This is a preprint of an article published in JASIST: (Stvilia, B., Gasser, L., Twidale M., B., Smith L. C. (2007). A framework for information quality Assessment. *JASIST, 58*(12), 1720-1733. http://www3.interscience.wiley.com/cgi-bin/abstract/115805126/ABSTRACT)

A Framework for Information Quality Assessment

Besiki Stvilia[1], Les Gasser[2], Michael B. Twidale[2], Linda C. Smith[2]

[1]College of Information, Florida State University

269 Louis Shores Building, Tallahassee, FL 32306, USA

bstvilia@fsu.edu

[2]Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, 501 E. Daniel Street, Champaign, IL 61820, USA

{gasser, twidale, lcsmith}@uiuc.edu

Abstract

One of the main components in information quality (IQ) assurance is an IQ measurement model design and operationalization. One cannot manage IQ without first being able to measure it meaningfully and establishing a causal connection between the source of IQ change, the IQ problem types, the types of activities affected, and their implications. A better understanding is needed of the roots of IQ change through the development of a systematic, predictive, reusable IQ assessment framework. The framework should enable effective IQ reasoning through the disambiguation of IQ problem sources, and through the rapid and inexpensive development of context-specific IQ measurement models. Here we propose a general IQ assessment framework. In contrast to context-specific IQ assessment models, which usually focus on a few variables determined by local needs, our framework consists of comprehensive typologies of IQ problems, related activities, and a taxonomy of IQ dimensions organized in a systematic way based on sound theories and practices. The framework can be used as a knowledge resource and as a guide for developing IQ measurement models for many different settings. The framework was validated and

refined by developing specific IQ measurement models for two large-scale collections of two large classes of information objects: Simple Dublin Core records and online encyclopedia articles. Some of the results of the collection-specific operationalization of the framework are reported.

1. Introduction

Information is increasingly becoming a critical resource in contemporary societies and organizations. For institutional and individual processes that depend on information, the quality of information (IQ) is one of the key determinants of the quality of their decisions and actions. The familiar "garbage in, garbage out" mantra of computing expresses the problem succinctly. The amount and diversity of information available, and the number and variety of information publishers have grown at an unmanageable rate. Unfortunately, as more information becomes available for use, it becomes increasingly difficult to identify "garbage." Historically, there have been culturally sanctioned mechanisms of IQ assurance, such as the peer review process for research, human screening and cleaning for database entries, and careful editing processes for books and magazines. However, these are breaking down for reasons of scale and cost (McCook, 2006).

One of the main components and cost drivers in IQ assurance is the development and operationalization of an IQ measurement model. One cannot manage IQ without first being able to measure it meaningfully. A review of IQ assessment frameworks proposed previously (Eppler, 2003; Wang & Strong, 1996), however, has shown that most of these frameworks are ad hoc, intuitive, and incomplete and cannot produce robust and systematic measurement models. Little work had been done to identify and describe the roots of IQ problems and link them consistently with information activity types. Although some of the frameworks have captured a particular organization's or group's perception of IQ well, they have not aligned well with the general

structure of IQ variance, which makes them brittle and limits their reuse. A better understanding of the causes of IQ change, and the development of a systematic, predictive, reusable IQ assessment framework are needed. The framework should enable effective IQ reasoning through disambiguation of IQ problem sources, as well as the rapid and inexpensive development of context-specific IQ measurement models.

2. Information Quality Assessment Framework

This paper proposes a general IQ assessment framework. In contrast to context-specific IQ assessment models, which usually focus on a few variables determined by local needs, this framework consists of comprehensive typologies of IQ problems, related activities, and a taxonomy of IQ dimensions organized in a systematic way based on sound theories and practices. The framework can be used as a knowledge resource and guide for developing IQ measurement models for many different settings.

Concepts

We start describing the framework by providing definitions for the basic concepts – data, information and quality. The word "information" has been used in many different ways to refer to different things in different circumstances. The Oxford English Dictionary (OED) defines information as "action of informing" or "an item of information or intelligence." Several attempts have been made by different researchers to define information in a conceptually sound way. Machlup (1983) uses a definition similar to the OED's: "action of telling or the fact of being told something." Bar-Hillel (1964) makes an attempt to provide a formal definition of semantic information using inductive logic, which can be applied to logical propositions or statements. Statements are either atomic, describing the properties of objects, or conjunctive

forms of atomic statements. According to his definition, the information provided by a logical statement increases with the number of propositions excluded by the statement. Thus, the greater the probability of a statement, the smaller its information content: cont(i) = 1 – m(i), where m(i) is the probability of the statement i. Belkin and Robertson (1976) too give a general definition of information in the spirit of the information theory (Cover & Thomas, 1991; Shannon & Weaver, 1949) - "the structure of any text [data] which is capable of changing the image-structure of a recipient" - where structure is defined simply as an order and text as a "collection of signs purposefully structured by a sender with intention of changing the image-structure of a recipient". Brookes' (1974) definition of information is similar to Belkin and Robertson's - "[information] adds to or modifies in any way the knowledge structure with which it reacts." Higgins (1999) continues the information theoretic tradition and defines information as data with "recognizable patterns of meaning" which may reduce uncertainty for a decision maker. Taylor (1986) on the other hand, attempts to draw a clearer distinction among data, information and knowledge. He defines data as a sequence of symbols; information as data plus relations; and knowledge as information selected, analyzed, judged and organized in a stock which then can be used for informing and decision making. Tague-Sutcliffe (1995) based on (Fox, 1983) defines information as follows "I is the information provided to a user Y by S in context C if (1) there exists a record such that (a) Y reads or otherwise perceives S in a context C; (b) Y uses his or her conceptualizing capacity to understand S; (c) I is the conceptual structure that Y understands by S. (2) Y views context C, in general, as a source of true sentences. Buckland (1991) lists three main definitions of information found in the literature: (1) information-as-process; (2) information-as-knowledge and (3) information-as-thing. Information-as-process denotes the process of informing or communicating some knowledge. Information-as-knowledge refers to

what is perceived in "information-as-process". That is, the information being interpreted within some context or some consistent network of other information. Finally, "information-as-thing" is an object, like data or a document that bears some information or is considered to be informative. Under this category Buckland attempts to place all information carriers. Clearly, this also assumes its application in an informing action. Indeed, the "information-as-process" category assumes or subsumes the other two classes of information definitions: Information as Process = Informing (Information Carrier Used, Interpretation of Information). Note that this more formal representation is the same as the OED definition of Information – Action of Informing.

Often data and information are used interchangeably. Redman (1992, 1996) defines data as a view triplet - <e, a, v> - where the value v is selected from the domain of the attribute a for the entity e. This definition is similar to Taylor's definition of information (Taylor, 1986), as it moves away from the traditional definition of data as a "raw" sequence of symbols and contains a reference to a schema – the context in which it needs to be interpreted. However, the definition does include pragmatics – a reference to the receiver's context. In this form, it echoes Bar-Hillel's definition of absolute semantic information (Bar-Hillel, 1964). It is important to note that confusion of terminology is not rare in the literature. As Redman (1992) astutely notes, what is data to one person can be information to another. For our purposes, and based on the definitions above, *we can define data as a raw sequence of symbols; information as data plus the context of its interpretation and/or use, and, finally, knowledge as a stock of information internally consistent and relatively stable for a given community*. Note the context of interpretation in the information definition includes both the underlying entity represented by the data, and its environment, including social, cultural and technological structures. Moreover, the contexts of

data creation and interpretation can be the same or different if the data is created in one context and later interpreted in a different context.

