

# QUALITY OF SEMANTIC STANDARDS

ERWIN FOLMER

# QUALITY OF SEMANTIC STANDARDS

Erwin Folmer

**Graduation committee:**

prof.dr. Ramses Wessel (chairman, secretary)	University of Twente
prof.dr. Jos van Hillegersberg (promotor)	University of Twente
dr.ir. Paul Oude Luttighuis (assistant promotor)	Novay
prof.dr. Robert Stegwee	University of Twente
prof.dr. Peter Apers	University of Twente
prof.dr. Sjaak Brinkkemper	Utrecht University
prof.dr.ir. Erik Fledderus	Eindhoven University of Technology / TNO
prof.dr.ir. Jan van den Ende	Erasmus University
dr. Boris Otto	University of St. Gallen

**This research was supported by :**

**TNO** innovation  
for life

**NOV**

**UNIVERSITY OF TWENTE.**

**CTIT**

ISBN: 978-90-365-3323-2

DOI: 10.3990/1.9789036533232

URL: <http://dx.doi.org/10.3990/1.9789036533232>

<http://semanticstandards.org>



Layout & Printing: Nicole Nijhuis, Gildeprint Drukkerijen - Enschede

# QUALITY OF SEMANTIC STANDARDS

PROEFSCHRIFT

ter verkrijging van

de graad van doctor aan de Universiteit Twente,

op gezag van de rector magnificus,

prof. dr. H. Brinksma,

volgens besluit van het College voor Promoties

in het openbaar te verdedigen

op donderdag 5 april 2012 om 14.45 uur

door

Erwin Johan Albert Folmer  
geboren op 10 juli 1975  
te Zevenaar, Nederland

This dissertation has been approved by:

prof.dr. Jos van Hillegersberg (promotor)

dr.ir. Paul Oude Luttighuis (assistant promotor)

## Contents

### Preface

<i>Chapter 1 – Introduction to the Research</i> .....	1
1.1 The concepts defined .....	1
1.1.1 Interoperability.....	1
1.1.2 Semantic standards.....	1
1.1.3 Quality.....	2
1.1.4 Measurement and instruments.....	4
1.2 Perspectives on interoperability and standardization .....	4
1.2.1 The Business perspective.....	4
1.2.2 The Government perspective.....	5
1.3 Research motivation.....	7
1.3.1 Examples .....	7
1.3.2 The problem.....	8
1.4 Research goal .....	10
1.4.1 The goal and main research question .....	10
1.4.2 Scope.....	11
1.5 Research design and outline .....	11
 <i>Chapter 2 – Introduction to Standards</i> .....	15
2.1 The standards domain.....	15
2.1.1 Standards: typology.....	16
2.1.2 Standards: the processes and the product.....	19
2.1.3 Standards: the organization.....	21
2.2 Standards and interoperability.....	23
2.2.1 Integration and interoperability.....	25
2.2.2 Other frameworks and maturity model for interoperability.....	28
2.2.3 The impact of interoperability.....	30
 <i>Chapter 3 – Problem Relevance</i> .....	31
3.1 Research approach .....	31
3.1.1 Research framework.....	31
3.1.2 Limitations .....	34
3.1.3 Survey population.....	34
3.2 Results .....	36
3.3 Discussion.....	37
3.4 Conclusions .....	38
 <i>Chapter 4 – Identification of Research Gap</i> .....	39
4.1 Research questions and method .....	39
4.2 Classification framework.....	41
4.3 Classification process and results.....	44
4.4 Findings .....	46
4.5 Conclusions and discussion.....	47

<i>Chapter 5 – Semantic Standards Literature Exploration</i> .....	49
5.1 Semantic standard terminology .....	49
5.1.1 Semantic approaches, languages and technology.....	50
5.1.2 Examples of semantic standards.....	51
5.2 Semantic standards: Development and adoption .....	53
5.2.1 Development.....	54
5.2.2 Adoption.....	56
5.2.3 Tactics for semantic standards.....	58
5.3 Quality .....	59
5.3.1 Quality in different domains.....	59
5.3.2 General standardization quality.....	62
5.3.3 Quality of semantic standards .....	64
 <i>Chapter 6 – Design Approach</i> .....	 69
6.1 Introduction.....	69
6.2 Research set up.....	69
6.2.1 Specify requirements.....	70
6.2.2 Build instrument.....	71
6.2.3 Evaluate instrument.....	72
6.2.4 Validate iQMSS.....	73
6.2.5 Summary of our research design .....	73
6.3 Research classification .....	75
6.3.1 Research type .....	75
6.3.2 Research epistemology.....	75
6.3.3 Research design .....	75
6.3.4 Research methods/approaches .....	76
6.3.5 Summary.....	76
6.4 Validation of research approach.....	76
6.4.1 Guidelines for design science research.....	76
6.4.2 Principles for conducting and evaluating interpretive field studies.....	78
6.5 Conclusions .....	78
 <i>Chapter 7 – Requirements</i> .....	 81
7.1 Research approach .....	81
7.2 The process.....	82
7.3 The results.....	83
7.3.1 Overview.....	83
7.3.2 Useful for semantic standards of different SSOs.....	84
7.3.3 Efficiently determine the quality and improvement suggestions.....	87
7.3.4 Have useable results for SSOs.....	87
7.3.5 General observations and discussion .....	87
7.4 Reflection on requirements.....	88
7.5 Conclusions .....	88
 <i>Chapter 8 – Design of the iQMSS</i> .....	 89
8.1 The reference model.....	89
8.1.1 Design constraints of the iQMSS.....	89
8.1.2 iQMSS use case.....	90
8.1.3 Design of model iQMSS.....	92

8.2 The iQMSS languages.....	96
8.2.1 Research approach.....	96
8.2.2 The quality languages.....	96
8.2.3 The selection of SMO as language .....	98
8.3 The development environment of iQMSS .....	102
8.3.1 Development environment for quality models.....	102
8.3.2 Export and visualization.....	105
8.4 Conclusions.....	106
<i>Chapter 9 – The Semantic Standards Model.....</i>	<i>109</i>
9.1 Research approach and concepts .....	109
9.1.1 Research process.....	110
9.2 SSM design.....	111
9.2.1 Design objectives.....	111
9.2.2 SSM overview .....	111
9.2.3 SSM in Detail.....	113
9.3 SSM application Case A: Dutch Ministry of Economic Affairs.....	116
9.3.1 Context .....	116
9.3.2 Standards selection process.....	117
9.3.3 SSM applicability .....	118
9.4 SSM application Case B: Siemens Corporate Technology .....	119
9.4.1 Context .....	119
9.4.2 Standards evaluation process.....	119
9.4.3 SSM applicability .....	121
9.5 SSM: Evaluation, discussion and conclusions .....	121
9.5.1 Evaluation .....	121
9.5.2 Discussion.....	122
9.5.3 Conclusions .....	123
<i>Chapter 10 – Explorative Evaluation Studies.....</i>	<i>125</i>
10.1 The iQMSS – build 0.3.....	125
10.2 Explorative field test – SETU.....	127
10.2.1 Research approach.....	127
10.2.2 SSM applied.....	127
10.2.3 QMSS applied.....	127
10.2.4 Evaluation .....	129
10.3 Lab experiment – XCRI .....	130
10.3.1 Experimental set up .....	130
10.3.2 Results of the use of SSM.....	131
10.3.3 Results of the use of QMSS .....	131
10.3.4 Evaluation .....	132
10.4 Conclusions .....	134
<i>Chapter 11 – The Quality Model of Semantic Standards.....</i>	<i>135</i>
11.1 Research methodology and structure.....	136
11.1.1 Overall approach for build iterations.....	136
11.1.2 Final build research approach (build 0.7).....	138
11.1.3 The structure of QMSS.....	141

11.2 The QMSS (final build 0.7) .....	143
11.2.1 Product quality.....	143
11.2.2 Process quality.....	148
11.2.3 Quality in practice.....	151
11.3 The set of measures for product quality.....	155
11.4 The iQMSS usage process .....	158
11.4.1 Preparation.....	158
11.4.2 Measuring.....	160
11.4.3 Analysis.....	161
11.4.4 Recipe.....	161
11.5 Conclusions .....	162
<i>Chapter 12 – Validation of the iQMSS.....</i>	<i>163</i>
12.1 Research approach.....	163
12.1.1 RQ1: Re-assessment of requirements .....	163
12.1.2 RQ2: Field test.....	163
12.1.3 RQ3: Expert workshops.....	165
12.2 Validation based on requirements.....	166
12.3 Field test IMS LRM.....	167
12.3.1 Preparation phase.....	167
12.3.2 Measuring.....	170
12.3.3 Analysis.....	171
12.3.4 Field test discussion.....	172
12.4 Validation sessions at OMG and NOiV.....	173
12.4.1 Discussion results.....	173
12.4.2 Survey results.....	173
12.5 Conclusions .....	174
<i>Chapter 13 – Conclusions and Future Research.....</i>	<i>175</i>
13.1 General conclusions.....	175
13.2 Conclusions: the contribution to practice and science.....	176
13.2.1 Contribution to practice.....	176
13.2.2 Contribution to science .....	177
13.3 Future research .....	177
<i>Appendix A – Problem Survey Design (chapter 3).....</i>	<i>183</i>
<i>Appendix B – Survey Results (chapter 3).....</i>	<i>185</i>
<i>Appendix C – Literature Search Strategy (chapter 4).....</i>	<i>189</i>
<i>Appendix D – Overview of Selected Studies (chapter 4).....</i>	<i>193</i>
<i>Appendix E – Set of Requirements for Instrument (chapter 7).....</i>	<i>197</i>
<i>Appendix F – Requirements Goal Tree (chapter 7).....</i>	<i>201</i>
<i>Appendix G – Graphical Presentation of SSM (chapter 9).....</i>	<i>205</i>
<i>Appendix H – Existing Models for Semantic Standards (chapter 9).....</i>	<i>207</i>
<i>Appendix I – Example of a Semantic Standard Classification in Case A (chapter 9).....</i>	<i>209</i>
<i>Appendix J – Example of a Semantic Standard Evaluation in Case B (chapter 9).....</i>	<i>211</i>
<i>Appendix K – SSM Grounded by Case Study Research (chapter 9).....</i>	<i>213</i>
<i>Appendix L – iQMSS Build 0.3 (chapter 10).....</i>	<i>215</i>
<i>Appendix M – SETU Case Result of Applying SSM (chapter 10).....</i>	<i>219</i>
<i>Appendix N – SETU Case Result of Applying QMSS (chapter 10).....</i>	<i>221</i>

<i>Appendix O – The Sources (chapter 11)</i> .....	227
O.1 The software quality domain .....	227
O.2 The IS quality & success domain .....	233
O.3 The data quality domain.....	241
O.4 The standards quality domain.....	246
O.4.1 General: The EU government perspective on quality.....	246
O.4.2 General: What is a good standard?.....	248
O.4.3 General: What is not a good standard?.....	249
O.4.4 Quality statements: Suggestions for better quality.....	249
O.4.5 Quality attributes: Suggestions for better quality & measures.....	251
O.5 The XML standards quality domain .....	253
O.6 Evaluation framework for standards.....	255
O.7 Other works.....	257
O.7.1 Requirements specification study.....	258
O.7.2 The product engineering quality domain.....	258
O.7.3 The Integrate sources.....	259
O.7.4 Quality in practice.....	260
O.7.5 Other quality instruments .....	261
<i>Appendix P – IMS Learning Resource Metadata (LRM) Description (chapter 12)</i> .....	263
<i>Appendix Q – Measurement Table IMS LRM (chapter 12)</i> .....	267
<i>Appendix R – Analysis Report of iQMSS Assessment on IMS LRM (chapter 12)</i> .....	275
 <i>References</i> .....	 277
 <i>Glossary of Terms</i> .....	 295
 <i>Abstract</i> .....	 297
 <i>Nederlandse Samenvatting</i> .....	 299
 <i>Dankwoord</i> .....	 301
 <i>About the Author</i> .....	 303

## Preface

### *Examples of semantic errors*

“In Desert Storm, an aerial observer located an enemy unit and sent a bombing request to the artillery headquarters. Using the enemy location’s coordinate received from the artillery headquarters, the Navy ship off the coast fired two rounds, but both missed the target by 527 meters, a distance way greater than expected precision. It turned out that the artillery headquarters and the Navy used different geo-coordinate systems with which the same coordinates represent different locations on earth” (Zhu & Fu, 2009).

“The NASA Mars Climate Orbiter was lost after messages between two different systems were misinterpreted: the first was sending values in US Customary units (lbf-s), the second assumed that the values arriving were measured in SI units (Ns); the result was an initial orbit 170km lower than planned—23km below survivable height” (Davies, Harris, Crichton, Shukla, & Gibbons, 2008). The result: the loss of US \$125 million.

