

IQ: Purpose and dimensions

Phyllis Illari

Department of Science and Technology Studies, University College London

Email address: phyllis.illari@ucl.ac.uk

ABSTRACT

In this article I examine the problem of categorising dimensions of information quality (IQ), against the background of a serious engagement with the hypothesis that IQ is purpose-dependent. First, I examine some attempts to offer categories for IQ, and a specific problem that impedes convergence in such categorisations is diagnosed. Based on this new understanding, I suggest a new way of categorising both IQ dimensions and the metrics used in implementation of IQ improvement programmes according to what they are properties of. I conclude the paper by outlining an initial categorisation of some IQ dimensions and metrics in standard use to illustrate the value of the approach.

1 Introduction

Understanding information quality (IQ) is a pressing task. Undertaking it involves two related aspects, one conceptual and the other implementational. This is because what is needed is a settled analysis (or analyses) of IQ that matches definitions of IQ measures and improvement programs as well as ways to implement them. Unfortunately, current literature on IQ offers no settled agreement on answers to at least four closely related questions:

1. What is a good general definition of IQ?
2. How should we classify the multiple dimensions of IQ?
3. What dimensions of IQ are there, and what do key features such as ‘timeliness’, ‘accuracy’ and so on mean?
4. What metrics might one use to measure the dimensions of IQ, bearing in mind that more than one metric may be required to yield an overall measure for a particular dimension?

These questions begin with the most clearly conceptual one, and descend to questions much more closely concerned with implementation. This dual nature of the problem of understanding IQ is recognised in the literature: ‘Both data dimensions and schema dimensions are usually defined in a qualitative way, referring to general properties of data and schemas, and the related definitions do not provide any facility for assigning values to dimensions themselves. Specifically, definitions do not provide quantitative measures, and one or more *metrics* are to be associated with dimensions as separate, distinct properties.’ (Batini & Scannapieco, 2006, p. 19.) Qualitative descriptions of the meanings of words or phrases such as ‘information quality’, or ‘timeliness’ are not the same as formal metrics required to measure them, and which are needed for implementation.

In this paper, I intend to address only the conceptual aspect of the question, not the implementational one. However, this will involve touching upon all four questions, because these four questions ultimately need to be answered collectively. On the one hand, trying to answer the questions sequentially question 1 first,

then moving forward to question 2 and so forth is tempting but unlikely to succeed because, without some understanding of sensible implementable metrics and measures, it seems impossible to give a really meaningful general definition of IQ. On the other hand, it is equally unlikely to be fruitful to try to answer question 4 first, and then attempt to move backward to the others, because designing effective metrics for measuring IQ requires grasping what IQ itself is. Since this set of questions needs to be answered collectively, anyone trying to answer any of these questions is in a way concerned with all four. This might sound paradoxical, but in fact it is simply realistic. The idea is that, just as it takes two to tango, it takes both conceptual understanding and implementation, in alliance, to succeed with regard to IQ. We need to improve our conceptual understanding, then implementation measures, then back to conceptual understanding, and so on, until we get it right.

This illustrates the challenge of understanding IQ: there's no one place to start in assessing, improving or understanding IQ: you can legitimately choose any one of these questions as the place to start. But the ensuing job is messy, because you can't answer any one of these questions adequately in complete isolation from answering all of the others, as an answer to any one of these questions constrains possible answers to all the rest. With this in mind, I shall proceed in this article by developing a conceptual framework for approaching these questions, and then seek to map available metrics on to the developing conceptual picture. In this way, I hope to show that much of the task of answering the question of what IQ is indeed requires conceptual effort, and indicate what can be achieved by mapping implementable metrics to the conceptual framework I develop. In the light of this, I will not attempt in this paper to make a novel study of IQ practice, nor to extend any formal IQ metrics, although those studies must ultimately complement the conceptual study I engage in here. The ultimate test of this conceptual work is forward-looking: it will succeed if it does prove useful in moving forward the overarching project of improving IQ.

Here is a quick outline of the article. In section two, I shall discuss question 1 above, explaining the first major challenge for IQ: being stuck between purpose-dependence and the need to re-purpose data. In section three I shall jump to question 4 above, to explain the second major challenge for IQ: the domain-specificity of successful IQ metrics. There is less to be said about this question, conceptually, as it is the most clearly implementational of the four questions. However, understanding the implementation challenges is important to the work of understanding IQ conceptually. In section four, having examined both end-questions to set up the challenges of IQ, I then move to the middle-ground, looking at the issue of dimensions and their classification, to address questions 2 and 3 above. I shall discuss existing efforts to classify dimensions, and identify a problem that is impeding convergence of these efforts. I shall then offer my own classification, in terms of what IQ is a property of, and give an initial mapping of some IQ dimensions to that classification. It shall become clear that this intermediate theorising is important to IQ. To anticipate, I shall attempt to clear up some of the current confusion, but I shall not attempt to offer a single answer to questions 2 and 3. I will return to this point in section four. In the conclusion, I shall summarise the results obtained. A final terminological note: throughout this article I shall confine myself to considering 'information quality' or 'IQ'. Much of the literature also writes of 'data quality' or 'DQ'. Yet in the following pages nothing theoretically significant depends on the distinction between IQ and DQ because, given the level of abstraction at which I am working, conceptual issues about IQ and DQ do not need to be distinguished.

