OPERATIONS QUALITY OF DATA AND INFORMATION:

TELEOLOGICAL OPERATIONS RESEARCH-BASED APPROACH, CALL FOR DISCUSSION

(Research-in-progress, IQ Concepts and Models)

Zbigniew J Gackowski
California State University Stanislaus
ZGackowski@csustan.edu

Abstract: This is a theoretical attempt to discover within the vast universe of data/information (D/I) quality dimensions the first invariants and principles of more lasting validity. It required a move from the internal (designers') toward the external (users'), from the ontological to the teleological perspective and a full recognition of the relativity of quality within the context of decision situations. This paper presents for discussion and challenge an outline of a theoretical framework that encompasses: a defined main point and main frame of reference, a model within which operations quality of data or information can be assessed in any type of purposive operations, a universal hierarchical result-oriented taxonomy of data/information operations quality attributes, identification of the universal primary quality requirements, identification of the essential differences in viewing the quality of data and the quality of information, and an economical sequence of examining at least the universal direct primary quality dimensions of data/information.

Keywords: Data, information, quality dimensions, relativity of operations quality; universal taxonomy, universal quality requirements, logical interdependencies, and examination sequence of operations quality dimensions of data and information

INTRODUCTION

The 10th Anniversary International Conference on Information Quality (ICIQ) is here. The progress is evident. Still, there is more uncertainty than certainty. The empirical approach to examination and assessment of data/information quality, while producing awareness of the problem and practical improvements, did not produce research results of lasting validity. This paper presents an outline of a theoretical approach to data/information quality in response to the opinions expressed during two panel discussions on IQ Research Directions (ICIQ-03, ICIQ-04) as a contribution to future discussions on the subject with a call for discussion and challenge.

In research and practice, there is a clear need to recognize the undeniable relativity of quality within the context of purposive operations. It requires a bold move from the solely internal towards the external, from the ontological towards the teleological perspective of quality. All attributes of data/information operations quality in any type of purposive operations acquire practical meaning and importance from the context of their use, hence should be assessed within the same context. The internal view leads towards defining only the logical and engineering requirements of quality, which are under the exclusive control of designers. The external view leads towards user requirements of quality that should be under the users' control and belong to the realm of operations research and management disciplines. For a correct assessment of data/information quality, 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 the use of data and information. This fact alone makes very questionable most of the attempts of empirical, in particular survey-based studies, which ask subjects unqualified questions about their view of data/information quality and how they rank their different attributes when the context of their use is undefined or unknown.

In 2002, Liu and Chi [6] categorized different approaches to data quality as intuitive, empirical, and theoretical. Initially, the intuitive and the empirical approaches were most prevalent, but these lack theoretical foundations on how DQ/IQ attributes are defined, grouped, and examined. 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*". They developed a concept of **evolutional and theory-specific data/information quality** that evolves along the stages of data collection, organization, presentation, and application.

To develop a full-fledged qualitative framework of DQ/IQ, one must reach beyond the empirical survey-based assessment of loosely defined attributes of data/information (D/I) operations quality, where the correctness and completeness of the results cannot be proven via fundamental principles. The effort devoted to the empirical studies should be applied to approaches with a higher potential of producing results of more lasting validity. 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 DQ/IQ dimensions in their multiple aspects.

The main contributions of this paper are:

- A definition of relativity of all aspects of data/information operations quality for purposive operations
- A definition of the main point and the main frame of reference for the subject
- A qualitative model based on the concept of decision situations in purposive operations, within which DQ/IQ is assessed
- A universal hierarchical result-oriented taxonomy of attributes of operations quality of data and information
- Identification of the first universal direct primary mandatory operations quality requirements for any information value that is examined for the first time within the context of its task-specific use
- Identification of the essential difference in assessing quality of data and quality of information
- Identification of an economic sequence of examining the universal direct primary operations quality dimensions of data/information
- Identification of several universal principles pertinent to operations quality of data information used in purposive operations
- Examples of some predictive and explanatory features of the proposed framework for DQ/IQ

This paper is a call for discussion, critique, and challenge of its main thesis with the purpose of arriving at a first outline of a more universal and sturdier framework for examining the operations quality of data and information.

BACKGROUND

DATA QUALITY VIEWED EMPIRICALLY

In the empirical survey-based approach, researchers made attempts to arrive at an empirical definition of data quality based on the data consumer's perspective: that is, of those who use the data. In 2005, Mende after a review of research processes described in Philosophy of Science, History of Science, and Research Methodology with respect to how productive they are producing explanatory theories stated that empirical methods in this respect are inherently useless [7, p. 189].

Beyond Accuracy: What Data Quality Means to Consumers

The most representative example of this kind or empirical survey-based research is the one published in 1996 by Wang and Strong [11]. It attempts to define data quality not be by providers or custodians of data, but instead, by data consumers. ***Data quality*** is defined as data *fit for use* by data consumers**.** A ***data quality dimension*** is defined as a set of data quality attributes that represent a single aspect or construct of data quality. "*Data is treated as a product. While most data consumers are not purchasing data, they are choosing to use or not to use data in a variety of tasks"* [11, p. 3].

They also say: "*Poor data quality (DQ) can have substantial social and* <u>*economic impacts*</u> *(emphasis added). Although firms are improving data quality with practical approaches and tools, their improvement efforts tend to focus narrowly on accuracy. We believe that data consumers have a much broader data quality conceptualization than IS professionals realize. The purpose of this paper is to develop a framework that captures the aspects of data quality that are important to data consumers.*

A two-stage survey and a two-phase sorting study were conducted to develop a hierarchical framework for organizing data quality dimensions. This framework captures dimensions of data quality that are important to data <u>*consumers*</u> *(emphasis added). Intrinsic DQ denotes that data have quality in their own right. Contextual DQ highlights the requirement that data quality must be considered within the context of the task. Representational DQ and accessibility DQ emphasize the importance of the role of systems. These findings are consistent with our understanding that high-quality data should be intrinsically good, contextually appropriate for the task, clearly represented, and accessible to the data consumer.*

Our framework has been used effectively in industry and government. Using this framework IS managers were able to better understand and meet their data consumers' data quality needs. The salient feature of this research study is that quality attributes of data are collected from data <u>*consumers*</u> *(emphasis added) instead of being defined theoretically or based on researchers' experience. Although* <u>*exploratory*</u> *(emphasis added), this research provides a basis for future studies that* <u>*measure data quality along the dimensions of this framework*</u> *(emphasis added)*" [5]

This study was later widely used and cited, and still is. In 1999, it was recommended for undergraduate curricula in IS and MIS [4] and inspired other empirically oriented research efforts published in 2002 such as: "AIMQ: A Methodology for Information Quality Assessment" [5].

The researchers conscientiously followed the selected methodology. The problem lies, however, in insufficient examination of the discrepancies between the real subject of the study and the assumptions used. The proposed and recommended Conceptual Framework of Data Quality is incomplete and inconsistent in the proposed grouping of the retained and suggested dimensions. The authors of the study rejected some quality dimensions inconsistently with the principles of information system design, the fundamental principles of business economics, and the generally accepted practice of handling data of real importance in business, public administration, and military operations.

