Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2012 Proceedings | Proceedings

# Data and Information Quality Dimensions in Engineering Construction Projects

Soffi Westin
*Information Systems, University of Agder, Kristiansand, Norway.*, soffi.westin@uia.no

Olav Birger Holen
*Private enterprise, Kristiansand, Norway.*, ol-birg@online.no

Tero Päivärinta
*Computer and Systems Science, Luleå University of Technology, Luleå, Sweden.*, tero.paivarinta@ltu.se

# Data and Information Quality Dimensions in Engineering Construction Projects

**Soffi Westin**
University of Agder, Norway
soffi.westin@uia.no

**Olav Birger Holen**
Private enterprise, Norway
ol-birg@online.no

**Tero Päivärinta**
Luleå University of Technology, Sweden
tero.paivarinta@ltu.se

## ABSTRACT

Poor data and information quality (DQ/IQ) causes delays and cost overruns in engineering construction projects. However, only little DQ/IQ research has been performed in this context. This paper explores quality dimensions in the context of engineering construction projects. The most important dimensions identified by Ge, et al., (2011) is used as a basis and compared with dimensions used in 12 large engineering construction projects in one organization. The findings show that six of these dimensions are in use in those projects: accessibility, security, relevancy, completeness, consistency, and timeliness. In addition, the findings indicate another dimension also very important in this context; logical coherence. The logical coherence dimension compares different data values and determines if there is any illogicality between them. Three dimensions are monitored by rules provided by a DQ/IQ tool, and we discuss about the contributions which such a tool can provide for an engineering construction firm.

### Keywords

Data quality, information quality, quality dimensions, engineering, construction, logical coherence

## INTRODUCTION

Large engineering construction projects include a variety of project types such as groundwater (Frimpong et al., 2003), high-rise (Faridi and El-Sayegh, 2006; Long et al., 2004), buildings (Assaf and Al-Hejji, 2006) and offshore development projects in the oil and gas industry (Dyrhaug, 2002). The characteristics for large engineering construction projects are their need for several different kinds of expertise, they cost millions of USD and the construction sites are usually far away from where the design is carried out. The teams carrying out these projects are of a transient nature and deliver their products through temporary project-based organizations which exist only for the single project (ibid). The main output from the engineering design phase (for example the design of an oil rig) to the assembly phase (for example the assembly of the oil rig) is drawings. The information on these drawings is mainly extracted from various data sources, such as engineering databases.

Delays and cost overruns are common phenomena in construction projects (Long, Ogunlana, Quang and Lam, 2004; Sambasivan and Soon, 2007; Toor and Ogunlana, 2008) and several researchers have reported delays as the most costly problem (Faridi and El-Sayegh, 2006; Ling, Pham and Hoang, 2009). Recent research identifies lack of information related to drawings as one of the most significant factor for delays and cost overruns (Dai, Goodrum and Maloney, 2009; Westin and Päivärinta, 2011). This implies poor data and information quality (DQ/IQ) in the data sources and systems in the construction engineering field.

DQ/IQ has become important for businesses also in general. The amount of business related data and the impact of poor DQ/IQ information are growing continuously (Ramaswamy, 2006; Strong, Lee and Wang, 1997). Assessment of DQ/IQ thus becomes equally important. DQ/IQ can be assessed from the viewpoint of many quality dimensions. A quality dimension includes a definition and a set of measures for DQ/IQ evaluation. Among the most frequently mentioned are accuracy, completeness, consistency and timeliness (Carlo Batini, Cappiello, Francalanci and Maurino, 2009). In addition, a recent review identified that .the most frequently mentioned dimensions are accessibility, security, relevancy, ease of understanding, reliability, ease of manipulation, and objectivity (Ge, et al., 2011).

Even though research on DQ/IQ assessment is performed in various contexts, little has been performed in the context of engineering construction. There are a few exceptions; among them are studies of DQ/IQ assessment in large engineering projects (ref. e.g. Blechinger, Lauterwald and Lenz, 2010; Westin and Päivärinta, 2011) and assessment of current DQ/IQ tools and their relevance for this context (ref. e.g. Neely, Lin, Gao and Koronios, 2006). In the light of the serious impact of poor DQ/IQ and the early stages of DQ/IQ research in this context more research is needed. Since quality dimensions are a fundamental concept in DQ/IQ research it could be useful to increase knowledge on the need and use of such dimensions in engineering construction projects. The research question of our study hence is:

*Which quality dimensions are needed in engineering construction projects and why?*

To answer this question we present a case study of 12 large engineering projects in one organization. Based on the most important DQ/IQ dimensions as identified by Ge, et al., (2011) we compare existing quality dimensions with dimensions used in these projects. The findings confirm the importance and use of several existing quality dimensions. In addition the findings demonstrate the use of a dimension not mentioned in previous research.