2 Purpose: the rock-and-a-hard-place of IQ

To begin at question 1, a major conceptual problem in the literature is the *purpose-dependence* of good information. The general idea is simple. For example, information is timely if it gets to you before you need to use it, and that depends on the purpose for which you intend to use it. Information that gets to

you soon after it is gathered is not timely if it is too late to use; while information that gets to you the day before you need it is timely even if that information has been held up by inefficient processing before it reaches you. Indeed, the obvious importance of purpose to IQ has gained so much currency that many working in, or influenced by, the MIT group accept ‘fit for purpose’ as a general definition of IQ. For example: ‘Quality has been defined as fitness for use, or the extent to which a product successfully serves the purposes of consumers ...’ (Kahn, Strong, & Wang, 2002, p. 185). More recently, definitions of quality dimensions in the ISO/IEC 25012:2008 all make reference to a ‘specific context of use’ (ISO, 2008). One important feature, included in a specific context of use, is normal purposes in that context of use.

However, further and deeper analysis of the purpose-dependence of IQ and the effective connection of such analysis to implementation have proven to be serious challenges: ‘While fitness for use captures the essence of quality, it is difficult to measure quality using this broad definition.’ (Kahn et al., 2002, p. 185). In particular, there is a need to understand how to lay out more specific IQ dimensions (questions 2 and 3) and specific metrics for these dimensions (question 4), against the background of a general definition of IQ (question 1) as broad as ‘fit for purpose’. Further, there is a limit to how much information can reasonably be tailored for a particular purpose, as re-purposing good quality information is becoming increasingly important. This is the rock-and-a-hard-place of IQ, which I examine in this section.

2.1 The rock of purpose-dependence

While the MIT group thinks IQ is best generally defined as information that is ‘fit for purpose’, both they and many others still think that at least some dimensions of IQ, and even some aspects of IQ itself, are purpose-independent. These might be called ‘inherent’ or ‘intrinsic’ dimensions of IQ. Consider for example: ‘Inherent information quality is, simply stated, data accuracy. Inherent information quality is the degree to which data accurately reflects the real-world object that the data represents.’ (English, 1999, p. 22.) Even the MIT group, which of course has done an enormous amount to gain recognition for the purpose-relativity of IQ, think that some dimensions are independent of purpose. Describing one of their fourfold classifications of dimensions, which is one of the most widely used, Lee *et al.* write: ‘Intrinsic IQ implies that information has quality in its own right.’ (Lee, Strong, Kahn, & Wang, 2002: 135.)

However, take accuracy. Accuracy for one purpose is not sufficient for accuracy for another purpose. The accuracy required for address data to be usable for a marketing campaign might very well not do if the purpose is more urgent and significant, such as vital security decisions. A reasonable response is to say that purpose changes how accurate information has to be to count as accurate *enough* – and so for the information to be of high enough IQ for the task. But purpose doesn’t change what accuracy itself means. This is understandably tempting, but is not wholly satisfactory for all cases. When gathering data to represent a worldly thing, only some aspects of that thing can be represented. To grasp the problem, consider recording heights of a population. The heights can be recorded to various decimal points, using various kinds of measuring devices. It might be natural to think that the more decimal points height is measured to, the more accurate that measurement is. But a moment’s reflection on measuring the height of a person as 163.467732452524677cm should undermine this. Most of the decimal points are positively a disadvantage for most purposes, if anything impeding the accuracy of the final result. The idea is that accuracy is affected by relevance. It is not merely that accurate *enough* is set by purpose, but that even accuracy itself is infected by relevance of this kind.

Ultimately, the only completely accurate model of the system is the system itself. But the system itself is no good to you – that is why you need to extract information about some aspects of the system, and store it in a database.¹ The aspects recorded are the relevant aspects, and accuracy in this context is determined also by relevance – relevance to the intended purpose. The general problem here is that all dimensions of IQ are infected with relevance – relevance for the purposes intended for the information. This is why I call this ‘the relevance problem’. The best interpretation of all dimensions of IQ is affected by purpose. This is true even though some IQ metrics can be *defined* independently of purpose – such as tuple completeness, which measures whether there are missing values in tuples in the data. Metrics are indicators of the quality of a dimension; they are not the dimension itself. I will return to this point below.

The same view is shared by others: ‘These considerations show that even a dimension such as accuracy, which is considered only from the inherent point of view in the ISO standard, is strongly influenced by the context in which information is perceived/consumed.’ (Batini, Palmonari, & Viscusi, 2012). However, there is no need to conclude from the purpose-relativity of IQ, that IQ is *subjective*. Purpose is a *relational* rather than a relative concept: something has (or fails to have) a purpose for something else. Consider food, for example, it is a relation, but not a relative concept/phenomenon: something as a type (e.g., grass) is food for a specific type of eater (e.g., a cow) but not for another type (e.g., a human). Likewise, IQ does not depend merely on the opinion of the user. The purpose is chosen by the user, but how well different metrics and dimensions fit the same purpose is a matter of objective assessment; the user is constrained by the chosen purpose, and it is the purpose that determines IQ, not the user. What must be concluded instead is that what IQ means, and the best interpretations of the various IQ dimensions, are all dependent on the purpose of the information in question. I shall refer to this as the purpose problem.

2.2 The hard place of re-purposable data

Severe as it is, the purpose problem is only the beginning. There is an important response to what I have called the relevance problem, which deserves careful consideration. Consider the following: ‘Quality is *not* fitness for purpose. The diagnosis code of “broken leg” was “fit for purpose” to pay a claim. But it was *not* fit to analyze risk. Quality is fitness for *all* purposes made of the data, including the *likely* future uses. Quality information will be used in many new ways in the intelligent learning organization. Information fit for one purpose but lacking inherent quality will stunt the intellectual growth of the learning organization.’ (English, 1999, p. 16.)