### *The Apgar analogy of our research*

The Apgar score, invented by Virginia Apgar in 1953, evaluates the health for newly born babies, and based on this instrument the mortality of babies is strongly reduced. The instrument looks at five aspects (**A**ppearance, **P**ulse, **G**rimace, **A**ctivity, **R**espiration), and scores them accordingly on a 3 point scale (0,1,2). The Apgar score is the sum of those five scores. The scoring is being done after 1,5 and 10 minutes of the birth of the child. Scoring less than 4 points means immediate attention, while scoring 7 or higher means normal. The main contribution of the APGAR score is that the subjective measurement has become objective, and viability is factor 6 improved.

Wouldn’t it be useful to have an Apgar score for information systems (IS)? (Glass, 2008) or even more specific for semantic standards?

In our research we go one step beyond the Apgar score by not aiming for a score, but aiming for an instrument to improve the quality of standards which might lead to improved interoperability. By using the instrument it is intended that quality will become a “control factor” for standard developers.

# Introduction to the Research

This chapter sets the scene and introduces the important concepts within this research before presenting the motivation, main research goal and the structure of this research.

### 1.1 The concepts defined

This section starts by introducing the concepts of interoperability and semantic standards, including the concepts of quality and measurement, the central concepts of this thesis.

#### 1.1.1 Interoperability

Although many definitions of interoperability are in use, one of the most common is the IEEE and ISO definition: “The ability of two or more systems to exchange data, and to mutually use the information that has been exchanged” (IEEE, 1990). The context of this research excludes many kinds of interoperabilities, such as intra-organizational, document and multimedia file formats. The focus is on inter-organizational interoperability, based on Internet-technology. The characteristics of the content level of Internet-based inter-organizational interoperability is based on (open) XML-based standards, low complexity and it is not that partner-specific, while the transportation level is based on open Internet communication protocols, high interoperability and low communication costs. And a broad trading partner scope (Zhu, Kraemer, Gurbaxani, & Xu, 2006).

The definition of (inter-organizational) interoperability used within this research is “The ability of two or more organizational systems to exchange information, to interpret the information that has been exchanged and to act upon it in an appropriate and agreed upon manner” (adapted from (Rukanova, 2005)). Semantic interoperability excludes the technical exchange but includes the content, meaning, processing and interpretation of the information that is being exchanged.

#### 1.1.2 Semantic standards

“Standards, like the poor, have always been with us” (Cargill, 1989; Cargill & Bolin, 2007). Information systems (IS) without standards are hard to imagine. A standard, in the simplest sense, is an agreed-upon way of doing something (Spivak & Brenner, 2001). Arguably the most used definition of a standard is the definition used by ISO and IEC (De Vries, 2006; Spivak & Brenner, 2001; Van Wessel, 2008): “A standard is a document, established by consensus and approved by a recognized body, that provides, for common and repeated use, rules, guidelines, or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context”.

This definition is broad and is in line with the broad use of the term standards, although the part “consensus and approved by a recognized body” is disputed. Regarding our focus on inter-organizational interoperability, there is only a limited group of standards that are relevant, including both technical and semantic standards. Technical standards, such as TCP/IP, HTTP, SOAP, XML, are a prerequisite for Internet-based inter-organizational interoperability.

In our research we focus on semantic information system standards (in short: semantic standards), a relatively new area of standardization. Semantic standards reside at the presentation and application layer of the OSI model (Steinfeld, Wigand, Markus, & Minton, 2007). They include business transaction standards, inter-organizational information system (IOS) standards, ontologies, vocabularies, messaging standards, document-based, e-business, horizontal (cross-industry) and vertical industry standards. The often used examples are RosettaNet for the electro technical industry, HealthLevel7 for the health care domain, HR-XML for the human resources industry and Universal Business Language (UBL) for procurement. Semantic standards are designed to promote communication and coordination among organizations; these standards may address product identification, data definitions, business document layout, and/or business process sequences (adapted from (Steinfeld et al., 2007)). Both point to point and hub IT architectures might facilitate this standards based communication and coordination between organizations (Steinfeld, Markus, & Wigand, 2011b).

In the world of inter-organizational communication there is a lot of semantic variety, which means that the same notifiers have different meanings, which might lead to misunderstandings: semantic ambiguity (Rebstock, 2009). Standards are signs (Brzezinski, 2010b), words, phrases and symbols. Semantics deal with the meaning of these notifiers in the sense of how these notifiers relate to reality, how they represent, designate and signify things (Rukanova, 2005). Semantic standards consist of semantics (meanings) and often syntax (a formal structure) and might include pragmatics (intentions) as well. Its value lies in defining the semantics not the syntax. Often, semantic standards involve XML representations (the syntax) of information, but again the key value of the standard lies in its description of the meaning of data and processes within a context to achieve semantic interoperability as part of inter-organizational interoperability.

Semantic standards differ from other type of standards (like technical standards) in many ways, in amongst other its development and maintenance approach and its context dependencies. Both the content and development approaches of semantic standards are highly dependent on the context. Examples of context factors are, among others, regulatory, governance structure, government participation, ICT maturity of an industry and the market situation.

A wide range of development and maintenance approaches are the result of the existence of different standard setting organizations for each industry. Many semantic standards are maintained by the industry itself, by setting up a dedicated Standard Setting Organization (SSO). The often used term Standard Development Organization (SDO) is avoided as many studies limit SDO to a formal organization like ISO, and exclude many organizations that develop and maintain standards for the ICT domain, the so-called industry fora or consortia. The importance of consortia will continue to rise in the ICT domain. SSO includes every organization that is involved in developing and maintaining standards, including formal bodies, industry consortia and anything else that can be present in practice.

There are different stakeholders regarding semantic standards. We have distinguished them as:

- Standard developers: develop and maintain the standard.
- Standard implementers: implement the standard in systems (and processes).
- Standard users: use the implemented standard (the system or process).
- Others, such as standard policy makers, standardization researchers, etc.

### 1.1.3 Quality

As standards are a means to an end — interoperability — a general assumption is that a good standard will improve interoperability. Surprisingly, the question as to what makes a standard good is relatively rarely explicitly expanded on in literature on standardization (De Vries, 2007). However, Markus, Steinfeld, Wigand & Minton (2006) note that the technical contents of the standards will have an impact on the standard's diffusion. However, diffusion and adoption involve acceptance and implementation, and it does

not necessarily mean that interoperability will be achieved. In other words not all successful standards (high adoption) are high quality standards that lead to interoperability.

We have learned from the data quality domain that the quality of data depends on the design and production processes involved in generating the data. Data standards improve data quality in dimensions such as consistency, interpretability, accuracy, etc. However, when data standards are too cumbersome, users may circumvent the standards and introduce data that deviate from these standards. Thus, research in this area also needs to study how protocols and standards impact data quality and how organizations can promote user compliance. In addition, the quality of the standards is also subject to quality evaluation (Madnick, Wang, Lee, & Zhu, 2009). This knowledge from the data quality domain highly resembles the semantic standards domain.

To design for better quality, it is necessary first to understand what quality means and how it is measured (Wand & Wang, 1996). Quality is defined by ISO as: “The totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs” (ISO/IEC, 2001). Quality is a complex notion; an important distinction can be made between:

1. The quality of the product and the process. The standard is in itself the document (the product), or as it is often called, the standard specification. A standardization process is in place for the development and maintenance of this product. The quality aspects deal with the product and process of standardization, meaning that it is aimed at the specification of the standard and the process of designing and maintaining that specification. This quality might be independent of the problem where and what the standard is used for in practice.
2. The quality of the solution to an interoperability problem. Standards are not individually goals; the main reason they exist is as solutions for real-life interoperability problems. A well-documented standard (high-product quality) does not imply that it is a good solution to every interoperability problem. The alignment of the standard to the interoperability problem is of special interest.

The latter addresses our research most appropriately: We are not looking for the gold-trimmed standard, we are focusing on standards themselves and their fitness for use in practice. We define the quality of a semantic standard as: its ability to achieve its intended purpose — semantic interoperability — effectively and efficiently. A high quality standard is, or has a high chance of becoming, an effective and efficient solution for an interoperability problem; a low quality standard does not solve the problem for which it is designed, cannot be implemented efficiently, or has little chance of being adopted. All the phases of the lifecycle of a standard may influence quality. Moreover, quality deals with both intrinsic aspects (the document) and situational aspects (environment) of the standard. This definition applies Juran’s definition of quality — fitness for use (Juran & Gryna, 1988) — to the semantic standards domain, and is in line with the earlier presented ISO 9126 software quality definition (ISO/IEC, 2001). In the end, high quality semantic standards may involve network externalities, avoid lock-ins, increase the variety of systems products, trade facilitation and reduce transaction costs (Blind, 2004). More importantly, they solve or lower economic and social problems, such as imperfect interoperability costs or they solve social related problems.

The concept of “use” in relation to standards has different perspectives (Burton-Jones & Gallivan, 2007), for instance the perspective of an individual organization or an industry sector, where the latter generally relates to interoperability research, and it is also quite difficult to study (Reimers, Johnston, & Klein, 2010). According to Reimers et al. it is difficult to study inter-organizational IS phenomena, such as semantic standards, because it is not enough to only study individual implementations, but studies of whole industries over periods of time are needed. Current theory might be insufficient to tackle this and more practice oriented approaches, including other data collection methods, are needed (Reimers et al., 2010; Reimers et al., 2011).

### 1.1.4 Measurement and instruments

“You can’t control what you can’t measure” is a famous quote by Tom DeMarco and many others before him. When discussing the quality of standards, some kind of measurement of the quality is needed. An instrument can be used to perform the measurement. ISO (ISO/IEC, 1984) defines a measuring instrument as “a device intended to make a measurement, alone or in conjunction with other equipment”. Wikipedia (2011a) describes instrumentation engineering as “the engineering specialization focused on the principle and operation of measuring instruments which are used in design and configuration of automated systems in electrical, pneumatic domains etc.”. In our context, however, neither the measured objects nor the instrument itself are predominantly physical. Therefore, we turn to the definition in Webster’s (Merriam-Webster, 2011) dictionary: “a measuring device for determining the present value of a quantity under observation”. For this research this definition can be put in context by defining an instrument as: A measuring device for determining the quality value of a semantic standard. Measuring devices can have many forms, such as hardware, software, models, questionnaires, etc. Even the combination of several measuring devices is an instrument. The core of a quality instrument is a quality model. According to ISO (ISO/IEC, 1984), a quality model is a set of characteristics and the relationships between them which provide the basis for specifying quality requirements and evaluating quality.

## 1.2 Perspectives on interoperability and standardization

The topic of standardization and interoperability is important for both industry and governments. The following sections will look at both perspectives in more detail. Most of the statements deal with general ICT standardization and interoperability, which contains semantics but also technical standards and interoperability. Literature, strategies and policies do not always makes distinctions between different kinds of standards. From section 1.3 on, we will focus on the problems related to semantic standards.

### 1.2.1 The Business perspective

The potential size of B2B e-commerce to the economy is vast, though somewhat difficult to pin down. Trillions of dollars had been forecast by Goldman Sachs and the Gartner Group for 2005 (Lucking-Reiley & Spulber, 2001). “Although there are different definitions of Electronic Commerce, it is generally acknowledged that B2B accounts for the largest dollar volume of Electronic Commerce” (Albrecht, Dean, & Hansen, 2005). Standards and interoperability are of key importance for B2B e-commerce (EC, 2008). “As a trade volume of several trillion US \$ is globally processed using inter-organizational standards each year, they are a business topic of high importance” (Löwer, 2005).

A less-obvious trend is that organizations of all sizes in a number of industries are getting together to develop standards, and sometimes also technologies and information services providers, to support joint business processes (Markus, 2011). This trend may accelerate the externalization of Information and Communication Technology (ICT), not just to ICT services providers, but also to industry consortia governed by industry members (Markus, 2011). On the negative side, innovation might be halted: a lack of standards slows down the process of outsourcing in huge companies, and might even explain why many are not satisfied with their outsourcing relation (Davenport, 2005). A lack, or a low level of adoption of standards may lead to interoperability problems (EC, 2008).

As early as 1993, a number of businesses and governments alike were aware of the importance of standards for ensuring interoperability (Rada, 1993). Today, in an increasingly interconnected world, interoperability is more important than ever, and interoperability problems are very costly. Studies of the US automobile sector, for example, estimated that insufficient interoperability in the supply chain adds at least one billion dollars to operating costs, of which 86% is attributable to data exchange problems (Brunnermeier & Martin, 2002). Later studies mention five billion dollars for the US automotive industry and 3.9 billion dollars for the electro technical industry, both representing an impressive 1.2% of the value of shipments in each

industry (Steinfeld, Markus, & Wigand, 2011a). The adoption of standards to improve interoperability in the automotive, aerospace, shipbuilding and other sectors could save billions (Gallaher, O'Conner, & Phelps, 2002).

