

Towards a quality model for semantic IS standards

Erwin Folmer

University of Twente & TNO

erwin.folmer@tno.nl

Joris van Soest

University of Twente

Abstract: This research focuses on developing a quality model for semantic Information System (IS) standards. A lot of semantic IS standards are available in different industries. Often these standards are developed by a dedicated organization. While these organizations have the goal of increasing interoperability, there is no way to determine the quality of such a standard. This research will provide quality attributes relevant to semantic IS standards. A theoretical grounded model is created and validated by 19 experts through a survey. Based on these findings a quality model to assess the quality of semantic IS standards has been constructed.

Keywords: Quality model, semantic IS standard, interoperability.

1. INTRODUCTION

With the introduction of XML and the Internet, e-business became available for many companies. Much focus is nowadays on the concept of inter-organizational interoperability: the ability of two or more social-technical systems to exchange information, to interpret the information that has been exchanged and to act upon it in an appropriate and agreed-upon matter [22].

Research has shown that a lack of interoperability costs the automotive industry in the USA an estimate of \$ 1 billion per year and a delay of two months in the introduction of new models [3]. Standardization is a way to achieve interoperability. A standard, in the simplest sense, is an agreed-upon way of doing something [24]. Semantic IS standards are used to communicate and cooperate with partners, suppliers or customers in an efficient and effective way. These semantic IS standards describe the meaning of information and syntax of messages that are exchanged. A semantic IS standard is a mean to achieve the goal of interoperability. The extent to which a semantic IS standards is capable of providing an effective contribution to this interoperability can be described as the fitness for use. So a qualitative good standard is able to achieve a high level of interoperability.

Although these standards are usually developed with the best intentions, they often have quality issues like difficult to understand, multiple interpretations, etcetera [12]. Hardly any study has been done to determine which quality aspects increase interoperability. This study will focus on developing a quality model for semantic IS standards.

Quality of semantic IS standards is strongly related to information quality. The main distinction is that the academic area of information quality often is focused on the quality of information within an organization, while on the other hand the quality of information exchanged between organizations is often related to the area of semantic IS standards. In other words quality of semantic IS standards deal with inter-organizational information quality. Semantic IS standards are the traditional mean for data integration within inter-organizational value chains. These inter-organizational value chains might be related to e-business, or more specific related to e-health, e-learning, etc.

1.1 Background

Most of the IT-standards are developed outside traditional standards developing organizations (like ISO or CEN), in so called industry specific consortia (like W3C or OASIS). Semantic IS standards are even a step further, they are often developed in a separate organization dedicated to one specific industry standard. An example is the HR-XML standard developed by the HR-XML Consortium.

Because there are so many different consortia, the quality of a standard can differ quite a lot between them. It is remarkable that little is known about quality of semantic IS standards. If standards are being developed to increase interoperability, the degree in which interoperability can be achieved will most likely be influenced by the quality of the standard. A research among 34 SDOs (standard developing organizations), including international standards like XBRL, HR-XML, ACORD, HL7 and national standards like SETU, StUF, Aquo, shows that more than 90 percent of these organizations think that the quality can be improved [13]. A large majority thinks an improvement of their standard will contribute to better interoperability. It is, however, difficult to improve quality if the quality is not known. More than 80 percent of the questioned SDOs would like to use a tool to assess the quality of their standard if it is available.

1.2 Problem statement

To date there does not exist a quality model to assess the quality of semantic IS standards. While most standards are developed to increase the interoperability in specific domains, there is a lack of methods to assess the quality of these standards. In a business environment where there is an increasing exchange of information, it becomes more and more important to develop standards of high quality to improve efficiency and effectiveness of inter-organizational data integration.

1.3 Research questions

Since there is a need for a quality model for semantic IS standards, more research on this specific area is required. The overall goal is *to build a quality model for semantic IS standards*. To achieve this goal, the following research questions are relevant:

- What structured set of quality attributes determine the quality of a semantic IS standard?
- What can we learn from other disciplines, like software engineering or product engineering?

The results will be a structured list of quality attributes that are applicable to the domain of semantic IS standards. A description will be made for each quality attribute including a definition. A validation will be performed to determine the extent of usefulness to practice of this model.

The outline of this research is as follows. In section 2 the research methodology will be explained. To create a model to assess the quality of semantic IS standards, we first take a look at the literature. A literature study has been conducted to find quality attributes which can determine the quality of a semantic IS standard (section 3). A model is constructed based on the findings in literature (section 5). This model is validated through a survey (section 6). Based on the finding is the survey, a second improved quality model has been constructed (section 7).

2. RESEARCH APPROACH

This research is conducted using the design science principle as explained by Hevner [16]. “Design science addresses research through the building and evaluation of artifacts [...]” [16]. This process is inherently iterative, and consists of build and evaluation steps. This cycle is repeated until the appropriate business needs are satisfied. The artifact is the quality model for semantic IS standards. Within the build phase we will use theories available in literature to create a first model, the artifact. We

will evaluate this model (the artifact) through a survey. These two steps complete the first iteration of the design cycle. The design artifact becomes more relevant and valuable with each iteration [16]. After our first iteration, a second iteration is started, consisting only of build phase. The evaluation results are used to refine the artifact, our end result.

3. LITERATURE REVIEW

Recent study have identified a research gap on quality of transactional standards [11]. A systematic literature study was conducted that covered the top 25 journals of information systems. Since there is no quality model specific for semantic IS standards, we need to look at literature that might have some parts in common. Three main research areas have been looked into; product quality, data- and information-quality and information systems/software quality. These areas all have a history in quality research and have commonalities with semantic IS standards.