It is natural to expect that the definition of information guides the definition of information quality. One may think of quality in general and information quality in particular as the degree of usefulness or the "fitness for use" (Juran, 1992) in a particular typified task/context, or more than one task/context. On the other hand, the definitions like "meeting or exceeding customer expectations" or "satisfying the needs and preferences of its users" (Evans & Lindsay, 2005; McClave & Benson, 1992) put more emphasis on user needs. Although at first these two dimensions may look different, in reality they are "two sides of the same coin". User information needs are largely shaped by the actions/tasks the user needs to perform and vice versa. Indeed, Taylor's theory of Information Use Environments posits that people's long term information needs are directly linked to their professions, that is, to their professional activities (Taylor, 1991). Redman (1992) attempts to develop a more formal definition of quality: "a product, service, or datum X is of higher quality than product, service, or datum Y if X meets customer needs better than Y." Strong, Lee and Wang (1997) and Wang and Strong (1996) on the other hand use a more general definition of quality –"fitness for use"- proposed earlier in manufacturing by (Juran, 1992). Marschak's (1971) definition of quality is more specific and refers to representation quality – how accurately information represents a particular event. Similarly, Wand and Wang (1996) define data quality as the quality of mapping between a real world state and an information system state. In a more recent work, Eppler (2003) adopts both definitions of quality - meeting the customer expectations and meeting the activity requirements - acknowledging the important duality of quality: subjective (meeting the expectations) vs. objective (meeting the requirements). All these definitions assume that there exist some shared

norms of quality, or quality expectations, and the ways of measuring the extent of meeting those norms and expectations. In addition, it becomes clear that information quality is contextual. For our purposes, however, we will use the general definition of quality – '*fitness for use*' - which we believe encompasses both aspects of quality. That is, it can accommodate quality evaluated as 'fitness' to both individual and communal/societal uses. It is essential to retain both views of IQ as often only one of these views may be available for grounding IQ assessments in.

We define an *IQ dimension* as any component of the IQ concept. Following the definition of Keeney and Raiffa (1976), the *measurability* of an IQ dimension is defined as the ability to assess the variation along a dimension within a reasonable cost. Clearly, to measure the IQ of an information entity along a particular dimension, the dimension needs to be grounded meaningfully in measurable attributes of the entity. *Measuring* is defined as the process of mapping the attribute-level distributions of real-world entities to numbers or symbols in an objective and systematic way. Accordingly, a *measure* is defined as a relation associating the attribute-level distributions of real-world entities or processes with numbers or symbols. We define a *measurement* as a symbol or number characterizing a particular attribute in an objective way. An *IQ problem* is defined as occurring when the IQ of an information entity does not meet the IQ requirement of an activity on one or more IQ dimensions. Finally, we define an *IQ assessment framework* as a multidimensional structure consisting of general concepts, relations, classifications, and methodologies that could serve as a resource and guide for developing context-specific IQ measurement models.

Sources of IQ Problems

The framework identifies the following four major sources of IQ variance: *mapping*, *changes to the information entity*, *changes to the underlying entity or condition*, and *context changes*. To

serve as an effective knowledge resource for the development of an IQ measurement model in different contexts, the framework needs to be well aligned and connected with this IQ variance structure. Note that the sources of IQ variance implicitly suggest both the types of IQ problems and the types of IQ assurance activities (i.e., inappropriate mapping vs. correcting the mapping, and changing the underlying entity vs. realigning the information entity with the underlying entity).

Mapping-related IQ problems arise when there is *incomplete*, *ambiguous*, *inaccurate*, *inconsistent*, or *redundant* mapping between some state, event, or entity and an information entity (Marschak, 1971; Wand & Wang, 1996). Figure 1 shows examples of IQ problems caused by ambiguous and redundant mapping. In the first example, two different date attributes of one entity are mapped into a single date attribute of another entity. There is insufficient information to do an inverse mapping and identify which original attributes the date values refer to. The other example contains two instances of the same element with exactly the same content - a clear redundancy (see Figure 1).

Figure 1: Examples of mapping-related IQ problems

Interestingly, whereas some types of mapping-related IQ problems, such as accuracy, ambiguity, redundancy, and inconsistency, are context independent and can be evaluated objectively, other problems, such as incompleteness, are contextual and may require a knowledge of the activity context to be measured.

IQ assessments are usually made with reference to some social equilibrium regarding what constitutes good IQ in a given activity context. Context itself consists of two main components: *culture* (language, norms) and *sociotechnical structures* (including economic relationships and standards). Any change in context, whether it be *temporal* or *spatial*, can change how IQ is understood and evaluated, and can lead to an IQ problem. An information entity can be of good quality in its original context but can become of lower quality once it is moved to a different context. Likewise, the social and cultural components of context can change in time, causing

changes in the needs, expectations, and economy of quality, which itself can translate into changes in IQ evaluations.

Changes may occur in the information entity itself or in the real-world entity it represents. Changes can be malicious — intended to make the entity unusable or to disrupt a related activity. Alternatively, a change can be made to eliminate or mitigate an IQ problem and to better align the IQ of the entity with the community's general understanding of IQ or the IQ requirements of a particular activity.

Thus, the process of IQ change for an information entity can be *passive* or *indirect*, caused by changes to the underlying entity or context (i.e., culture, sociotechnical structures, and domain knowledge). In general, these changes are not intended to affect the IQ of the information entity. The context can also be changed *actively* to affect the quality of the information entity. New sources can be introduced or planted, or the existing ones can be modified with the intention of reinforcing or contradicting the information presented by the entity (Garfinkel, 1967; Gracy, 2002).

Clearly, there can also be *active* or *direct* quality degradation through malicious changes to the information entity (see Figure 2). Quality degradation actions (reducing an IQ level on a particular dimension) may not necessarily be ill intended, however. Often content administrators must remove or restrict access to information (i.e., reduce its accessibility but protect its integrity), such as when there is a threat of frequent vandalism. Information can also be abridged (reduced in completeness) to meet the needs of a certain audience or for certain uses.

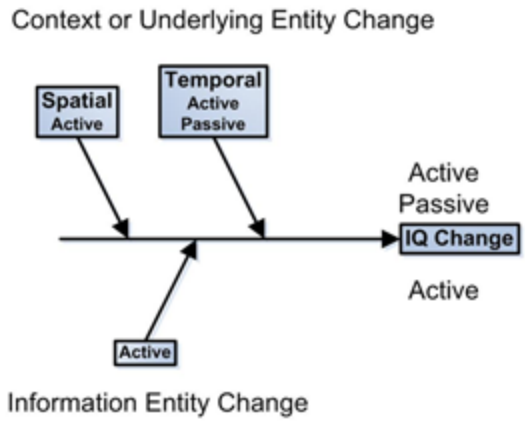

Figure 2: Sources of dynamic IQ problems

Taxonomy of IQ Dimensions

The central part of the framework is a taxonomy of IQ dimensions. Once we identified the sources of IQ variance, we were able to develop a taxonomy of IQ dimensions that would allow us to evaluate the IQ variance caused by these sources in a systematic and meaningful way. The original taxonomy of IQ dimensions in the framework was based on an analysis of 32 representative items from the IQ literature containing a quality assessment model or a framework (Gasser & Stvilia, 2001). The taxonomy consisted of 22 IQ dimensions organized into three categories based on the IQ variance structure discussed in the preceding section (see Figure 3, Table 3):

1. **Intrinsic IQ:** This category includes dimensions of IQ that can be assessed by measuring internal attributes or characteristics of information in relation to some reference standard in a given culture. Examples include spelling mistakes (dictionary), conformance to formatting or representational standards (HTML validation), and information currency (age with respect to a standard index date, e.g., “today”). In general, intrinsic IQ

attributes persist (as long as the reference culture does not change often) and depend little on context. Hence, these can be measured more or less objectively.

2. **Relational or contextual IQ:** This category of IQ dimensions measures *relationships* between information and some aspects of its usage context. One common subclass in this category includes the *representational* quality dimensions. Those dimensions measure how well an information entity reflects (maps) some external condition (e.g., actual accuracy of addresses in an address database) in a given context. Clearly, since external entities and conditions can change independently, relational or contextual characteristics of an information entity are not persistent with the entity itself. The usage context refers to the context of an activity system, which, as discussed in the preceding section, can change in time and space.
3. **Reputational IQ:** This category of IQ dimensions measures the position of an information entity in a cultural or activity structure, often determined by its origin and record of mediation.