The already huge importance of standards and interoperability will continue to grow. Networked business models are becoming an indisputable reality in today's economy (Legner & Lebreton, 2007), and a recent Capgemini study concludes that to be ready for 2020 companies need to "significantly increase their degree of collaboration as well as their networking capability" (Falge, Otto, & Österle, 2012). Due to the increasing demand for plug-and-play business within supply chains it will gain further significance (Löwer, 2005). Standards and interoperability are a continuous subject by nature: business environments are changing, and standards have to be continuously adapted to the changed environment (Löwer, 2005). Standardization is also one of the main issues often found to be high on a CIO's objective list within multinational enterprises, as for instance within Siemens (Weitzel, Beimborn, & König, 2006).

Interoperability and standards are not a new issue: their importance to welfare and economic growth are widely acclaimed and centuries old. Famous examples include railway gauges, power plugs, the battle between VHS and V2000, and the different DVD+/- standards. The Internet has flourished thanks to standards, and without Internet there would not have been e-business. Within the ICT domain new developments, such as open data and cloud computing, are largely driven by standards.

### *1.2.2 The Government perspective*

The importance of standards and interoperability has been noted. At regional (for instance Asia and Europe) and national levels a great deal of attention is paid to standards and interoperability at the policy level.

The European Commission released several policy studies on standardization; one of its policy goals is: "Increase the quality, coherence and consistency of ICT standards" (EC, 2009). In 2010, immediately after becoming responsible for Europe's digital agenda, Mrs. Kroes made strong statements about Europe's ambition. The first key action in Europe's digital agenda is "to have more and better standards recognized and created in Europe" (Kroes, 2010). As a rationale she repeated her statements from other speeches (Kroes, 2010): "Interoperability boosts competition and we need more of that".

The EC addresses slow standardization as being a weakness and aims to set interoperable standards in its flagship initiative: the Innovation Union (EC, 2010c). It re-emphasizes the important role standards play for innovation. It addresses the challenge for the European standardization system, and expresses the need for a dynamic and efficient standardization system. Europe's standard-setting framework must catch up with fast-moving technology markets because standards are vital for interoperability (EC, 2010b).

Based on the wide criticism of Europe's standards setting framework, many reports have been released related to reforming EU standardization. Many have the vein of self-evaluations from European Standards Organizations (ESOs) and are trying to minimize the change to the existing situation, and are focusing on the value of the current ESO system, thereupon suggesting improvements (EP, 2010; Pindar, 2010). Some of these improvements concern the role of governments (more active participation in standards development), some concern ESOs (improved access to the standards), and some concern the quality of standards.

On Feb. 4<sup>th</sup>, 2011, the European Parliament invited the European Commission to: "make proposals to accelerate, simplify and modernize standardization procedures notably to allow standards developed by industry to be turned into European standards under certain conditions" (EC, 2011a). The Commission's answer, released in July 2011, consists of a vision and a regulation (EC, 2011c). The vision aims to adapt Europe's standards activities to a quickly changing global economic landscape. In order to respond rapidly to evolving needs in all areas, a comprehensive, inclusive, efficient, and technically up-to-date European

standardization system will be required. Chief among the measures announced by the Commission are the following initiatives (EC, 2011c):

- The European Commission will enhance its cooperation with European Standardization Organizations (ESOs) in order to speed the availability of standards. ESOs should reduce the average time to develop European standards or European standardization deliverables requested by the Commission by 50% by 2020. This means a reduction from 36 to 18 months by 2020.
- Standards for information and communication technology (ICT) will play a more prominent role in the EU in an effort to stimulate innovation, cut administrative costs, and encourage interoperability between devices, applications, data repositories, services, and networks. The Commission will demand that European standards for innovative products and services be quickly elaborated and adopted, in such fields as ecodesign, smart grids, energy efficiency of buildings, nanotechnologies, security, and eMobility.

The European standardization strategy is focusing on its three European Standardization Organizations (ESOs: CEN, CENELEC and ETSI). An action plan should focus the directions of these ESOs by aiming at certain domains such as health care, e-business, e-government, and Internet of things (EC, 2010a).

With regard to e-Government, the European Interoperability Framework (EIF) and Strategy (EIS) have been set on a vision of e-Government interoperability. The release of the second version of the EIF shows the value put by the European Commission on semantic standards (EC, 2010e): "Public administrations should support the establishment of sector specific and cross-sectoral communities that aim to facilitate semantic interoperability and should encourage the communities to share results on national and European platforms".

The European e-Government Action Plan 2011-2015 (EC, 2010d) sees open specifications and interoperability as pre-conditions for developing e-government. The action plan addresses the importance of standards for cost-effective interoperability. Specific actions are set for carrying out EIF and EIS, organizing exchange of expertise, and aligning the national interoperability frameworks to the EIF (EC, 2010e). A focus on mandating open standards exists within e-governments, particularly in national policies such as the Dutch policy named "Netherlands Open in Connection" (NOiV, 2007). Other examples include the UK government (CabinetOffice, 2011), but also India has set a policy (Gol, 2010) which has similarities to the Dutch and UK government policies on promoting open standards (Mutkoski, 2011).

On a national level, the US health care program is exemplary. In its ambition to achieve quality and efficient health care, former President George W. Bush declared an executive order, stating a commitment to standards to achieve quality and efficient health care (US, 2006). It should reduce the calculated 98,000 losses of life caused annually by a lack of interoperability in care ICT systems (Venkatraman, Bala, Venkatesh, & Bates, 2008).

Within the US Government, standardization is an essential part of the "America Competes Reauthorization Act of 2010" (US, 2010). The NIST (National Institute of Standards and Technology) has gained additional funding to work together with the private sectors on developing standards for key technologies like cloud computing, emergency communication, and green ICT (Cooney, 2011). The act also includes the appointment of a new government function addressing the importance of the topic at a government level: The undersecretary of Commerce for Standards and Technology.

The US standardization system is highly decentralized, and the US administration does not intervene in the process, nor does it mandate any standards, which is contrary to European governments, but it requires US government agencies to participate in standardization (Ernst, 2010). The only requirement of the United States Standards Strategy is that "The process encourages coherence to avoid overlapping and conflicting standards" (ANSI, 2010).

Within the ICT standardization arena the influence of Asian countries, most notably China, is increasing (Jakobs, 2009a). China's latest plan for standardization defines standardization as an enabling platform for indigenous innovation: using standards as a tool for economic development (Ernst, 2010). Based on lessons learned from different ICT standards projects (Fomin, Su, & Gao, 2011; Steen, 2011; Stewart, Shen, Wang, & Graham, 2011), China's policy has moved from regulation to promotional activities, taking a more flexible and pragmatic approach and moving in the international domain from being a standards user, to a co-shaper and in some areas the lead shaper (Ernst, 2010).

Although oversimplified, and not covering the changing role of China, Ken Krechmer (SIIT Mailinglist, August 13<sup>th</sup> 2011) summarizes it as: "The EU funds their standards, seeing them as a governmental issue. America ignores their standards, seeing them as a commercial issue. China enforces their standards, seeing them as a policy issue".

### 1.3 Research motivation

The previous section shows the importance of the topic in general to both industry and government. This section narrows the scope and describes the problem situation, and that is the starting point of our research goal.

#### 1.3.1 Examples

We started this chapter with two examples of semantic errors. The following examples will show the impact of semantic ambiguity and will introduce the SETU semantic standard as an example of the hundreds of semantic standards that exist.

##### An example of semantic ambiguity: the mass-casualty incident

Interoperability is not only essential for economic reasons, but also for well-being. For instance, interoperability between all aid organizations (for example fire brigades, first-aid teams, hospitals, police and government officials) is essential for saving lives, but it is challenging because of the complex context. An example of where semantic ambiguity might lead to disastrous social grief is triage. Triage is a simple process of sorting victims into several categories (using color codes) when medical assistance is scarce. However different countries have slightly different triage categories; several use 4 categories, including the category used for victims who are dead, or whose injuries make survival unlikely. In the Netherlands for example it is category T4 and in Japan it is category 0. Semantically the distinction between "dead" and "will die" is a major difference, especially when one realizes that this triage classification is the only medical information used as a means of communication with relatives. Other organizations in other countries have realized that the distinction between dead and will die is medically not interesting, but that a distinction is needed because this information is used for other goals as well. They decided to add a fifth category, T5, meaning "dead". Nonetheless, some have implemented the four-category system, while others have implemented the five category system; the following statement says it all: "The German triage system also uses four, sometimes five color codes to denote the urgency of treatment" (Wikipedia, 2011b). T4 is now a classical example of semantic ambiguity as misunderstandings might easily arise about the meaning. Major disasters in countries like the Netherlands often make use of international medical assistance; e.g. victims of the New Years Eve disaster in Volendam were taken to hospitals in Belgium amongst others. Semantic ambiguity about (nearly) dead will hurt social well-being by causing social grief.

##### An example of a semantic standard: the SETU standard for reporting time & expenses

The SETU standard is a semantic standard supporting the processes of hiring temporary staff. It is a Dutch localization of the international HR-XML standard (Van Hillegersberg & Minnece, 2009), and is maintained by the SETU organization, a not-for-profit foundation ([www.setu.nl](http://www.setu.nl)). It has been acclaimed by the Dutch government as achieving interoperability within the process of hiring temporary staff through temporary staffing organizations. Since May 2009, SETU has been listed on the "Comply or Explain" list, which means

that every (semi) public organization in the Netherlands has to comply with the SETU standard when ordering temporary staff electronically.

SETU is a set of specifications that define the data and some processes that need to be exchanged between buyers and suppliers, including XML Schemas, for, amongst others, assignments, timecards, and invoices related to temporary staffing.

When using SETU, efficient interoperability is achieved between buyers and suppliers in relation to temporary staffing, without any vendor lock-in. The same SETU-based message exchange can be used continuously when for example the buyer replaces its supplier, or deals with multiple suppliers.

Without SETU the buyer and supplier would have to start an ICT project to discuss the communication related to electronic timecards, etc. This implementation project needs an investment that can only be used for this project for a certain time period (the contract period). The achievement of interoperability for this single project is costly and not efficient. The investment will be gone by the time the contract has ended and will not be continued. This is called the vendor lock-in as switching costs occur when changing supplier.

The selection by the Dutch government suggests an achievement of perfect interoperability, implying the SETU standard is high quality. However even the SETU standard could do a better job to achieve interoperability. Questions arise at the borders of the functional domain, for instance related to invoicing. Also the relationship with HR-XML is somewhat questionable (Chapter 10).

### 1.3.2 The problem

In 2009, the European Commission recognized the importance of the quality of standards and set a policy to “increase the quality, coherence and consistency of ICT standards”, albeit that it is more focused on speed (EC, 2009). Although not part of the policy, several suggestions have been made to improve the quality, for example (Pindar, 2010):

- Standards should be comprehensible, simple and easy to use so that they can be implemented better by users.
- Reduce the excessive number of cross-references between standards.
- Provide user-friendly guidelines for the use of the standards, free online abstracts, better online access to consultation drafts and simple electronic search functions.
- Standards should be designed and adapted to take account of the characteristics and environment of SMEs.

It shows that governments are experiencing problems related to the quality of standards and are setting up policies or guidelines to improve the quality. These problems are related to the lack of adoption of standards, lack of SME involvement and standards that are too complex. The later problem is especially related to the quality of standards.

Although relatively late, the importance of interoperability is now also understood by academics: there has been a large growth in publications in recent years on interoperability (Legner & Lebreton, 2007). Research is limited to the adoption of such standards (Steinfeld et al., 2007). However not all (adopted) standards will lead to interoperability. Not all standards are successful in achieving their goal, not all standards are adopted, not all standards are the best solution to the problem, etc.

The above shows the problems related to standardization in general, but the situation for semantic standards is more complex. Electronic data exchange along the supply chain has been discussed in the IS academic literature for many years and remains a practical problem for enterprises worldwide (Frick & Schubert, 2011). From the early 90's (since the rise of Internet-technology) significant budgets and efforts have been raised for the development of semantic, XML-based standards resulting in hundreds of such standards. Some of them are quite successful, judging from the adoption of those standards. Although these interoperability standards have been created for a range of industries (Zhao, Xia, & Shaw, 2005), problems seem to persist,

which might be related to the quality of the standards themselves, and the processes by which they are developed. The billions spent on standardization might have more of an impact on the desired level of interoperability if there was more focus on quality in both current research and practice.