Most notable authors in the field of product quality we looked in to are; Crosby [5] (1979) and Garvin [14] (1984).

There is lots of research about quality conducted in the field of Information systems/software quality. Arguably most famous is the work of McCall [4, 19] (1977), Boehm [2] (1978), ISO 9126 [26] (2001), DeLone & McLean [8] (2003). But many others have deliverable valuable work as well like Delen & Rijsenbrij [7] (1992), Grady [15] (1992, FURPS model), Dromey [9, 10] (1995-1996), Dedeker [6] (2000) and O'Brien & et al. [21] (2007).

Authors in the field of data quality and information quality are, amongst others, Wand & Wang [27] (1996), Wang & Strong [28] (1996), Katerattanakul [18] (1999), Alexander & Tate [1] (1999), Shanks [23] (1999), Naumann & Rolker [20] (2000), Zhu et al. [29] (2000) Kahn et al. [17] (2002), Stvilia [25] (2007).

While their application domains might differ, all used definitions or classification of quality are more often similar, then different. We summarized the quality attributes from this vast amount of literature. Table 1 shows this summary including the originating discipline (the 3 columns).

<i>Quality attribute</i>	<i>Product Quality</i>	<i>Information Quality/ Data Quality</i>	<i>Information Systems/ Software Quality</i>
Accessibility		Kahn [17], Shanks [23], Stvilia [25], Wang [28]	Dedeker [6]
Accuracy		Alexander [1], Katerattanakul [18], Naumann [20], Delen [7], ISO 9126 [26] Shanks [23], Stvilia [25], Wand [27], Wang [28]	
Adaptability			Dedeker [6], DeLone [8], O'Brien [21], ISO 9126 [26]
Aesthetics	Garvin [14]	Katerattanakul [18]	Dedeker [6]
Amount of Info		Kahn [17], Naumann [20], Stvilia [25], Wang [28]	
Analysability			ISO 9126 [26]
Attractiveness		Katerattanakul [18]	ISO 9126 [26]
Authority		Alexander [1], Stvilia [25], Zhu [29]	
Availability		Naumann [20], Wang [28], Zhu [29]	Dedeker [6], Delen [7], DeLone [8], O'Brien [21], ISO 9126 [26]
Believability		Kahn [17], Naumann [20], Wang [28]	
Changeability		Kahn [17]	Boehm [2], ISO 9126 [26]
Clarity		Wand [27]	Boehm [2], ISO 9126 [26]
Co-existence			ISO 9126 [26]

Completeness	Kahn [17], Naumann [20], Shanks [23] Stvilia[25], Wand [27], Wang [28]	Dedeke [6], Delen [7], DeLone [8]
Complexity	Stvilia [25]	
Compliance		ISO 9126 [26]
Conformance	Crosby [5], Garvin [14]	ISO 9126 [26]
Consistency	Kahn [17], Katerattanakul [18], Naumann [20], Shanks [23], Stvilia [25], Wand [27], Wang [28]	Dedeke [6]
Correctness	Shanks [23], Wand [27]	Boehm [2], McCall [4, 19], Delen [7], Dromey [9, 10]
Cost-effectiveness	Naumann [20], Wang [28]	Boehm [2], DeLone [8]
Customisability		DeLone [8], O'Brien [21], ISO 9126 [26]
Decision making	Katerattanakul [18]	Delen [7]
Degradability		Delen [7], ISO 9126 [26]
Diversion possibility		Delen [7]
Durability	Garvin [14]	
Ease of operation	Wang [28]	
Efficiency	Naumann [20] Wand [27]	Boehm [2], McCall [4, 19], Dedeke [6], Delen [7], Dromey [9, 10], ISO 9126 [26]
Explicitness		ISO 9126 [26]
Extensibility		O'Brien [21]
Fault tolerance		Dedeke [6], ISO 9126 [26]
Features	Garvin [14]	Katerattanakul [18]
Flexibility	Wand [27], Wang [28]	Boehm [2], McCall [4], Delen [7]
Free of error	Kahn [17], Naumann [20], Wang [28]	Dedeke [6]
Functionality	Alexander [1], Kahn [17]	Dromey [9, 10], FURPS [15], ISO 9126 [26]
Helpfulness		Boehm [2], ISO 9126 [26]
Installability	Naumann [20]	ISO 9126 [26]
Interoperability		ISO 9126 [26], O'Brien [21]
Integrity		McCall [4, 19], Delen [7], O'Brien [21], ISO 9126 [26]
Learnability		Boehm [2], McCall [4, 19]
Maintainability		Dedeke [6], Dromey [9, 10], ISO 9126 [26]
Manageability		Boehm [2], McCall [4, 19], Delen [7], ISO 9126 [26]
Maturity		ISO 9126 [26]
Navigation		ISO 9126 [26]
Objectivity	Alexander [1], Katerattanakul [18]	Dedeke [6], DeLone [8]
Openness Specification	Alexander [1], Kahn [17], Naumann [20], Wang [28]	
Operability		ISO 9126 [26]
Perceived Quality		ISO 9126 [26]
Performance	Garvin [14]	DeLone [8], FURPS [15]
Portability	Garvin [14]	Naumann [20] DeLone [8], Dromey [9, 10], O'Brien [21], ISO 9126 [26]
Recoverability		Boehm [2], McCall [4, 19], Delen [7], ISO 9126 [26]