I call this the ‘multiple purposes response’. It is important because it identifies a crucial worry: if you design a system to give you maximal IQ for one particular purpose, you might very well design it so that the information is too fragile to be turned easily to another purpose. This is a familiar point in design – the more carefully a tool is honed for one purpose, the more limited it becomes in terms of reapplication. Consider trying to eat soup with a fork, or spaghetti with a spoon.

This problem is exacerbated by the fact that good data costs money, and is very valuable. If the government or a company or a research institution is to invest a substantial amount to improve the quality of its information, it is a reasonable requirement that the improved information still be usable at least for some time into the future. In all these organizations, repurposing of data is pretty important. In science,

¹ The only exception to this point is when data itself is a creation of a process, and so the data is all there is. There is no distinction between data about the system and the system itself, which is what generates the problem in other cases. Even so, in most cases, accuracy is infected by relevance in the ways I have argued.

there are various movements afoot to maintain data in a reusable form, particularly data from medical trials, such as that led by the FDA in the US, or Health Level Seven in Europe.

The challenge now is to recognise the need to repurpose data, *without* ignoring the real depth of the purpose-dependence problem. This is where IQ is: stuck between the rock and the hard place. To address this, return to the idea that some metrics used to help assess IQ can be defined independently of the purpose to which the information is to be put. But, recall, these metrics can only be used as indicators of IQ once they are interpreted in the light of that purpose. Nevertheless, this shows the possibility of disentangling indicators that can be defined on your information – or more precisely, defined on your information system – from metrics that measure different aspects of the relation of your information to the purposes for which it is to be used. An overall assessment of IQ will always require metrics of the second type.

This offers a practical solution. There will always be deficiencies of some sort in information that is actually available, but deficiencies can be managed so long as you know what they are. One wishes to avoid being faced with information that looks good, but isn't, or information where one cannot tell whether it is any good. One also wants to avoid information that looks bad, but is good, as one risks throwing away a valuable resource. But ultimately, information that looks bad, and is bad, isn't as big a problem as information that looks good, but isn't.

The metric or measure we get when we succeed is merely an estimate or indicator of IQ: 'Although it is common in the IQ literature to talk of "measuring", "evaluating" or "assessing" the quality of information, in practice the best we can hope for is to compute a close *estimate* of quality. ... At the end of all this, the best we can achieve is to combine the results from the various checks to make a defensible guess at the quality of the data, rather than a definitive, absolute measure of its quality.' (Embury, 2012). The result of making IQ indicators available to the user is to empower the user. This is in broad agreement with the following observation: 'unless systems explicitly track their information quality, consumers of the information they provide cannot make judgments and decisions with high confidence. Information providers don't have to provide perfect IQ, but they need to be explicit about what IQ they do provide.' (Keeton, Mehra, & Wilkes, 2009 p. 28.) This, then, is how IQ improvement or assessment is often done, although the strategy is not always clearly articulated. Clear articulation will help, alongside a clear understanding of the nature of the problem that requires such a strategy to be adopted.

Recognising this tension between the rock and the hard place should help to avoid misunderstanding, particularly the mistake of looking at metrics that have been designed to look purpose-independent, and taking them to be truly purpose independent, in spite of the fact that they have to be allied with purpose-dependent metrics to give an overall indication of IQ itself, and any IQ dimension.

3 Domain specificity

Now I have discussed the first major challenge of IQ, which enters at question 1, the most obviously conceptual question. The integration of the questions is hopefully already very clear: purpose-independent metrics are going to be crucial to help address the purpose-dependence problem. To continue to lay out the major challenges of IQ, I jump to the other end of the list, question 4, the most clearly implementational question. I am not going to make any attempt at implementation, but question 4 is relevant to the conceptual project of understanding IQ, because the conceptual questions can't be answered without understanding the severity of the domain specificity problem for implementation.

The domain specificity problem can be stated fairly simply. Successful metrics to estimate IQ can be defined, but they tend to be very specific to the context for which they are designed. When the ISO standard talks about the importance of a ‘specific context of use’ (ISO, 2008) for IQ, one other thing it means is that successful IQ metrics are designed for a specific domain of application. This takes two forms. First, metrics are designed to cope with the particular structure the data is maintained in. Most metrics are designed for highly structured data, such as that maintained in severely restricted databases. Such metrics do not transfer to data structured in a different way, or to unstructured data, such as data found sprawling on the internet. The second element is that successful metrics are frequently designed with domain knowledge in mind. For example, a metric for estimating how current address data is might use information about how often, on average, people move house in the population of interest. Such a metric would not transfer to other populations, without adjustment.

There is less for a philosopher to say about question 4, as of course much of the work on metrics is highly technical. But there are two points worth noting. First, the problem domain specificity creates for IQ is that it impedes the building up of a common resource for IQ academics and practitioners. It is hard to build a library of well-understood metrics that can be seized on and used in many different situations. As it is, practitioners have to do a great deal of their work designing metrics from scratch. They build up expertise in such design, of course, but not in the form of a library of metrics. Second, this is, like purpose-dependence, a relational problem. Domain specific metrics are dependent on a domain. This problem, however, seems to be dealt with much better by computer scientists. This is perhaps because domain specificity does not appear to create a subjectivity problem. However, the two problems are closer in nature than may appear.

4 Dimensions and their classification

Having laid out the major challenges of IQ, I move now into the middle-ground, to questions 2 and 3, i.e. the theorising between the more conceptual understanding of IQ and its implementation. This mid-ground theorising should, hopefully, more clearly connect the conceptual understanding of IQ and the design of metrics that allow implementation of IQ improvement measures. In particular, answering questions 2 and 3 should enhance understanding of how the metrics used to measure IQ meet the major challenges I have identified. I will follow the tradition current in the computer science literature of working top-down, trying to reach from the more conceptual questions such as question 1, down to the metrics of question 4. However, my most important aim is to work on the *connection* between the conceptual and the implementational. I do not mean to imply that I take question 1 to be in any way privileged. Working bottom-up from question 4, working out what successful metrics might imply about the nature of IQ, would be a perfectly acceptable project, although I do not pursue it here.