Intrinsic IQ measurements reference general *cultural* norms and conventions. Relational measures reference the *immediate context or object of IQ assessment*, and reputational measures reference some culture- or community-related *reputation network- or merit-based order* (see Figure 3).

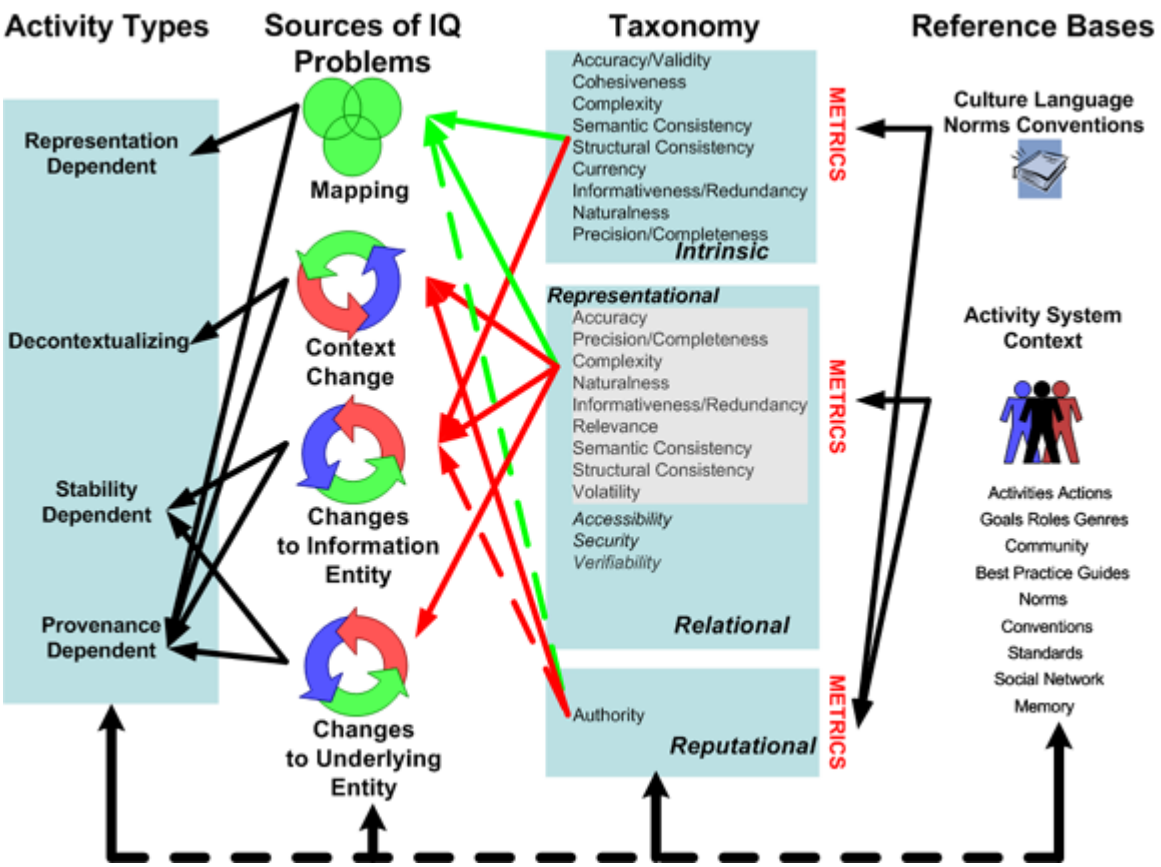

Figure 3: Conceptual model of IQ measurement

Metrics

In addition to a taxonomy of IQ dimensions, the framework contains a set of 41 general metric functions implemented as Java codes, which can be reused to develop context-specific IQ metrics. Most of the metrics in the set were obtained from the IQ literature. Of the 41 metrics in the current version of the metric set, 30 are object or collection attribute based. The other 11 metrics are process metadata based (Gasser & Stvilia, 2001).

Process metadata can provide valuable knowledge about the information entity, which can be used to develop indirect IQ measures. In an ideal form, process metadata can inform us about *who* did *what* to the entity, *when*, *where*, *how*, and *why*. It can be used for a number of purposes. It can help to evaluate the quality of individual contributions or transactions through the

evaluation of information provided by the *what* part and the resultant quality of the entity. The *when*, *where*, *how*, and *why* parts provide contexts for the contribution, and they help to interpret it unambiguously and with less effort. The *how* and *why* parts of process metadata may also contribute to knowledge accumulation, exchange, and learning, including IQ-related knowledge (norms, standards, and best practices). Finally, the *who* part, in combination with the quality evaluations of contributions and resultant product quality, can facilitate the establishment of merit-based orders. It can allow indirect evaluation or predict the quality of an information entity based on the reputation of a contributor. Alternatively, the quality of a contribution can be used to evaluate the intentions or goals of the contributor and, ultimately, to predict or indirectly evaluate the quality of his or her other contributions.

Likewise, the success or failure (*quality*) of the *process* of entity use may allow us to estimate the *quality* of *the information entity*. At the same time, information about the *quality* of the *information entity* can help in estimating the *quality* of the *process* — the likelihood of its failure or success (see Figure 4).

Thus, entity-based measures can be used to estimate both *entity* and *process* quality problems, and process-based measures can be used to estimate or predict the likelihood of *entity* and *process* quality problem incidents. Which metric to use for a given IQ dimension will depend on the *availability*, *cost*, and *precision* of the metric and the *importance* of the dimension itself.

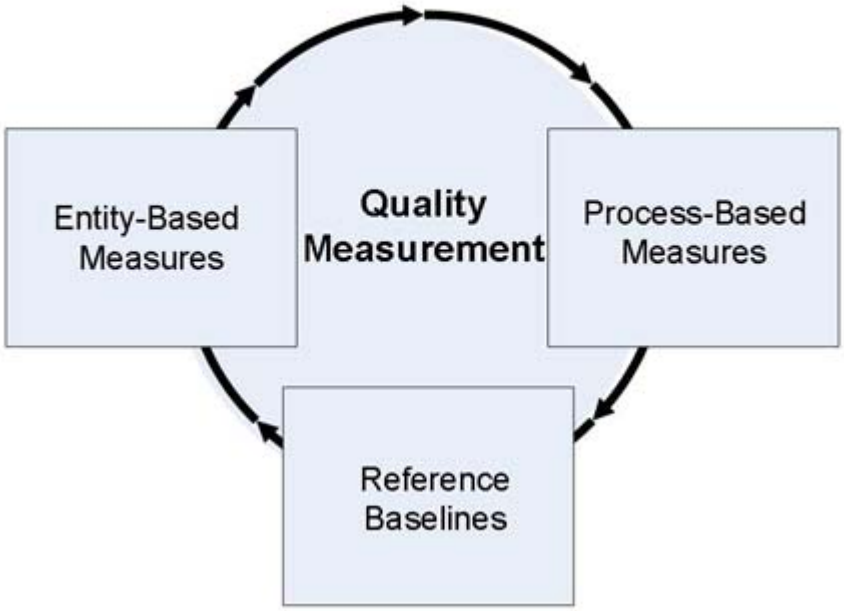

Figure 4: Duality of entity- and process-based measures

Types of Activities Affected by IQ Problems

Understanding the causes of IQ problems helps us to predict which kinds of activities could be prone to frequent problems. Based on the sources of IQ variance we have identified, we can organize information activities into four clusters: (1) *Representation Dependent* — activities that depend on how well one information entity represents another entity or some condition; (2) *Decontextualizing* — activities that use information outside its original context of creation (for instance, an activity may remove information entities from their original contexts and aggregate them into a new collection to support specific information needs or tasks); (3) *Stability Dependent* — activities that depend on how stable the information or its underlying entity is; and (4) *Provenance Dependent* — activities that depend on the quality of metadata of the information's provenance, mediation, and upkeep (see Figure 3).