Relevancy			Delen [7], ISO 9126 [26]
Reliability		Wang [28], Naumann [20], Stvilia [25], Kahn [17], Dedeker [6], DeLone [8], Dromey [9, 10], Wand [27]	FURPS [15], ISO 9126 [26]
Replaceability	Garvin [14]	Alexander [1], Naumann [20], Wand [27]	Boehm [2], McCall [4, 19], DeLone [8], O'Brien [21], ISO 9126 [26]
Reputation			ISO 9126 [26]
Resource behaviour		Wang [28], Shanks [23], Naumann [20], Kahn [17]	
Reusability			Dromey [9, 10], ISO 9126 [26]
Robustness			Boehm [2], McCall [4, 19], Delen [7], ISO 9126 [26]
Scalability			Delen [7]
Security			O'Brien [21]
Serviceability		Kahn [17], Naumann [20], Stvilia[25], Wang [28]	Dedeker [6], Delen [7], DeLone [8], FURPS [15], O'Brien [21], ISO 9126 [26]
Stability	Garvin [14]	Naumann [20]	Delen [7], DeLone [8], O'Brien [21]
Suitability			ISO 9126 [26]
Testability			ISO 9126 [26]
Time behaviour			Boehm [2], McCall [4, 19], Delen [7], O'Brien [21], ISO 9126 [26]
Timeliness			ISO 9126 [26]
Traceability		Alexander [1], Kahn [17], Naumann [20], Shanks [23], Stvilia [25], Wand [27], Wang [28], Zhu [29]	Dedeker [6], Delen [7]
Understandability		Katerattanakul [18], Naumann [20], Stvilia[25], Wang [28]	Boehm [2], Dedeker [6], Delen [7], ISO 9126 [26]
Usability		Alexander [1], Kahn [17], Naumann [20], Shanks [23], Wand [27], Wang [28]	Boehm [2], Dedeker [6], DeLone [8], Dromey [9, 10], FURPS [15], ISO 9126 [26]
User-friendliness		Shanks [23]	Boehm [2], McCall [4, 19], Delen [7], DeLone [8], FURPS [15], O'Brien [21], ISO 9126 [26]
Value-Added			Delen [7], ISO 9126 [26]

Table 1: List of quality attributes including sources

4. DRAFT QUALITY MODEL

Based on all the quality aspects we found in the previous section we continued the first build phase. A first selection of attributes that are relevant in assessing quality of semantic IS standards was done within the Integrate project, by having expert sessions select and discuss the most relevant attributes. The outcome, the draft model, is heavily inspired on the ISO 9126 model, especially the categorization. The ISO 9126 model is a popular framework and is commonly used to assess the quality of software. The attributes present in the ISO 9126 model are well defined and provided us with the base of our quality model for semantic IS standards. All attributes have been selected that were labeled relevant in relation to quality of semantic IS standards within the Integrate project. However within this project the attribute “acceptance” was added, although no traces of this attributes in the literature was found. Both the use of ISO 9126 as well as the selection of the quality aspects is somewhat arbitrary, but this limitation has been overcome by the survey as evaluation.

4.1 Categorization

The categories, in line with ISO 9126, included in the draft model are: Functionality, Reliability, Usability, Portability and Maintainability. The categorization makes it possible for the user to select parts

of the quality model based on its specific needs. Adoptability and Openness are two categories that were added as category. The Openness category was added because this is nowadays seen as important aspect of a standard although it is related to the standard development organization. However it is also seen as indicator for the quality of the specification. The model is depicted in Figure 1.

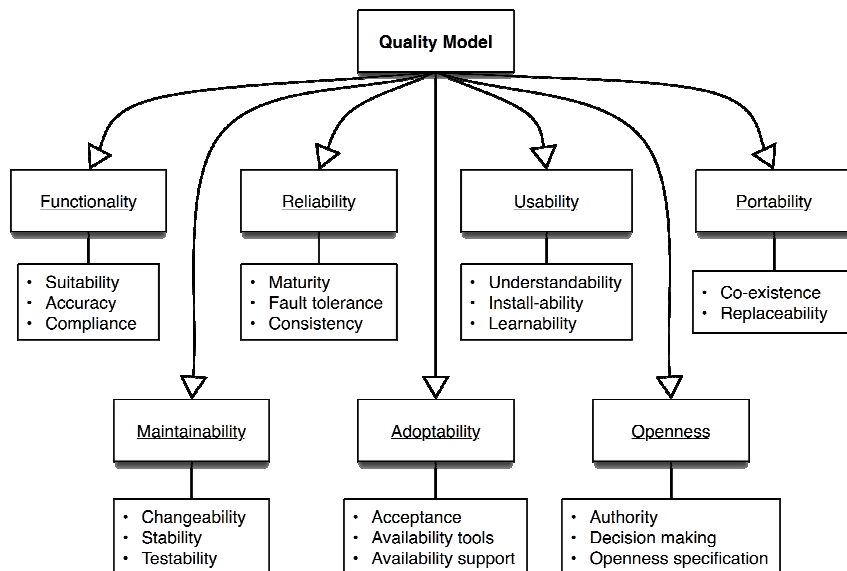


Figure 1: Draft quality model

4.2 Elements of draft model

The definitions of each attribute will be followed by an application of that attribute to the field of semantic IS standards.

Suitability (definition adapted from ISO 9126 [26]):

The capability of the standard to provide an appropriate set of functions for specified tasks and goals. Standards are being used to overcome interoperability issues.