I shall now try to show what can be achieved by keeping in mind that the process of improving IQ, including defining it, defining and categorizing its dimensions, and designing metrics to measure those dimensions, involves identifying metrics that can be defined on the data, and combining them with metrics that pay specific attention to purpose, and to the domain of interest.

In this section, I shall look at existing attempts to classify IQ dimensions, diagnose what may be wrong with them, and identify a fruitful approach. I shall then map some existing IQ metrics discussed by Batini and Scannapieco (2006) onto that approach. To anticipate, the main goal of this section is to show how important it is to understanding IQ that we can be precise about what IQ itself and what various IQ dimensions and metrics are actually properties of. For example, are they properties of the data held by a single information producer? Or are they properties of the dynamic relationship between a whole information system, which is changing through time, and long-term users of that system?

The importance of answering such questions is a direct result of the purpose-dependence of IQ, and of the fact that a great deal of work designing and improving IQ involves trying to find a purpose-independent, intrinsic feature of the data itself to measure and use as an indicator of what is in fact a complex purpose-dependent feature of a relationship between data and user. Increased precision on these matters will help us understand how to think in a usefully clearer way about categories, dimensions and metrics. At core, the aim is to allow greater precision and visibility about those features of the data that travel with it, as purposes change during repurposing, and which have to be reassessed. Ultimately I will argue for moving from a hierarchical organization of IQ dimensions and metrics to a relational model linking IQ dimensions and purpose.

4.1 Why existing classifications of IQ dimensions won't converge

An important feature of the literature on IQ is an attempt to classify IQ dimensions, to answer question 2. These attempts are proliferating, and there seems to be little convergence so far in the classifications produced. In this section, I shall examine some of the best known attempts at producing such categorisations of dimensions, and seek to diagnose the problem that is impeding a useful convergence in the debate on this issue.

I begin with the categorisation of Wang (1998), which is one of the earliest and most influential categorisations of IQ dimensions, and is still frequently cited. Table 1: Wang's categorisation (Source: Wang (1998)), below is the table given in the original paper (Wang, 1998, p. 60):

IQ Category	IQ Dimensions
Intrinsic IQ	Accuracy, Objectivity, Believability, Reputation
Accessibility IQ	Access, Security
Contextual IQ	Relevancy, Value-Added, Timeliness, Completeness, Amount of data
Representational IQ	Interpretability, Ease of understanding, Concise representation, Consistent representation

Table 1: Wang's categorisation (Source: Wang (1998)),

There are now quite a few dimension arrangements in this style. Indeed, Lee et al. (2002) even give us two comparison tables of classifications of IQ dimensions, one for academics and one for practitioners, reproduced in Table 2 (Lee et al., 2002, p. 136), laid out according to the Wang (1998) categories:

	Intrinsic IQ	Contextual IQ	Representational IQ	Accessibility IQ
DoD [10]	Accuracy, completeness, consistency, validity	Timeliness	Uniqueness	
MITRE [25]	Same as [39]	Same as [39]	Same as [39]	Same as [39]
IRWE[20]	Accuracy	Timeliness		Reliability (of delivery)
Unitech [23]	Accuracy, consistency, reliability	Completeness, timeliness		Security, privacy
Diamond Technology Partners [24]	Accuracy			Accessibility
HSBC Asset Management [13]	Correctness	Completeness, currency	Consistency	Accessibility
AT&T and Redman [29]	Accuracy, consistency	Completeness, relevance, comprehensiveness, essentialness,	Clarity of definition, precision of domains, naturalness, homogeneity, identifiability, minimum	Obtainability, flexibility, robustness

		attribute granularity, currency/cycle time	unnecessary redundancy, semantic consistency, structural consistency, appropriate representation, interpretability, portability, format precision, format flexibility, ability to represent null values, efficient use of storage, representation consistency	
Vality [8]			Metadata characteristics	

Table 2: Classification for practitioners (Source (Lee et al., 2002))

This is enough to illustrate a lack of convergence that should be cause for concern to those interested in the project of categorising dimensions. The problem is explicitly noted: ‘In comparing these studies two differences are apparent. One is whether the viewpoint of information consumers is considered, which necessarily requires the inclusion of some subjective dimensions. The other is the difficulty in classifying dimensions, for example, completeness, and timeliness. In some cases, such as in the Ballou and Pazer study, the completeness and timeliness dimensions fall into the intrinsic IQ category, whereas in the Wang and Strong study, these dimensions fall into the contextual IQ category. As an intrinsic dimension, completeness is defined in terms of any missing value. As a contextual dimension, completeness is also defined in terms of missing values, but only for those values used or needed by information consumers.’ (Lee et al., 2002, pp. 135-136). Here, they are commenting only on part of the overall comparisons they make, but the concern is clear: there is no settled agreement even on the most deeply embedded dimensions. Now, lack of convergence, of itself, may not be a problem. However, the particular form of lack of convergence currently impedes the building of intermediate theory and so progress in IQ, in ways I shall describe.

