IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB12698

This paper appears in the book, *Emerging Trends and Challenges in Information Technology Management, Volume 1 and Volume 2* edited by Mehdi Khosrow-Pour © 2006, Idea Group Inc.

Intrinsic and Internal vs. External View of DQ/IQ (A Case of Relativity)

Zbigniew J. Gackowski, California State University Stanislaus, College of Business Administration,
Dept of Computer Information Systems, 801 W. Monte Vista Ave., Turlock, CA, 95382, T (209)667-3275, ZGackowski@csustan.edu

ABSTRACT

This qualitative inquiry into the universe of data/information quality (DQ/IQ) dimensions presents a rationale for a move from the intrinsic and internal toward the external, from the ontological to the teleological view of DQ/IQ. The focus is on approaches that derive attributes from established theories. The relativity of quality requirements is discussed including how quality requirements are changing at different levels of controlling purposive operations.

INTRODUCTION

The inquiry contributes to the discussion on research directions on data and information quality (DQ/IQ). In research and practice, there is an urgent need to recognize the undeniable relativity of quality requirements within the context of purposive operations. Subsequently, one needs to move from the solely intrinsic and internal towards the external, from the ontological towards the teleological view of quality. All quality dimensions of data/information values acquire practical meaning and importance from the context of their uses, hence are relative to, and should be assessed within the same context.

Liu and Chi (2002) categorized approaches to quality as intuitive, empirical, and theoretical. Initially, the intuitive and the empirical approaches dominated, but lack theoretical foundations on how DQ/IQ attributes are defined and grouped. They identified four theories: mathematical theory of communications, information economics, ontological mappings, and operations research. Nevertheless, they concluded, "Existing theoretical approaches are limited in their ability to derive a full-fledged measurement model" and a "generally accepted model has not yet appeared."

This inquiry stays within the theoretical approaches, which promise results of a more lasting validity. The theoretical approaches derive attributes from established theories. In business, one distinguishes different levels of information support: operations support and management support. Relativity of quality requirements is discussed at different levels of controlling operations.

To arrive at a full-fledged framework of DQ/IQ, one must reach beyond the empirical survey-based assessment of loosely and haphazardly defined attributes of data/information quality. Empirical studies are not discussed here, because their correctness and completeness cannot be proven via fundamental principles. The effort devoted to the empirical studies should be at least partly redirected to approaches with the potential of producing broader, profounder, and more lasting results. Similarly, efforts spent on developing better metrics for assessment of empirically derived quality attributes, but not well founded, are secondary to the importance of a stronger qualitative framework for assessing the basic quality dimensions in their multiple aspects. Empirical studies are useful for immediate, particularly one time, improvements. When seeking results of more lasting validity one needs to develop a solid qualitative model first before embarking upon empirical confirmatory studies.

The main contributions of this paper are:

- A demonstration of the advantages of the operations view in identifying the major quality requirements for data/information values
- A demonstration how dramatically quality requirements change when ascending from the lower to the higher levels of controlling operations.

PERTINENT[1] LITERATURE REVIEW

1. In 1996, based on **ontological foundations**, Wand and Wang (1996) proposed four data quality dimensions (complete, unambiguous, meaningful, and correct) that are intrinsic to system design and operations. Within the confines of the assumptions used[2], "those attributes have crystal-clear definitions and theoretically sound justification, but they constitute only a small subset of known attributes leaving the rest unspecified" (Liu and Chie, 2002). However, they are mislabeled as "intrinsic data quality dimensions" instead of quality dimensions *intrinsic to system design and operations.*
2. In 2005, anchoring the concept of data/information quality in **operations research**, **management science, and decision science,** Gackowski (2005) defined a theoretical teleological content-focused framework of **operations** quality requirements of data and information values viewed from the perspective of purposive operations. This approach made possible the definition of:
   - A universal[3] taxonomy of the entire universe of quality requirements of data/information values into direct and indirect ones, the direct into primary and secondary ones, and the primary into universal ones and task-specific ones
   - Sufficient conditions for attainment task-specific usability of single data/information values and for task-specific effective operationally complete sets of usable and only then useful data and information values with a clear distinction of their only effective or even economically effective completeness
   - An economical examination sequence of at least the direct primary quality requirements

RELATIVITY OF QUALITY REQUIREMENTS

A good theory must have a rock solid:

- ***Main point of reference*** — here, the main **purpose** of the operations
- ***Main frame of reference*** — here, the **circumstances** of operations.

From the teleological perspective one assumes:

- A qualitative cause/effect diagram of operations is available. Description of the situation comprises a decision situation matrix, the actions required, and the expected results. Decision makers use available data and obtained information values to decide to act or not to act.
- Results of operations are functions of the operations quality of data/information values used.
- The relative strength of each factor represented by a data/information value is determined by its impact on a measurable main purpose of operations.