Accuracy (definition adapted from ISO 9126 [26]):

The capability of the standard to provide the right or agreed results or effects with the needed degree of precision. Does the implementation of the standard do what it is supposed to do? Does it live up to the expectations?

Compliance (definition adapted from ISO 9126 [26]):

The capability of the standard to adhere to standards, conventions or regulations in laws and similar prescriptions. This can come from government or the industry. Financial reports are a good example. To what extent are these aspects covered within the standard?

Maturity (definition adapted from ISO 9126 [26]):

The capability of the standard to avoid failure as a result of faults in the standard. When there are not many bugs in the standard, errors will not likely occur. The amount of unsolved bugs or the amount of changes in a release might be a good indicator for this. If the standard is mature, often there is a stable release schema for new versions.

Fault Tolerance (definition adapted from ISO 9126 [26]):

The capability of the standard to maintain a specified level of performance in cases of faults occurring in the implementation. The amount of manual work needed for correcting an error can be a good indicator. Can the implementation continue to work with the error?

Consistency (definition adapted from Stvilia [25]):

The extent to which similar attributes or elements of an information object are consistently represented

using the same structure, format, and precision. Inconsistency will most likely lead to errors in use or implementation.

Understandability (definition adapted from ISO 9126 [26]):

The capability of the standard to enable the user to understand whether the standard is suitable, and how it can be used for particular tasks and conditions of use. Is all the information easy to read? Complex documents will not lead to better implementations. Readability scores can be a good indicator.

Install-ability (definition adapted from ISO 9126 [26]):

The extent to which the standard can be implemented easily. Is the standard easily installed into existing information systems or organizations?

Learnability (definition adapted from ISO 9126 [26]):

The capability of the standard to enable the user to learn its application. The time needed for a user to learn the use or implementation of the standard.

Co-existence (definition adapted from ISO 9126 [26]):

The capability of the standard to exist next to other standards. Will a standard function properly next to another standard, set up for the same goal? Is it possible to access the same information, or does the information use different naming for example?

Replaceability (definition adapted from ISO 9126 [26]):

The capability of the standard to be used in place of another specified standard for the same purpose in the same environment. Is it possible to replace the current standard with a newer version without much hassle? Does the standard provide backwards compatibility?

Changeability (definition adapted from ISO 9126 [26]):

The capability of the standard to enable a specified modification to be implemented. Does the standard provide possibilities for committing changes to the standard? Does the standard provide the option to add localization functions or code-lists? How long does it take to change something in the standard?

Stability (definition adapted from ISO 9126 [26]):

The capability of the standard to avoid unexpected effects from modifications of the standard or environment. New versions emerge over time, as well as new hardware or infrastructure. Does the standard keep its level of function after changes?

Testability (definition adapted from ISO 9126 [26]):

The capability of the standard to enable implementations to be validated. Is there a way to test an implementation? The availability of reference implementation might help. Is there certification?

Acceptance:

The extent to which the standard is used and supported by different kind of stakeholders. How well is the standard used in the target domain? A measurement can be the market share of a standard.

Availability tools:

The extent to which the standard provides tools for implementation. Implementing the standard should be as easy as possible. Additional tools to support the implementation should increase its use. Does the standard developing organization provide methods to let the standard communicate with other software products?

Availability support:

The extent to which knowledge and support is available. To use a certain standard knowledge is needed to implement it successfully. Is there enough knowledge and support available? How fast do you get response from the support department? Is there some external consultancy available for this standard?

Authority (definition adapted from Stivilia [25]):

The degree of reputation of the standard in a given community or business area.

Some standards are highly valued by certain users. This might be because the standard is of better quality, or because of the reputation of the standard development organization.

Decision making (definition adapted from Delen et al. [7]):

The organizational characteristics of the standard developing organization and the way decisions are being made. Is there consensus decision making, or majority voting, or anything else?

Openness specification:

The extent to which the standard provides free to use specifications. Is the specification available for everyone without additional costs or efforts?

5. SURVEY

Evaluation is a crucial component of the research process [16]. It provides valuable feedback to the development of the artifact. "A design artifact is complete and effective when it satisfies the requirements and constraints of the problem it was meant to solve" [16]. The evaluation is being conducted as final step in the first design cycle iteration. Evaluations can take place in different ways, some examples of evaluations are; surveys, experiments, simulations or case studies. In this research a survey was selected. A survey has some major advantages for our purpose of getting feedback on our quality model. A survey is conducted in the field, not in a laboratory. This will ensure the information gathered will be relevant for the next iteration in the design cycle as it contains information about the business context it is meant to be used in. Another advantage of a survey is that you do not have control over the participant. There is little interference which will provide us with honest, unbiased answer to the questions.

This survey had two main goals; first to check what the experts think are important quality attributes for semantic IS standards, second to check if our chosen quality attributes are relevant for assessing the quality of semantic IS standards. These two parts were clearly separated in the survey to ensure our model did not bias the respondents. The model was only introduced after the first part of the survey was finished. The first part of the survey consists of our large list of quality attributes. The following question was asked:

Q1. Which elements do you think are in some way relevant for assessing the quality of a semantic IS standard?

The choices were our previously found 70 quality attributes, presented with a checklist where multiple answers were possible. A pop-up window was provided with a list of all quality attributes and definitions.