AIMQ: A Methodology for Information Quality Assessment [5]

Here data quality is also viewed from the information <u>consumers</u>' (emphasis added) perspective, and is based on the Total Quality Management (TQM) principles. Probably it contains the broadest overview of academics' and practitioners' views on IQ dimensions. In 2002, in [5], Lee, Strong, Kahn, and Wang say: "*Information quality (IQ) is critical in organizations. Yet, despite a decade of active research and practice,* <u>*the field lacks comprehensive methodologies for its assessment and improvement*</u> *(emphasis added). Here, we develop such a methodology, which we call AIM Quality (AIMQ) to form a basis for* <u>*IQ assessment and benchmarking*</u> *(emphasis added). The methodology is illustrated through its application to five* *major organizations. The methodology encompasses a model of IQ (emphasis added), a questionnaire to measure IQ, and analysis techniques for interpreting the IQ measures. We develop and validate the questionnaire and use it to collect data on the status of organizational IQ. These data are used to assess and benchmark IQ for four quadrants of the model. These analysis techniques are applied to analyze the gap between an organization and best practices. They are also applied to analyze gaps between IS professionals and information consumers (emphasis added). The results of the techniques are useful for determining the best area for IQ improvement activities*."

On its surface, it appears to be a strong model within the confines of TQM principles until; however, one realizes its inherent limitations. It is limited to products or services, and to given specifications or preferences of information users. These limitations are substantial when one becomes aware of the consequences:

- Products or services are not identical with purposes, goals, and objectives of business operations.
- Specifications provided by a contracting entity may be sacred to the contractor, but they may be substantially deficient in meeting the actual purpose of operations.
- Preferences of information users within any organized entity may deviate considerably or even be in conflict with purposes of the entity they serve or work for.

DATA QUALITY VIEWED THEORETICALLY

1. In 1996, based on **ontological foundations**, Wand and Wang [10] proposed four intrinsic data quality dimensions: complete, unambiguous, meaningful, and correct. Within the confines of the assumptions used[1], "*those attributes have crystal-clear definitions and theoretically sound justification, but they constitute only a small subset of known attributes leaving the rest unspecified*" [6]. The above is true when assuming a perfect interpretation of the authors' intent. In reality, casual readers of secondary sources, in particular, are lured into believing that for instance "*incompleteness*" of data has been defined, whereas in actuality, the designer of an information system did not provide enough states for proper representation of the selected states of reality. All the intrinsic dimensions are logical or engineering requirements of an information system to assure its *correct* internal functioning. They are *necessary* but *insufficient* requirements for a successful information system. Their ultimate success depends on a purpose-focused selection of those real-world states that needs to be reckoned with in decision-making. Thus, results come only from a *sufficient* combination of necessary requirements defined from both the teleological and the ontological perspective.

2. In 2002, in an **evolutional and theory-specific approach** to data quality, Liu and Chi [6] try to overcome the weaknesses of the product analogy based approach and the narrowness of the ontological approach that is limited to the internal view only. They claim that data have meaning only through a theory. As data evolve through the stages of the data evolution life cycle (DELC), they undergo a sequence of transformations and exist independently as different morphons (**captured** data; **organized** data, **presented** data, and **utilized** data – see Figure 1.). Each transformation introduces independent errors such as measurement errors during data collection, data entry errors during data organization, or interpretation biases in data presentation. Different theories apply to different stages of the DELC; hence, different definitions are needed to measure the quality of those morphons. Instead of a single universal concept of DQ, four hierarchical quality views are used for data collection, organization, presentation, and application. They measure respectively the quality of **collected** data, **stored** data, **presented** data, and **utilized** data. The evolutionary nature of the four views implies that the quality of data at earlier stages of DELC contribute to the data quality at later stages. The authors suggest a monotonically increasing order of specificity of the four hierarchical views of DQ.

[1] "**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" [9] (emphasis added).

Certainly, the concept of evolutional and theory specific data quality (ETSDQ) is a landmark. It addresses new issues, never or rarely discussed before, but it is also burdened with serious drawbacks. One may argue that data derive their meaning, relevance, and utility value only from their use for task-specific purposes. A theory may only explain its role for us. Another drawback is that data/information is not treated

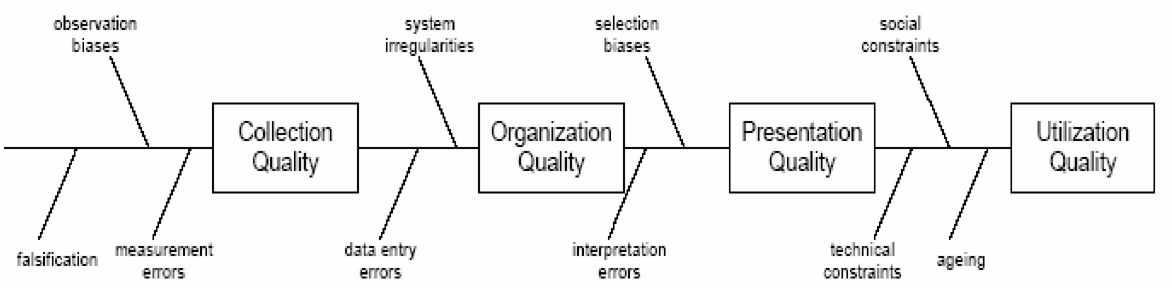

Figure 1. Fishbone diagram for evolutional data quality. Source: Liu and Chie [6]

as other resources that require similar attention as the 4 M's (Methods, Machines, Materials, Manpower). They all, in their own way, serve the same operations' purposes. Why incur the cost of acquisition, storage, presentation, and utilization of data for no defined purposes of higher value? Data collection should be guided by the specific purposes the data are used to serve. It is a postulate of business economics. Sound systems analysis also teaches that the hierarchical examination of DQ should be top down and dataflow up; the requirements of data application quality should determine the requirements for data presentation quality, those in turn the data organization quality, and finally the latter the data collection or acquisition quality.

3. In 2004/2005, anchoring the concept of data/information quality in **operations research**, **management science, and decision science,** Gackowski [1, 2] proposed a **purpose-focused perspective** derived from business needs. It was defined in four steps: (1) Development of a relatively complete qualitative cause/effect diagram, in the form of a fishbone diagram, to identify the major factors impacting the desired business results. (2) Impact analysis of the relative strength of each factor identified before. (3) Development of an informational model of the decision situation by taking inventory of what is already known (***data***), and what is unknown (***information***) and must be acquired by business intelligence, and then ranking each data/information value by its impact on the operational outcomes using any agreed measure[2]. (4) Examination of each data/information value with regard to quality requirements.

This approach enables a hierarchical result-oriented taxonomy, which divides the entire universe of DQ/IQ attributes separating them into direct and indirect ones, the direct into primary and secondary ones, and the primary into mandatory always or mandatory only in specific circumstances. It defines the necessary and sufficient conditions of task-specific, effective, operationally complete sets of data or information and the two main levels of their usefulness when they may be *only effective* or they are also *economically effective* [2]. It also facilitates a result-oriented, simplified, time saving and economical examination sequence of the direct primary mandatory quality dimensions.