Sherif, Egyedi, and Jakobs (2005) claim that their paper on the quality of standards was the first to address this topic, albeit only for technical standards. Semantic standards are important for inter-organizational interoperability and for solving data exchange problems. Communication between applications is mainly governed by technical and semantic standards (Bedini, Gardarin, & Nguyen, 2011). The advent of XML and web services, and more generally service-oriented architectures, has contributed greatly to the development of such standards-based integration solutions (Bedini et al., 2011). This has led to a problematic situation with the existence of countless (XML-based) dialects and languages defined by standard bodies, covering data dictionaries, messages, business processes, code lists, partner profiles, web service descriptions and so on (Bedini et al., 2011).

Lack of interoperability is often observed between different data exchange implementations in e-business, even when the systems implement the same standard specification (Brutti, De Sabbata, Frascella, Novelli, & Gessa, 2011), suggesting a lack of standard quality. Other research in the financial reporting domain show the same lack of interoperability, related to the implementations of the standard and its quality (Bovee, Ettredge, Srivastava, & Vasarhelyi, 2002; Zhu & Fu, 2009; Zhu & Wu, 2010). Also issues related to health care standards that tamper interoperability are widely-acclaimed (Figure 1), but have broad and even contradicting origins.

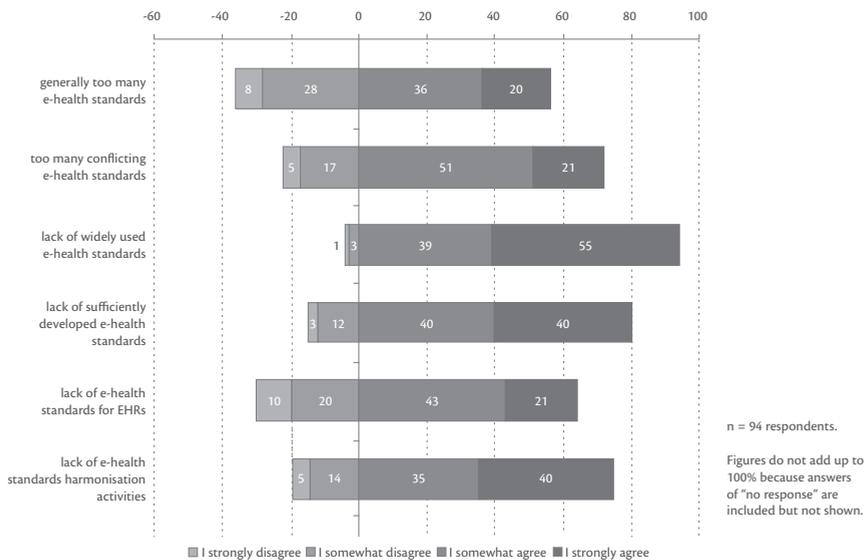


Figure 1 – Issues related to e-health standards (EC, 2008)

In the ICT domain, industry consortia set the vast majority of important standards, in contrast to formal standards organizations (Rada & Ketchell, 2000). However, governments prefer the official status that formal standards have to offer (Rada & Ketchell, 2000). This results in a mismatch especially if a European government wants to select ICT standards and prefers the formal ESOs. Although a fast track procedure for standards developed outside the ESOs is foreseen within ESOs, it does not solve the issue (EC, 2011c). In contrast to popular belief, government preference for ESOs is not related to openness: many of the industry consortia have open characteristics while the democratic process and free availability of formal standards is overestimated (Egyedi, 2003; West, 2007).

Industry consortia, such as SSO, are however growing in number and importance; they cannot be neglected anymore in government policies (EC, 2011c; Kroes, 2010). In the semantic area there is often one dedicated consortium that maintains one specific semantic standard for a specific domain. However both large industry consortia and formal bodies are aiming for the inclusion of more of these semantic standards. For instance OMG, The Open Group and W3C, are all industry consortia involved in semantic standards for different domains, and currently W3C is offering a free online platform that can be used by semantic standards initiatives. The formal bodies bring their formal status into play, while the industry consortia offer their expertise and flexible processes.

The SSOs are relatively young and the development of semantic standards is relatively new and complex, and mainly done by volunteers from different industries, many without relevant ICT and standardization knowledge. These SSOs are releasing standards with unproven quality and contribution to interoperability because many of the SSOs are process-oriented, not product-oriented. This means that for SSOs it is more important to monitor the timing of process steps than to monitor the quality of the outcome. Also because several SSOs are evaluated on their output numbers, which explains why SSOs are focusing on delivering as many standards as possible. In the long run, the low quality will have a boomerang effect on the SSO, as standards maintenance will become costly since updates are required to correct faults.

From the standards' users viewpoint other issues related to the unknown quality arise. For instance organizations are choosing (open) standards with unknown impacts on interoperability instead of high-quality standards. But also high and unexpected costs of standard implementation projects may occur because of all kinds of unknown issues relating to the standard. But in the end the most important aspect for users is that they do not achieve effective and efficient interoperability resulting in a reduction in economic welfare and social life.

In summary, the main problem is that the full potential of interoperability is not achieved because of sub optimal standard quality. Unfortunately without an understanding of quality, it will be hard to systematically improve the quality. Chapter 3 will gather more in-depth knowledge and understanding, and will validate our problem of the situation in practice.

## 1.4 Research goal

This section will describe the research intentions of this study. Based on the previously described motivation, the research goal and question are presented. Some scoping and context is given to improve the understanding of the research goal.

### 1.4.1 The goal and main research question

Based on the previous section, we have formulated two problem hypotheses:

1. The quality of semantic standards can be improved.
2. It is not possible to have a clear understanding of the quality of semantic standards.

The goal of this research is to create a better understanding of the quality of semantic standards as a starting point for improving standards' quality. Subsequently the potential of standards to achieve interoperability is likely to be improved as well. A detailed view of the quality of standards is needed before quality can become a leading managing instrument for SSOs. A quality measurement instrument will help to get this view on standards.

Our main research objective is to design an instrument to measure the quality of semantic standards that will aid standard developers in improving their standards. The resulting main research question is: *What are the characteristics of an instrument to measure the quality of semantic standards that will aid standard developers in improving their standards?*

### 1.4.2 Scope

To clarify the research goal this section will elaborate on the scope of this research. Standardization and interoperability are the central themes of this research. Their relation can be put simply as: standards are (one of the) mean(s) to achieve interoperability. The assumption is that high quality standards have (amongst others) a positive impact on interoperability, which as a statement is not a subject of validation within this research. However as the goal of standards is to improve interoperability, the improvement of the contribution to interoperability by a standard is certainly part of this research.

The introduction already stated that economic and social life is hardly possible without a certain level of interoperability. The general assumption is that a higher level of interoperability will lead to improved value chains and improved economic welfare and social life. This assumption is neither a subject of nor is it validated by this research.

Figure 2 structures the relevance of the instrument to economic and social life and also scopes the research.

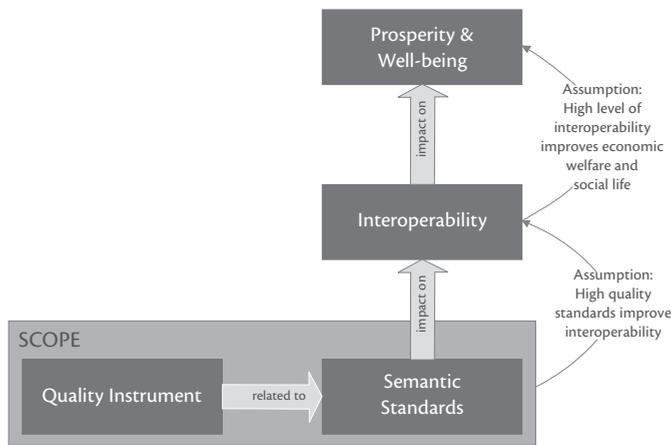


Figure 2 – The scope of the research

The primary scope of this research is to investigate the feasibility and the realization of an instrument to measure the quality of a semantic standard. The quality is related to the impact of the semantic standard on interoperability, since quality is the fitness for use and standards are used to achieve interoperability. The scope is on semantic standards, and not on technical standards amongst others, because these are crucial for inter-organizational interoperability. The instrument is a device, meaning that some kind of tooling might be part of the instrument; however the main emphasis will be on the models of the instrument. Professional software development resulting in (commercial) tools is not within the scope of this research.

The scope is deliberately limited to standard developers and SSOs, who are in general very open to improvements of their standards. There are other useful possibilities, such as rating or comparing standards, but these might result in a different instrument.

### 1.5 Research design and outline

In general, research can qualify as a behavioral science or design science (Hevner, March, Park, & Ram, 2004). While behavioral science is aimed at the prediction and explanation of behavior, the design science paradigm is aimed at the creation of new artifacts. The intention of our research is to generate knowledge to develop a solution for a real-life problem; a prescription-driven research approach (van Aken, 2004). Design

science is described as (Hevner et al., 2004): Design as an artifact, Design as a process and Design as a wicked problem. Different types of artifacts are common: constructs, models, methods, and instantiations. Design as an artifact means that the result is best described as a utility. The quality measurement instrument is a combination of artifacts and is a utility to be used by standard developers.

The process focus of design is related to two processes: Build and Evaluate. These will be used in the iterative design cycles. Finally design science is also seen as a problem solving paradigm. In this situation the problem is the unknown quality of semantic standards.

Our research question contains all three aspects of design science research, and therefore design science has been chosen as a primary research methodology. To be able to answer this question the research has been broken down into the following parts, based on the fundamentals of design science research, and are presented in the synchronous chapters in this thesis, in line with the methodology for design science research (DSRM) (Peffer, Tuunanen, Rothenberger, & Chatterjee, 2007).

Chapter	Objective	Research Methodologies
<b>1. Introduction: Research</b>	Set the scope for this research by describing the problem, research question and the approach of this research.	Design Research
<b>2. Introduction: Standards</b>	Introduction to the domain of standardization.	Literature Study
<b>3. Problem Relevance</b>	To verify if our research is solving a real-life problem.	Survey
<b>4. Research Gap</b>	Identify if there is a research gap related to this research and to verify the new and innovative characteristics of this research.	Systematic Literature Review
<b>5. Literature Review</b>	Describe the current knowledge in literature that can be used to build upon.	Literature Study
<b>6. Design Approach</b>	Setting up a design approach for building and evaluating the artifacts.	Design Research
<b>7. – 11. Build &amp; Evaluate</b>	Building and evaluation of artifacts according to the depicted design method.	Described in Chapter 6
<b>12. Validation</b>	The validation of this research and in particular the final version of the artifacts.	Field Test, Workshops, Survey
<b>13. Conclusions &amp; Further Research</b>	To answer the main research question & to define suggestions for further research.	Design Research

Table 1 – Chapters, objectives and research methodologies

Part 1, this introduction chapter, contains the scope of this research and the research outline, which will be followed by an introduction to the standards domain (Chapter 2). Part 2, the background section, will verify if important design science criteria are met, related to real-life problem relevance (Chapter 3), and its new and innovative characteristics (Chapter 4). It will also include a study of the current literature (Chapter 5). After part 2 both the problem relevance and the scientific state of the art will be known. Based on that knowledge the design method to build and evaluate iterations will be decided. The design approach and its initiation (carrying out the design approach) are captured within the chapters of part 3, but will be elaborated within Chapter 6. This chapter will discuss the research methodology in more detail and will describe how the fundamentals of design science research have been implemented within this research. Part 4 will finalize this thesis by validating the artifacts and this research in general (Chapter 12), before presenting the main conclusions and suggestions for future research (Chapter 13).

This study combines both quantitative and qualitative research in a logical way. First a quantitative study (the problem survey in Chapter 3) was carried out to establish the contours of the field, before a qualitative oriented study was designed and more in depth knowledge (Silverman, 2006) was acquired. During the

whole research process, including the build and evaluation phases, we have combined several qualitative and quantitative design approaches, a so called mixed method (Mingers, 2001).

This research structure is presented in Table 1, including the research methods and is graphically depicted in Figure 3. More information on the objectives and the structure of the Chapters 7 to 12 is presented in Chapter 6.