The remaining sections of this paper present related research, the research approach and case description, the findings, the discussion of the findings, and finally the conclusion.

## RELATED RESEARCH

DQ/IQ has been defined broadly through its “fitness for use” (Wang and Strong, 1996), and the definition is adopted in this research. Whereas early research focused on query techniques on multiple data sources and data warehouses in the end of 1980s, the research field has later on spread to a number of new application areas, such as customer resource management, knowledge management, supply chain management and enterprise resources planning (Madnick, Wang, Lee and Zhu, 2009), including contexts such as data warehousing (e.g. Blake and Mangiameli, 2011; Wixom and Todd, 2005), health care registers (e.g. Pipino and Lee, 2007; Vician, 2011), retailing and internet related systems (e.g. de Corbiere, 2009; Helfert and Hossain, 2010). Recently also engineering construction has become a field of interest for DQ/IQ researchers (e.g. Tribelsky and Sacks, 2011; Westin and Päivärinta, 2011).

An important aspect of DQ/IQ assessment is the use of “quality dimensions”. Quality dimensions make it easier to define and discuss issues related to DQ/IQ without referring to specific data values. Rather it is groups of data values with the same qualities (e.g. accurate values; consistent values, etc). Batini et al. (2009) identify altogether 28 quality dimensions in their review of data quality methodologies. Many of the dimensions have had varying definitions and measures for data quality evaluation. However, four core dimensions; accuracy, completeness, consistency and timeliness have been emphasized frequently throughout the methodologies. Other potential quality dimensions in previous research have been currency, volatility, uniqueness, appropriate amount of data, accessibility, credibility, interpretability, usability, derivation integrity, conciseness, maintainability, applicability, convenience, speed, comprehensiveness, clarity, traceability, security, correctness, objectivity, relevancy, reputation, ease of operation, and interactivity (ibid.). In their study Ge, et al., (2011) have assessed 17 dimensions commonly used and tailored them to a set of 9 dimensions they perceive as the most important dimensions (see Table 1).

A dimension of special interest in our case is coherence or logical coherence. Previous research have discussed on a few definitions on this dimensions for example in terms of consistency in dates: “like you cannot die in the future, you cannot be born in the future” (Piprani and Ernst, 2008, p. 5); coherence in statistical data: “……includes coherence between different data items pertaining to the same point in time, coherence between the same data items for different points in time, and international coherence” (Brackstone, 1999, p. 16); or more general: “coherence implies that two or more values do not conflict with each other” (Singh, Park, Lee and Rao, 2009, p. 6). Similar examples are presented in other studies (e.g. Bahorich and Farmer, 1995; C. Batini and Scannapieco, 2006; Berti-Équille, 2007; Ochoa and Duval, 2009; Peralta, 2008; Stadler and Kolbe, 2007; Vassiliadis, Bouzeghoub and Quix, 2000). In the field of engineering construction one study addresses this issue as a type of inconsistency using examples of various value mismatches: between adjacent items (e.g. mismatch between a pipe diameter and its belonging valve diameter); between attributes of one item having conditional functional dependencies (e.g. the type of valve constrains the allowed nominal diameter range); between items and corresponding catalogs (e.g. the valve from a specific supplier has to satisfy modeling ranges of the according catalog), and, between items in the corresponding environment (e.g. a cooling system should have cooling item types) (Blechinger, et al., 2010).

In this study we use logical coherence dimension in addition to the most important dimensions as identified by Ge, et al., (2011) as a basis for our discussion.

| Quality Dimensions | Attributes of Items |
|---|---|
| Accessibility | accessible, obtainable, retrievable, available |
| Security | secure, protected, authorized access |
| Relevancy | useful, relevant, applicable, helpful |
| Accuracy | correct, accurate, free of error, precise |
| Completeness | sufficient, complete, comprehensive, include all necessary values, detailed |
| Consistency | consistent meaning, consistent structure, presented in the same format |
| Timeliness | current, up to date, delivered on time, timely |
| Ease of understanding | easy to understand, easy to comprehend, easy to identify the key point |
| Reliability | reliable, dependable |
| Ease of Manipulation | easy to manipulate, easy to aggregate, easy to combine |
| Objectivity | impartial, unbiased, objective, based on facts |

**Table 1. Quality Dimensions and Attributes of Items (Ge, et al., 2011)**

In Table 1 we have listed these dimensions together with "Attributes of Items". "Attributes of Items" are measuring items that are suitable for measuring the related dimension (ibid).