This approach views data information items within a universal pragmatic context that illuminates their specific role and significance. It is applicable not only at the operational, but also at the strategic levels of management. Where necessary, it requires a clear distinction of data (the known, given or assumed) and information defined in a utilitarian way[3]. It is, however, still research in progress without a sufficient

[2] Examples: net income after taxes, retained earnings, return on investment, return on equity, cost effectiveness of services, etc.
[3] For instance, when it qualitatively changes the decision situation itself and/or quantitatively changes the results, and/or changes the actions that implement the decisions made.

validation of its main thesis. Nevertheless, the soundness of the presented rationale for this kind of approach to DQ/IQ has not yet been challenged. If extended by a *dependency map of IQ dimensions* (in progress), it could facilitate a fast and economical sequence of their examination. Similarly, a further extension of the proposed result-oriented taxonomy by overlaying it with a *cognitive taxonomy of factors impacting business results* seems to be promising. The latter may elevate the assessment of IQ from mainly the operational level to the *strategic level* of applications in business, public administration, and military operations.

PURPOSE: THEORETICAL FRAMEWORK FOR OPERATIONAL QUALITY

RELATIVITY OF QUALITY

Accepting the widely adopted definition of ***quality*** as "fitness for use," one notices that it is the *use* of data/information that determines what *fits*. Thus, quality is determined by the use of data or information. Some authors see it as a problem [10], others as an opportunity. It is a problem when one limits the view of data/information quality to the information system designers' view and the faithful representation of the known aspects of the world. This may be proper, to a limited degree, in the design of general-purpose systems. In business, public administration, and military systems analysts and designers usually custom-tailor their designs to meet specific needs of the organizations the information systems serve. Otherwise, off-the-shelf information systems would cover the majority of needs.

The assessment of how well data/information fits their intended use is determined by the prevalent view of what is important in its use. Again, the use is determined mainly by its **purpose** and the **circumstances**. Both define something similar to a force field that determines all aspects of quality of whatever is used under those conditions. In tightly run, cohesively bonded organizations, views focused on tactical or strategic operations' 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.

Only under the assumption of a "tightly run ship" will all aspects of data or information quality be determined by the purpose and circumstances of operations within which they are used. In theory, innumerable authors [7, p.202] refer to it as "**unobservable first cause**" despite the fact that nearly two centuries ago Schopenhouer trashed it a contradiction in adjective. He replaced it by the "**principle of sufficient reason of acting**" or in short by the "**law of motivation**" [11]. Indeed the triggering motives for action by a single decision maker are unobservable and may not be known by those who act except for the decision makers themselves by introspection. Sometimes they may be inferred only with enormous difficulties. Let us not be fooled with officially announced mission statements, statements of intent, declarations, denials, etc., as it frequently and naively is assumed.

Thus, ***relativity of quality of data or information*** is defined here as a functional dependency of all its aspects on the **purpose** and **circumstances** of operations where those data or information values serve as resources. Since this definition refers to human purposes and relativity manifests itself in human actions, teleological methods [7, p. 202] are natural in such situations.

METHOD: POINT AND FRAME OF REFERENCE AND A THE MODEL OF DQ/IQ

A fruitful theory must be based on the identification and/or definition of a rock solid point of reference and frame of reference for further deliberations. In this case, as the ***main point of reference*** serves the main purpose of the operations and as the ***main frame of reference*** serve the circumstances under which the operations take place. Once the above are defined, a teleological operations research-based approach is natural for defining the outline of a model of reasoning about operations quality of data and information. One assumes:

- A relatively complete qualitative cause/effect diagram of operations, known also as a fishbone diagram, is available or can be drawn. It identifies the major factors affecting the purpose of operations described by the situation itself, the expected results, and the required actions to implement the decisions made. Based on the data/information available, decision makers (here data/information users or clients) make pertinent decisions and take subsequent actions. The results of the operations can be determined as a function of the data and information used.

- An analysis can reveal the relative strength of each factor, here data or information, by its impact on the main purpose of operations. In business, various criteria are used to measure the main purpose of its operations as a function of the data/information used such as net income after taxes, retained earnings, return on investment, return on equity, etc. In public administration, measurable/observable results can be derived from the entity's mission. In military operations, they may be described by the expected tactical or strategic objectives.

- From the view point of data/information only, one can develop an informational model of the decision situation under consideration by taking inventory of what is already known, and what still must be acquired. That, which is known, given, or available constitutes the **data component** of the model. Anything that is not yet available, is still missing, and thus must be acquired by proper intelligence or other means, constitutes the **informational component** of the model.

This approach requires a rigorous distinction between data and information, which otherwise may not be required. Here, this distinction is made within the context of decision situations. Decision makers or acting agents already know some aspects of the situation, but some other aspects remain yet unknown. Hence, ***data values*** are symbolic representations of aspects of reality that are known, given, or assumed true. ***Reality*** encompasses organizations and their environments. Within reality, one distinguishes ***entities***, which are objects or events represented symbolically by their identifiers and values of their attributes. ***Information values*** refer to those symbolic representations of things, events, and unknown states of the environment that must yet be acquired, and which may change the decision situation by itself and/or the operations results, and/or the actions that are necessary to implement the decisions made.

From the viewpoint of the theory of communications, any symbolic representation already known contains or conveys zero (0) bits of the amount of information. Shannon's formula for the amount of information $\mathbf{A_I = -log_2\, p_I}$ associates $\mathbf{A_I}$ bits of information with any symbolic representation of reality that is yet unknown as a function of its probability $\mathbf{p_I}$. The amount of information measures the rarity or the surprise effect associated with the object, event, or state. Thus, symbolic representations of objects, events, or states, which are very unlikely, with probability $\mathbf{p_I}$ close to zero, are associated with nearly an infinite amount of information for their recipients. ($\mathbf{A_I = -log_2}\, 0 = \mathbf{-\, log_2}\,(1/\infty) = \mathbf{log_2}\,\infty \approx \infty$ [bits])

RESULT-1: UNIVERSAL TAXONOMY OF ATTRIBUTES OF OPERATIONS QUALITY OF D/I

Most textbooks and the cited empirical studies list under different names a plethora of attributes or dimensions of data/information quality for consideration. The major question is, however, how to examine those attributes in real life situations. Which of them affect the situation results directly or indirectly, are primary or secondary, mandatory or optional, should be examined first, or are not fully attainable and therefore one must learn to act with only some acceptable level of quality. This leads us to a universal taxonomy of all known and not yet known data/information quality attributes. Such a ***universal hierarchical result-oriented taxonomy*** is defined as follows [2]:

1. One can subdivide the universe of all operations quality attributes into direct attributes and indirect or subordinate attributes. Changes to the values of ***direct attributes*** affect the decision situa-

tion itself, and/or the results of operations, and or the actions to implement the decisions made. Changes to values of ***indirect attributes,*** as the name suggests only indirectly affect the entire situation, for they only determine or contribute to the values of direct attributes.

2. The direct attributes can be subdivided into primary and secondary attributes. Changes of the values of ***primary attributes*** result in *qualitative* changes to the decision situations under consideration, while changes to values of ***secondary attributes*** only *quantitatively* change the results of operations. The latter are mostly of economic nature. If not only effectiveness, but also economy of operations is required, they become mandatory too.