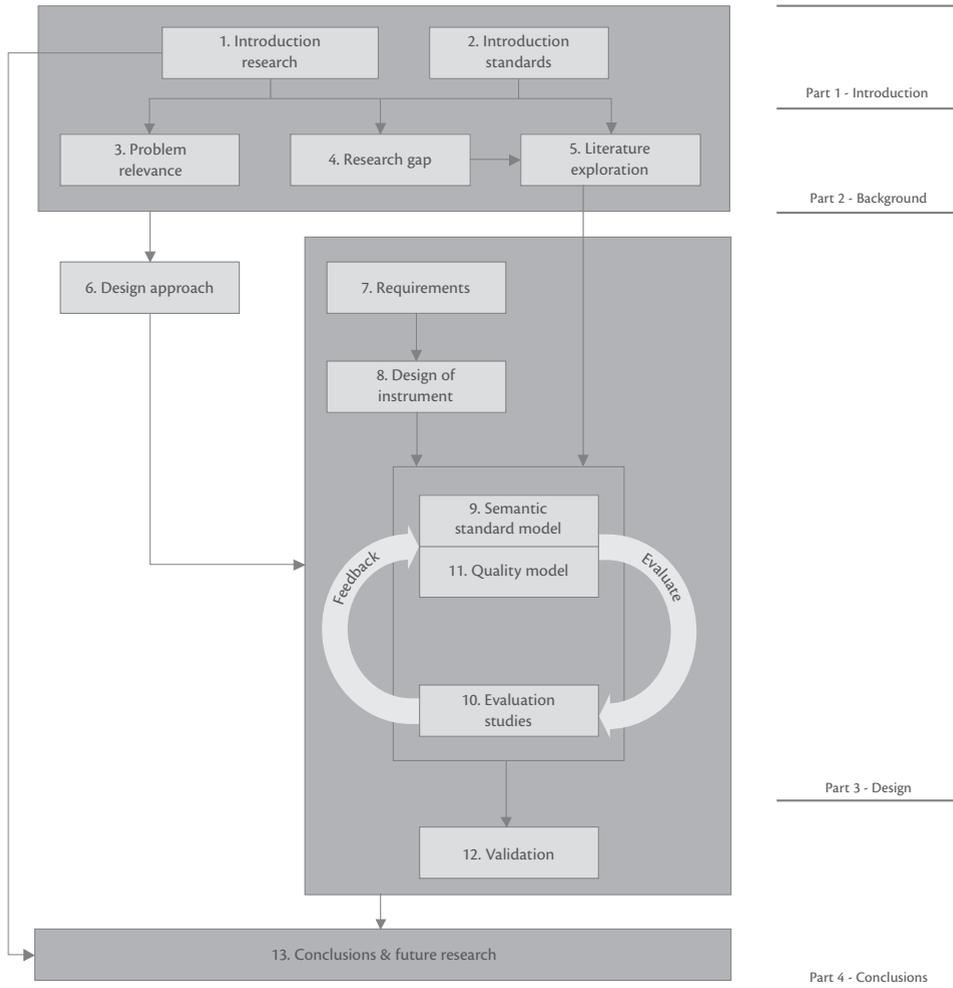


Figure 3 – Research and chapter outline



# Introduction to Standards

This chapter will provide a general introduction in standards and interoperability, for basic understanding of this domain. The scope is broader than semantic standards, and goes even beyond ICT standards. In-depth coverage of semantic standards will be described in Chapter 5.

### 2.1 The standards domain

Many studies describe standardization examples from recent and past times. Simons and De Vries (2002) include an extended list from McDonalds 'Hamburger', credit cards, light bulbs, petrol, paper formats up to screw threads, voltage, etc. Spivak and Brenner (2001) go even further back in time with examples starting from 3000 BC, but also include dramatic examples like the Baltimore fire (1904) where equipment from neighboring cities did not work because of a difference in hose couplings. Even older examples from the ancient Greeks (500,000 to 700,000 year ago) are present in literature (Anh, 2007).

Often used examples include ISO 9000 (and ISO 14000), AC/DC voltage (McNichol, 2006), and railway gauges (Spivak & Brenner, 2001), and more recently the VHS/betamax case (Park, 2006) and different DVD standards (Gauch, 2008; Van Wegberg, 2006). Regarding information technology the most common example studied in the nineties is the use of EDI (Electronic Data Interchange). EDI systems provide such widely cited benefits as reductions in paperwork, personnel and inventory costs, order lead time, and data errors (Wang & Seidmann, 1995). 75% of those studies, based on a systematic literature review, focused on the benefits of data exchange (Elgarah et al., 2005). These promised significant benefits by facilitating the exchange between business partners, reducing errors, increasing speed, cutting cost, and building as a competitive advantage, were not completely met since EDI standards failed to capture the requirements of the shared context (Damsgaard & Truex, 2000). EDI standards lacked a clear and complete lexicon, did not have fully specified grammar, and had virtually no semantics (Rukanova, Slooten, & Stegwee, 2006).

Although much attention has been given to technical tools (communication software) in the EDI-time span (Rukanova et al., 2006), the community expressed that "EDI is 90 per cent business and 10 per cent technology" (Swatman, Swatman, & Fowler, 1994). In practice, it is difficult to make a distinction between the technical aspects of integration and the organizational issues of implementation and integration (Swatman et al., 1994).

The arrival of XML, a standard foundation, has boosted the development of B2B standards (Zhao, Xia, & Shaw, 2007). Nowadays, XML-based standards are common, since XML-based standards involve fewer costs in comparison with EDI standards (Wigand, Steinfeld, & Markus, 2005). Many of the latest trends like web services, service oriented architectures, cloud computing, etc. are dependent on standards to fulfill their promise (Kreger, 2003; Zur Muehlen, Nickerson, & Swenson, 2005).

---

This chapter is based on:

(Folmer & Verhoosel, 2011): State of the Art on Semantic IS Standardization, Interoperability & Quality, Enschede: TNO, University of Twente, CTIT, NOiV.

### 2.1.1 Standards: typology

The famous quote by Tanenbaum (1989) says it all: “The nice thing about standards is that you have so many to choose from”. And “Researchers working on standards still struggle to order and understand existing standards” (Rukanova, 2005). Probably Cargill (1989) was among the first with a classification of voluntary and regulatory standards. There are major differences between different kinds of standards, for instance between pure technical standards and applied EDI standards for inter-organizational communication (Damsgaard & Truex, 2000). Therefore many studies have been performed to create some sort of order in the standardization domain, but several authors question definitions given by others, resulting in many different typologies. Arguably the most used definition of a standard is the definition used by ISO and IEC (De Vries, 2006; Spivak & Brenner, 2001; Van Wessel, 2008) as presented within Chapter 1, however this definition is arguable since it is too focused on traditional formal standardization bodies such as ISO (Van Wessel, 2008).

Several other definitions are used and discussed as well, for instance De Vries (2006) questioned the definition used by Jakobs: “A publicly available definitive specification of procedures, rules and requirements, issued by a legitimate and recognized authority through voluntary consensus building observing due process, that establishes the baseline of a common understanding of what a given system or service should offer”. And De Vries also questioned the definition used by Tassej, who defines an industry standard as “a set of specifications to which elements of products, processes, formats, or procedures under its jurisdiction must conform”. We will discuss different typologies based on different perspectives: General, Economic, Technical/ICT.

#### General perspective

Since there are many typologies, De Vries has set up a classification framework for those typologies; De Vries (2006) and also Van Wessel (2008) use the view of the subject matter in their own work:

1. Subject matter related classifications
  - a. Related to differences in entities
  - b. Related to requirements (basic, requiring, measurement)
2. Classifications related to standards development
  - a. Related to actors that are interested or involved
  - b. Related to organizations that set the standard
  - c. Related to the process of developing standards
3. Classifications related to standards use
  - a. Functional classification of standards
  - b. Standards related to business sectors
  - c. Classifications related to business models
  - d. Classification by extent of availability
  - e. Classification by degree of obligation

Another useful classification based on three axes comes from Spivak and Brenner (2001):

1. Level (from company, industry, to national, regional, international (voluntary), international (mandatory))
2. Subject (electrical equipment, clothing, transportation, food, ICT, etc)
3. Aspect (legislation, products standards, testing, inspection, environmental, etc)

Many authors, including Updegrove (1995) use defacto and dejure standards as a classification, based on the organization which develops and maintains the standard involved. Dejure standards are released by formal bodies like ISO, while defacto standards can be released by industry consortia or any kind of organization. As well as defacto and dejure, regulation and consortium standards are also commonly used (Updegrove, 2007).

Another classification is based on the organization that drives the process (De Vries, 2006; Van Wessel, 2008): Governmental, Formal, Consortium or Company. Or classification based on characteristics of the process (De Vries, 2006; Van Wessel, 2008): Anticipatory or participatory or responsive, open or closed, consensus or non-consensus. Or geographical classification (national, regional, international) and so on (De Vries, 2006).

On a higher level, Rukanova (2005) also made an attempt to classify standards on their abstraction level: Method, Meta-model, Concrete model, Operational standard.

All these different classifications can be mapped onto the earlier presented framework of classifications. The one to use depends on the intended goal and purpose of the classification; e.g. if you want to select standards that are obligatory by law then a classification based on the degree of obligation would make most sense. If you want to select standards for the healthcare industry, then a subject matter related classification seems obvious.

### Economic perspective

A classical definition by David and Greenstein described by Van Wessel (2008), and being reused by De Vries (2006), distinguishes:

1. Reference standards (or requiring standards, interference standards). Reference standards set requirements for entities or relations between entities.
2. (minimum) Quality standards (or measurement standards). Quality standards set requirements for entity characteristics to assure a certain level of quality (De Vries, 2007), where measurement standards provide methods to be used to check whether requiring standards' criteria have been met.
3. Interface or compatibility standards (or basic standards). These standards provide a structured description of (aspects of) interrelated entities, where compatibility deals with the fitting in order to function together.

David and Greenstein (1990) use the following for the classification of literature on compatibility standards in economics, as shown in Figure 4 (Reinstaller, 2008).

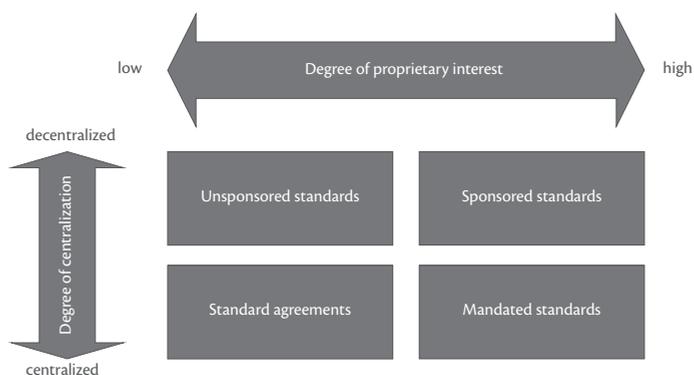


Figure 4 – Economic classification (Reinstaller, 2008)

Perera (2007) breaks down compatibility standards into horizontal (two functional equivalent objects (e.g. Telephones) and vertical (functionally different: Tracks and Trains or hardware and software) or backwards and forwards, but also introduces customer interface standards, in addition to the interference, quality and compatibility categories. Other economists also adapt the categories slightly by applying an economic subject classification, where one standard might fit in multiple classes (Blind, 2004; Swann, 2000, 2010): Compatibility/interface (e.g. USB interface), Minimum quality/safety (e.g. ISO 9000), Variety reduction (e.g. clothing sizes), Information standards (e.g. tax reporting).

## Technical / ICT perspective

The earlier mentioned typologies are valid for all kinds of standards. Our research scope is within the ICT domain, which justifies a look at specific technical and ICT typologies that exist as well. A typology based on the timing of the standard in relation to ICT products and services can be differentiated by anticipatory standards, enabling (participatory) standards and responsive standards (Sherif, 2006). For example SMS is an example of a responsive standard (the GSM system was already mature), while WAP is an example of a failing anticipatory standard.

Sherif (2006) continues with the introduction of a layered architecture for technical standards (Figure 5). The reference standards include well known examples like Volt, Watt, ASCII, the OSI-model, while examples of similarity standards are encryption algorithms and operating systems. Compatibility standards are usually profiles or implementation agreements to reduce the amount of options in a standard in order to achieve interoperability. Flexibility standards focus on compatible heterogeneity, that is, the capability of a single platform to interoperate with different systems and its upward and downward compatibility (Sherif, 2006).

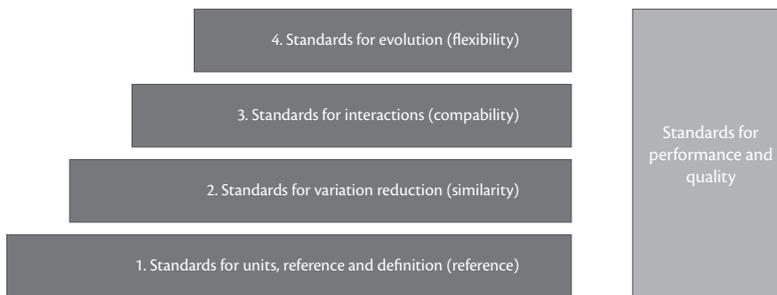


Figure 5 – Layered architecture for technical standards (Sherif, 2006)

Within the ICT domain, Cargill (1989) did some pioneering work by introducing the distinction between:

- Implementation and conceptual standards
- Product and process standards

There is a major distinction between e-business standards and traditional ICT standards (Zhao et al, 2007), which might explain why there are several typologies specific for e-business standards. An example of a classification needed for e-business is a pyramid construction with technology at the bottom (Albrecht et al., 2005):

- Foundation technology standards as fundament:
  - Data type standards
  - Scheme expression languages
  - Common communication methods
- On top of the fundament, the marketplace standards for defining the information exchange:
  - Business categorization
  - Product and service representation schemes
  - Shared transaction templates
- On top of the information, the commerce services and applications for defining the interaction:
  - Discovery technology
  - Transaction execution technology

Another more sophisticated classification for e-business has been made by Chari & Seshadri (2004), who use a layered approach: industry domain (domain independent or dependent), application domain (data, business or presentation logic) and integration level (transport, data format, process). And then use color codes to distinguish de jure standards from consortium standards.