Accepting the widely adopted definition of ***quality*** as "fitness for use," one notices that the *use* of data/information determines what *fits*. Thus, quality is determined in the most pragmatic manner by the improved results of using data or information values. All aspects of operations quality are relative. Some authors see it as a problem, others as an opportunity. It is a problem, when one limits the view of quality of data/information values only to the information system designers' view[4] with regard to the faithful representation of the known aspects of the world as states of information systems (Wand and Wang, 1996). From the teleological perspective of operations, the relativity of operations quality requirements is an advantage on all counts; it literally guides examiners toward effective solutions. In business and public administration, system analysts and designers usually take the teleological view and custom-tailor their designs to meet critical needs of the organizations the information systems serve. Otherwise, off-the-shelf information systems would cover the majority of needs.

Purpose, if measurable, guides examiners like a guiding star. It provides examiners with an unquestionable point of reference and a well-defined yardstick – a unit of measurement. The latter, enables measuring the results of whatever one does in the realm of operations quality. It indicates whether one proceeds in the right direction, by how much, and how far one is from the destination point – the optimum level of quality. Purpose, as the selected point of reference, illuminates all aspects of quality requirements so that with a relative ease one can see which of quality aspects:

- Affect the outcome *directly* and which one only indirectly
- Are of *primary* importance because they qualitatively change the decision situation and the actions to implement the decisions made
- Are *secondary* because they only qualitatively change the actions and/or the results of operations
- *Indirectly* impact the operations by affecting the direct quality requirements first – via them.

Assuming that mapping of functional dependencies of the direct quality requirements on the indirect ones sooner or later becomes available; it literally will assist examiners in moving around in the little explored universe of 179+ quality attributes identified by Wang and Strong (1996).

How well a data/information value fits its intended use is determined by the prevalent view of what is important in its use. Purpose and circumstances define a type of a *force field* that determines all aspects of quality of whatever is used, not only of data and of information. In tightly run cohesively bonded organizations, views focused on strategic business purposes dominate. Otherwise, the local and particular views, whether individual or group, have the upper hand. The latter are guided by local interests and preferences. When this happens, a gradual disintegration of organizations begins.

Relativity of quality requirements of data or information values means that quality requirements are determined by the purpose and circumstances of operations where the values are used. This is the very first fundamental universal principle of operations quality requirements. It pertains to all types of representations of the states of the real world as data values, information values, and rules of reasoning. All known and yet unknown dimensions of quality are subject to this law (Gackowski, 2005).

In operations, there is no room for *intrinsic* aspects of quality of anything for their use. They may be intrinsic to something else, e.g. design of the delivery and distributions system, but not to data/information values themselves. Nevertheless, even then the external view of the market ultimately determines by trial and error the adequate level of quality intrinsic to the design of products and services. These, however, ignore the content of data fields that is the data and information values. The teleological approach to quality is *content-focused;* not how the data fields (their containers not values) are formally organized; not how data or information *values* are perceived, but how they *impact* corresponding decision situations, necessary actions, and results.

It is time to move attention toward a definition of quality requirements that directly responds to the needs of purposive operations. They are under the control of decision-makers and users and belong to the realm of management disciplines. For the ultimate results, both views are important but they are not of the same weight. One may strongly argue that the external view is the dominant one, for it is derived from the ultimate purpose of information systems applications. This can be explained well when one analyzes quality requirements at different levels of controlling operations.

LEVELS OF VIEWING DQ/IQ AND RESEARCH PRIORITIES

One can distinguish several perspectives of viewing data/information quality. On one hand, there is the internal view of the designers, which is process and system centric. This view is clearly independent of the data content, use, and purpose. On the other hand, there is the external teleological view that is use centric, application centric or better purpose centric; it focuses on the quality of data and information values for their use.

The internal view has two basic levels:

1. The process centric one is focused on the quality dimensions of faithful mapping of a specific subset of real world states into data values in their delivery systems.
2. The delivery and distribution system centric level is focused on a reliable implementation of the data mapping requirements, a secure storing of data values, and making them easily available to authorized users.

Users, who view things externally, assume that the data in the delivery and distributions system adequately represent the pertinent fragments of reality, are securely stored, available reliably and timely to authorized users, when needed. They focus their attention, however, on the purpose to be served, on the acquisition of the necessary situation-specific data and information values, its convenient presentation, and effective use.

Any full-fledged research framework should reflect the relative nature of the quality requirements of data and information values. Hence, the most fitting approach to examining data/information quality is the operations research view. In MIS, there is a generally accepted categorization of information systems into **operations support** information systems, and **management (support)** information systems (MIS) (O'Brien, 2004a and 2004b). Quality requirements for the operations support category differ substantially from those required by management support. The latter support differs distinctively at the different levels of decision support such as operations or **non-managerial** decision level, **supervisory** level, **tactical** level, and **strategic** level.