3. Within the primary attributes, one must also distinguish those of universal validity versus the situation-specific ones determined by the circumstances of operations. The ***universal direct primary requirements*** apply to each data/information value under all possible circumstances; the ***situation-specific ones*** pertain only to some attributes of some data/information values under specific circumstances.

RESULT-2: UNIVERSAL OPERATIONS QUALITY REQUIREMENTS FOR ANY D/I VALUE

The argument for universal operations requirements is easier to comprehend when one distinguishes examination conducted (A) For the first time, and (B) Routinely during regular operations. One may gain insight into how operations quality of data/information manifests itself by conducting a thought experiment [7, 202] when decision makers face a previously not yet encountered data/information item.

A. FIRST TIME OR INITIAL EVALUATION OF OPERATIONS QUALITY OF DATA/INFORMATION

(1) Interpretable during acquisition

It seems certain that at first information values must be interpretable. Messages or composite data/information may consist of one or more values. For them to be later usable at all, they must be **interpretable** within the process of their acquisition[4]**.** Pragmatically, interpretability means the received information value matches any state with some attributed or associated meaning in the mind of the receiving individual, or any state that automatically triggers a designed sequence of state transitions in the receiving numerically controlled device. It is how pragmatically that meaning manifests itself in operations. When for any reason the targeted individual or the receiving device is unable to interpret a data or information value, it is lost and it must be *excluded* from further examination.

Interpretability is contextual; a more educated receiver, a conditioned one, a trained one, or a different receiving device may be able to interpret it even in adverse conditions. The information carrying signals must be noticeable to senses or sensors, discernible, recognizable, or identifiable. The latter are indirect attributes that determine or contribute to the direct attribute of quality - interpretability. Those considerations are of major concern for users and disseminators of information. The interpretability of information values during their acquisition should not be confused with presentational interpretability for users/clients, to be discussed later.

More complex conditions and circumstances must be considered by information disseminators. A plethora of factors here comes into play. Many of them are of a very subtle psychological nature dealing with how to effectively and gainfully reach the targeted client by employing the many findings, skills, and tricks offered by the art of communications, marketing, advertising, etc, which intersect here. They are the subject of studies of how to effectively attract others' attention and how to persuade them to take the desired actions by changing the state of their minds.

[4] Typical current and important examples represent the volumes of documentation in Arabic available to the CIA and the military intelligence is frequently useless due to shortage of reliable translators from Arabic into English.

(2) Of significant impact, relevance by Gleim [3] for CPAs, and by Wang and Strong [11]

Once a data/information value becomes meaningful for decision makers, it can be examined further for its situational and contextual importance. In purposive operations, the content or meaning of individual data/information values or any combination thereof must make a *significant* impact on the situation under consideration. In operations, ***impact*** should be assessed by the scope of changes it causes in the decision situation itself and/or in the results of operations, and/or the actions required implementing the decisions made. If the impact is insignificant from the perspective of decision makers, the remaining attributes of the information item are irrelevant too. Here one arrives at one of the first universal principles of operations quality of data/information.

Use of any data/information value may impact the decision situation itself and/or the results of operations and/or the actions required to implement the decision. The said impact lends importance to all the remaining attributes of operations quality of any data/information value under consideration.

In other words, it is the **principle of the supreme role and precedence of relevance of data and information values (measured by their impact on the results of purposive operations) over all other attributes of DQ/IQ that lends them their relative situation-specific importance among other quality attributes within the case.**

Impact, however, can be quantified or at least ranked either from the viewpoint of information disseminators or information users. Now, one may ask also how the payoff, added value or better *added main effect* that defines the main purpose depends on the use of any specific data/information value, whether its impact is ***significant*** enough to warrant the decision makers' consideration. In a more rigorous manner, one may say that in a specific situation a data/information value may be qualitatively relevant but quantitatively irrelevant when its impact is negligible. If so, one should *cease* further examination of its remaining quality attributes.

It may also happen that the size of the impact may depend on other factors such as type of its availability: whether restricted only to a specific decision-maker, fully unrestricted or anything in between. Restricted availability of data/information gives advantage to some decision makers over others. Unrestricted availability may reduce that advantage to insignificance. These, however, are examples of *situation specific* requirements. (See the later discussed dimension "operationally timely available")

One must be also aware of the very frequent case when a data/information value of a *zero added main effect* or payoff might still significantly impact the ultimate outcome. This takes place when it is a mandatory required companion of another information value that is associated with a significant **added main effect**. (The latter label reflects better those situations when the main purpose is not defined in monetary units, for instance in rescue missions.) For instance, emergency calls for roadside assistance with a well-defined payoff must be accompanied by information values about the location. Without the latter, such calls cannot be effectively handled on their own. The definitions of relevance referred to in the subtitle are much too narrow from the operations research view.

(3) Operationally timely available, timeliness by Gleim [3], accessibility by Wang and Strong [11]

Once decision makers determined that they are dealing with a data/information value of significant impact, they should ask whether in regular operations it will be timely available, hence the label "operationally timely available"; means before it loses its capacity to make a significant impact. It pertains to individual data or information values and any combination thereof. If the data are not available in time to meet the user's need, why bother about other requirements? In ever-changing reality, time is of the essence. Even with all remaining requirements met perfectly, if timely availability cannot be assured, the impact of late data/information values may be null (again, see Footnote 4).

The "operational timely availability" requirment may also be viewed differently. For instance, whether the data or information value under consideration is available exclusively to a single interested individual. Here, one deals with two exremes: with restricted or unrestricted availability. From the viewpoint of logical interdependencies among DQ/IQ dimensions, one can see here an interesting case of circular interdependence between two quality dimensions. In order to consider at all the "operationally timely availability" of any data/information value, it must be of "significant impact". In a competetive envirnoment, however, frequently a significant impact of data/information may depend on its restricted availability. The latter is an example of situation-specific requirements that depend on the circumstances, are task-specific and do not pertain to all values, hence cannot be considered universal.

(4) Actionably credible, believable by (Wang and Strong, 1996), reliable by (Gleim, 2004)
Messages declared of significant impact and operationally timely available, which implies they were interpretable during acquisition, must be tested by decision makers for **credibility**, that is, whether they are true or can be relied on. The adjective **true** means consistent with reality. While probing for veracity, users/clients face dramatic options: whether they received valid information, misinformation, or disinformation

- ***Valid information*** faithfully represents or reflects reality. To this end, it should be objective (unbiased), accurate (error free), precise (sufficient number of digits or points per inch), and current (up to date). Usually validity is assumed, when information is of proven authorship, from a reputable source; replicable, confirmed otherwise, or traced back to the responsible originator, where the level of his/her responsibility should be proportionate to the potential consequences of possible errors.

- ***Misinformation*** unintentionally misrepresents reality. It may be distorted at its acquisition, communication, storing, processing, presentation, and the interpretation by itself.