Due to a rising star called “services”, Blind (2009) defines empirically-based taxonomies for services and for e-business. Although both taxonomies contain a second more detailed level, only the main items will be mentioned within Table 2.

<b>Taxonomy of standards for services:</b>	<b>Taxonomy of standards for e-business:</b>
• Service Management	• Environmental, Health and Safety Management
• Service Employee	• Customer Interaction
• Service Delivery	• Service Delivery
• Customer Interaction	• Data Flows and Information Systems
• Data Flows and Security	• Data Security

Table 2 – Taxonomies for services and e-business (Blind, 2009)

More specific for the e-business domain is the hierarchical and functional stratification (Zhao et al., 2005):

- e-Business standards (e.g. RosettaNet, MISMO, Papinet, STAR, etc.)
- Interaction standards (e.g. BTP, SAML, BPEL4WS, WSDM, etc.)
- Communication protocols (e.g. UDDI, WSDL, SOAP, etc.)
- Internet standards (e.g. HTTP, TCP/IP, XML, etc.)

This shows some resemblance to the Open System Interconnection (OSI) model, that consists of the following layers:

- Application: interacts with software applications.
- Presentation: establishes context between Application layer entities.
- Session: controls the dialogues (connections) between computers.
- Transport: transparent transfer of data between end users.
- Network: functional and procedural means of transferring data between networks.
- Data-Link: transfer data between network entities.
- Physical: electrical and physical specifications for devices.

Standards for the presentation and application level are often called semantic standards (Steinfeld et al, 2007), while the standards on the levels below are called syntactical or technical standards. The classification used by Steinfeld et al. (2007) decomposes the semantic standards into horizontal (cross-industry) and vertical (industry-specific) standards.

Although we have shown a broad range of classifications, many more classifications are possible, for instance based on interoperability levels, resulting in technical, semantic and organizational interoperability standards. Which we implicitly use in this research by calling it “semantic standards”. Other classifications might be the technological level (for instance OSI stack), or based on the philosophy of the approach.

### 2.1.2 Standards: the processes and the product

Based on the ISO booklet The Aims and Principles of Standardization, Spivak and Brenner (2001) mention the following generic aims of standardization:

- Simplification for society, prevents unneeded variation in products.
- Interchangeability: When varieties are limited interchangeability will increase.
- Standards as a means for communication: Communication between producer and consumer.
- Symbols and codes to reduce the effects of different languages.
- Safety: As well as specific safety products, a uniformity of product failure conditions.
- Consumer and community interest: Product labels like energy consumption, flammability.
- Reduction of trade barriers: to avoid the imposition of unique standards by nations to exclude the products of others.

Another aim of standardization can be commercial gain by vendors. Especially in ICT the commercial value of setting a standard that becomes defacto standard and is protected by intellectual property rights, can be enormous. ISO continues by defining the process of standardization, including two notes (Spivak & Brenner, 2001; Van Wessel, 2008): “The activity of establishing with regard to actual or potential problems, provisions for common and repeated use, aimed at the achievement of the optimum degree of order in a given context.

Note 1: In particular, the activity consists of the processes of formulating, issuing, and implementing standards.

Note 2: Important benefits of standardization are improvement of the suitability of products (including services) and processes for their intended purposes, prevention of barriers of trade and facilitation of technical cooperation”.

Another De Vries definition used by several others (Hanseth, Jacucci, Grisot, & Aanestad, 2006; Van Wessel, 2008) is: “Standardization is the activity of establishing and recording a limited set of solutions to actual or potential matching problems, directed at benefits for the party or parties involved, balancing their needs and intending and expecting that these solutions will be repeatedly or continuously used, during a certain period, by a substantial number of the parties for whom they are meant.”

From an economic perspective, the aim of a standardization process, and the criteria by which it needs to be judged, is twofold (Van Wegberg, 1999):

1. Develop and select the best standard, that is, the one that (over its lifetime) will generate the highest value to society as a whole (the stakeholders).
2. Organize this process of standards development and selection at the lowest transaction costs.

When transaction costs (of the development of the standard) are decreased, more parties try to get involved in the standardization process (Van Wegberg, 1999) since organizations only participate in the standardization process when the expected benefits are higher than the expected costs of participation. Zhao et al. (2007) mention three main reasons for participation in standards development:

1. Orient the standard to their own business practices and systems.
2. The better the standard and the faster it is developed, the greater the benefit there is for the developers who are also standard users.
3. Companies also benefit from in-depth discussions in the development process with their peers.

Considerable literature on standards lifecycles exists. Amongst others are Cargill (1995), De Vries (2007) and Egyedi and Blind (2008). Söderström (2004) compared seven different standards life cycle models, and built a new model based on that (Figure 6).

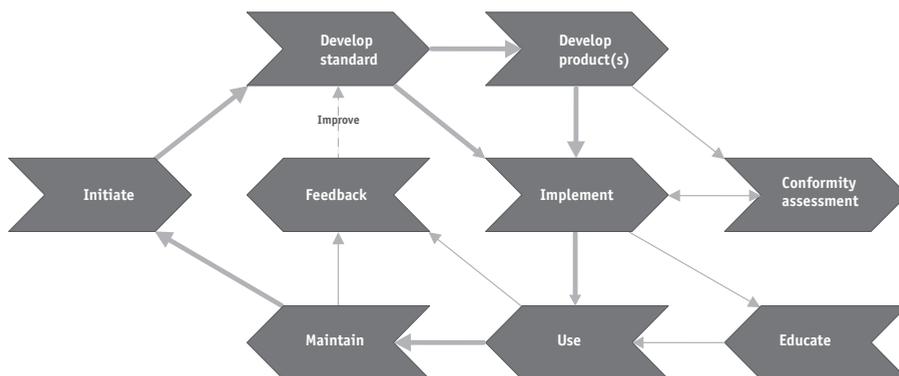


Figure 6 – Extended general lifecycle (Söderström, 2004)

From a standardization organization perspective, the life cycle of a standard is often simplified to a development and maintenance phase, each having its own process. Research often focuses on the development process, resulting in useful knowledge when involved in the understanding of the dynamics of standardization. A study on web services choreography standards (Nickerson & Zur Muehlen, 2006), showed that:

- Working groups in Internet standard development function as a population ecology, i.e. a living organism that lives and eventually dies.
- Standard developers function as part of an interactional field, whereby their actions are interdependent with those of other standard makers. (Standard makers are professionals who sometimes switch jobs but remain involved in standard making within the same workgroup.)
- The bylaws of the organization are the source of institutional stability in Internet standard making.

This contribution shows the importance of the standards organization, which will be discussed in the next section.

### 2.1.3 Standards: the organization

Different terms are used, but the most common is the Standards Development Organization (SDO), the organization that develops and maintains standards. More recently, the terms Standards Setting Organization (SSO) (Cargill & Bolin, 2007; Krechmer, 2006; Simcoe, 2007; West, 2007) and Standards Setting Body (SSB) (Jakobs, 2009b) or informal standards development organization (Song, Jiang, & Wu, 2007) are used. Often the term SDO is reserved for the formal/traditional development organizations (Cargill, 1989; Spivak & Brenner, 2001), while SSO includes all the organizations that develop standards, like OASIS, W3C and IETF. Since our research is not limited to formal standards, we use the term SSO for all organizations involved in standards development and maintenance.

The formal international SDOs include (Cargill, 1989; Frenkel, 1990; Simons & Vries, 2002; Song et al., 2007):

- Global: ISO, IEC en ITU
- Regional (Europe as an example): CEN, CENELEC, ETSI
- National: ANSI, NEN, DIN, BSI, etc.

Many authors describe the process of national, European and international formal standardization, most probably because it is fairly complex (Blind, 2004; Cargill, 1989; Cargill & Bolin, 2007; De Vries, 2007; Hesser & Czaya, 2007; Jakobs, 2009b; Simcoe, 2007; Spivak & Brenner, 2001).

However the world has changed, which many studies (Branscomb & Kahin, 1995; Cargill, 1995; Updegrave, 1995; Wagner, Cargill, & Slomovic, 1995) have shown, but was accurately described by (Hawkins, 2009): “By the late 1980s, spurred largely by the burgeoning Internet phenomenon, most of the significant standardization activity in computing and much of the telecom activity (especially in the higher value-added segments) was occurring in a rapidly expanding array of independent consortia that were dominated by major ICT vendors”.

Although ISO created a special committee for Information Technology (JTC1), consortia that have no relation to JTC1 are increasingly producing the important ICT standards (Rada, 1998). The result is that important ICT domain standardization organizations are not part of the formal SDO world, including organizations like W3C, OMG, OASIS, OAGI, GS1, and more specifically, all sector specific standardization organizations. This consortia movement has led to the fragmentation of standardization (Van Wegberg, 2006), and consortia now dominate the world of ICT standardization (Rada & Ketchell, 2000).

Different terms are used for these organizations including SSO, but also industrial consortia or fora, to stress the voluntary characteristics of contributing to the development of these standards. One of the reasons why ICT standards have been developed outside the traditional SDOs is the need for fast development times,

which is possible within SSOs (Rada, 2000; Simons & Vries, 2002; Van Wegberg, 2006), although the need for faster development times and the assumption that SDOs are slow is questionable (Mähönen, 2000).

Also mentioned is the role of consensus decision making which differs between formal SDOs (consensus) and consortia, which has an impact on the speed, and might have an impact on openness as well. This could be to the advantage of formal SDOs (Rada, 1995; Rada, Cargill, & Klensin, 1998). However this might be overtaken in practice (Egyedi, 2003). Especially since many consortia have copied procedures from formal SDOs like consensus decision making and the use of extensive voting, into their own procedures.

Other reasons that ICT standards are developed outside traditional SDOs may be confidentiality and intellectual property rights (De Vries, 2007; Simons & Vries, 2002). Others suggest economic motives:

- Van Wegberg (1999) states that to enable the development of a standard with low transaction costs, an increase in division of labor is needed, leading to specialized standardization bodies, which explains the growing number of highly specialized standardization bodies.
- “One indication of the perceived private and social gains from standardization is the increasing effort – much of which centers on information technology industries – to improve the performance of existing standards-setting bodies and, where that appears infeasible, to form new organizations” (David & Greenstein, 1990).

Many consortia have copied parts of the procedures from the formal bodies, for instance the use of voting.

Although these organizations appear to be growing in number and are influencing information technologies which are playing an increasingly important role in advanced economies (David & Greenstein, 1990), this has not been picked up accordingly in policies and research. Far less attention has been devoted by e.g. economists and political economists to examine the workings of standards-writing organizations (fora) (David & Greenstein, 1990). Consequently, not many studies are performed on how SSOs work in practice, with the exception of IETF (Simcoe, 2007). It is also not picked up in formal policies, for instance the European Union's policy, which did not keep pace with the market developments and stick to the old world: “The commissioners favor the adoption of a unified worldwide terminology, and consider that standards are only those developed by recognized standardization organizations. At the international level, ISO and IEC are such organizations; at the European level they are CEN, CENELEC, ETSI” (Bucciarelli, 1995).

The existing SSOs differ enormously in nature. Their credibility should not only depend on producing sound standards, but also on avoiding the temptation to abuse standards in making them a cash cow for the organization (Samuelson, 2006). In order to compare different SSOs, especially for the selection of an organization to support a standardization process, a framework has been set up, which has been tested on several SSOs, including OASIS, OMG, W3C and others (Jakobs & Kritzner, 2009).

Although it is impossible to state which SSO is the best, some think that IEEE is the best SSO (Cole, 2004), and others mention IETF as a good example of an open SSO (Krechmer, 2008). Related aspects are the speed of the process, consensus in decision making, and free or sold standards, all of which are addressed in the Communications of the ACM (Rada, 1995; Rada & Berg, 1995; Rada et al., 1998). The latter requires changes within the standardization world. Although one formal SDO does release their standards for free on the Internet (ITU-T). However these comparisons are mainly based on pragmatic aspects but more fundamental issues like the vision of the SSO are often not taken into account.