## RESEARCH APPROACH AND CASE DESCRIPTION

We have chosen the case study approach for our research. The reasons for that are multiple. First, DQ/IQ research in this context is rather new and still very much under-investigated. In such new and little explored fields case studies are especially appropriate (Eisenhardt, 1989). Second, to be able to better understand DQ/IQ in a specific context, a case study is suitable (phenomenon examined in a natural setting is a key characteristic of case studies (Benbasat, Goldstein and Mead, 1987). Third, since access to suitable organizations can be hard to achieve (Walsham, 2006), we want to exploit the opportunity of access we have been provided by our target organization.

Our target organization is a European, multi-discipline engineering construction company with capabilities related to global management, design, procurement, completion and generally execution of complex installations for the oil and gas industry. The organization delivers engineering design for construction projects and possesses a significant share of global markets in its product and project domains. The most employees are involved in engineering and construction projects. The biggest projects may typically cost more than 100 million Euros and take up to three years of calendar time.

Each project involves several engineering disciplines. Due to tight schedules engineering tasks are performed in parallel (concurrent engineering) even if the tasks are interdependent. In an environment like this, it is possible that some data values are omitted, or preliminary values used, when engineering item data is inserted for the first time in a record. The intention from the engineers is to insert the missing or correct values as soon as they are known. Meanwhile, thousands of records are inserted and it gets difficult or even impossible to manually identify incorrect values or missing values. An in-house developed DQ/IQ assessment tool is hence used by the projects to help identify any DQ/IQ insufficiencies. The tool is rule-based which means that rules are developed for assessment of data and information in various data sources.

The result is presented to the engineers in a report. Amongst other information the report contains a descriptive name of each rule and the number of related errors. Project participants can click the number of errors reported on a specific rule, and the report will display all information needed to correct the error (e.g. ItemID, which values are missing or incorrect etc). By identifying and reacting to observed DQ/IQ errors earlier the projects hope to increase DQ/IQ in extracted drawings which previously have been identified as insufficient (Westin and Päivärinta, 2011). Increasing the level of DQ/IQ in drawings should then lead to a reduction in delays and cost overruns.

The rules are defined and implemented based on internal organizational requirements, customer requirements and suggestions from the engineers. We have investigated the rules used by 12 projects within the organization. Based on Ge, et al.'s (2011) quality dimensions and measuring items, every rule have been investigated to determine its type of measuring item and hence the related dimension. A total of 246 rules have been examined. The number of rules per project depends on the size of the project and the customer requirements. Amongst the 12 investigated projects the number of rules varied from 49 to 138.

## FINDINGS

When comparing the dimensions used by the projects with dimensions identified by Ge, et al. (2011) we found that not all of Ge, et al.'s dimensions or measuring items were in use. In addition we found a dimension in use not mentioned by Ge, et al. In Table 2 column one we first list the dimensions common for the projects and Ge, et al., and then we have added the dimension only identified in the projects. Column two contains the attributes of items (measuring items) the projects used to assess the belonging dimension. The third column contains the number of rules used to assess each dimension, or an 'X' to indicate that the dimension is handled by other means (e.g. organizational policies).

| Quality Dimensions | Attributes of Items | Number of rules applied |
|---|---|---|
| Accessibility | available | X |
| Security | secure, protected, authorized access | X |
| Relevancy | relevant | X |
| Completeness | include all required values | 131 |
| Consistency | consistent meaning | 43 |
| Timeliness | current, up to date, delivered on time, timely | X |
| Logical Coherence | two or more values do not conflict with each other | 72 |

**Table 2. Quality Dimensions and Attributes of Items in use, measured by rules (indicated by a number) or handled by other means (indicated by an X)**

In the next section we explain how the different dimensions are understood and used by the projects. Further we discuss whether this is in line with Ge, et al. (2011) and how it possibly differs.