Information quality in *Decontextualizing* activities is affected mostly by *spatial* changes in the context of use. Consequently, guided by this framework, we may expect IQ problems in the Relational and Reputational dimensions (see Figure 3). Note that we have assumed that measurements are made within the same culture. If that is not the case, then all reference baselines could change, including for the Intrinsic dimensions; consequently, all measurements on all dimensions would need to be reevaluated. *Stability-Dependent* activities can be affected both by changes in the information entity and by temporal changes in its underlying entity. Hence, we might expect that any of the Intrinsic and Representational measurements may change, leading to IQ problems. Likewise, *Representation-Dependent* activities rely on the quality of mappings between the information entity and the real-world condition or entity it represents. A gap between the demanded and actual quality levels on any Intrinsic or Representational IQ dimension can lead to an IQ problem. *Provenance-Dependent* activities, on the other hand, will be affected by any quality deficiencies in the creation and mediation process metadata of the information entity. Since *Provenance-Dependent* activities often require one to trace and examine the whole life cycle of the information entity starting from its origins, one might encounter IQ problems in each temporal snapshot that might be available to the entity. This means that a *Provenance-Dependent* activity could be prone to IQ problems from the taxonomy on any dimension.

Interestingly, although the Reputational and the infrastructure-related Relational dimensions (Accessibility, Security, Verifiability) may not directly evaluate representation-related IQ problems, those dimensions still can be used to predict Intrinsic and Representational IQ levels (see Figure 3). The predictions are usually based on the assumption that highly reputable information entities may exhibit a high IQ on other IQ dimensions as well. Obviously, this

assumption may not always hold, and a highly secure or reputable data store may still contain "garbage."

Clearly, the categories in the activity typology are not independent and there are substantial overlaps among them. The activity categories are intended to focus attention on certain shared IQ-related activity characteristics that are consistently easy to grasp, remember, and associate with a relevant IQ problem structure and measurement objectives. The typology allows an easy and quick decomposition of an IQ measurement task, IQ problem prediction and detection, and identification and activation of relevant IQ measurement variables and relations (see Figure 3). For instance, if the activity system of an information entity (activities involving the entity) includes a *Decontextualizing* activity, then the IQ measurement model of the entity needs to include all Relational and Reputational dimensions. In addition, all measurements on those dimensions may be invalidated and may need to be recalculated every time the evaluation context of the information entity changes.

Context-Dependent IQ Measurement

The IQ assessment framework not only helps us to understand the structure of an IQ assessment task, but also helps us to reason about context-specific variations in IQ assessments and to construct context-specific IQ metrics and baseline representations.

First, the activity system of an information entity needs to be analyzed for activity types using the framework (see Figure 3). Identifying the activity types from the framework leads to identifying the possible dimensions involved, along with general metric functions.

Next, the activity system is decomposed into actions, operations, roles, and tools. We can model and analyze each action and operation analytically. In addition, we can develop and brainstorm different use scenarios (Carroll, 2000) of the information entity and its components,

and use these scenarios to identify key relations among different actions and entity characteristics within the activity. At the same time, we can empirically examine information entities in the domain for variant characteristics and their relations to the types of IQ variance and problems suggested by the framework (see Figure 3). The identified relations, variant characteristics, IQ problem incidents, and general metric functions from the framework are used to design activity-specific IQ metrics and estimate critical values for them.

If the genre of an information entity is known or identified with some certainty, an IQ analyst can use the shared expectations for structure, functionality, and representation associated with the genre to identify important IQ-related variables and the entity attributes or characteristics for grounding IQ metrics. The task model (Find, Identify, Select, and Obtain) proposed by the IFLA Study Group on the Functional Requirements for Bibliographic Records (FRBR, 1998) and a library catalog record are good examples of a typified activity and genre pair. High-level computer languages such as C++ also have the equivalent of genres, called *types*, which are used to enforce certain predictable behaviors of input and output data and objects in software modules.

In addition, when the representative collection of a particular information class for a particular community is available, the statistical analysis of information entity attribute profiles and related process metadata can help to estimate the context-specific critical and target values of the IQ metrics.

IQ Measurement Aggregation

IQ measurement should be done not just for the sake of measurement but also to enable and support effective decision making. IQ measurements taken at different levels, in different

dimensions, and even in different contexts may need to be aggregated and presented in a tractable and actionable way to enable IQ-based selection and reasoning.

IQ measurement aggregation is a complex process, and here we present only a short summary of the steps involved. The process begins with aggregating and reconciling the *concept trees of the IQ measurement models*, *metric functions*, and *measurement representations* with regard to scale, precision, and formatting. The process ends with aggregating the measurements according to the value structure and constraints of the activity system for the information entity's IQ.

We propose two approaches for identifying the IQ value structure and aggregating IQ measurements. The first is an analytic approach that begins with a logical analysis of the information entity, the activity system context, the IQ measurement context(s) (which can be different from the activity system context), and the cultural norms and rankings of the IQ dimensions to determine the IQ requirements for the aggregation. The aggregation IQ requirements are then decomposed into component requirements for each level of the aggregation. The analysis builds on and contemplates different scenarios of information entity use, with different levels and values for its characteristics and related metrics, and then analyzes their potential impacts on the activity outcome value. This process helps to identify IQ value curves and trade-offs for the characteristics and for the measurements based on those characteristics. The IQ value curves are then combined into transfer functions to aggregate the measurements up through the hierarchy (see Figure 5). Note that the IQ *requirements* and the *measurements* flow not only *top down* and *bottom up*, but they may also move *horizontally* (see Figure 5). The aggregate entity may use individual components from different collections, and even from different domains and geographical regions, which may lead to both *spatial* and *temporal* variations in the reference baselines of the measurements. A financial company may

have branches in different countries, in different time zones and the same information or the same classes of information at each branch can be a subject of different legislation and regulations on information collection, storage, transfer, representation, copyright, security and privacy. One may need to modify the security requirements for the artifact if some of the information it aggregates has higher security requirements than the one it currently has. That is, the aggregated information may meet the artifact's security requirements, but the artifact's security requirements may not be high enough to allow the use of that information. A dollar amount in 2006 may not represent the same value as the same amount in 1996. Likewise, an IQ judgment made a year ago may not be understood and valued in the same way today as the community's value structure may have changed.

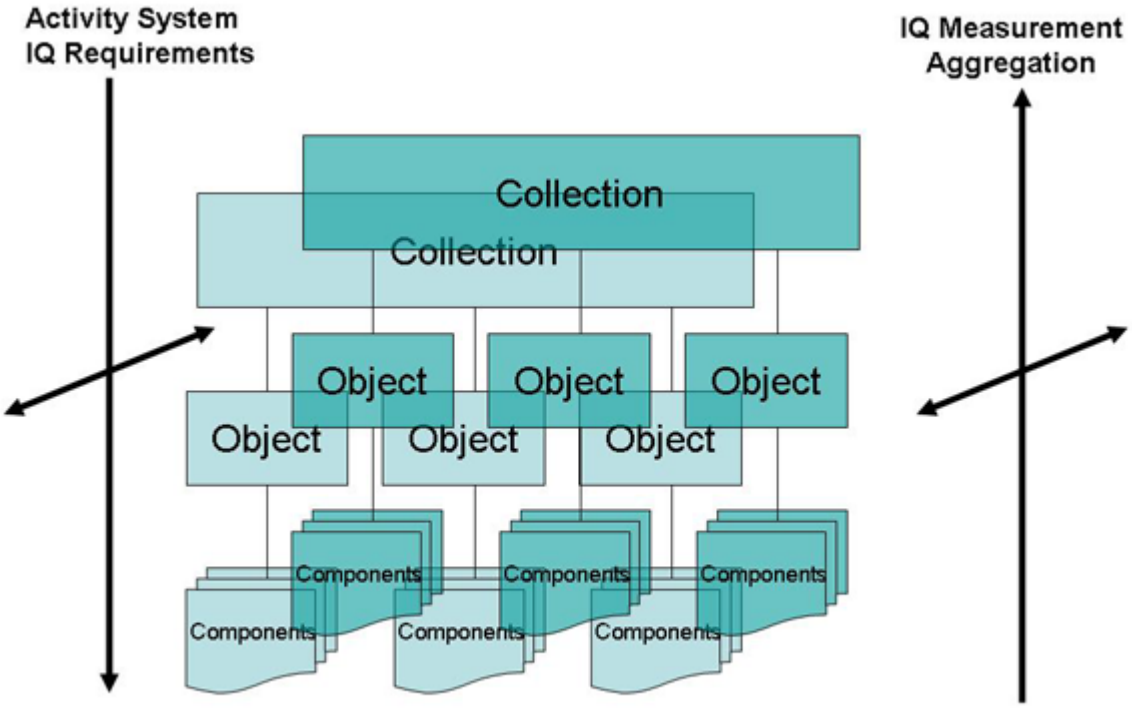

Figure 5: IQ measurement aggregation in an activity system

The second approach, which is much simpler, is an empirical, grounded approach. However, it implies access to end-user evaluations of IQ. The approach includes both qualitative and quantitative analyses of both the information entity and the user or community assessments of IQ embodied in IQ artifacts, organized activities, and actions. Statistical profiles of IQ measurements are generated, and user IQ evaluations and statistical learning techniques are used to extract the user or community value structure for IQ measurements (Stvilia et al., 2005b).