- ***Disinformation*** intentionally misinforms. On one hand, in simple cases, it may not be clear who the originator is due to omission of contact addresses, when it was originated or updated, what methods of collection or acquisition were used, etc. On the other hand, all the above may be present and available. Now, however, the user or client faces two extremes of deception with many possibilities in between. All the above listed indicators of validity are given: (1) To appear legitimate, but one or more of them are false, or (2) Are true, and usually presented in a very motivational manner, but actually the intent of the message is deliberately malicious, criminal, aimed at trapping the gullible.

Credibility data/information values is a complex function of their many indirect attributes of quality such as their the **credibility of the source** they are derived from, the **variety of independent sources** available, the **quality of their mapping** within the delivery system and the **credibility of their presentation** in indirect informing. These attributes are referred to as **indirect attributes of the first order**. Most of them again are functionally dependent on many other **indirect attributes of the second order**. For instance, on one hand, the reputation of data/information sources is a again a complex function of such favorable factors as whether they offer any **means of verification**, **replication**, and any type of **warranty** with respect to the veracity of D/I values they provide. On the other hand, the reputation of the source, as well the credibility of data/information presentation to their users will always be impaired by imperfections of their **definition**, **objectivity**, **accuracy**, **precision**, **currency**, and **reliability of delivery**. All this will be subject of a separate presentation within the promised task of mapping of such functional interdependencies of all remaining universal direct primary attributes of DQ/IQ.

Since credibility is rarely-to-never fully attainable, in many situations users have to learn to act with only an acceptable level of credibility labeled **actionably credible.** For practical purposes, ***actionably credible*** can be defined as the degree of credibility at which the user is willing to take action. The definition is precise, but the actionable level of credibility again is a function of the decision situation and all the circumstances, including in particular the personality of the decision maker.

"Actionable credibility" completes the list of the four universal operations quality attributes of data or information that defines by enumeration the ***task-specific universally necessary operations quality requirements of usability of single data or information values*** in all situations of **direct informing;** when the same individual acquires and uses information with no intermediaries between the source and the user/client [2]. It is usable from the senders and/or the targeted clients' viewpoint. Usability indicates that a data/information value may be used but not necessarily effectively. In order to arrive at ***task-specific sufficient operations quality requirements of usability of a single data or information values*** one must complement the list with the situation-specific direct mandatory operations quality attributes, for instance the previously mentioned exclusivity of availability.

In other words, this is the principle of defining the universally sufficient operations quality requirements of situation-specific usability of a single data or information value.

(5) Task-specific effectively operationally complete

Completeness of data/information pertains to a set of identified task-specific factors of significant impact. Once decision makers arrive at such a set of usable data or information values, they must test it for its completeness with regard to the situation under consideration. Completeness of data or information values in decision-making, in contrast to their mapping while storing, processing and presenting them is more complex than it appears on its surface. Completeness is strongly related to the data set's impact as defined earlier. One must distinguish between at least two types of completeness: **operational completeness** and **congnitive completeness**.

Within the context of decision situations, ***operational completeness*** measures the degree to which the data/information values of significant impact are available. Operational completeness may be measured in percentage points (0 - 100%) as the ratio of the sum of all results that can be attributed to the corresponding values available and the sum of results attainable when all values were available, which is an ideal situation. In real-life situations, usually, some residual operational results remain unaccountable. This means it is not possible to attribute them to any previously identified factors. They may be used as a relative or absolute measure of how incomplete the impact analysis is.

Murkier is the qualitative or **cognitive** aspect of **completeness** of data/information values. In real life situations, in the fight for survival, on a battlefield or in global business competition, one may never be certain whether all relevant success factors or dangers are identified and evaluated. Prudence requires gathering more information to inform interested decision makers so that they may assess all the maybe not yet perceivable but potentially critical factors for planning of counter measures and contingency provisions. The critical blow most frequently comes from a danger or direction not identified and recognized in time. It cannot be considered mandatory on its own merit, for it is rarely-to-never fully attainable. This, however, is a separate complex issue that deserves a separate inquiry.

Like credibility, completeness is measured by a continuum of degrees. From the pragmatic viewpoint, there is at least one important degree of task-specific completeness, when it becomes **effectively operationally complete.** It should be a useful completeness. Gleim [3] reminds us, "Without usefulness no benefits are provided." When economy of operations is secondary to operational effectiveness, one may be satisfied with effective operational completeness. A ***minimal*** e***ffective operational completeness*** is attained only when, within the set of task-specific mandatory usable data or information values, there is at least one value of *significant added main effect*, added value, or payoff of operations. It pertains to operations conducted with an *all out effort* according to their purpose when at least temporarily economy is of secondary concern.

In other words, it the **principle of defining the universally sufficient operations quality requirements for a situation specific set of usable data/information values of minimal effective operational completeness.** Of course, the more situations-specific usable data/information values of added main effect are included in such a set the more effective the operations under consideration may become.

Completeness closes the list of the five universal direct primary mandatory requirements of data or information quality, which pertain to all values, their sets, and situations. They are *mandatory*. Changes to the values of the respective quality attributes result in <u>qualitative</u> changes in the decision situations under consideration. Sometimes, there may be other additional situation-specific quality requirements such as the mentioned exclusive or restricted availability.

Thus, ***acquisition interpretability, of significant impact, operationally timely availability, actionable credibility, and task-specific effective operational completeness constitute the first five universal mandatory quality requirements that are prerequisites for successful operations.*** Of course, cognitive completeness of even known potential threats does not guarantee success when ignored by decision makers.

The sequence in which the above direct primary quality requirements are presented has been obtained by examining which one "is a prerequisite of" another one. If it is not the case, then one should test which one is easier to test. Thus, the presented ***sequence of examining the universal direct primary mandatory attributes of data/information quality seems to be the only logical, and at the same time, the most economical one.***

What follows is a discussion of the direct but *secondary* attributes of operations quality of data or information. The secondary attributes are of economic nature.

WHEN ECONOMY OF OPERATIONS IS AN ISSUE

In most cases, economy of operations is an issue. If this is so, one must reexamine the above direct primary mandatory quality requirements from the perspective of whether they are *economically* attainable. At this point in the examination, one knows already how sensitive the results of the decision situation are to the use of any specific data/information value. The quantified and *ranked* impact indicates the entire *added main effects* (gross benefits, payoffs, added values) that are possible to attain and provides the decision makers with a **situational reference scale**. It suggests how much attention one should pay to each data or information value relative to the others when reexamining them for economy of the remaining universal direct primary mandatory quality requirements that make those values usable.

Thus, one must test the economic level of acquisition interpretability, operational timely availability, actionable credibility, and ultimately the task-specific effective operational completeness of the entire set of task-specific usable data/information values. Changes in values of the respective direct secondary data or information quality attributes result only in *quantitative* changes in operations results. These changes are additive. Therefore, the sequence of their examination is formally irrelevant. When economy matters, an effective operationally complete set of task-specific usable data/information values must be not only effectively but also economically useful. Then the direct secondary quality attributes are mandatory too. For instance, when a usable data/information value cannot be economically acquired, one should exclude it from further examination of its other quality attributes. Therefore, the attributes should be examined in a sequence determined by the ascending amount of effort required to test their economic aspect, which means starting with the easiest and finishing with the most difficult one.

The following three requirements are also mandatory and pertain to all situations and to all single data/information values of significant impact on the main effect, if economy of operations is an issue.