The reason for this is that there is a particular source of this problem, holding up any successful mapping of IQ dimensions onto categories. Batini and Scannapieco (2006, p. 39) note: ‘According to the definitions described in the previous section, there is no general agreement either on which set of dimensions defines data quality or on the exact meaning of each dimension. In fact, in the illustrated proposals, dimensions are not defined in a measurable and formal way. Instead, they are defined by means of descriptive sentences in which the semantics are consequently disputable.’ The first important point is the descriptive, qualitative understanding of both categories such as ‘intrinsic’ and ‘contextual’, and dimensions such as ‘timeliness’ and ‘accuracy’, however disputable, are performing a useful role in our conceptualisation of IQ. Categories such as ‘intrinsic’ and ‘representational’ and so on have an intuitive meaning, easy to understand and use, that is helpful to IQ practitioners and academics alike. The concepts of these categories are performing some kind of useful function in the academic literature, and in practice. Similarly for the concepts of IQ dimensions themselves, such as ‘accuracy’, ‘completeness’ and ‘timeliness’. They have intuitively understood meanings that are functioning usefully in the thinking of both practitioners and academics (see Batini and Scannapieco (2006, p. 19))

This is problematic because the IQ dimensions, defined according to the intuitively meaningful words that are generally used for dimensions, do not map onto the IQ categories, defined in turn according to the intuitively meaningful words that are commonly used for categories. I will spell this out in much more detail in the next subsection, by trying to offer a mapping between IQ *metrics* and categories, showing how the dimensions are built up, that will work, which will require adapting both categories and dimensions. Before, let me indicate the problem as briefly as possible. The heart of it is that the current meaningful dimensions have to be *split*, and split into the metrics used as indicators, to map properly onto existing meaningful categories. ‘Accuracy’, ‘timeliness’, ‘completeness’ and so on do not fit onto categories like ‘intrinsic’ and ‘contextual’ – only parts of these dimensions fit into each of these categories.

This is difficult to get clear, and so I shall illustrate the problem here very crudely (see Table 3: Dimensions fall into multiple categories), using the intrinsic-accessibility-contextual-representational categories of Wang (1998), and the well-known dimensions of accuracy and completeness. The core idea is that accuracy has aspects that are intrinsic, but may also have aspects that fall under accessibility, contextual *and* representational features, as does completeness. Accuracy itself is not entirely intrinsic or representational, and so on, but shows aspects of all of the categories. Ultimately, as I have argued, all dimensions are purpose-dependent.

intrinsic	accessibility	contextual	representational
Metrics that measure elements of accuracy, defined only on the data.	Information about such 'intrinsic' metrics, concerning availability to user	Features of some or all of the 'intrinsic' metrics, relevant to the purpose for which information to be used	Features of presentation of the 'intrinsic' metrics and information that allow the user to use it effectively for his/her purpose
Metrics that measure elements of completeness, defined only on the data.	Information about such 'intrinsic' metrics, concerning availability to user	Features of some or all of the 'intrinsic' metrics, relevant to the purpose for which information to be used	Features of the presentation of the 'intrinsic' metrics and information that allow the user to use it effectively for his/her purpose

Table 3: Dimensions fall into multiple categories

I hope the intended point is clear: aspects of *all four columns* in Table 3 feed into an overall measure of the accuracy, and the completeness, of the information, in so far as these are dimensions of IQ itself.

This means that, while useful, this fourfold categorisation does not categorise dimensions themselves, but something else. Dimensions do not map onto these categories, using intuitively understood words that do seem to have a function in the IQ literature and practice 1-1: they do not map in such a way that each dimension can be allocated to one, and only one, category. This is what creates a problem. And although there may be other difficulties, this one by itself is already so significant to be sufficient to explain the lack of convergence in the debate on categories of IQ dimensions. Different scholars, with different intuitions about the most important *aspect* of accuracy or completeness, or different metrics in mind, will naturally allocate these dimensions to different categories.

This at least drives lack of convergence, but note that the problem is more serious than this. There are not multiple competing, but coherent and sensible options for middle-ground theorising, but instead significant muddle. Meaningful terms like 'intrinsic' and 'contextual', which are highly relevant to the severe challenges of IQ that I have identified, cannot be used effectively. This is a severe impediment to developing this kind of badly needed middle-ground theorising.

The search for categories continues despite this problem, because there is a real need for something intervening between dimensions of IQ, and IQ itself, to give structure for thinking about IQ and its dimensions. Those engaged in this project are absolutely right that there is a need for something in this middle ground, given how far apart the two ends are: the conceptual understanding of IQ as fit for purpose but repurposable, and domain-specific task-specific metrics. At the moment, the major sustained attempt has been better to understand dimensions of IQ, question 3, and offer dimension categories, question 2. But such approaches are not likely to succeed, since they all attempt to map each dimension to a single category. The risk is that, in order to fit square pegs in round holes, the relations between the two are made increasingly loose, until fit is achieved only by means of irrecoverable vagueness.

I shall attempt to use the insights developed here to make a positive suggestion to move the debate forward by splitting the dimensions. Initially, this will make both categories and dimensions less intuitively meaningful, but I hope to show how the overall framework ultimately recovers the meaningful

aspects of both category and dimension terms currently in use, while still clearing away some of the current confusion. It is worth noting two purposes here. Initially, my aim is to help the IQ field in computer science – primarily in the academic literature – move forwards in building intermediate theory, by making a suggestion to them for where to work. I hope that this might ultimately also be of use to help in the practice of IQ improvement programmes in terms of both theoretical knowledge and practical tools available as standard, but of course that is a much more distant goal. I shall comment further in section 4.4 on what I take my theoretical contribution to be.

4.2 What is IQ a property of? Towards a classification for IQ dimensions

I shall now try to get more precise about the lesson learned from the discussion above, and begin the task of designing something like a classification of IQ dimensions that can generate settled agreement. I shall argue that what is vital to understanding IQ is the answer to the question what *exactly* IQ itself, its dimensions and its metrics are properties of. Note that what I offer is not a classification in the same spirit as existing ones, but more like a representation of elements worth representing in any particular classification, put together for a particular job. Further, while dimensions are represented, the *basic* elements in the classification are not dimensions, but metrics, for reasons I shall explain.

