Hovenga, E. J. S., Kidd, M. R., & Cesnik, B. 1996. *Health Informatics. An Overview* . Melbourne: Churchill Livingstone.

Juran, J. M., & Godfrey, A. B. 1999. *Juran's Quality Handbook* (5 ed.). New York: McGraw-Hill.

Kerr, K., 2006 *The Institutionalisation of Data Quality in the New Zealand Health Sector*. Unpublished PhD Thesis, University of Auckland.

Klein, B., & Rossin, D. F. 1999. Data errors in neural network and linear regression models: An experimental comparison, *Data Quality* vol. 5, p. 25

Lee, Y. W., Pipino, L. L., Strong, D. M., & Wang, R. Y. 2004. Process-embedded data integrity, *Journal of Database Management* . vol. 15, pp. 87-103

Lee, Y. W. 2004. Crafting rules: context reflective data quality problem solving. *Journal of Management Information Systems*, vol 20, no. 3, pp. 92-119.

Long, J. A., & Seko, C. E. 2002. A new method for database data quality evaluation at the Canadian Institute for Health Information, *7th International Conference on Information Quality*. Boston: Massachusetts Institute of Technology.

Loshin, D. 2001. *Enterprise Knowledge Management. The Data Quality Approach* California: Academic Press. p. 493

Malcom, A. (1998). *Poor data quality costs 10% of revenues, survey reveals*. Viewed 18 Oct 2005 http://computerworld.co.nz/news.nsf/UNID/CC256CED0016AD1ECC25684C000E0278?OpenDocument&Highlight=2,Poor,data,quality,costs

McFadden, F. R., Hoffer, J. A., & Prescott, M. B. 1999. *Modern Database Management* (Fifth ed.). Reading: Addison-Wesley.

Mumford, E. 2001. Action Research: Helping Organisations to Change. In E. M. Trauth (Ed.), *Qualitative Research in IS: Issues and Trends* (p. 298). London: Idea Publishing Group.

Olson, J. E. 2003. *Data Quality. The Accuracy Dimension*. San Francisco: Morgan Kaufmann

Pauleen, D. J., & Yoong, P. 2004. Studying human-centred IT innovation using a grounded action learning approach. *The Qualitative Report*, vol 9, no. 1, pp. 137-160.

Pringle, M., Wilson, T., & Grol, R. 2002. Measuring "goodness" in individuals and healthcare systems, *British Medical Journal*. vol. 325, pp. 704-707.

Redman, T. C. 2001. *Data Quality. The Field Guide*. Boston: Digital Press.

Saba, V. K., & McCormick, K. A. 2001. *Essentials of Computers for Nurses. Informatics for the New Millennium* (Third ed.). New York: McGraw-Hill.

Shankaranarayan, G., Ziad, M., & Wang, R. Y. 2003. Managing data quality in dynamic decision environments: an information product approach. *Journal of Data Management*, vol 14, no. 4, pp. 14-32.

Shortliffe, E. H., & Barnett, G. O. 2000. Medical Data: Their Acquisition, Storage and Use. In L. M. Fagan (Ed.), *Computer Applications in Health Care and Biomedicine* (2nd ed), pp. 37-69 Springer.

Strong, D. M., Lee, Y. W., & Wang, R. Y. 1997. Data quality in context, *Communications of the ACM*, vol. 40, pp. 103-110.

Strauss, S.E., Scott Richardson, W., Glasziou, P., & Haynes, R.B. 2005. *Evidence based medicine* (third ed.) London: Churchill Livingstone

Tayi, G. K., & Ballou, D. P. 1998. Examining data quality, *Communications of the ACM*, vol 41, pp. 54-57.

Wang, R. Y. 1998. *A product perspective on total data quality management*, Communications of the ACM. New York: Association for Computing Machinery.

Wang, R. Y., Lee, Y. W., Pipino, L. L., & Strong, D. M. 1998. Manage your information as a product, *Sloan Management Review*, pp. 95-105.

Wang, R. Y., Strong, D. M., & Guarascio, L. M. (Eds.). 1996. Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems, vol 12*, no. 4, pp. 5-33.

Copyright