(E1) Economically interpretable during acquisition

Interpretability of incoming information values during their acquisition is the first universal direct primary mandatory requirement that must be met in order to trigger the chain of further examinations. In certain situations, however, it may be attained only at a prohibitively high cost in comparison to the associated added main effects so that it is not worth the effort. Similarly, it also pertains to rescue operations when it does not make sense to lose more lives to save another live, except for bodyguard duties. Such cost may entail the cost of decoding, translation, maintaining a system of early detection and warning about dangers (hurricanes, earthquakes tsunami, missile attack, etc.). Therefore, while examining it, the first question to ask is whether it is technologically possible at all. The next mandatory question becomes whether its other mandatory requirements can be met economically in its general meaning, as explained before. Only when interpretability of information values during their acquisition is economically attainable, does the door open for examination of the remaining direct secondary requirements. On the other hand, when the stakes of national security are higher than the established economical criteria, the concerned data or information values remain effectively usable, even though it is not economical in its conventional meaning.

(E2) Economically operationally timely available

Meeting mandatory requirements usually does not add value; it only makes the data/information's values usable. The cost of making values operationally timely available should not exceed the associated added payoffs. In addition, timely availability is scalable. One may receive the necessary data or information not only on time, but also more or less in advance[5]. The cost of rendering them sooner "operationally available" should also not exceed the associated additional benefits. The additional time may be used for making decisions with less haste, and/or for better preparation of actions. Hence, one may also obtain better results when additional time is available. Additional time may add cost. Excessive additional time, however, may also cause deterioration of results due to human forgetfulness or possible distractions between the time of early warning and the time for action. There is no analytical formula to estimate optimal timing; however, in specific situations it may be possible to determine the best timing experimentally. Whether it is worth the effort depends on the difference in results it makes and on how much it will cost to inform sooner.

(E3) Economically actionably credible

Actionable credibility may be compromised by deficiencies in several indirect quality attributes, for instance for instance by imperfection in their **objectivity, accuracy, precision** and/or **currency**.

Objectivity, here meant as free from bias, may be compromised in the process of data/information acquisition due to the approaches and methods used in selecting the primary sources, measuring points, observation points, and when collecting, processing and presenting data. The resulting bias may be either unintended due to ignorance or introduced intentionally. In both cases, the results of such distortions may be significant, and in the latter case, deceptive and damaging. To rectify the bias and/or compensate for it may require engagement of substantial additional resources. That it is economically justified can be estimated only when the size of its impact on the results is significant enough.

Another problem is **accuracy** meant as free from errors, including random errors. One encounters errors in all situations. Usually accuracy is indicated indirectly by inaccuracy of data or information values, which is the complement to one (1) of accuracy. Inaccuracy or error rate equals one (1) minus accuracy. A typical gross measure of **inaccuracy** in this sense is the **error rate**. One calculates it by dividing the

[5] How important it is and how neglected it may be demonstrates the aftermath of the recent hurricane Katrina.

number of values in error by the total number of data or information values gathered. In practice, a more useful measure of inaccuracy due to different kind of errors is the **expected cost of dealing** with their consequences. One may calculate it by multiplying the number of data/information values by the probability or frequency of each type of error by the average cost of dealing with each type of them. This measure of **inaccuracy** provides the users with a better idea how serious the consequences of each type of error are. One may reduce many errors by using check digits, error self-detection codes, error self-correcting codes, etc. A good example is the use of barcode readers, which considerably reduce many errors. Clients of information systems, even business systems analysts, need not to be experts in dealing with such situations, but they should be taught to recognize the need for those measures.

One may also encounter low **precision** in the representation of reality. For numerical data, ***precision*** is measured by the number of significant digits used. One measures the precision of pictures and images by the number of dots per inch. This unit is commonly used to describe the precision of printers, computer screens, scanners, etc. Insufficient precision of data/information presentation may compromise the results obtained.

There is a trap associated with accuracy and precision. Generally, they are overrated [11]. U*nchecked* efforts to increase the level of accuracy and/or precision of any data or information value can become counterproductive. The ultimate determination of the indispensable and economically justified level of any of them depends on how significant the impact is.

Currency of data/information values means here whether they are adequately up to date for their intended use. It was labeled timeliness by Wang and Strong and defined in [11], as "t*he extent the age of the data is appropriate for the task at hand.*" The label "timeliness" used here by some authors is in conflict with the terminology used by CPAs for "timely availability," causing unnecessary confusion. The frequency of updates should be optimized, where possible, since either insufficient frequency or too frequent updates are detrimental to the cost effectiveness of operations.

Maximum operations benefits from using data/information values can be attained only at the *optimum level of objectivity, accuracy, precision, and currency*. Finding this optimum is not easy, but the truth is that it lies somewhere between the low and high level. For instance, whenever information technology professionals tempt clients with higher accuracy, precision or currency than they had before, they should ask bluntly "What will be the additional benefits and at what additional cost?" When one has no indication that their increased level leads to higher cost effectiveness, forget it. To the surprise of many, one thing is sure; the *examination of their economic level should be postponed nearly until the very end.*

(E4) Task-specific economically operationally complete

The ultimate test for economy of data/information quality has to be performed on task-specific effective operationally complete sets of usable data/information values. Now, as mentioned before, those sets must be not only effectively operationally complete but also *economically operationally complete*. Thus if economy takes precedence over other considerations, all the before discussed direct secondary quality requirements become mandatory as well. Here, one may distinguish three types of task-specific effective operational completeness:

- ***Economical task-specific operational completeness*** - attained when the sum of all payoffs or added values exceeds the sum of all costs of the operations under consideration, if they can be expressed in the same units
- ***Cost effective task-specific operational completeness*** - attained when the ratio of the sum of all added main effects divided by the sum of all operations costs exceeds the required level.

- ***Expected cost effective task-specific operational completeness*** - computed as above, but as a function of expected effects and expected costs

The most economical sequence of examining the direct secondary quality requirements seems to be from the easiest to the most difficult to test.

B. ROUTINE EVALUATION AND ASSURANCE OF OPERATIONS QUALITY OF DATA/INFORMATION

In repetitive *routine* operations, data/information quality is viewed and perceived differently by decision makers and/or users of data/information values. Once the role of the data/information items is relatively well known, in particular when information technology is used, then the useful data items are stored in common databases to be shared among many users/clients. One deals here with ***indirect informing*** that takes place when data and information are stored and processed on their way between information sources and users/clients.