I first note the complexity of the problem. Batini and Scannapieco (2006) write: ‘definitions do not provide quantitative measures, and one or more metrics are to be associated with dimensions as separate, distinct properties. For each metric, one or more measurement methods are to be provided regarding ... (i) where the measurement is taken, (ii) what data are included, (iii) the measurement device, and (iv) the scale on which results are reported. According to the literature, at times we will distinguish between dimensions and metrics, while other times we will directly provide metrics.’ (Batini & Scannapieco, 2006, p. 19) In order to answer the four questions I began with, and so lay out a framework for consistent settled thinking about IQ, it is not just dimensions that I need to map onto the categories I have in mind: ultimately I also need to lay out the relations between dimensions, their categories, and metrics and measures.

Consider what IQ could be a property of. Naturally, it is a property of information, but what information, exactly? There is a surprisingly large number of candidates:

- Single data item;
- Set of data about a particular worldly item;
- All data about a particular class of worldly items;
- All data in a database;
- Whole information system, even if it accesses multiple databases;
- Single data source;
- Whole information system, even if it accesses multiple databases, some or all of which use multiple sources;
- Whole dynamically evolving information system, so including IQ improvement measures which operate over time;
- Relation between entire (dynamically evolving) information system and a data consumer with a particular purpose (possibly a long-term one) in mind.

This list is probably not exhaustive. It may seem odd to count the later possibilities as possible bearers of IQ. But data is usually a collective. We do not usually worry about the quality of a datum, although we might, of course. However, clearly multiple data, or a collective of information, are legitimate bearers of information quality. As soon as that is noticed, the question of what collective we have in mind when assessing IQ is a natural one, and a question that is important for understanding IQ. It matters for what

we count as, most obviously, completeness, but it also matters for other dimensions. If we think of the collective as the whole functioning information system, then *dynamic* properties of that system, such as correction mechanisms, also become legitimate parts of the bearer of IQ.

Recall what I have indicated as the fundamental problem: that defining, modelling, and implementing good IQ requires transforming purpose-dependent features of a whole information system into, as far as is possible, proxy indicators of IQ. These proxy indicators are, as far as is possible, intrinsic features qualifying only parts of the system itself, rather than properties of the relationship between the system and its context. This means that they are features that can be defined on, and are properties of, the system itself, isolated from the world and from the purposes of any user. Now, a settled classification of standard IQ dimensions and metrics along the lines of what they are properties of would seem likely to help in the enterprise that engages with the fundamental problem.

This idea offers a way of categorising IQ dimensions that might lead to considerably more agreement and so convergence. I also hope to show that it will maintain some of the intuitive notions already in use, such as ‘intrinsic’ and ‘contextual’, which are already functioning usefully in the debate, as these notions will be recoverable from the end result.

4.3 A new classification

The idea of the new classification is to look carefully at the information system, and identify parts of it that are different bearers of properties relevant to IQ, creating a diagram with spaces for each. Then start identifying the elements of the IQ improvement program: IQ itself, dimensions and metrics that you want to map. Then map the elements of the IQ improvement program onto the spaces representing the bearers of the property. Note that the mapping from dimension to category is not 1:1 but 1:N. Note also that there are two *kinds* of things that might be bearers of properties relevant to IQ, and the two must be distinguished:

- 1) Parts of the information system before you:
 - a. in which case the important thing is to get clear on which parts, as there may be several that are useful to distinguish.
- 2) Relations between the information system and something external to it, its ‘context’. This most notably includes:
 - a. the relation (deployment) between the information system and the purpose of the user, and,
 - b. the relation (reference) between the information system and the external world, particularly aspects of the world represented somewhere in your information system.

The difference between these two can no doubt be represented successfully in a myriad of ways. In our example below:

- 1) Properties of parts of the information system itself fall into columns, headed ‘Data, or the data in a particular population’, ‘a particular source of information’ ‘information in the single information system in front of you’, and ‘information across several information systems’ to discriminate different parts of an information system that may well be worth distinguishing.
- 2) Relations between the information itself and the two crucial features of its context are represented by the ‘open’ columns on either side of the columns for the information system:
 - a. The left hand one ‘relation between the information system itself and the world’ allows representation of relations between the proxy indicators that can be defined on the information system, and features of the external world that are *not* the user or the purpose of use.
 - b. The right hand one ‘relation between information system and the purpose of the user’ allows representation of the other relational features of IQ.

I have made an initial mapping of some existing dimensions and metrics into this space, beginning with timeliness and associated metrics. CAPITALISED words represent IQ dimensions, while words in lower case represent metrics or measures. A single row of the table contains metrics and measures that are related to the dimension also contained in that row – specifically, they are used as proxy indicators of the quality of the dimension. But they are metrics defined on the data, so they can also be used as proxy indicators of the dimension – suitably reinterpreted – if the purpose shifts.

This kind of mapping could usefully be done with any kind of element of IQ, including entirely new metrics, which may require more elements of the information system and its context than I illustrate below to be identified as bearers of the properties measured. However, I will illustrate the idea of the mapping rather crudely and briefly using dimensions and metrics discussed by Batini and Scannapieco (2006), and using abstract descriptions of some of the kinds of things that I might want to identify as the bearers of the properties we are interested in when defining and constructing measures for IQ improvement. I begin with the dimension of timeliness in Table 4 below.