## DISCUSSION

*Accessibility* is in the projects understood as the availability of needed information. This understanding is in line with Ge, et al., (2011), since "available" is one of their identified attributes of items. However, availability is not covered by any used rules, rather unavailable information result in errors identified by other rules. For example: imagine an engineer waiting for information related to an instrument or a cable or any type of part s/he is responsible for. The engineer is about to insert the record for this part into the database, but will not be able to fill in values for all the required fields since s/he is still waiting for that information. The record will hence be inserted with some missing values in some fields. If this record is checked by a rule concerning completeness, an error will be reported due to the missing values in some fields. From previous research we know that lack of information in drawings extracted from the databases is a common phenomenon in this context (Dai, et al., 2009; Westin and Päivärinta, 2011). Available information in time could be one of the problems.

*Security* is only regarded through the fact that the used DQ/IQ assessment tool provides reports only to the individual project and only project participants have access to the reports. This is in line with the understanding of this dimension as presented in Lee, Strong, Kahn and Wang, (2002, p. 135), where they state that "the system must be accessible but secure". (Ge, et al. also refers to this study for the definition of the security dimension). This dimension is hence not covered by any rule implemented in the DQ/IQ control system; rather it is covered by organizational policies.

*Relevancy* is understood as relevant information. "Relevant" is also one of Ge, et al.'s identified measuring item for this dimension. Whether the information is perceived as relevant or not is based on customer and internal requirements. An assessment of these requirements is performed prior to selecting rules for use in each project. The total set of chosen rules will hence be considered assessing all relevant data in the project. This means that there are no rules *checking* for relevancy; rather data relevance is determined based on the prior assessment of requirements and the following choice of rules.

*Accuracy* is in the projects understood as a combination of completeness and consistency. This understanding is also pointed out by Ge, et al. (section 3.1); "users presume that inaccurate information consists of incomplete and inconsistent information". The rules identified for the dimensions completeness and consistency are hence covering the accuracy dimension.

*Completeness* is understood as all required values should be inserted in the records. Again, this is in line with one of the identified measuring items for this dimension. The number of rules assigned to this dimension could indicate that this dimension is perceived as very important which is no surprise; without values important information is missing. Completeness can also be viewed in conjunction with the accessibility dimension. If the engineers have to skip some values from the record while the values are not yet available, the completeness dimension and its measuring items will provide an overview of missing values. If available information in time is a common problem in our context, the rules for assessing completeness are important to make sure that every value is there by the time the drawings are extracted.

*Consistency* is in the engineering projects understood as whether data values representing the same issue are identical across different tables or different data sources. This is also a common understanding of the consistency dimension (Pipino, Lee and Wang, 2002). This means that the measuring item "consistent meaning" can be used because identical values would imply consistent meaning. The number of rules assigned to this dimension (43) illustrates the need for this type of quality assessment also in this context.

*Timeliness* is understood as whether the data is delivered on time to the customer or to internal receivers. "Delivered on time" is also one of the measuring items for this dimension (Ge, et al., 2011). In the projects this dimension is covered by delivery dates set for each record. The delivery dates are used as a grouping indicator when reporting DQ/IQ status. This means that for each delivery date the users are able to see if there are any errors reported by the rules on items that belong to a certain date. The delivery date grouping makes it easier to prioritize which errors to attend to at the moment.

Despite of their popularity in other domains (Ge, et al., 2011), *Ease of understanding, Reliability, Ease of manipulation* and *Objectivity* are dimensions that seem to be not very relevant for these projects. Maybe the reason lays in the nature of engineering construction; data values are considered understandable, objective and easy to manipulate (in the sense of inserting and editing records). However, there are some exceptions where the data is not reliable. For example when the information is not available and therefore missing or presented with a preliminary value. Then the rules for other dimensions such as completeness, consistency or logical coherence will intercept these instances and report them as DQ/IQ errors.

*Logical coherence* is understood as if two or more values do not conflict with each other. This is in line with Singh et al.'s (2009) definition of coherence and similar to Blechinger et al.'s (2010) understanding of a type of inconsistency. Even if the terms logical coherence and inconsistency in this sense have been discussed to a certain degree, our case seems to highlight the importance of this dimension more than previous research.