Operationalizing the Framework

Thus, to develop a specific IQ measurement model from the framework, we must begin with an analysis of the activity system of the information entity. We first need to identify the types of activities from the activity typology of the framework. We then use those types to select a set of relevant IQ dimensions from the framework and to identify potential metrics and reference sources (see Figure 3). We need to decompose the information entity and analyze it and its activity system to identify key IQ-related variant characteristics and relations of the entity and activities. The characteristics and relations can then be used to develop context-specific IQ metrics. Finally, we can use either an analytical or an empirical approach to aggregate the IQ measurements into a single IQ index or into multiple indices for each IQ dimension in the model. The decision on which approach to use will depend on the *availability* of the end-user IQ assessment data, the *cost* of analysis, the *precision* of the IQ index needed, and the *criticality* of the IQ measurement information to an individual or an organization.

As with any other model, the performance of the IQ measurement model must be evaluated. *Precision* and *recall* are two previously established measures used to evaluate an information retrieval model (Salton & McGill, 1982). Similarly, the performance of an activity-specific IQ measurement model can be evaluated by two metrics: its ability to identify *all* information

entities from a high-quality class or set (i.e., the set of information entities that meet or exceed the IQ requirements of an activity) (*Complete*), and its ability to identify *only* those information entities that belong to the high-quality set (*Sound*). Obviously, the number of IQ classes in a model can be more than two; a two-class model is simply easier to illustrate and comprehend.

Let us assume that to measure the IQ of information entities in a certain collection, we have developed an IQ measurement model based on the framework. We have identified relevant IQ dimensions and related measures from the framework and grouped them into four independent categories: Accuracy, Completeness, Consistency, and Redundancy. A model like the one shown in Figure 6 could be used to evaluate the performance of the IQ measurement model.

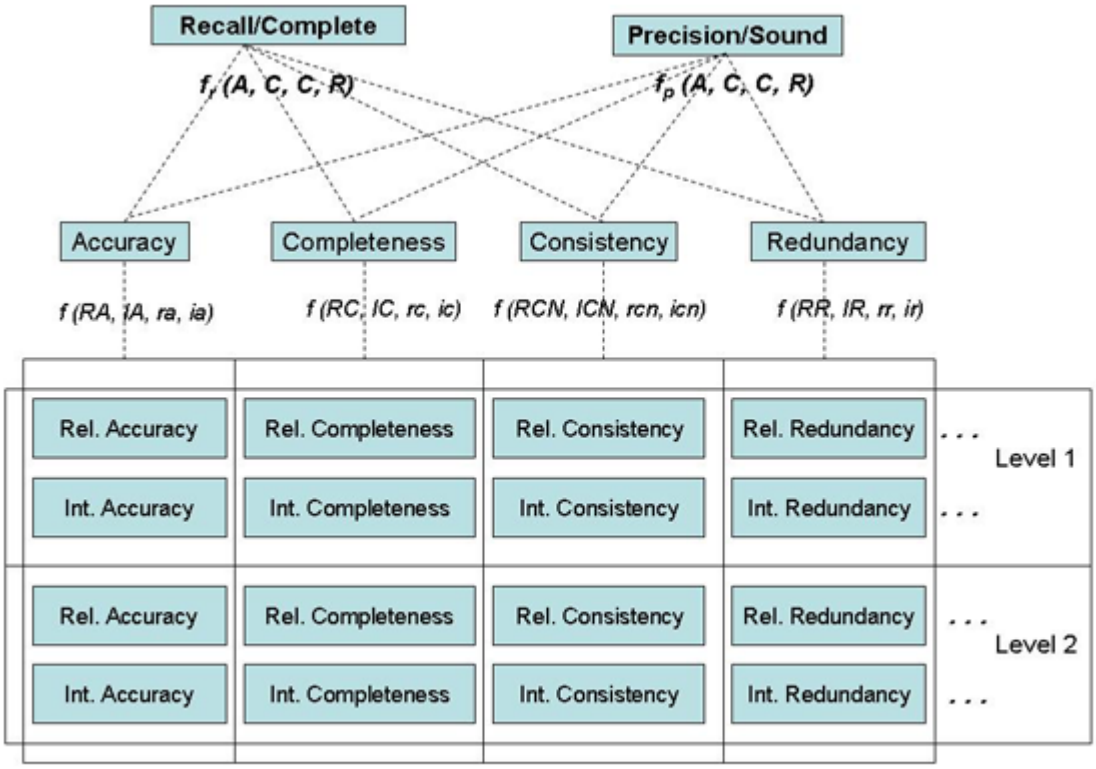

Figure 6: Evaluation of the IQ measurement model (Int. = Intrinsic; Rel. = Relational)

3. Case Study 1

The framework was operationalized for and tested on two large-scale collections of two large classes of information objects. The goal was to challenge the framework on the consistency and completeness of its logic, handling context-sensitivity, and the ease of operationalization. The local IQ measurement models were constructed and metrics were computed.

First, we applied the framework to an aggregated metadata collection of the Colorado Digitization Program of Cultural Heritage (CDP)[1]. The collection aggregated Simple Dublin Core (DC)[2] objects from more than 30 different data providers. The types of participating organizations ranged from small museums, public libraries, and school libraries to large academic libraries and museums. The types of original objects digitized by the local providers were also diverse and covered the gamut of objects, from archeological artifacts and biological specimens to photographs of early settlers (pioneers) and sheet music. The total size of the collection at the time of the analysis was 27,444 records.

The Oxford English Dictionary defines metadata as "a set of data that describes and gives information about other data." A more sophisticated definition of metadata would be "a structured expression of specific knowledge about documents that allows some specific functions." Indeed, Taylor (2004), synthesizing the various definitions of metadata found in the literature defines metadata as "structured information that describes the attributes of information packages for the purposes of identification, discovery, and sometimes management." Thus, in addition to *being an information object* itself, a metadata object *represents* another object or process and may serve as a tool by *providing certain functionalities* in typified activities. Early models for measuring metadata quality and methodologies for identifying metadata quality priorities

[1] http://www.cdpheritage.org/cdp/index.cfm
[2] http://www.dublincore.org/documents/dcmi-terms/

were proposed in (Basch, 1995). Quality assurance problems of online database aggregators were discussed in (Williams, 1990). The recent surge of large-scale digitization initiatives, adoption of campus-wide digital document repositories and archives by large academic institutions led to an increase in the number of studies that looked at metadata quality problems in those projects (Barton et al., 2003; Bruce & Hillman, 2004; Dushay & Hillman, 2003). Although these research studies provide rich empirical account of metadata quality problem incidents, the proposed models are context specific. They lack an integrated approach and may not be readily generalizible and operationalizable. The goal of this IQ assessment framework, on the other hand, is to provide a logically sound and comprehensive knowledge structure that could be reused for developing context specific information quality (including metadata quality) measurement models in a systematic and inexpensive way.