These distinctions provide a helpful insight into the potential problems. For simplicity, the discussion of the operations support systems is limited to transaction processing systems (TPS) and excludes process control systems (PCS), computer-aided design (CAD), computer-aided manufacturing (CAM) and enterprise collaboration systems (ECS) (O'Brien, 2004a and 2004b). Transaction processing systems provide the *foundation* for practically all the remaining types of business information systems. No other business information system can be developed and operated when routine, daily business transaction processing has not been successfully implemented. The rationale for considering TPS as fundamental lies in the fact that ***transaction-processing systems*** capture, store, process, and retrieve transaction data for input to *other* types of business information systems. MIS requires and depends heavily on properly functioning transaction-processing systems. (They were the main domains of traditional data processing systems). The way of developing higher-level MIS is to begin with the deployment of transaction-processing systems first.

In their pure version, transaction-processing systems should faithfully keep track of the current state of the business reality and all the changes

taking place. They use as input deterministic data values documented according to the Generally Accepted Accounting Principles (GAAP). At the operational level, there is rarely a need for distinguishing data from information. In routine applications, one does not yet ask inquisitive questions why the data are needed, how they will be used, or to what business purpose they contribute. Simply they are required to satisfy GAAP. Those principles determine all data quality requirements including their materiality as required by the Financial Accounting Standards Board (FASB), reliability (credibility), timeliness, currency, etc. The Generally Accepted Auditing Standards (GAAS) determine the auditing procedures, etc.

At the level of business or administrative transaction processing, such a view is adequate. Similarly, the empirical approaches may help to some degree. This also explains why most of transaction processing can be accomplished with off-the-shelf application software packages. The same assumptions for assessing the quality requirements of data/information values, however, when applied to situations pertinent to **management support information systems**, are valid but insufficient. At this level, a qualitative change in thinking about quality is necessary. Any upward movement towards the management support information systems that assist in attaining business goals requires a bold paradigm shift in thinking from the:

- Ontological towards the teleological view,
- Engineering requirements towards the user requirements view,
- Intrinsic and internal towards the external situation- or task-specific view,

The form and format of data/information presentation is usually differentiated for end users at different levels of decision-making, namely the **non-managerial** level, and the **supervisory**, **tactical**, and **strategic** managerial level. When viewing quality requirements from different levels of decision-making, one immediately notices a strong shift in emphasis on different of quality requirements. At the strategic level, when using the operations view (Gackowski, 2005), the **task-specific effective operational completeness**[5] of **actionably credible**[6] information about the most urgent and dangerous threats and opportunities are the first ones to be examined with regard to their direct primary quality requirements. At the tactical level, the likely most important are: **significance of impact** and **operational timely availability.** At the lower level, however, **effective operational completeness** and all the four **direct secondary quality requirements**, which only quantitatively affect the business results, are the current concern.

With regard to the system development life cycle (SDLC), the internal systems engineering quality requirements are strictly under the designers' control. Here, however, one must again take exception to the position taken by some authors that the quality of data generated by information systems depends on their design viewed from the intrinsic and internal perspective of DQ/IQ (Wand and Wang, 1996). It depends on the design only partially. It depends much more on the user specifications the designers should follow. Designers may deliver a perfectly functioning system, which, however, from the viewpoint of the desired business outcomes may be a useless system.

Wand & Wang (1996) stated even more that the actual use of data is outside of designer's control. Again, this is only partially true. It depends on how the scope of the information systems was defined, viewed, and understood; whether it ends with technical printouts or displays only, or whether they should be designed to actually support the attainment of the stated purposes. It is a sad state that too many designers of business information systems act as if their responsibility ends with the computer generated outputs. This purely technical attitude, proper in the development of general-purpose software, is not what business owners and managers expect. Even in accredited schools of business or management, the systems analysis and design courses are mostly taught by young inexperienced instructors. They may have the proper terminal doctoral degree; however, in most cases, they never managed information system development, never participated in such projects, and never had any intimate contact with a business environment where such systems are used. Therefore, they limit themselves to teaching formal techniques and technical aspects of information system analysis and design only.

In contrast, instructors with practical experience know that in business environments the system designers' responsibilities are broader. They know how to design a report or message with a built-in feedback, which literally forces the addressee to read it, even more, to act upon it. They may insist on documenting and notifying the system about the corrective actions the end users actually triggered or neglected to trigger. If the feedback does not meet the requirements of the organizational policies then in effect, the system may also be designed to automatically notify and alert the immediate higher level of management. Notification may be about absence of action, excessively delayed action, about inadequate or ineffective action, etc. In addition, they may build in stringent auditing requirements, which, if not satisfied may again trigger other notifications and alerts. Experienced and conscientious designers with proper business experience may incorporate many subtle tricks into their designs. Those features may nearly assure the use of the values generated by information systems.