The second part of the survey is intended to validate the model, by determining if the selected quality attributes should be included in the model. Furthermore a question was asked about how to categorize the quality attributes that were selected in the first part of the survey but were not included in the model. . First a definition of the category was given, and then the selection of quality attributes was listed. After giving a definition of the selected quality attribute we asked the following question per attribute, on a 5-point likert scale:

Q2. Do you agree that SUITABILITY should be included in the model?

We repeated the same type of question for every quality attribute in each category, including a definition of that specific attribute. The final question for each category was:

Q5. Which of the your previously selected attributes should be added to this category?

Here the possible answers were the selected attributes from the first part of our survey (Q1), presented in a multi-selection checkbox, again with the possibility to have a look at the definitions. These questions were repeated for each of the 7 categories. Finally several questions were asked about the respondents, including their experience. The complete set of information about the survey setup and the results is due to space limitations not listed within this paper but is available by contacting the authors.

5.1 Results

We selected 27 experts mainly from TNO, University of Twente and Novay, who have participated in research about semantic IS standards. These have been invited to participate in this survey. A total of 19

complete responses were gathered, which resulted in a response rate of 70%. Most of the respondents work at research institutions or universities, while others have been involved in the creating process of a semantic IS standards. The years of experience in the field of semantic IS standards varied between 1 year and 25 years. Of the respondents, 73,7% had more than 3 years experience. Among them, 31,6% considered themselves to be an ‘expert’, 57,9% ‘average’ and 10,5% a ‘beginner’.

Quality attribute	Count	Percent %	Present in draft model
Consistency	16	84.2%	Yes
Interoperability	14	73.7%	No
Openness Specification	13	68.4%	Yes
Adaptability	12	63.2%	No
Correctness	12	63.2%	No
Reusability	11	57.9%	No
Completeness	11	57.9%	No
Accessibility	11	57.9%	No
Maintainability	11	57.9%	Yes
Availability	11	57.9%	Yes
Accuracy	11	57.9%	Yes
Understandability	10	52.6%	Yes
Usability	10	52.6%	Yes
Efficiency	10	52.6%	No
Free of error	10	52.6%	No
Testability	10	52.6%	Yes
...	

Table 2: Survey results, Part 1

Part 1

At the first part of the survey, consistency, interoperability and openness specification were considered the most relevant for assessing the quality of semantic IS standards with a score of respectively 16, 14 and 13 at the first question. See Table 2 for a summary of the results. All attributes were selected at least one time, except ‘Attractiveness’, which was not selected a single time. The top 16 answers contained 8 quality attributes which are also present in our draft model.

Part 2

Table 3 shows the (most significant) results gathered from the second part of the survey. A Cronbach’s alpha was calculated on the questions regarding the specific quality attributes used in the model. Cronbach’s alpha is a measure of how well individual variables vary, indicating the reliability of the single factor representing the multiple individual variables. Since the individual variables all measure the same construct, namely the quality of semantic IS standards, it is possible to calculate this alpha. The Cronbach’s alpha is 0,846 which is considered as very good.

Question	Mean	Std. Deviation
Q2. Suitability	4,21	1,032
Q3. Accuracy	3,63	1,212
Q4. Compliance	3,95	1,177
Q6. Maturity	3,53	1,172
Q7. Fault tolerance	3,00	1,085
Q8. Consistency	4,68	0,582
Q10. Understandability	4,16	1,259
Q11. Install-ability	3,68	1,108
Q12. Learnability	2,89	1,197

Q14. Co-existence	3,47	1,124
Q15. Replaceability	3,61	0,979
Q17. Changeability	4,00	0,745
Q18. Stability	3,63	0,895
Q19. Testability	3,68	1,376
Q21. Acceptance	3,89	0,875
Q22. Availability Tools	3,58	1,071
Q23. Availability Support	3,89	0,937
Q25. Authority	2,74	0,991
Q26. Decision Making	4,05	1,177
Q27. Openness Specification	4,21	1,273

Table 3: Survey results, Part 2

At the Functionality category, 6 respondents chose Completeness and Accuracy as to be added to this category. Other attributes, selected more than 3 times, were the attributes already present in the model.

Free of error and Correctness scored respectively 7 and 4 times within the Reliability category. Remarkably ‘Fault tolerance’ has one of the lowest means (3,00) and only 1 respondent thinks it should be added to the Reliability category. Fault tolerance is also once selected in relation to the Functionality category, and it is absent at all other categories.

At the Usability category Learnability has a mean of 2,89 which is also one of the lowest, only Authority (within Openness category) scores even lower. Install-ability scored a mean of 3,68, but was only selected one time in the question about adding it to this category. Accessibility, not present in the model, was selected 5 times to be added to this category, the same amount as Understandability and Usability which were already present in the model.

Co-existence and Replaceability both scored quite well at their individual questions, respectively 3,47 and 3,61. Remarkable to note was that the respondents that selected those two attributes in the first part of the survey, a minimum number of respondents chose Co-existence (2) and Replaceability (1) to be added to the category of Portability. Adaptability (4) and Interoperability (3) have been more often selected.

Within Maintainability, Changeability scored one of the highest with a Mean of 4,00 and a low standard deviation (0,745), but was only selected 5 times at the first part of the survey, and 2 times to be added to this category.

Openness Specification was chosen 5 times to be added to the Adoptability category and 4 times to be added to the Openness category, with 50% and 60% of the respondents selecting that attribute.