The main purpose of the aggregated collection was to support the search and retrieval of cultural information collected from diverse sources. In addition, the nature of the collection assumed the use of an aggregation activity, as the collection was created through harvesting and aggregating metadata from individual providers. That information-seeking activity belongs to both the *Representation*- and *Stability-Dependent* activity categories, whereas the aggregation activity belongs to the *Decontextualizing* activity category. Hence, one might expect IQ problems on any of the Representational and Intrinsic IQ dimensions (see Figure 3). Following this insight, we selected the first 18 dimensions (Intrinsic and Representational) from the framework (see Figure 3) to develop an IQ measurement model for the collection.

After selecting the IQ dimensions, we needed to examine the metadata objects and analyze them to translate the object level IQ requirements (to be Accurate, Consistent, Complete, etc) into component level requirements at each level of object decomposition. That would allow each IQ measure to be grounded in metadata object characteristics and each IQ dimension to be measured meaningfully. We manually analyzed a random sample of 150 records to look for

incidents of mapping-related IQ problem types suggested by the framework (see Figures 1 and 3, Table 1).

From the content analysis, we identified instances of all five types of mapping problems (see Figure 1, Tables 1 and 2) involving 8 out of the 18 dimensions suggested by the framework. This does not mean, however, that there were no IQ incidents on the other 10 dimensions. We simply did not or could not evaluate those dimensions, either because it was too expensive, because we did not have the required expertise to make such IQ judgments, or because the information we possessed about the activity system context was insufficient to construct context-specific IQ reference baselines. For instance, assessing the intrinsic accuracy of some of the metadata records required specialized subject knowledge or knowledge of the named entities, which in many cases we did not have. In addition, metadata objects were analyzed in isolation from each other. That is, the problem incidents were identified with regard to individual record contexts. The statistics presented in Table 1 do not include the IQ problems that could have been found if we had analyzed metadata encoding across the whole sample (i.e., at the collection level), which was not done because of the high cost. Note that identifying and measuring the mapping problems of Ambiguity, Redundancy, and Inconsistency required only an examination of the information entity and not a knowledge of the measurement context.

Problem clusters	% of sample	Quality problem incidents counted
Ambiguity	56	Contradicting values of the same elements
Inaccuracy	25	Broken links to related objects
Incompleteness	100	Empty elements or element tags; less precision or completeness than expected (such as circas, i.e., "known imprecision") for an element; elements missing from a recommended set of elements
Inconsistency	82	Inconsistent formatting or representation of the same elements
Redundancy	54	Repeated elements containing the same values (duplicates)

Table 1: Metadata quality problems (the numbers represent the percentages of sample records having a particular quality problem)

Dimensions / Problems	Incompleteness	Redundancy	Ambiguity	Inconsistency	Accuracy
1. Intrinsic Completeness	x				
2. Intrinsic Redundancy/Informativeness		x	x		
3. Intrinsic Semantic Consistency			x	x	
4. Intrinsic Structural Consistency				x	
5. Relational Accuracy					x
6. Relational Completeness	x				
7. Relational Semantic Consistency				x	
8. Relational Structural Consistency				x	

Table 2: Quality problem–dimension mapping

Based on a logical analysis of the attributes and structural properties on which the objects varied and their relations to the IQ problem incidents, we developed 13 IQ measures. All these measures (with the exception of the one for Relational Semantic Consistency) could be calculated automatically, which is a clear advantage of using them for large data collections when a manual evaluation of quality is unrealistic. Eight of the 13 measures in Case Study 1 were based on metrics from the framework. Hence, the metric reuse factor for Case Study 1 was: (number of metrics from the list)/(total number of metrics used) = 8/13 ≈ 0.62.

No community criteria or end-user evaluations of quality were available for the collection. Therefore, we could not construct comprehensive baseline or target models with which to compare the IQ of the records. However, the community had developed a best practices guide (the Western States Dublin Core Metadata Best Practices Guide (WSDCMBP)[3]), which included a number of recommendations for the structure and representation of metadata objects. We also had access to Dublin Core Metadata Initiative (DCMI) recommendations for use of the DC element set and for controlled vocabularies. We used the list of required DC elements from the WSDCMBP guide to assess the Relational Completeness of record schemas. The set represented the community's consensus on the schema for the aggregated collection. In addition, based on

[3] http://www.cdpheritage.org/index.cfm

the Library of Congress MARC to DC Crosswalk,[4] we developed a measure assessing record completeness relative to the FRBR activity, which consisted of four actions: *Find*, *Identify*, *Select*, and *Obtain* (IFLA Study Group on FRBR, 1998; see Table 4). The measure evaluated how well a particular record supported each of these actions and the activity as a whole, based on the number of relevant elements for each action present in its schema. The International Organization for Standardization (ISO) standards ISO 639 and ISO 3166 for language and country code representations, guides for media types, and date and time formatting recommended by the DCMI[5] were used to assess the Relational Consistency of the metadata.

Because we did not have access to creation or use process information for the metadata, the IQ measures were mainly based on three sources: an object's *schema*, its *content (term) vector*, and the *values of coded elements* such as <date> and <language>. Consequently, one might expect the same source-based measures to be highly correlated. Furthermore, data-modeling problems at the schema level, such as repeated use of coded elements, could lead to quality problems at the content level. Hence, the measures reflecting those problems could be correlated as well.

To make different IQ measurements comparable and avoid counting the effects of the same variable more than once, we needed to make the IQ metrics as independent as possible (Michell, 1990). We used the technique of exploratory factor analysis with a principal component extraction method and Varimax rotation (Johnson & Wichern, 1998) to explore the dependencies and variance structure of the proposed measures. The factor analysis identified seven groupings of the related measures, which were then used to form IQ metrics for the collections.

---

[4] MARC to DC Crosswalk available from http://www.loc.gov/marc/marc2dc.html
[5] http://www.dublincore.org/documents/dces/

4. Case Study 2

In the second case study we examined the quality of English Wikipedia,[6] an online, open, collaborative encyclopedia (Stvilia et al., 2005a, Stvilia et al., 2005b).

As with any general-purpose encyclopedia, Wikipedia is mainly used to provide a starting research point on a particular topic. Crawford (2001) has identified three categories of questions that can be best answered by using encyclopedias: (1) ready reference questions ("What is it?" and "Where can I find it?"); (2) general background information questions; and (3) preresearch information leading to more targeted and detailed sources. Hence, as in the aggregated metadata collection analyzed earlier, one of the main activities for which Wikipedia could be used is an information-seeking and information-retrieval activity. Along with directing the user to information resources on a particular topic, another important function of a Wikipedia article is to provide an overview of the topic itself. To achieve that successfully, the article needs to aggregate and summarize relevant information from different sources. Hence, aggregation is a major activity in the article's creation process.

Furthermore, Wikipedia is an open system. It allows anyone to create or modify its content. Although this open policy has helped Wikipedia to attract thousands of good contributors, it has also made its content vulnerable to vandalism and fraudulent edits.

Finally, in contrast to the metadata collection, the Wikipedia editorial community is not restricted to one culture but is distributed around the world. All these factors suggest that the content of Wikipedia articles could be affected by all four sources of IQ variance - Improper Mapping, Context Changes, Information Entity Changes, and Underlying Entity Changes - and that we could expect IQ problems on any of the 22 IQ dimensions of the framework.

---

[6] http://en.wikipedia.org

Because Wikipedia articles are not structured, fewer object attributes were accessible in which to ground IQ measures. However, Wikipedia preserves and provides open access to a substantial amount of metadata on the creation and maintenance process of articles in the form of article edit histories, discussions, and vote logs. We used those process metadata to evaluate an article's IQ indirectly but inexpensively. We developed article profiles consisting of descriptive measurements of 19 measures corresponding to 8 IQ dimensions from the framework (Stvilia et al., 2005b). Eleven of the 19 measures were based on article edit history metadata, and 8 measures represented article attributes and surface features (see Table 5). In addition, 9 of the 19 measures were based on general metrics from the framework. Hence, the metric reuse factor for Case Study 2 was 0.47.