In technologically advanced environments, data are usually organized in databases or data warehouses. These may be run by information providers or by client organizations for themselves. In such situations, data/information acquisition, entry, verification, validation, storing, making them available, and converting them into a presentation interpretable format develops into separate specialized business processes. Then also, testing data or information values for *actionable credibility*, called “reliability” by Gleim [3], or “believability” by Wang and Strong [11] is mostly not performed by users/clients, but by specialized personnel during acquisition using specialized techniques on behalf of all authorized sharing users. In routine operations, the dominant problem is the operations quality of *data*, much less than the operations quality of *information* (as defined before). Now the assessment of data or information quality is viewed differently. In contrast to the previous situation when the quality of newly acquired information values had to be assessed, one deals mainly with acquisition of data values, much less of information values. In routine operations, most data have already a well-established role. Now:

1. The requirement of *acquisition interpretability* of data is much less of a problem, considerably reduced, and not as acute as with information.
2. The requirement of significant impact lost its punch for the need for those data/information values was well established before by many users/clients or applications.
3. The requirement of operational timely availability of data stored in shared databases is reduced mainly to the problem of authorized technical accessibility, not the real-life problem whether the necessary information will become available on time.
4. The requirement of actionable credibility of data values in such an environment is much less the concern of individual internal data/information users, but subject to established rigorous procedures that assure their integrity during acquisition, entry, storing, making available, and even their presentation on behalf of the data sharing users.
5. The requirement of task-specific effective operational completeness, in most cases boils down to careful design of a corresponding *subschema* for application processing and predefined inquiries.

All this is valid in the well-established division of labor between those who acquire, store, process, and make available data/information values to various applications, individual decision makers, users, and clients. Data or information acquisitors are separated from data users by complex data and information delivery systems. In such a situation, data/information that was interpretable during acquisition may no longer be interpretable when presented to decision makers or users/clients that differ from data acquisitors. They may be of different mindsets, languages, cultural backgrounds, conventions, etc. Thus, new problem emerge. The problem of presentation interpretability comes into play and a new aspect of actionable credibility arises. Again, one should distinguish here two distinctively different levels: operationally presentation interpretable and economically presentation interpretable.

Operationally presentation interpretable

In most organizations, due to the division of labor, acquisition of data/information and their use are separated, thus ***indirect informing*** takes place. In such situations, users must be presented with data and information that is interpretable and understandable by them. (For instance, it must be legible, in their preferred language, measurement units, conventions, etc.) At least it must be at its minimal level operationally usable however, not necessarily economically or conveniently. Thus, **operational presentation interpretability** of data/information values becomes another direct primary mandatory quality requirement, but not a universal one. It is imposed by the organizational circumstances within which data are used. Presentation of data/information may be done more (but not less) conveniently. This presents a new economical issue, how far one should go when improving convenience of data/information presentation.

Economically presentation interpretable or conveniently presented

Ease and comfort, even enjoyment of the use of data or information in educational services or computer games is related to its form, format, and mode of delivery. It may affect how fast users read, perceive, interpret, comprehend, analyze, absorb, draw conclusions, react, and finally act upon it. Within this category, one considers **clarity, consistency, order, media used, level of summarization, user-preferred type of presentation such as text, graph, diagram, picture, esthetics,** etc. In the case of composite information, these properties are determined not only by its components but also by the way, the components are combined. Some deficiencies in this respect rarely preclude use of the presented data/information values. These indirect quality attributes may increase or decrease the convenience of the data/information use and their procurement cost, hence subsequently their expected cost effectiveness.

When data/information is presented to users/clients via a complex delivery system a new aspect of actionable credibility arises. Besides, of the previously mentioned objectivity, accuracy, precision, currency, another indirect quality requirement of faithful mapping, as precisely defined by Wand and Wang in [10] becomes a mandatory assumption and prerequisite.

Rarely does anything useful come at no cost and at no risk; hence, the real cumulative measure of usefulness should be the **expected cost effectiveness** of the operations assessed from the viewpoint of the purpose they serve. It can be evaluated by either objective criteria, when an adequate model of the decision situation exists or post-*facto*, after deployment of an information system. The ultimate goal of examining data/information values for the above discussed quality requirements is to arrive at a **set of task-specific usable data/information values that are at least effectively operationally complete**. When not only effectiveness but also economy is required one wants such sets to be also effective economically, cost effectively, or even wants them to exceed a required level of expected cost effectiveness. Then of course, all the direct secondary quality requirements, which are of economical nature, are becoming mandatory, as well.

When combined and summarized, all the direct secondary quality attributes determine the economic usability of the data or information concerned, and ultimately, when combined, the cost effectiveness of the delivery system and the entire set of operations. The same aspects, when considered from the viewpoint of information disseminators carry more weight. Even subtle differences in these aspects may decide whether clients respond at all, and/or how they respond. One must also avoid a one-sided assessment of the cost effectiveness of informing that is only from the disseminators or only from the clients' viewpoint. In an ideal solution, it should be cost effective for both sides. This is what makes a business relationship lasting and successful.

DISCUSSION: PREDICTIVE AND EXPLANATORY ASPECTS OF THE FRAMEWORK

"The purpose of science is to develop theories, which can be defined as sets of formulations designed to explain and predict phenomena" [8, p. 3]. For the sake of brevity only a few examples of some of the predictive and explanatory aspects of the proposed theoretical framework and model for operations quality of data or information will be given.

It seems it will be difficult, if possible at all; to find any example or argument to the contrary (indirect proof) that *all identified and also not yet identified attributes of data/information quality* are subject to the following principles

1. The presented universal hierarchical result-oriented taxonomy can accommodate all the operations quality attributes of data and information.
2. The impact a relevant data/information value exercises upon the main effect, which measures the main purpose of operations, lends its importance to all remaining operations quality attributes of the same value, and ranks the importance of that value among all the other values.
3. A failure to meet any one of all the universal direct primary mandatory operations quality requirements with respect to any data/information value renders irrelevant further examination of the remaining quality requirements for the same value.

The inherent relativity (with regard to the purpose and circumstances of operations) of how attributes of operations quality of data/information values are viewed, perceived, and evaluated explains why even highly educated subjects of empirical, in particular, survey-based studies are inherently incapable of coherently and consistently responding to questions about those attributes, if the decision situation of the operations is not defined (as it was the case in [11]), for instance:

- The universal direct primary quality requirements are of the highest importance; hence, always all of them should be ranked 1.
- Effective operational completeness is an extreme example of a universal requirement of the highest importance, and fifth by priority in a logical and economical sequence of examination, was ranked only 10 out of 20 in [11].
- The mandatory requirement of acquisition interpretability of data/information was not even considered. For instance, consider an example when the stakes are high, as the USA is experiencing just now, the anguish of the three Polish mathematicians who cracked the Nazi Germany military code and passed it successfully to the British at the brink of WWII. Those without such experience find it difficult to develop a feeling for the enormous difference between *acquisition of data* that, in most cases, practically are waiting to be collected, and *acquisition of information* of real utility value that changes our thinking about a situation, and/or the results of operations, and or the conduct of the operations, also in business (see also Footnote 4). After reviewing the content of many the current MIS textbooks, one can see that they are unhelpful in this respect [1].
- Similarly, "accessibility" of data is mandatory for any action, is unquestionably of highest importance and third by priority in a logical and economical sequence of consideration was ranked a low 7 by importance by the subjects of the empirical study in [11]. Thus, if rational thinking could be applied, although the questions in the empirical survey made it impossible, accessibility should come before anything else, except for relevancy, hence should be ranked 4 instead of 7, because if not met, it renders irrelevant bothering about any other quality attributes of the same value, etc. Usually in empirical studies, no one expects that the answers given by subjects will be derived by rational thinking. The actual responses, here again, demonstrate that the subjects did not experienced the anguish of extracting important information from the real world of undeclared war when on must race against time to avert disaster, etc because their main experience was based on accessing of carefully prepared data from a database, which is an oversimplification of real-life situations.