5.2 Discussion

Half of the attributes present in our model have also been selected by the participants. The attributes which were selected by more than half of the participants are candidates to include in the second quality model. The results of the second part of the survey learned us that four attributes (Fault tolerance, Learnability, Authority and Co-existence) scored a mean lower than 3,5 (see Table 3). All other attributes were higher with a peak of 4,65 of Consistency. This is an indicator that our selected attributes (minus those four) present in the model are contributing to the goal of assessing the quality of semantic IS standards, at least according to the experts.

Remarkably not all attributes selected in the first part of the survey, and present in our model, returned at the specific categories. An explanation of this can be that some respondents did not find it necessary to add the already present attribute to the category. It can be seen logical that when a respondent agreed to include an attribute to the model just a few questions before, he did not want that attribute to add somewhere else, so that might have been a reason not to select that answer. Other remarks that were given by the respondents shows the need of avoiding complexity:

“Please keep your model small and simple” and “Don’t make the model too complex.”

Furthermore a remark was given about our categorization:

“[...] All attributes are correct and need to be considered, but from different perspectives. My advice is to structure them according to these perspectives [specification, organizational aspects, adoption, implementation aspects].”

A possible explanation to why the respondents did not select a certain attribute at the first part of the survey, and valued the attribute quite high at the individual question within the second part, might be that the respondent did not read the definition. At the first part of the survey the definition list provided by a button. A list of 70 attributes with definitions emerged in a pop-up window. Although no evidence was found, it could be that some participants did not press the button to view the definitions. It is possible the participants saw the definition at the corresponding page for the first time. This might explain the lack of choices in the first part, and high score at the individual questions. For example ‘Decision support’ might have an unknown definition to the respondents at the beginning of the survey and based on that it was not selected, but when they were forced to read the definition in the second part, it was considered a good attribute. This line of thought is supported by several remarks from the respondents:

“The first list contains much too much overlap in definitions”

“[...] So many [attributes] and some seem to be overlapping.”

“Due to the large list of possible quality aspects its sometimes difficult to remember their definitions.”

6. FINAL QUALITY MODEL

Based on the results and feedback of the survey, we started a second iteration of the design cycle. This resulted in an adjusted list of definitions, specific for semantic IS standards. Some attributes were combined to reduce the ambiguity and overlap. After the new definitions-list was created, a second model was built. Feedback from the survey was included in the new model. This resulted in a reduction from the initial 70 attributes used in our first model, to 35 newly defined attributes.

In the process of grouping and re-defining the attributes, we categorized the new attributes into three new and different categories inspired by the respondents; Specification, Organizational aspects, Implementation. It provides a separation of concerns which can be useful in practice. If someone wants to compare quality attributes associated to the implementation of the standard in different products, you only have to look at the implementation category.

- The Specification category is everything that is about the specification. A good rule of thumb is looking at it as a manual for the standard. The Specification handles all the elements which can be seen as ‘the product’.
- The Organizational aspects category is about the control of the standard. It defines how the standard is originated and how the process of development and maintenance is arranged.
- The final category is the Implementation category. Here are all the attributes related to the implementation of the specification. It is related to practice, when a specification is used and a standard is functioning in a certain (business) environment.

After the new categorization and definitions were made, the next step was to update our initial model. The quality attributes from our draft model were included in the second model, except Fault Tolerance, Learnability, Authority and Co-existence. These four attributes were left out based on the results of the survey. Results of the first part of the survey provided us with the addition of the following attributes to the model: Interoperability, Correctness, Completeness, Adaptability, Reusability, Accessibility, Availability, Free of error, Extensibility.

These attributes were selected by more than half (52,9%) of the respondents. The next step was to look at the newly created definitions list and merge some attributes to lower the requested complexity of the

model. ‘Correctness’ and ‘Free of error’ is an example of two attributes which was merged into one. The resulting quality model is presented in Figure 2.

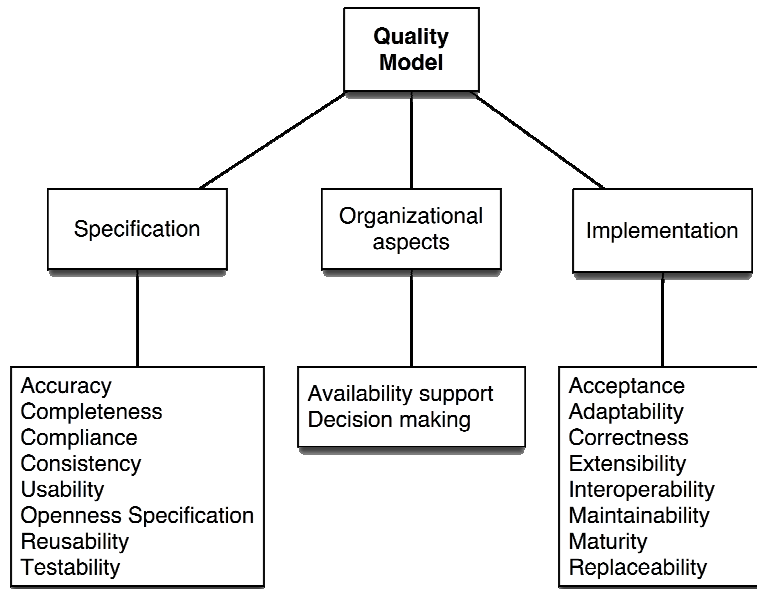


Figure 2: Quality model

The grouping and re-defining of the attributes was done in small iterations. Attributes with similar meaning were grouped together. Each group (or single attribute) was assigned to a category: specification, organizational aspects, or implementation category. This process was repeated until every attribute was assigned to the new categories. For example, the attributes Maintainability, Flexibility, Changeability and Customizability were combined into one attribute ‘Maintainability’. The definition was adapted accordingly and shows traces to the broadened meaning of maintainability: *The capability of a standard to provide a flexible way to modify, change or customize the implementation of a standard for use in different specified environments.*

All quality attributes and their classification are listed in Table 4. The column ‘Similar meaning’ is representing the quality attributes that were combined. Bold items are included in the model.