The IQ measurement model was required to capture a substantial portion of IQ variance in an article, and at the same time avoid redundancy and bias. This implied that the IQ measurements would be correlated as little as possible. However, edit history-based measures in the profiles were highly correlated. To lose the least possible amount of IQ variance embedded in the article profiles, we decided not to discard the correlated measures but to group them together into less-correlated IQ metric functions. To identify variable groupings, we applied an exploratory factor analysis (Johnson & Wichern, 1998) to the IQ measure profiles of 834 randomly selected articles. The factor analysis suggested seven IQ metrics based on the variable groupings identified by the first seven extracted components. Since the profile measures used different scales and were not normalized, we decided to retain the extracted component score coefficients when defining the IQ metric functions (see Table 6).

In contrast to Case Study 1, Wikipedia provided access to the IQ judgments of the community in the form of Featured Articles (FAs)—articles that had been voted as Wikipedia's best.

Because FAs embodied the community's IQ value structure for high-quality articles, these articles, along with their selection criteria, were valuable resources for designing and validating individual IQ metrics as well as for testing the entire IQ measurement model following the method suggested in Section 2.8.

To evaluate the performance of the measurement model, we computed seven IQ measurements for each article in the pooled set belonging to two sets: the FAs (236 articles) and the Random Articles (RAs) (828 articles; 6 articles with a zero edit history were not used). The profiles were then clustered and classified, and the resultant clusters and classes were compared with the IQ judgments of the community, as embodied in two IQ classes: FAs and non-FAs (RAs).

Comparison of the classes to the clusters produced by density-based clustering (Ester, Kriegel, Sander, & Xu, 1996) showed that only 152 articles (14%) were placed into incorrect clusters. After clustering, the pooled set was classified using a C4.5 decision-tree classifier (Quinlan, 1993). With a 10-fold cross-validation, the precision and recall for the FA set were 90 and 92%, respectively. For the RA set, these numbers rose to 98 and 97%, respectively.

Thus, the IQ measurement model was shown to be successful in discriminating high-quality articles in the Wikipedia collection.

5. Conclusion

In this paper we introduced a general framework for IQ assessment. The framework consists of the typologies of IQ variance; the activities affected; a comprehensive taxonomy of IQ dimensions, along with general metric functions; and methods of framework operationalization. The framework establishes causal connections among the sources of IQ variance attributable to potential IQ problem structures and types of activities, and provides a simple and powerful

predictive mechanism to study IQ problems and reason through them in a systematic and meaningful way.

The framework can serve as a valuable knowledge resource and guide for the rapid and inexpensive development of specific IQ measurement models in many different settings by suggesting relevant IQ dimensions, trade-off relations, relevant general metric functions, and methods of operationalization. The framework has been successfully applied to develop IQ measurement models for two large-scale collections of two large classes of information objects, DC records and encyclopedia articles.

In future research we will continue examining the work organization of IQ in different communities of practice. We will study the cost, dynamics, and intervention structures of IQ for different classes of information entities and the problems of automatic IQ metadata generation and reasoning.

Appendix

	Dimension		Definition
Intrinsic	1.	Accuracy/ Validity	The extent to which information is legitimate or valid according to some stable reference source such as a dictionary or set of domain constraints and norms (soundness)
	2.	Cohesiveness	The extent to which the content of an object is focused on one topic
	3.	Complexity	The extent of cognitive complexity of an information object measured by some index or indices
	4.	Semantic Consistency	The extent of consistency in using the same values (vocabulary control) and elements to convey the same concepts and meanings in an information object. This also includes the extent of semantic consistency among the same or different components of the object
	5.	Structural Consistency	The extent to which similar attributes or elements of an information object are consistently represented using the same structure, format, and precision
	6.	Currency	The age of an information object
	7.	Informativeness /Redundancy	The amount of information contained in an information object. At the content level, it is measured as a ratio of the size of the informative content (measured in word terms that are stemmed and stopped) to the overall size of an information object. At the schema level it is measured as a ratio of the number of unique elements over the total number of elements in the object
	8.	Naturalness	The extent to which the model or schema and content of an information object are expressed by conventional, typified terms and forms according to some general-purpose reference source
	9.	Precision/ Completeness	The granularity or precision of the model or content values of an information object according to some general-purpose IS-A ontology such as WordNet

Relational/ Contextual	10.	Accuracy	The degree to which an information object correctly represents another information object, process, or phenomenon in the context of a particular activity or culture
	11.	Accessibility	Speed and ease of locating and obtaining an information object relative to a particular activity
	12.	Complexity	The degree of cognitive complexity of an information object relative to a particular activity
	13.	Naturalness	The degree to which the model and content of an information object are semantically close to the objects, states, or processes they represent in the context of a particular activity (measured against the activity- or community-specific ontology)
	14.	Informativeness /Redundancy	The extent to which the information is new or informative in the context of a particular activity or community
	15.	Relevance (Aboutness)	The extent to which information is applicable in a given activity
	16.	Precision/ Completeness	The extent to which an information object matches the precision and completeness needed in the context of a given activity
	17.	Security	The extent to which information is protected from harm in the context of a particular activity
	18.	Semantic Consistency	The extent of consistency in using the same values (vocabulary control) and elements required or suggested by some external standards and recommended practice guides to convey the same concepts and meanings in an information object
	19.	Structural Consistency	The extent to which similar attributes or elements of an information object are consistently represented with the same structure, format, and precision required or suggested by some external standards and recommended practice guides
	20.	Verifiability	The extent to which the correctness of information is verifiable or provable in the context of a particular activity
	21.	Volatility	The amount of time the information remains valid in the context of a particular activity
Reputational	22.	Authority	The degree of reputation of an information object in a given community or culture

Table 3: IQ dimensions

	Dimension	Kinds of IQ problems counted	Possible metrics	Cost
1.	Intrinsic Precision/ Completeness	Empty elements or element tags; less precision or completeness than expected for an element	Count of empty tags; count of incomplete values (circas); number of distinct elements	Automatic
2.	Intrinsic Redundancy	Repeated schema elements; repeated element values	Count of instances of repeated schema elements; Information Noise [content = 1 – (size of the term or token vector after stemming and stopping)/(object size before processing)]	Automatic
3.	Intrinsic Semantic Consistency	Contradicting values for the same elements	Count of instances of the same elements having different values	Automatic
4.	Intrinsic Structural Consistency	Inconsistent formatting or representation of the same elements	Count of instances of the same elements using different formatting	Automatic
5.	Relational Accuracy	Broken links to related objects	Counts of broken links	Automatic
6.	Relational Completeness	Missing elements from a recommended set of elements	Number of elements present from the WSDCMBP set of required elements (Title, Creator, Subject, Description, Date, Format, Identifier, Rights); FRBR Support Index for the DC schema, defined as follows: $$FRBR\_RC = \prod_{t=1}^{n} \frac{(e_{Ct} - e_{It})^2 - (e_i - e_{It})^2}{(e_{Ct} - e_{It})^2}$$ where t is the number of tasks in the activity; e_{Ct} is the critical number of relevant distinct elements for task t; e_{It} is the ideal number of relevant distinct elements for task t; and e_t is the number of relevant elements for task t in the record. The largest value ($FRBR\_RC = 1$) is achieved when $e_t = e_{It}$ for all t, and	Automatic

			FRBR_RC = 0 when $e_t = e_{Ct}$. Consequently, the values of *FRBR_RC* take the range of [0,1]. In the case of FRBR activity, the number of tasks or actions, *t*, equals 4: *Find, Identify, Select, Obtain.*	
7.	Relational Semantic Consistency	Elements containing inappropriate values according to a standard	Counts of instances of element misuse	Semiautomatic
8.	Relational Structural Consistency	Elements containing value codes that are not in a standard	Counts of instances of element formatting not matching recommended guidelines	Automatic
9.	Relational Verifiability	Original or related objects that are inaccessible or unrecoverable	(number of identifier + number of source + number of relation)/3	Automatic

Table 4: IQ dimensions and possible measures (Case Study 1)