### Standards development

Other than the SSOs, some expert organizations exist to try to professionalize the process of standards development, including SES (Standards Engineering Society), IFAN (International Federation for the Application of Standards) and EURAS (European Academy for Standardization). The SES developed a standard on standards (Spivak & Brenner, 2001), and at the moment those are ANSI/SES standard ANSI/

SES-1-2002 - Recommended Practice for the Designation and Organization of Standards and SES 2:2006 - Model Procedure for the Development of Standards. Concomitantly, ISO has availed its ISO/IEC Directives Part 2, Rules for the structure and drafting of International Standards. The British Standards Institution (BSI) released a standard for standards as guidance in the development process of standards.

To professionalize the volunteers involved in standards making, several organizations developed guidelines for the development process (Freericks, 2010), some of which are specific for service standards:

- CEN: CHESSS: Guidance document for the preparation of service standards.
- ISO/IEC: Guide 76: Development of service standards.
- IFAN: Guide 3: Guidelines to assist members of committees in preparing user-oriented European standards.

One of the key challenges in the standardization process is to achieve active participation of different stakeholders. Different kinds of standards users exist (Jakobs & Kritznier, 2009):

- Direct users: users of standards; e.g. ICT vendors service providers.
- Mediators: e.g. consultants.
- Indirect users: users of standards implementations.

Hawkins (2009) describes the stakeholder triad, with ICT vendors, ICT Consumers and ICT Appliers as stakeholders that dominate the standards arena. Especially achieving consensus between different kinds of stakeholders within standardization workgroups is challenging, because business interests may be conflicting.

## **2.2 Standards and interoperability**

Standards are important for ensuring interoperability (Rada, 1993). “Standards are necessary both for integration and for interoperability” (Dogac, Kabak, Namli, & Okcan, 2008). “Adopting standards-based integration solutions is the most promising way to reduce the long-term costs of integration and facilitate a flexible infrastructure” (Chari & Seshadri, 2004). Some go even further: “Inter-organizational collaboration requires systems interoperability which is not possible in the absence of common standards” (Gerst, Bunduchi, & Williams, 2005). And the potential of standards, in relation to the problematic introduction of proprietary solutions, is shown in a case study from the automotive industry (Steinfeld et al., 2011a).

And although it is generally accepted that standards are needed to achieve interoperability: “Setting and adopting a common standard for B2B transactions, therefore, is a natural step to enhance compatibility or interoperability among companies, generating great value for individual firms and the industry overall” (Zhao et al., 2007), there is little evidence for that (Wybo & Goodhue, 1995).

Although much standardization literature describe standardization challenges or problems (for instance the adoption problem), real critical studies are scarce. One empirical study (Wybo & Goodhue, 1995) does not show the theoretical expected interdependence with the level of usage of semantic standards. One possible explanation is that data standards are not the only solution, e.g. some simple semantic inconsistencies might be easy to solve by mapping or transformation. Or the problems caused by semantically inconsistent data are smaller than presumed (Wybo & Goodhue, 1995). Thus, a semantic standard may not be the optimal solution (too complex/expensive) for a simple interoperability goal.

Other solutions might be found in the area of data fusion and information integration: a topic on which a lot of time is spent within large enterprises. Integration activities cover any form of information re-use, such as moving data from one application’s database to another’s, translating a message for business to business e-commerce, and providing access to structured data and documents via a web portal (Bernstein & Haas, 2008).

Common Semantics	Vertical Industry Language 1	Vertical Industry Language 2	...	Vertical Industry Language n
	Horizontal Language			
Common Syntax				
Common Messaging Mechanism				
Common Communication Mechanism				

Figure 7 – Framework for interoperability standards (Jain & Zhao, 2003)

A framework for interoperability containing different kinds of standards is presented by Jain & Zhao (2003). Figure 7 and Figure 8 contain the framework and the framework including exemplary standards. This research is focused on the common semantics: what we call semantic standards.

Common Semantics	Insurance (ACORD)	Human Resource (HR-XML)	...	Healthcare (HL7)
	Horizontal Language (ebXML)			
Common Syntax (XML)				
Common Messaging Mechanism (Web Services)	Service Composition (BPEL4WS)			
	Service Discovery (UDDI)			
	Service Description (WSDL)			
	XML Messaging (SOAP)			
Common Communication Mechanism (Internet)	Transport (HTTP, SMTP, FTP, BEEP)			
	Networking (TCP/IP)			

Figure 8 – Framework for interoperability including standards (Jain & Zhao, 2003)

Like standards, interoperability is a concept with many different meanings. A study on interoperability definitions found 22 different meanings (Kosanke, 2006). An often used definition is the definition from IEEE we presented in Chapter 1: “Interoperability is the ability of two or more systems or components to exchange information and to use the information that has been exchanged” (Legner & Lebreton, 2007; Rukanova et al., 2006). Another used definition is used by the U.S. Department of Defense in their LISI (Levels of Information Systems Interoperability): The ability of systems, units, or forces to provide services and accept services from other systems (Legner & Lebreton, 2007).

Based on a comparison of different definitions, Van Lier (2009) concludes that interoperability deals with the making of agreements on three levels:

- Technical (technical exchange)
- Semantic (content and meaning)
- Context (interpretation, processing, apply)

This seems in line with the European Interoperability Framework (EIF); it agrees that interoperability is more than a pure technical subject. The EIF version 1 divides interoperability into three layers (EC, 2004):

- Technical: Interconnecting computer systems and services on a technical level (e.g. data integration, message transfer, and network).
- Semantic: Creating a common understanding and guaranteeing process ability of exchanged information in a “meaningful manner” (e.g. data processing, data standards).
- Organizational: Definition of cross-organizational business goals and business process modeling (e.g. administrative issues, collaboration agreements).

The second version of the EIF added a new layer called legal interoperability for aligned legislation for cross border information exchange (EC, 2010e). Based on the original EIF, but with an additional distinction between technical and syntactic in line with ETSI, Kubicek and Cimander (2009) arrived at a four level interoperability approach which is similar to ETSI’s approach (Van der Veer & Wiles, 2006):

- Technical: Technically secure data transfer (signals).
- Syntactic: Processing of received data (data).
- Semantic: Processing and interpretation of received data (information).
- Organizational: Automatic linkage of processes among different systems (processes).

Pragmatic interoperability, the effect of data exchange, is sometimes used in combination with semantic interoperability as well (Asuncion & Van Sinderen, 2010).

### 2.2.1 Integration and interoperability

Interoperability is defined by coexistence, autonomy and a federated environment, whereas integration refers more to the concepts of coordination, coherence and uniformization (Chen, Doumeingts, & Vernadat, 2008). A fully integrated system is tightly coupled indicating that components are interdependent and cannot be separated. Interoperability means loosely coupled implying that components are connected and can interact but still contain their own logic of operation (Chen et al., 2008).

#### **A different, more sophisticated and focused view on interoperability**

A starting point for a more sophisticated view on interoperability might be the earlier presented OSI model (from physical connectivity, data link, network, transport, session, presentation, to the application level). The first four can be called “Bit Streams” while the upper three are called “Message Streams” (Libicki, 1995). Unfortunately the top layer (application) contains subjects like FTP or X.400 implying that semantic standards are much higher in the stack than can be expressed. Rukanova (2005) uses Stamper’s semiotic framework to define interoperability. This semiotic framework involves signs; organizations communicate in

signs, and for signs to have a meaning they need to be interpreted at six different levels: physical, empirical, syntax, semantic, pragmatic, and in the social world. Based on this fundament a distinction is made by Stegwee & Rukanova (2003) between interworkability, interoperability and interchangeability (Table 3), while the fundament is also used to define the concept of inter-organizational interoperability as “the ability of two or more socio-technical systems to exchange information, to interpret the information that has been exchanged and to act upon it in an appropriate manner” (Rukanova, 2005). According to Gerst, Iversen & Jakobs (2009), the distinction between “e-business” and “infrastructure” is artificial, and they state that any assessment of the effect of standards on e-business has to take all the standard layers into account. Rukanova’s definition takes this into account.

Type	Purpose	Technical	Human	Process
Interconnectivity	Enables two systems to communicate with each other	Communication standards, like TCP/IP or X.25	Communication systems like speech and writing	Providing for external inputs and outputs
Interchangeability	Enables two systems to exchange information	Data representation standards, like ASCII or HTML	Language systems like natural language and vocabularies	Displaying the same behavior in terms of input/output
Interoperability	Enables two systems to operate together as one	Interaction standards like SMTP or SOAP	Behavioral scenarios and procedures, attached to e.g. military orders	Providing for external controls on process behavior

Table 3 – Interconnectivity, Interchangeability & Interoperability (Stegwee & Rukanova, 2003)

Kosanke (2006) shows that it gets complicated when these terms are also used in an IEC study, albeit differently. Kosanke describes the levels from IEC TC 65/290/DC, with increasing compatibility (Figure 9).

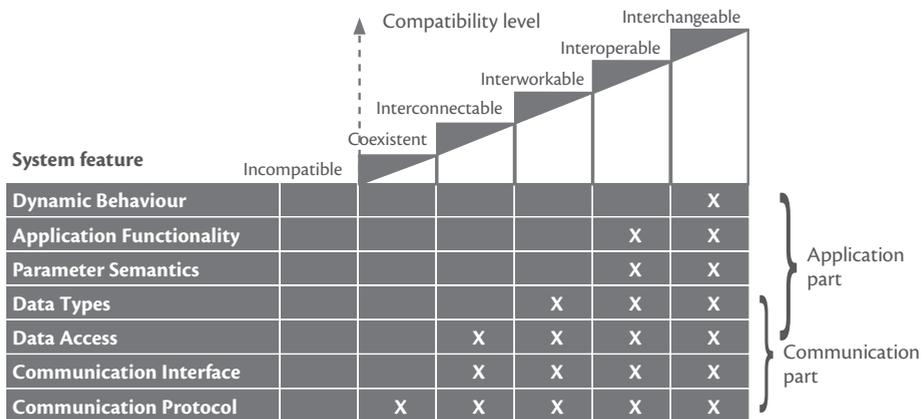


Figure 9 – IEC 65/290/DC compatibility levels (Kosanke, 2006)

The three most interesting top level definitions (from IEC) for the three terms are (Kosanke, 2006):

- Interworkability: ability of two or more devices to support transfer of device parameters.
- Interoperability: ability of two or more devices to work together in one or more applications.
- Interchangeability: ability of two or more devices to replace each other in working together in one or more application.

And Kosanke (2006) maps both models on each other that shows, interestingly, that both have a complete different opinion about the definition of interchangeability:

IEC TC 65/290/DC	Stegwee and Rukanova
	interconnectivity
interworkability	interchangeability
interoperability	interoperability
interchangeability	

Table 4 – The mapping of categories (Kosanke, 2006)

The term networkability is recently introduced and defined as “the ability of any number (m) of suppliers to speak the “same language” with any number (n) of customers at the interfaces between business processes and systems” (Falge et al., 2012).

We stick to the term inter-organizational *interoperability* which is a contrast to other terms like interchangeability more commonly grounded. We use inter-organizational to stress the automated communications between organizations (Rukanova, 2005), in line with a distinction based on the organization perspective (Benders, Batenburg, & Van Der Blonk, 2006):

- Intra-organizational standardization: Common reporting routines for example. However, in practice standardization often occurs at a system level (e.g. SAP for everything).
- Inter-organizational homogenization: “Homogenization between organizations is considerably more complex than the explicit motive of achieving common working procedures within an organization” (Benders et al., 2006).

Another interesting view is the direction of integration (Frick & Schubert, 2011), which is also applicable to interoperability:

- *Vertical* Integration: The partners are in the same industry sector but at different positions in the supply chain.
- *Horizontal* Integration: The partners are in the same industry sector and at the same position in the supply chain.
- *Diagonal* Integration: The partners are in different industry sector and at different position in the supply chain.

Inter-organizational interoperability refers also to the often used term Inter-Organizational (Information) System (IOS), for example used by (Lu, Huang, & Heng, 2006; Rukanova, Wigand, & Tan, 2009). IOS is defined as an automated information system shared by two or more companies (Cash Jr & Konsynski, 1985). Johnston & Vitale (1988) add: “to facilitate the creation, storage, transformation and transmission of information”.

Johnston and Vitale (1988) made the distinction in the IOS between content platform, delivery platform and trading partner base, and categorize different types of IOS based on:

- Business purpose.
- Relationship between the sponsoring organization and the other participants.
- Information function.