Category	Attribute	Definition	Similar meaning
Specification	Completeness	The extent to which a standard provide an applicable and appropriate set of functions with sufficient depth and scope to overcome interoperability needs.	Functionality, Relevancy, Suitability, Value-Added
	Usability	The extent to which the standard provides the user a clear and understandable overview of the concept and possibilities, and enables the user to understand whether the standard is suitable for use in a particular task,	Understandability, Clarity
	Aesthetics	The degree to which the standard is appealing or attractive to the customer.	Attractiveness
	Security	The capability of the standard to protect information and data so that unauthorized persons or systems cannot read or modify them and authorized persons or systems are not denied access to them.	Integrity
	Accuracy	The capability of the standard to provide the right or agreed results or effects with the needed degree of precision.	
	Complexity	The degree of cognitive complexity of a standard relative to a particular activity.	
	Compliance	The capability of the standard to adhere to other standards, conventions or regulations in laws.	

	Conformance	The degree to which the standard meet established software, hardware, and communication standards.	
	Consistency	The extent of consistency in using the same values (vocabulary control) and elements to convey the same concepts and meanings in a standard.	
	Durability	A measure of the standards product life.	
	Features	The extent to which the standard provides secondary/supplementary/more advanced functions and technologies.	
	Helpfulness	The extent to which the standard provides help functions.	
	Reusability	Extent to which a standard can be used in other applications.	
	Testability	The capability of the standard to enable modifications to be validated.	
	User-friendliness	The extent to which the standard is adjusted to the knowledge and experience of the users.	
	Amount of Info	The amount of information provided by the standard.	
	Openness Specification	The extent to which information about the standard is available, is easily retrievable and can be used freely.	Accessibility, Navigation
Organizational aspects	Stability	The capability of a standard to avoid unexpected effects from changes of the specification of a standard.	Robustness
	Decision making	The organizational characteristics of the standard developing organization and the way decisions are being made.	
	Explicitness	The capability of a standard to provide insight in the current operating status of the product	
	Objectivity	The extent to which the standard is operating unbiased (unprejudiced) and impartial.	
	Serviceability	How well the standard handles users' enquiries.	
	Availability	The degree to which the standard is available for every user when and where he or she needs it.	
Implementation	Correctness	The extent to which an implementation of a standard is correct and reliable in use and satisfies its specification.	Free of error
	Acceptance	The extent to which the standard is used and supported by different kind of stakeholders.	
	Maintainability	The capability of a standard to provide a flexible way to modify, change or customize the implementation of a standard for use in different specified environments.	Flexibility, Changeability, Customisability
	Extensibility	The extent to which a standard provides possibilities to extend the capabilities without affecting other parts of the implementation, without degradation of other quality attributes.	Scalability
	Reliability	The capability of the standard to maintain a specific level of performance when used in a certain context.	Performance, Perceived Quality, Resource behaviour, Efficiency
	Adaptability	The capability of the standard to be adapted for different specified environments without applying actions or means other than those provided for this purpose for the standard considered.	
	Operability	The extent to which standard is easily operationalized and kept in operation.	Manageability, Ease of operation
	Portability	The capability of the standard to be transferred from one environment to another.	Diversion possibility
	Time behaviour	The capability of the standard to provide appropriate response and processing times and throughput rates when performing its function, under stated conditions.	Timeliness
	Analysability	The capability of the standard to be diagnosed for deficiencies or causes of failures in the standard, or for the parts to be modified to be identified.	Traceability
	Fault tolerance	The extent to which the standard remains a certain level of performance in case of failures.	Degradability, Recoverability
	Installability	The capability of the standard to be installed in a specified environment.	
	Interoperability	The capability of the standard to interact with one or more specified systems.	
	Maturity	The capability of the standard to avoid failure as a result of faults in the standard.	
	Replaceability	The capability of the standard to be used in place of another specified standard for the same purpose in the same environment.	
	Cost-effectiveness	The extent to which the cost of collecting appropriate knowledge and implement the standard is reasonable.	

Table 4: Quality attributes for semantic IS standards

7. CONCLUSIONS

We take one step back and recapitulate the research questions:

- What structured set of quality attributes determine the quality of a semantic IS standard?

We have used the process of design science to construct a quality model which has been partly validated and improved based on expert opinions gathered through a survey. This structured set, or model, contain quality attributes that can be used to determine the quality of a semantic IS standard.

- What can we learn from other disciplines, like software engineering or product engineering?

A comprehensive list of quality attributes collected from different fields of research has been created and used for setting up the model specific for semantic IS standards.

The importance of this study lays in the fact that no quality model for semantic IS standards was present to date. This research provides a first step towards constructing such a quality model. The results show that quality attributes from different areas of research are not always compatible with semantic IS standards, but still gives valuable input and starting point for setting up a quality model for a new domain, like semantic IS standards.