The rules used for measuring this dimension compare different values of different fields. The assessment is important to avoid delays and extra costs. Let us discuss about an example:

The main assembly site (e.g. a yard) of an oil rig is usually far away from the location of design. Smaller parts could be put together at another site before shipment to the Yard. A rule of thumb in the projects is that all *main* assets should be assembled on the Yard. In this case the rules will assess the values in the site code field and the value of main/not-main asset field. If the asset is main asset, the site code should be Yard meaning that the data values would be logically coherent between themselves. If not, if the site code is incorrect, it could in practice result in assets being shipped to wrong locations. Shipping assets is costly and takes time. To return them and re-ship them will be even more costly. By attending the possible problem by using DQ/IQ assessment rules while the assets are still at their location of origin, these extra costs are avoided.

We compared the logical coherence dimension to Ge et al.'s (2011) dimensions and measuring items to see if it was possible to use any of them. The most similar dimensions were accuracy and consistency. If the measuring items for the accuracy dimension were to be used, they would probably not indicate any error. For example, since there is no real world equivalent to compare the values (yet), the site code value would return as correct as long as it is one of the legal site codes in the project. The main/not main asset field would always return as correct since the only two possible values to insert is Y/N, and they are both legal. Hence, the two inserted values would *each* be accurate, free of error and precise. Only by assessing the combined result; i.e. logical coherence, an error could be revealed.

In some sources, the consistency dimension has included a dimension named as "(logical) coherence". However, the consistency dimension can be viewed from a number of perspectives (Pipino, et al., 2002). Common perspectives are representational consistency (data is consistently presented in the same format (Wang and Strong, 1996)), semantic consistency (consistency of the same data values across tables (Pipino, et al., 2002)), and structural consistency (consistency in the representation of similar attribute values (Levitin and Redman, 1995)). From other perspectives it is possible to imagine that logical relations could be perceived as consistency between different values. For example: if a buyer was to choose a car model from an online system, and the car model only was produced in the color red, it should be impossible to choose the color green from the color field. The system would probably prevent any possibility of choosing any color but red as soon as the buyer had chosen a model. This is a common situation for several types of systems (you choose a flight date and then the available flight times, or you choose an author and then the available titles). But these situations are different because they are commonly dealt with by the time of insertion of data. This is not possible in engineering construction context since the data is inserted in several different ways; manually, imported from other data sources, and copied and inserted from other data sources. Also, due to a probability of not yet having received the predefined values of for example site codes, the engineers still have to be able to insert partial information without validation towards such predefinitions. Hence, the quality assessment has often to be performed *after* insertion, and some of the constraints materialize after initial data has been already put into the system. This causes the need for periodical control of DQ/IQ in engineering projects using the approach of concurrent engineering.

From previous literature we already know that lack of information related to drawings is one of the most significant factors for delays and cost overruns in engineering construction projects (Dai, et al., 2009). Knowing that this information is extracted from various data sources implies poor DQ/IQ in these data sources. So how can our findings contribute to this field? We see that the target organization, in their effort to increase the level of DQ/IQ in the projects, mainly lean on three DQ/IQ dimensions for monitoring. These dimensions are completeness, consistency and logical coherence. The result of their effort could be measured by assessment with regard to the drawings in those projects actively using the DQ/IQ tool versus drawings extracted by those projects not using the tool. We have not been able to identify research in this area pointing at these three DQ/IQ dimensions and providing suggestions of how to determine whether DQ/IQ efforts impact the extracted drawings. Whereas this paper is proposing such issues and actions, further research of ours will investigate the impact of using the implemented DQ/IQ tool, embodying the data quality rules, over time. We also hope to be able to show significant decrease in delays and cost overruns in the case organization as an ultimate result of our development.

## CONCLUSION

This paper aimed at identifying quality dimensions in DQ/IQ assessment performed in the context of engineering construction projects. For the assessment and discussion of each dimension we used the most important dimensions as identified by Ge et al., (2011) in addition to a dimension we termed “logical coherence”. 246 DQ/IQ assessment rules, implemented in a DQ/IQ assessment tool and used by 12 projects in one organization were investigated to determine which quality dimensions were in use. A total of seven quality dimensions were identified; accessibility, security, relevancy, completeness, consistency, timeliness, and, logical coherence. However, only three dimensions; completeness, consistency, and, logical coherence were directly handled by these rules. Accessibility was indirectly handled by rules from other dimensions (completeness and consistency) since these rules will reveal errors such as missing or incorrect values due to non-available information. Security was handled by organizational policy. Relevancy was handled by assessment of internal requirements and customer requirements prior to defining/selecting rules for the individual project. Timeliness was handled by delivery dates and these dates are used as a grouping indicator when reporting DQ/IQ status.