What is IQ a property of?					
The relation between information system and world	Data, or the data in a particular population	A particular source of information e.g. database or informant	Information in the single information system in front of you	Information across several information systems	The relation between information system and the purpose of a user
Rapidity of change in the target population	Volatility		Currency	Currency	TIMELINESS

Table 4: Timeliness and associated metrics

The idea is that timeliness is the dimension of IQ, which is relative to the purpose of use as already explained above. Currency is a metric which can be defined on the information itself, using something as simple as an update date, and it can be defined on information in one system or several, so that it falls into multiple columns. Currency does not yield timeliness, though, because whether an update date of two months ago is ‘recent’ depends on the volatility of the data in question – how rapidly the values of the data change. If your information is a house address, then 2 months ago is recent. If your information is levels of glucose within a metabolising cell, it is thoroughly obsolete. Volatility measures change in data, and of course this depends on the rapidity of change in the real-world target population.

With this simpler example in mind, I add other dimensions of usable accuracy and completeness in Table 8 below. The mapping is very far from complete or exhaustive. It is meant merely to illustrate. I suspect that this kind of mapping may be useful in many attempts to improve and better understand IQ, but that different aspects of the information system, on which different more specific metrics may be defined, will be more or less useful to identify in different cases.

What is IQ a property of?					
The relation between information system and world	Data, or the data in a particular population	A particular source of information e.g. database or informant	Information in the single information system in front of you	Information across several information systems	The relation between information system and the purpose of a user

Rapidity of change in the target population	Volatility	Sources may be characterised by usual quality	Currency	Currency	TIMELINESS
Semantic accuracy	Semantic accuracy	Sources may be characterised by usual quality	Syntactic accuracy Comparison functions Edit distance	Syntactic accuracy Comparison functions Edit distance	USABLE ACCURACY
Open World Assumption versus Closed World Assumption	Population completeness	Sources may be characterised by usual quality	Attribute completeness Entity completeness Column completeness	Attribute completeness Entity completeness Column completeness	COMPLETENESS

Table 5: Other dimensions and their associated metrics

As for timeliness, usable accuracy, and completeness with respect to purpose are the true dimensions of IQ, and, as I have argued above, they are dependent on the purpose of the user. Well-known metrics that are used as indicators of these dimensions can be defined on a single information system, and on multiple information systems. Some can be defined on a single attribute, such as attribute completeness. In both cases, again, there is also an important relation to the world. Semantic accuracy concerns whether the information in your system matches worldly values, while choosing between closed or open world assumptions involves making a big assumption – which should be marked – about the relation between the information in the system and the world. Again, useful relations between metrics as indicators of quality dimensions, the purpose of the user, and the nature of the world can be seen laid out in this manner.

The simplified mapping above was achieved conceptually, by examining the definitions and measures to pick out precisely what aspects of the information system they are defined on. Nevertheless, some quite interesting conclusions can be drawn. First, it is worth putting quite a few different elements of the information system into the columns for this mapping, and it is not difficult to think of more things that could usefully be represented. Second, many of the elements of IQ are properties of relations. Even some, such as semantic rules and integrity constraints, which can be defined on the information system itself, are properties of quite complex relationships. They remain properties of the information system itself, because those complex relationships are themselves internal to the information system. But note that semantic rules are often, if not always, constructed *successfully* using world-knowledge, and they will not transfer to data structured differently. Third, as expected, even though the dimensions of IQ themselves are properties of the relation between the whole information system and the user, some elements of all of them, particularly metrics used to measure them, can sensibly be defined just on the information system itself, so allowing such metrics to be properties of that system. This allows them to be used when data is transferred for a different purpose, as indicators that can be used to construct new estimates of the IQ of the data when used for that purpose.

Finally, the domain-specificity of metrics is also made clear. If metrics depend on domain-knowledge, it is worth representing that explicitly, so that it not be forgotten in the case of worldly change – perhaps trying to transfer a metric for currency of address data to a more volatile population.

4.4 Discussion of the classification

The idea has been to move from a hierarchical organization of IQ dimensions and metrics to a relational model linking IQ dimensions and purpose. To this end, the previous mapping offers several advantages, including the possibility of convergence of a single classification of IQ metrics and dimensions, or multiple non-competing classifications, classifications sensitive to what IQ improvement programs are really trying to do, a clear indication of potential pitfalls, and finally a valuable recovery of important concepts like ‘intrinsic’ and ‘contextual’. I shall briefly comment on each of them in turn.

First, convergence should be encouraged by this mapping, because it should be possible to map metrics and dimensions onto this kind of space, and useful in sharpening up their definition, and their interrelations. Deciding what such things are properties of – what they can be defined on – is a matter of considerably more objective assessment and should be much easier to agree on than whether entire IQ dimensions are, for example, ‘intrinsic’. This mapping also completely avoids the muddle at the heart of current attempts to map dimensions themselves onto categories.

Second, this kind of mapping lays out the tools of IQ improvement in a way that is sensitive to what IQ improvement programmes try to do. It lays out the relationship between metrics that are genuinely objective measures of the data itself, domain-specific metrics, and highly purpose-dependent features of the whole system. The place of such metrics as mere indicators of the relational IQ dimensions is clear. The tables give a representation of the scale of the challenge of IQ, and what is being done to meet it.

Third, as a complement to the table laying out useful features of tools, it also represents the gaps. These mappings visually represent where the enterprise of finding intrinsic features of the information to act as proxy indicators of properties of relational features is forced, where the metric or dimension is a property of a relation. The forced nature of proxy indicators of the quality of the information for the purposes of the user will not be blurred or easily forgotten with such maps in mind.