What does it mean? With hind sight, of course, one may dare to guess that those subjects actually were answering a different question – "How do you perceive the importance of the problems you were experiencing at your workplace with regard to the following dimensions: a list follows with respective definitions in an appendix" instead of the actually asked question "to rate the importance of each data quality attribute for their data on a scale from 1 to 9, where 1 was extremely important and 9 not important" [11, p 12]. Otherwise, in real life, accessibility and completeness could never end up with importance 7 and 10 respectively. One simply cannot act when pertinent data are not accessible and complete as required by the task unless one gambles.

For a framework to be sound and complete, much more has to be taken into account than the experience gathered from using data taken from shared common databases. There are problems, but these are rather technical problems. The above list of comments could be much longer, but for the sake of brevity, these few examples must suffice.

CONCLUSIONS AND LIMITATIONS

This paper constitutes an attempt at formulating a theoretical framework for examining and assessing the operations quality of data and information with the purpose of arriving at an explanatory theory and results of more lasting validity. It is based on the fundamental principles of human purposes and subsequent actions, which are the subject of *teleological methods* [7, p. 202]. The method of *thought experiments* was applied to generate the outline of the model by imagining the situation and using its features as premises for deductive arguments [7, p. 202] that enables the formulation of the first seemingly universal principles in this domain. The presented theoretical framework is a teleological, purpose-focused, operations research-based approach to a rational inquiry into the universe of all quality attributes of data/information attributes determined relatively by the purpose and circumstances of purposive operations. Those attributes are in many ways logically interdependent. The presented framework:

1. Is based on the rather unquestionable assumption of relativity of all aspects of quality as perceived by humans engaged in purposive operations. This perception changes with the changing purposes and circumstances of the situation-specific use of any data or information for any kind of operations.
2. Is anchored on the following assumptions:
   - The main point of reference is the defined purpose of operations.
   - The main frame of reference is the set of circumstances of operations.
   - The results of operations are a function of data and information used in operations.
   - The main subject perceiving the quality of data and information is the decision maker - the primary user/client.
   - The operations are a "tightly run ship" entirely focused on their main purpose.
   - The existence of a strict pragmatic distinction of "data" and "information."

Under the above assumptions, it was possible to:

- Define a universal hierarchical result-oriented taxonomy of all possible even only potentially possible aspects of data/information quality.
- Assert by mere substitution that the decision maker perceives the lost or unavailable data values exactly the same way as the lack of equivalent information values.
- Identify in operations five universal mandatory situation-independent direct primary information quality requirements: *acquisition interpretability, of significant impact, operationally timely availability, actionable credibility, and task-specific effective operational completeness of sets of usable data/information values.* Any changes to those attributes *qualitatively* change decision situations as defined in decision science.

- Identify separately the *presentation interpretability* of data and information as another direct primary mandatory quality requirement of applicability limited to indirect informing. Nowadays it is the most common case when nearly complete division of labor and separation of responsibilities between data/information acquisitors and data/information users is the established mode of operations; where common databases serve as the main source of data and information values.
- Identify four direct secondary quality requirements of economic nature, which only quantitatively change the operations results, but are becoming mandatory also, if economy of effective operations is of primary concern.
- Identify probably the economic sequence of examination of the universal direct primary quality attributes. The entire plethora of the remaining quality aspects ultimately ended up in the category of indirect attributes of data/information quality. The latter conclusion made obvious the need for developing a map of how the direct primary and direct secondary quality requirements depend on the indirect attributes of data/information quality even more urgent than it was originally perceived.
- Identify several universal principles in using any data and information values such as
    - The principle of the supreme role and precedence of relevance of data and information values (measured by their impact on the results of purposive operations) over all other attributes of DQ/IQ that lends them their relative situation-specific importance among other quality attributes.
    - The principle of defining the necessary situation-specific quality requirements of usability of any data/information value, and
    - The principle of defining the sufficient quality requirements of minimal effective operational completeness of sets of situation-specific usable data/ information values.

The most interesting conclusion, however, emerged only at the very end and become the subject of this request to all interested colleagues in academia for a challenge and discussion. This conclusion was not even considered, the less expected at the onset of this inquiry.

It seems, that

- the universal hierarchical result-oriented taxonomy of all quality attributes of data and information used in purposive operations,
- the universal requirements with regard to the operations quality attributes of data/information,
- their mutual interdependencies, such as being prerequisites, determining or contributing factors to other attributes, and
- the identified universal principles pointed out in the enclosed inquiry.

remain valid not only in information systems and informing science, but literally *in all disciplines dealing with operations conducted by autonomously purposively acting humans, their organizations; systems controlled by artificial intelligence (robots), and any combinations thereof*. Among others, all those disciplines have at least one thing in common that pertains to the presented subject - they study effectiveness of *actions based upon available data and obtainable information*. Without claiming completeness, one may list here: operations research in all its main areas (business, public administration, and military operations), decision science, management science, economics, medical science, political science, sociology, psychology, etc, and of course systems science/cybernetics.

Of course, this first attempt of outlining and presenting such a framework is simplified, because it implicitly assumes a single main purpose with no conflicting requirements and constraints imposed upon decision makers. This fact immediately opens many further research opportunities to expand and refine the framework so that it can accommodate cases that are more complex.

References

[1] Gackowski, Z. J., “Logical Interdependence of Data/Information Quality Dimensions – A Purpose-Focused View on IQ,” *Proceedings of the Ninth International Conference on Information Quality (ICIQ-04), Cambridge, Massachusetts* 2004.

[2] Gackowski, Z. J., "Informing Systems in Business Environment: A Purpose-Focused View," Informing Science Journal, Volume 8, 2005.

[3] Gleim, I. N., *CPA Review Financial.* Gleim Publications Inc, 2004, 12th edition.

[4] Khali, O. E. M., Strong D. M., Kahn, B. K., Pipino, L. L., Teaching Information Quality in Information Systems Undergraduate Education, Informing Science, Volume 2 No 3 1999.

[5] Lee, Y., Strong, D., Kahn, B., Wang, R. "AIMQ: A Methodology for Information Quality Assessment," *Information & Management*, December 2002, Volume 40, Issue 2, pp. 133-146.

[6] Liu, L., Chi, L. N., “Evolutional Data Quality: A Theory-Specific View,” *Proceedings of the Seventh International Conference on Information Quality (ICIQ-02), Cambridge, Massachusetts*.

[7] Mende, J., “The Poverty of Empiricism,” Informing Science Journal, Volume 8, 2005, pp. 180-210.

[8] Mason, E. J. and Bramble, W. J., *Understanding and conducting research,* McGraw-Hill, New York, 1978.

[9] Schopenhouer, A., *The Fourfold Root of the Principle of Sufficient Reason, trans. E.F.J. Payne*. La Salle, Illinois: Open court, 1974.

[10] Wand, Y., and Wang, R. Y. “Anchoring Data Quality Dimensions in Ontological Foundations,” *Communications of the ACM*, 39(11), 1996, pp. 86-95.

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