	Measures	Featured article set (236 articles)	Random article set (834 articles)	Set of misclassified instances from the random set (10 articles)	Dimensions	Source
1.	Number of anonymous user edits	82	2	87	Authority	Edit history
2.	Total number of edits	257	8	251	Authority	Edit history
3.	Number of registered user edits	171	6	149	Authority	Edit history
4.	Number of unique editors	108	5	109	Authority	Edit history
5.	Article length (in number of characters)	24,708	1,344	20,949	Intrinsic Completeness	Article
6.	Currency (time between the dump date and the date of the last update of the article, in days)	3	46	2	Currency	Edit history
7.	Number of internal links	206	17	176	Intrinsic Completeness	Article
8.	Number of reverts	12	0	8	Authority	Edit history
9.	Number of external links	9	0	12	Verifiability	Article
10.	Article median revert time (in minutes)	9	0	17	Volatility	Edit history
11.	Number of internal broken links	6	0	8	Intrinsic Completeness	Article
12.	Article connectivity (number of articles connected to a particular article through common editors)	836	154	826	Authority	Edit history
13.	Number of images	5	0	2	Intrinsic Redundancy/Informativeness	Article
14.	Article age (in days)	1,153	388	1,153	Intrinsic Consistency	Edit history
15.	Diversity (number of unique editors/ total number of edits)	0.4	0.7	0.4	Intrinsic Redundancy/Informativeness	Edit history
16.	Information noise [content = 1 – (size of the term or token vector after stemming and stopping)/(document size before processing)]	0.52	0.32	0.54	Intrinsic Redundancy/Informativeness	Article
17.	Flesch	36	27	36	Intrinsic Complexity	Article
18.	Kincaid	12	13	12	Intrinsic Complexity	Article
19.	Administrator edit share (number of administrator edits/total number of edits)	0	0	0.037	Intrinsic Consistency	Edit history

Table 5: Descriptive statistics of the article profiles (medians of the article IQ profile values)

(Case Study 2)

	IQ metrics	Definitions	Featured	Random
1.	Authority/Reputation	0.2 · number unique editors + 0.2 · total number edits + 0.1 · connectivity + 0.3 · number of reverts + 0.2 · number external links + 0.1 · number registered user edits + 0.2 · number anonymous user edits	198.1	19.8
2.	Completeness	0.4 · number internal broken links + 0.4 · number internal links + 0.2 · article length	5,014.2	275.6
3.	Complexity	Flesch readability score	36	27
4.	Informativeness	0.6 · information noise – 0.6 · diversity + 0.3 · number of images	1.4	0.2
5.	Consistency	0.6 · administrator edit share + 0.5 · age	576.5	194.0
6.	Currency	Currency	3	46
7.	Volatility	Median revert time	9	0

Table 6: Median values of the IQ metrics (Case Study 2)

References

Bar-Hillel, Y. (1964). *Language and information*. London, UK: Addison-Wesley.

Barton, J., Currier, S. & Hey, J. (2003). Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice. In: *Proceedings of the International DCMI Metadata Conference and Workshop*.

Basch, R. (1995). Introduction: An overview of quality and value in information services. In: R. Basch (Ed.), *Electronic information delivery*. (pp. 1-13). Brookfield, VT: Gower.

Belkin, N. & Robertson, S. (1976). Information science and the phenomena of information. *Journal of the American Society for Information Science, 27*(4), 197-204.

Brookes, B. (1974). Robert Fairthorne and the scope of information science. *Journal of Documentation, 30*(2), 139-152.

Bruce, T. & Hillman, D. (2004). The continuum of metadata quality: defining, expressing, exploiting. In: D. Hillman, E. Westbrooks (Ed.), *Metadata in practice*. Chicago, IL: ALA Editions.

Buckland, M. (1991). Information as thing. *Journal of the American Society for Information Science, 42*(5), 351-360.

Carroll, J. (2000). *Making use: Scenario-based design of human–computer interactions*. Cambridge, MA: The MIT Press.

Cover, T. & Thomas, J. (1991). *Elements of information theory*. New York, NY: Wiley.

Crawford, H. (2001). Encyclopedias. In R. Bopp & L. C. Smith (Eds.), *Reference and information services: An introduction* (3rd ed., pp. 433–459). Englewood, CO: Libraries Unlimited.

Dushay, N. & Hillman, D. (2003). Analyzing metadata for effective use and re-use. In: *Proceedings of the 2003 Dublin Core Conference*. Seattle, WA.

Eppler, M. (2003). *Managing information quality: Increasing the value of information in knowledge-intensive products and processes*. Berlin, Germany: Springer-Verlag.

Ester, M., Kriegel, H. P., Sander, J. & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases. In *Proceedings of the Second International Conference on Data Mining KDD-96* (pp. 226–231). Portland, OR.

Evans, J. & Lindsay, W. (2005). *The management and control of quality* (6 ed.). Mason, OH: Thomson.

Fox, C. (1983). *Information and misinformation: An investigation of the notions of information, misinformation, informing and misinforming*. Westport, CT: Greenwood Press.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Upper Saddle River, NJ: Prentice-Hall.

Gasser, L. & Stvilia, B. (2001). *A new framework for information quality (ISRN UIUCLIS-2001/1+AMAS)*. Champaign: University of Illinois at Urbana Champaign.

Gracy, D. (2002). What you get is not what you see: Forgery and the corruption of recordkeeping systems. In R. Cox & D. Wallace (Eds.), *Archives and the public good: Accountability and records in modern society* (pp. 247–264). Westport, CT: Quorum Books.

Higgins, M. (1999). Meta-information, and time: Factors in human decision making. *Journal of the American Society for Information Science*, *50*(2), 132-139.

IFLA Study Group on the Functional Requirements for Bibliographic Records (FRBR) (1998). *Functional requirements for bibliographic records: Final report*.

Johnson, R. & Wichern, D. (1998). *Applied multivariate statistical analysis* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Juran, J. (1992). *Juran on quality by design*. New York: The Free Press.

Keeney, R. & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: John Wiley & Sons.

Machlup, F. (1983). Semantic quirks in studies of information. In: F. Machlup, U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 641-671). New York, NY: John Wiley & Sons.

Marschak, J. (1971). Economics of information systems. *Journal of the American Statistical Association, 66*(333), 192–219.

McClave, J. & Benson, P. (1992). *A first course in business statistics* (5th ed.). New York, NY: Macmillan.

McCook, A. (2006). Is peer review broken? *The Scientist, 20*(2), 26.

Michell, J. (1990). *An introduction to the logic of psychological measurement*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Quinlan, J. (1993). *C4.5: Programs for machine learning*. San Francisco: Morgan Kaufmann.

Redman, T. (1992). *Data quality: management and technology*. New York: Bantam Books.

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

Salton, G. & McGill, M. (1982). *Introduction to modern information retrieval*. New York: McGraw-Hill.

Shannon, C. & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Strong, D., Lee, Y. & Wang, R. (1997). Data quality in context. *Communications of the ACM*, *40*(5), 103-110.

Stvilia, B., Twidale, M. B., Gasser, L. & Smith, L. C. (2005a). Information quality in a community-based encyclopedia. In: *Proceedings of the 2005 International Conference on Knowledge Management*. Charlotte, NC. 101-113.

Stvilia, B., Twidale, M. B., Smith, L. C. & Gasser, L. (2005b). Assessing information quality of a community-based encyclopedia. In: *ICIQ 2005: Proceedings of the International Conference on Information Quality.* Cambridge, MA. 442-454.

Tague-Sutcliffe, J. (1995). *Measuring information: An information services perspective*. San Diego, CA: Academic Press.

Taylor, A. (2004). *The organization of information (library and information science text series)* (2 ed.). Westport, CT: Libraries Unlimited.

Taylor, R. (1986). *Value-added processes in information systems*. Norwood, NJ: Ablex Publishing.

Taylor, R. (1991). Information use environments. In: *Progress in Communication Sciences*. (pp. 217-255)

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

Wang, R. & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5–35.

Williams, M. (1990). Highlights of the online database industry - the quality of information and data. In: *Proceedings of the National Online Meeting*. Medford, New Jersey.