7.1 Future work

Future work can be focused on multiple parts. First, our final model and new definitions should be validated. Another focus of future work could be directed to finding suitable measurements for each quality attribute. To use this model in practice you can always ask someone for example how complete the specification is. But how do you measure completeness in a specific situation? Maybe it's complete if it provides the minimalistic functions to operate. But other standards might consider all possible functions thinkable as a complete standard. Also different environments where the standard is implemented can differ quite a lot. A qualitative good standard in one industry might be organized by a closed organization, while a qualitative good standard in another industry is organized by an open community. The measurement itself is not enough; the scales are of equal importance, with different meanings for different uses. However we should not forget the goal of increasing interoperability. A generic quality model for semantic IS standards should be seen as a guide to improve interoperability. When application of the model leads to some improvements into the standard that will lead to interoperability improvement, we are one step closer to our goal.

REFERENCES

- [1] Alexander, J.E. and M.A. Tate. 1999. *Web Wisdom; How to Evaluate and Create Information Quality on the Web*. 1st ed. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.
- [2] Boehm, B.W., et al. 1978. *Characteristics of Software Quality*. North Holland.
- [3] Brunnermeier, S.B. and S.A. Martin. 2002. Interoperability costs in the US automotive supply chain. *Supply Chain Management*. 7(2): p. 71-82.
- [4] Cavano, J.P. and J.A. McCall. 1978. A framework for the measurement of software quality, in *Proceedings of the software quality assurance workshop on Functional and performance issues*. ACM. p. 133-139.
- [5] Crosby, P.B. 1979. *Quality is free: The art of making quality certain*. New York: McGraw Hill. 309.
- [6] Dedeker, A. 2000. A conceptual framework for developing quality measures for information systems. in *Proceedings of 5th International Conference on Information Quality*.
- [7] Delen, G.P.A.J. and D.B.B. Rijsenbrij. 1992. The specification, engineering, and measurement of information systems quality. *The Journal of Systems and Software*. 17(3): p. 205-217.
- [8] DeLone, W.H. and E.R. McLean. 2003. The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*. 19(4): p. 9-30.
- [9] Dromey, R.G. 1995. Model for software product quality. *IEEE Transactions on Software Engineering*. 21(2): p. 146-162.

- [10] Dromey, R.G. 1996. Cornering the chimera. *IEEE Software*. 13(1): p. 33-43.
- [11] Folmer, E., et al. 2009. Top IS research on quality of transaction standards: A structured literature review to identify a research gap. in *Proceedings of the 6th International Conference on Standardization and Innovation in Information Technology SIIT: Verlagshaus Mainz GmbH Aachen*.
- [12] Folmer, E., et al. 2010. Requirements for a quality measurement instrument for semantic IS standards. in *EURAS Proceedings 2010 Service Standardisation: Verlagshaus Mainz GmbH Aachen*.
- [13] Folmer, E., P.O. Luttinghuis, and J.v. Hillergersberg. 2010. Do Semantic IS standards lack Quality? A survey among 34 semantic IS standards [unpublished].
- [14] Garvin, D.A. 1984. What does "Product Quality" really mean? *Sloan management review*. 26(1): p. 25-43.
- [15] Grady, R.B. 1992. *Practical software metrics for project management and process improvement*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
- [16] Hevner, A.R., et al. 2004. Design Science in Information Systems Research. *Mis Quarterly*. 28(1): p. 75-105.
- [17] Kahn, B.K., D.M. Strong, and R.Y. Wang. 2002. Information quality benchmarks: product and service performance. *Commun. ACM*. 45(4): p. 184-192.
- [18] Katerattanakul, P. and K. Siau. 1999. Measuring information quality of web sites: development of an instrument, in *Proceedings of the 20th international conference on Information Systems*. Association for Information Systems: Charlotte, North Carolina, United States. p. 279-285.
- [19] McCall, J.A., P.K. Richards, and G.F. Walters. 1977. *Factors in Software Quality*. Nat'l Tech.Information Service. 1, 2 and 3.
- [20] Naumann, F. and C. Rolker. 2000. Assessment Methods for Information Quality Criteria. in *In Proceedings of the International Conference on Information Quality (IQ)*.
- [21] O'Brien, L., P. Merson, and L. Bass. 2007. Quality Attributes for Service-Oriented Architectures, in *Proceedings of the International Workshop on Systems Development in SOA Environments*. IEEE Computer Society. p. 3.
- [22] Rukanova, B. 2005. Business transactions and standards : towards a system of concepts and a method for early problem identification in standa[r]d implementation projects: Enschede. p. 262.
- [23] Shanks, G. and B. Corbitt. 1999. Understanding data quality: Social and cultural aspects. in *Proceedings of the 10th Australasian Conference on Information Systems*.
- [24] Spivak, S.M. and F.C. Brenner. 2001. *Standardization Essentials: Principles and Practice*. New York: CRC Press.
- [25] Stvilia, B., et al. 2007. A framework for information quality assessment. *Journal of the American Society for Information Science and Technology*. 58(12): p. 1720-1733.
- [26] Van Zeist, R.H.J. 1996. Specifying software quality with the extended ISO model. *Software Quality Journal*. 5(4): p. 273-284.
- [27] Wand, Y. and R.Y. Wang. 1996. Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*. 39(11): p. 86-95.
- [28] Wang, R.Y. and D.M. Strong. 1996. Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*. 12(4): p. 5-34.
- [29] Zhu, X. and S. Gauch. 2000. Incorporating quality metrics in centralized/distributed information retrieval on the World Wide Web. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)*: p. 288-295.