In schools of business (less in engineering schools), in research and teaching, there is a real need for moving beyond the internal view of information systems. Information systems designed and functioning correctly are the necessary but insufficient preconditions of success. Nevertheless, quality defined this way is unrelated to the desired business results, which should be the main point. The inherent relativity of quality of data/information values demands that the quality must be assessed from the natural viewpoint of defined business purposes. Data and information values to be of quality must fit the business purpose, and only then, it makes sense to consider the quality of their presentation, their processing, and their delivery system.

Some are concerned that the requirements of operations quality are quite complicated, which may hinder their usability in practice. The recommended approach compressed operations quality requirements to only 5 – 9 mandatory ones of **direct impact** on operations. Most of the plethora of 179+ dimensions is of indirect nature. Two separate cases of practical nature illuminate how convenient the operations quality approach is (Gackowski, 2006b).

The last question is **how to prioritize future research**. It seems that only the universal impact-determined hierarchical taxonomy of operation quality requirements (Gackowski, 2005) suggests rational answers to this question. In general, one may argue that:

1. Operations quality requirements that *qualitatively* that is radically affect decision situations certainly deserve the highest priority. These are the universal direct primary quality requirements (acquisition interpretable, of significant impact, operationally timely available, actionably credible, effectively operationally complete), which are valid for any value of data/information input.
2. Operations quality requirements that only *quantitatively* affect business outcomes deserve the second level of priority. These are the direct secondary quality requirements (economically acquisition interpretable, operationally timely available, presentation interpretable, actionably credible, and ultimately situation-specific economically effective operationally complete), which gain on importance when economy is an issue in data/information processing, delivery, and presentation,
3. Operations quality requirements that only *indirectly* affect the operations results via the direct quality requirements should be assigned the third level of priority.

When conducting research and planning, one must take into consideration additional questions, such as where the researchers' interests lie, where the resources are or any funding available, where are the immediate interests of business or administrative entities backed by willingness of funding such research, etc. If thinking along these lines is approximately acceptable, it may offer a starting point for identifying and planning specific research directions on DQ/IQ to widen their present scope.

CONCLUSIONS

This paper presents a compelling rationale for *extending* the current research perspective of quality of data and information values. It requires a decisive move from viewing quality from the:

- Ontological toward the teleological perspective,
- Intrinsic and internal toward the external situation or task-specific perspective
- Only formal correctness of databases towards a more content-focused perspective.

Such a move does not exclude the internal view applicable to the design and operations of general-purpose software and hardware products and services.

REFERENCES

FASB - Financial Accounting Standard Board (1983). *Accounting Standard: Statement of Financial Accounting Concepts.* Distributed by McGraw-Hill Book Company.

Gackowski, Z. J. (2005). Operations quality of data and information: Teleological operations research-based approach, call for discussion. *Proceedings of the 10th Anniversary International Conference on Information Quality—ICIQ-05* at Massachusetts Institute of Technology (MIT), Cambridge, MA.

Gackowski, Z. J. (2006b). Amount and utility of information values: Two cases of the two most misunderstood quality attributes. Proceedings of the 17th Conference IRMA2006. Washington, D. C.

Nauman, F., & Roth, M. (2004). Information quality: How good are off-the-shelf DBMS? *Proceedings of the Ninth International Conference on Information Quality (ICIQ-04), Cambridge, MA.*

O'Brien, J. A. (2004a). *Introduction to information systems* (12th Ed.). New York, NY: Mc Graw-Hill/Irwin.

O'Brien, J. A. (2004b). *Management information systems* (6th Ed.). New York, NY: Mc Graw-Hill/Irwin.

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM, 39*(11), 86–95.

ENDNOTES

[1] Limited to research that focuses explicitly on the internal vs. the external perspective of quality.

[2] "**The Internal View assumptions**: Issues related to the external view such as why the data are needed and how they are used is not part of the model. We confine our model to system design and data production aspects by excluding issues related to use and value of the data" (Wand and Wang, 1996) (emphasis added).

[3] It is impact-determined, hierarchical, and disjoint, of the already known and still unknown quality requirements. (18 formal and 12 semi-formal reviewers have not found any example to the contrary).

[4] To some degree, the internal view may be proper in the design of general-purpose software and hardware systems, where there is room for quality dimensions intrinsic to their design and operations.

[5] Generally, ***effective operational completeness*** of data/information is attained at least when a set of interpretable, operationally timely available, actionably credible information values has been gathered and among them there is at least one item of significant impact with a payoff or added value (Gackowski, 2005)

[6] ***Actionable credibility*** of information can be defined for practical purposes as the degree of credibility at which the decision maker is willing to take action in response to it (Gackowski, 2005)