Logical coherence implies that two or more values do not conflict with each other. This dimension seems to appear as a very important dimension in this context, despite the fact that it is rarely discussed in previous DQ/IQ research. The number of rules applied to this dimension (72 of 246) indicates the importance of this dimension and that several sets of data values are in need for such assessment.

So, three dimensions, completeness, consistency and logical coherence, were identified as the most relevant dimensions for assessment in our target organization. If errors reported are corrected in the data sources, it is believed that extracted drawings will increase in level of quality. To determine if this is correct these drawings could be compared with drawings produced by comparable projects that do not use the DQ/IQ tool. If the level of DQ/IQ increases it should be possible to identify impact on delays and cost overruns. These three dimensions is also the most important dimensions to assess by a DQ/IQ tool, since the other dimensions in use are either used for determination of which rules to use, are handled by organizational policy, or is indirectly handled by dimensions already assessed by rules.

We believe our findings contribute to research on DQ/IQ in the field of engineering construction by highlighting the importance of the three dimensions, the potential of controlling these dimensions better in the projects, and the suggestion of how to determine whether the use of the DQ/IQ tool resulted in the wanted impact on the drawings. Even if this research has been performed within one organization, we believe our findings can be used as guidance for similar organizations in their effort to increase their level of DQ/IQ. We will continue our research in the target organization aiming at presenting future suggestions for DQ/IQ assessment concepts needed in the context of engineering construction, and hopefully be able to identify impact on delays and cost overruns.

## REFERENCES


1. Bahorich, M. and Farmer, S. (1995) The coherence cube, *The leading edge*, 14, 10, 1053-1058.
2. Batini, C., Cappiello, C., Francalanci, C. and Maurino, A. (2009) Methodologies for data quality assessment and improvement, *ACM Computing Surveys*, 41, 3, 1-52.
3. Batini, C. and Scannapieco, M. (2006). Data-Centric Systems and Applications: Springer
4. Benbasat, I., Goldstein, D. K. and Mead, M. (1987) The Case Research Strategy in Studies of Information Systems, *MIS Quarterly*, 11, 3, 369-386.
5. Berti-Équille, L. (2007) Data quality awareness: a case study for cost optimal association rule mining, *Knowledge and information Systems*, 11, 2, 191-215.
6. Blake, R. and Mangiameli, P. (2011) The Effects and Interactions of Data Quality and Problem Complexity on Classification, *Journal of Data and Information Quality*, 2, 2, 1-28.
7. Blechinger, J., Lauterwald, F. and Lenz, R. (2010). Supporting the Production of High-Quality Data in Concurrent Plant Engineering Using a MetaDataRepository. *Proceedings of the Americas Conference on Information Systems*, August 12 -15, Lima, Peru.
8. Brackstone, G. (1999) Managing data quality in a statistical agency, *Survey methodology*, 25, 2, 139-150.
9. Dai, J., Goodrum, P. M. and Maloney, W. F. (2009) Construction Craft Workers' Perceptions of the Factors Affecting Their Productivity, *Journal of Construction Engineering & Management*, 135, 3, 217-226.
10. de Corbiere, F. (2009). Data Quality and Interorganizational Information Systems: The Role of Electronic Catalogues. *Proceedings of the Americas Conference on Information Systems*, August 6 - 9, San Francisco, California.
11. Eisenhardt, K. M. (1989) Building Theories from Case Study Research, *Academy of Management Review*, 14, 4, 532-550.
12. Faridi, A. S. and El-Sayegh, S. M. (2006) Significant factors causing delay in the UAE construction industry, *Construction Management & Economics*, 24, 11, 1167-1176.
13. Ge, M., Helfert, M. and Jannach, D. (2011). Information quality assessment: Validating measurement dimensions and processes. *Proceedings of the European Conference on Information Systems*, June 9 - 11, Helsinki, Finland.
14. Helfert, M. and Hossain, F. M. Z. (2010). An approach to monitoring data quality - product oriented approach. *Proceedings of the Americas Conference on Information Systems*, August 12 -15, Lima, Peru.
15. Lee, Y. W., Strong, D. M., Kahn, B. K. and Wang, R. Y. (2002) AIMQ: a methodology for information quality assessment, *Information & Management*, 40, 2, 133-146.
16. Levitin, A. and Redman, T. (1995) Quality dimensions of a conceptual view, *Information Processing & Management*, 31, 1, 81-88.
17. Ling, F. Y. Y., Pham, V. M. C. and Hoang, T. P. (2009) Strengths , Weaknesses , Opportunities , and Threats for Architectural , Engineering , and Construction Firms : Case Study of Vietnam, *Journal of Construction Engineering and Management*, 135, 10, 1105-1113.
18. Long, N. D., Ogunlana, S., Quang, T. and Lam, K. C. (2004) Large construction projects in developing countries: a case study from Vietnam, *International Journal of Project Management*, 22, 7, 553-561.
19. Madnick, S. E., Wang, R. Y., Lee, Y. W. and Zhu, H. (2009) Overview and Framework for Data and Information Quality Research, *Journal of Data and Information Quality*, 1, 1, 1-22.
20. Neely, M. P., Lin, S., Gao, J. and Koronios, A. (2006). The Deficiencies of current data quality tools in the realm of engineering asset management. *Proceedings of the Americas Conference on Information Systems*, August 4 - 6, Acapulco, México.
21. Ochoa, X. and Duval, E. (2009) Automatic evaluation of metadata quality in digital repositories, *International Journal on Digital Libraries*, 10, 2, 67-91.
22. Peralta, V. (2008). *Data quality evaluation in data integration systems.* Université de Versailles (chair) Raúl RUGGIA Professor, Universidad de la República, Uruguay.
23. Pipino, L. L. and Lee, Y. (2007). Applying IT to Healthcare: Humans, Errors, and a Data Quality Perspective. *Proceedings of the Americas Conference on Information Systems*, August 10 - 12, Keystone, Colorado.
24. Pipino, L. L., Lee, Y. W. and Wang, R. Y. (2002) Data Quality Assessment, *Communications of the ACM*, 45, 4ve, 211-218.
25. Piprani, B. and Ernst, D. (2008) A Model for Data Quality Assessment, *Lecture Notes in Computer Science*, 5333, 750-759.
26. Ramaswamy, M. (2006) On the phenomenon of information dilution, *Issues in Information Systems*, VII, 2, 289-292.
27. Sambasivan, M. and Soon, Y. W. (2007) Causes and effects of delays in Malaysian construction industry, *International Journal of Project Management*, 25, 5, 517-526.