Finally, this mapping allows the recovery of some important intuitive terms in the literature, but in more precise form. I suggest that intrinsic IQ metrics are those that can be defined solely on the information system itself, such as some specific completeness metrics. These are properties of the information stored, and our mapping still has the advantage of encouraging continuous attention to exactly what feature of the information stored they are properties of. Note, though, that it tends to be only metrics, and only some of them, which are intrinsic in this sense. And in so far as such metrics relate to IQ, they are always proxy indicators of a more complex relational property. Contextual features of IQ are those which attempt to measure something about the relationship between the information system and its context. I have now identified the two crucial features of that context: a) the relation between the information system and the purpose of the user, b) the relation between the information system and the world, including of course features of the world explicitly represented, such as birth dates, but also features of the world used to construct appropriate semantic rules for checking consistency. Ideas of ‘representational’ and ‘accessibility’ relations are less easy to define precisely. But I suggest they are thought of explicitly as themselves features of the relationship between the information and the user, which is an idea that requires future work. Further, here it is particular characteristics of the users that are relevant, such as the language they speak, and what technical skills and theoretical understanding they have, rather than merely their purpose.

Ultimately, our mapping has many advantages, and recovers the intuitive usability of terms that are performing a useful role in both the literature and practice.

5 Conclusion

I have briefly summarised my reasons for thinking that the purpose problem for IQ is serious, and that much of the work on IQ responds by looking for proxy indicators of IQ that can be defined on features of the information system itself. I have offered my approach to mapping elements of all major concepts engineered for IQ improvement onto a space designed to represent what they are properties of. This is my first attempt to address the four interrelated questions with which I began:

1. What is a good general definition of IQ?
2. How should we classify the multiple dimensions of IQ?
3. What dimensions of IQ are there, and what do key features such as ‘timeliness’, ‘accuracy’ and so on mean?
4. What metrics might one use to measure the dimensions of IQ, bearing in mind that more than one metric may be required to yield an overall measure for a particular dimension?

My mapping offers a way of seeing the problems laid out collectively, showing how much in common they have. Fitness for purpose is vital to IQ, and should inform understanding of the purpose of a classification, and also identification of dimensions and the design of metrics. It is due to the difficulty of addressing the fitness for purpose problem that metrics are used, as they are, as proxy indicators of purpose-dependent dimensions. This research can continue by examining further metrics and adding to the mapping above, and expanding understanding of how they are designed to meet the purpose problem.

I finish by commenting on the challenges I began by identifying. They are indeed serious. But properties of relations are not in themselves intractable. Relational properties internal to the information system itself are frequently defined very well, such as integrity constraints. The purpose problem is just that the bearer of some features of IQ is the relation between system and purpose of user. But there is nothing here that can’t be measured in principle. The relation might be imperfectly measured, perhaps, but no more imperfectly than some relational features internal to the information system itself are measured. If the purpose requires speed more than accuracy, this trade-off can be assessed, proxy measures found and implemented. If the purpose requires completeness, this too can be assessed, measures created and implemented, then tested and adjusted, and so on. From another point of view, we could track user choices, given stated purpose, and learn how to improve measures of the relation between the system and purpose that way. This is not very different from the domain-specificity of many metrics, which require the relation between the domain and the information system to remain unaltered.

To summarise, there are two major challenges of IQ. The first is that IQ itself is purpose-dependent, while we need to be able to repurpose data. The second is the domain-specificity of successful metrics. To succeed in IQ improvement and assessment, one side of the problem is just that we have to relate the information system to the world. This is probably going to mean that some measures will remain ineliminably domain-specific. The other side is that we have to relate the information system to the purpose of the user. So some measures will remain ineliminably purpose-specific. These two are both ineliminably contextual – but tractable – features of IQ.

Acknowledgements

Research for this article was supported by a two-year project, entitled “[Understanding Information Quality Standards and their Challenges](#)”, currently funded (2011-2013) by the British Arts and Humanities Research Council (AHRC). I owe thanks to Suzanne Embury and Carlo Batini for giving me very useful comments on earlier versions of this article, and to Luciano Floridi both for specific comments on this piece, and for the opportunity to collaborate with him on the wider project. Remaining errors are, of course, my own.

References

- Batini, C., Palmonari, M., & Viscusi, G. (2012). *The Many Faces of Information and their Impact on Information Quality* Paper presented at the AISB/IACAP World Congress, University of Birmingham.
http://philosophyofinformation.net/IQ/AHRC_Information_Quality_Project/Proceedings.html
- Batini, C., & Scannapieco, M. (2006). *Data Quality: Concepts, Methodologies and Techniques*: Springer.
- Embury, S. (2012). *Forget Dimensions. Define Your Information Quality Using Quality View Patterns*. Paper presented at the AISB/IACAP World Congress, University of Birmingham.
http://philosophyofinformation.net/IQ/AHRC_Information_Quality_Project/Proceedings.html
- English, L. (1999). *Improving data warehouse and business information quality*. New York: John Wiley & Sons Inc.
- ISO. (2008). *IEC FDIS Software Engineering - Software Product Quality Requirements and Evaluation - Data Quality Model* (Vol. 25012).
- Kahn, B. K., Strong, D. M., & Wang, R. Y. (2002). Information Quality Benchmarks: Product and Service Performance. *Communications of the ACM*, 45(4), 184-192.
- Keeton, K., Mehra, P., & Wilkes, J. (2009). Do you Know your IQ? A Research Agenda for Information Quality in Systems *ACM SIGMETRICS Performance Evaluation Review*, 37(3), 26-31.
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A Methodology for Information Quality Assessment. *Information & Management*, 40(2), 133-146. doi: 10.1016/s0378-7206(02)00043-5
- Wang, R. Y. (1998). A Product Perspective on Total Data Quality Management. [Article]. *Communications of the ACM*, 41(2), 58-65. doi: 10.1145/269012.269022