28. Singh, P., Park, I., Lee, J. K. and Rao, H. R. (2009) Information sharing: a study of information attributes and their relative significance during catastrophic events, *Cyber-Security and Global Information Assurance: Threat Analysis and Response Solutions, IGI Publishers*.
29. Stadler, A. and Kolbe, T. H. (2007). Spatio-semantic coherence in the integration of 3D city models. *Proceedings of the International ISPRS Symposium on Spatial Data Quality*, June 13 - 15, Enschede, The Netherlands.
30. Strong, D. M., Lee, Y. W. and Wang, R. Y. (1997) Data quality in context, *Communications of the ACM*, 40, 5, 103-110.
31. Toor, S.-U.-R. and Ogunlana, S. O. (2008) Problems causing delays in major construction projects in Thailand, *Construction Management & Economics*, 26, 4, 395-408.
32. Tribelsky, E. and Sacks, R. (2011) An empirical study of information flows in multidisciplinary civil engineering design teams using lean measures, *Architectural Engineering and Design Management*, 7, 2, 85-101.
33. Vassiliadis, P., Bouzeghoub, M. and Quix, C. (2000) Towards quality-oriented data warehouse usage and evolution, *Information Systems*, 25, 2, 89-115.
34. Vician, C. (2011). A Resource Review for Health Care Quality Execution: Business Intelligence Strategy. *Proceedings of the Americas Conference on Information Systems*, August 5 - 7, Detroit, Michigan.
35. Walsham, G. (2006) Doing interpretive research, *European Journal of Information Systems*, 15, 320-330.
36. Wang, R. W. and Strong, D. M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers, *Journal of Management Information Systems*, 12, 4, 5-33.
37. Westin, S. and Päivärinta, T. (2011). Information quality in large engineering and construction projects: A Delphi Case Study. *Proceedings of the European Conference on Information Systems*, June 9 - 11, Helsinki, Finland.
38. Wixom, B. H. and Todd, P. A. (2005) A theoretical integration of user satisfaction and technology acceptance, *Information Systems Research*, 16, 1, 85-102.