The value of an IOS is expressed in the following quote (Lu et al., 2006): “The strategic value of IOS has been well recognized for its real-time interaction, higher transaction security, more efficient and quicker payments, rapid response, reduced search costs, reduction in inventory and tighter link to customers. These benefits enable all parties to have high operational efficiency and capability, and more and more corporations tend to adopt IOS in order to gain competitive advantages.” The above definition of IOS encompasses many systems such as extranets, EDI, Internet EDI, B2B e-commerce and e-SCM. Zhu et al. (2006) also use IOS, and makes

a distinction with EDI through the use of the term Internet-based IOS. Internet-based IOS is characterized as being, on the content side: based on open XML based standards, low complexity and not that partner-specific; while on the delivery side: based on open Internet communication protocols, highly interoperable and low communication costs. It also has a broad trading partner scope. Based on these characteristics, this can also be called an *open standards IOS*. In summary, IOS is a broad term including concepts like data integration, but it differs from normal internal distributed systems by its ability to exchange information with the outside world (Johnston & Vitale, 1988). Inter-organizational relationships discriminate themselves by having the following characteristics (Löwer, 2005):

- Goal: Efficiency
- Direction: Vertical
- Resources: Coordinated
- Contract: Neo-classical
- Activities: Primary
- Formalization: High

### 2.2.2 Other frameworks and maturity model for interoperability

Interoperability is seen as an extremely important topic for an organization's ICT strategy and it is on the top of every CIO's wish list (Park & Ram, 2004), which might explain the abundance of interoperability frameworks. Architecture frameworks are often used in ICT, like for instance the Zachman Framework (Zachman, 1997), and these frameworks can also be used to look at interoperability. There are also dedicated interoperability frameworks as, for example, LISI (Kasunic & Anderson, 2004) from the American Department of Defense and the Athena framework (Berre et al., 2007) developed within a European Union funded project. Based on the work of Athena, a framework for Enterprise Interoperability has been developed, which is in the progress of becoming a CEN/ISO standard 11354-1 (Naudet, Latour, Guedria, & Chen, 2010).

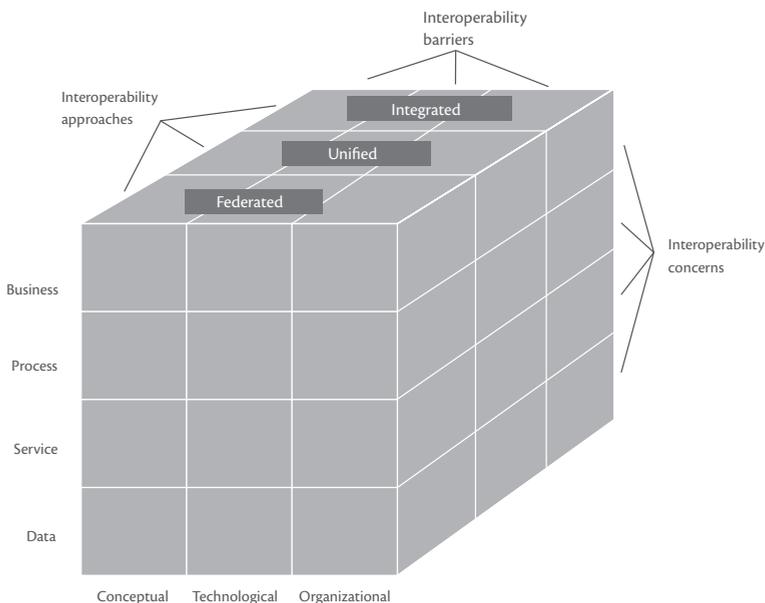


Figure 10 – Framework for enterprise interoperability (draft CEN/ISO 11354-1) (Dogac, Pattenden, & Zelm, 2010)

The *interoperability approach* is the desired level of integration; these levels are standardized in ISO 14258 (Kosanke, 2006). An *interoperability barrier* viewpoint has been identified to capture the incompatibilities and mismatches that obstruct the sharing and exchanging of information and other entities. Three categories of barriers are defined: conceptual, technological and organizational. *Interoperability concerns* defines the content of interoperation that may take place at various levels of the enterprise (data, service, process, business) (Ullberg, Chen, & Johnson, 2009).

The FnES report sums up several interoperability frameworks (Dogac et al., 2010), including the CEN/ISO 11354 framework as presented (Table 5).

Organisation	Name/Description
ISO 15745	Framework for Application Intergration
CEN/ISO 11354	Requirements for establishing manufacturing enterprise process interoperability
ATHENA FP6 IP	BIF: Business Interoperability Framework
CEN-ISSS EBIF	CEN eBusiness Interoperability Roadmap
UN/CEFACT	UN/CEFACT e-Business framework
OMG	Service Driven Architecture
iDABC	European Interoperability Framework for Pan-European eGovernment Services

Table 5 – Interoperability frameworks (Dogac et al., 2010)

### Interoperability maturity model

A maturity model exists for the measurement of the level of enterprise interoperability and it is similar to the CMMi model for software engineering. The LISI interoperability maturity model was set up in 1993, and it is also made up of five levels (Kasunic & Anderson, 2004), with a technical focus. LISI is much more than 5 interoperability levels. It contains several models, and an assessment process containing interoperability metrics. It contains a questionnaire for the identification of the appropriate interoperability level (Tolk, 2003) and an interoperability scorecard including quality attributes associated with interoperability (Kasunic & Anderson, 2004). These attribute measures are: connectivity, capacity, system overload, underutilization, under capacity, data latency and information interpretation and utilization, showing the technical emphasis.

However development has begun for an Enterprise Interoperability Maturity Model (EIMM) that builds upon the framework of enterprise interoperability (ISO 11354-1) as presented earlier. The EIMM (Berre et al., 2007; Knothe & Jochem, 2007) or MMEI (Maturity Model for Enterprise Interoperability) (Guedria, Chen, & Naudet, 2009) as it is known nowadays, contains 5 levels: unprepared (level 0), defined (level 1), aligned (level 2), organized (level 3) and adapted (level 4), and it includes metrics as well. Since the model is fairly new, usage is limited, but this might change when this model is given an ISO (11354-2) status.

Less enhanced, but potentially better resembling practice is the B2B Integration Maturity Model (BIMM), containing three perspectives on integration (Frick & Schubert, 2011):

- *Technical* integration describes how information is processed and shared electronically within and across organizations.
- *Organizational* integration refers to the organizational structures and processes, which are put in place to improve the efficiency and effectiveness of the supply chain.
- *Institutional* integration describes the formal and informal agreements, which govern inter-organizational relationships (governance structures).

### 2.2.3 The impact of interoperability

Very few publications address the impact of interoperability (Legner & Lebreton, 2007). Probably the first and most used is the US automotive case, suggesting low interoperability costs US \$1 billion per year (Brunnermeier & Martin, 2002). This study separates costs into:

- Avoidance costs (e.g. Investments to avoid future costs)
- Mitigation costs (e.g. Additional coordination costs)
- Delay costs (e.g. Loss of market share because of late entry)

Follow up studies by NIST researchers showed that the US \$1 billion was too optimistic and the new result showed a waste of US \$5 billion annual due to lack of interoperability resulting in order delays amongst others. For the electronics industry this is calculated to US \$3.9 billion annually (Steinfeld et al., 2011a).

Another study within the capital facilities industries contains a conservative estimate of US \$15.8 billion on inadequate interoperability costs (Gallaher, O'Connor, Dettbarn Jr, & Gilday, 2004). The case of the electro technical industry (Nelson, Shoonmaker, Shaw, Shen, & Wang, 2002) does not quantify, but shows a return on investment of less than 2 years (both sides), a reduction of transaction costs and cycle time. Based on the work within the European Framework project Athena, an interoperability costs breakdown is presented (Legner & Lebreton, 2007):

- Connectivity costs (per partner): Costs to establish or improve partner relations.
- Coordination costs (per transaction): Costs to enable and execute transactions.
- Control costs (per transaction): Costs to monitor transactions.

This work has led to the Interoperability Impact Assessment Model (IIAM) which shows the direct and strategic impact of investments in interoperability (Lebreton & Legner, 2007).

The healthcare domain also demonstrates the importance of interoperability and standardization to society. Venkatram et al. (2008), highlighted the relevance by citing reports from the Institute of Medicine about the errors in healthcare. The figures are impressive: 98,000 people die in hospitals due to errors (1999), and these errors costs hospitals US \$29 billion every year, while three out of four errors can be eliminated by better use of information technology. The lack of standardization and integration among the systems has made it difficult to reduce the medical errors. Lack of integration and data standardization is making health care services inefficient and costly (Venkatraman et al., 2008).

Although the impact of interoperability is most often related to effectiveness and efficiency, interoperability has also impact on innovation. It fosters innovation and opens the way for new business opportunities, in particularly businesses related to aggregators. These aggregators combine data from different sources, that have become accessible by having interoperability, and offer new services based on that data.

# Problem Relevance

In this chapter we take a closer look at the quality of semantics standards and their development processes, and survey the current state of quality and adoption of semantic standards through a survey of standards developers. The result of the survey is the evidence and indicates the practical relevance of the problem, the quality of semantic standards. The remainder of this chapter will present the research approach, the survey population and the results of this survey, before answering the main research question in the concluding section.

### **3.1 Research approach**

#### *3.1.1 Research framework*

As the starting point for studying the problem relevance, a research framework was developed. The starting point was the research question defined in generic terms: *Is there a need, based on the current situation and experienced problems, for a solution?* In relation to the subject of semantic standards quality, this can be formulated as: *Is there, based on the current standards development processes and experienced interoperability and adoption problems, a need to elicit the quality of semantic standards?* In this research, such elicitation involves a quality measurement instrument for semantic standards.

The main research question contains four concepts each giving rise to a proposition in our research framework. The research framework including propositions is depicted in Figure 11.

#### **Proposition for the current standards development process**

One of the issues in this research is whether the quality of semantic standards can be improved. Therefore it is interesting to know whether the current standards development processes include steps focused on quality aspects of standards, and whether there is room for further improvement.

*Proposition 1: The quality of standards can be increased by improving the standards development process.*

#### **Proposition for the interoperability problem**

If there is room for quality improvement, that would only make sense if it leads to better semantic interoperability.

*Proposition 2: Improved quality of standards leads to improved interoperability.*

#### **Proposition for the adoption problem**

Even if high-quality semantic standards improve semantic interoperability, such potential only materializes when the standard is actually being adopted in processes and systems.

---

This chapter is based on:

(Folmer, Oude Luttighuis, & van Hillegersberg, 2011a): Do semantic standards lack quality? A survey among 34 semantic standards. *Electronic Markets*, 21(2), 99-111.

*Proposition 3: High-quality standards will have a better chance of being adopted.*

### **Proposition for the desired assessment and visibility of quality**

If the interoperability and adoption problems (addressed in propositions 2 and 3) are influenced by the quality of standards, we need to verify whether the proposed solution could contribute to solving these problems. Is transparency of the quality of the standard valuable for standards developers? And, if so, do standards developers value an instrument for this?

*Proposition 4: There is a need to make the quality of the standard visible by assessment.*

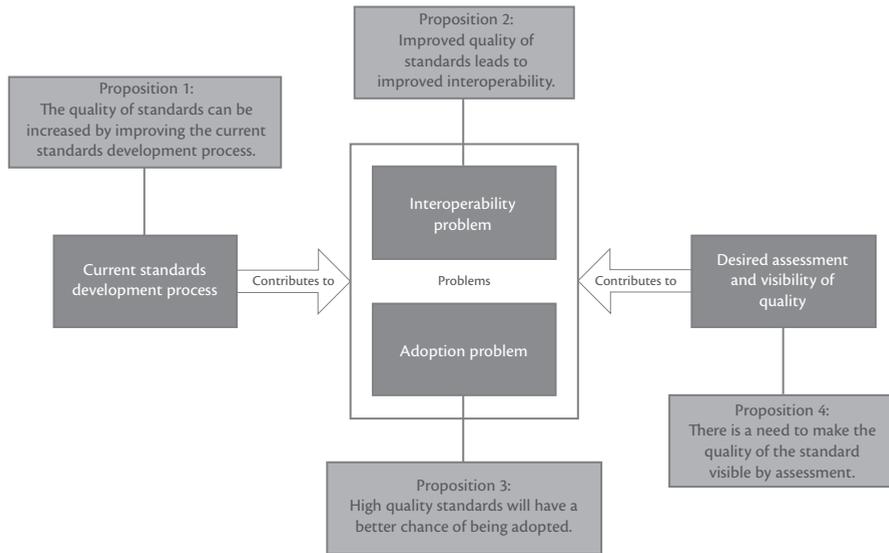


Figure 11 – Research framework - propositions

We conducted a survey to test these propositions. Surveys are often appropriate for problem clarification and verification of problem relevance. They offer a way of getting more stakeholders involved and getting structured and comparable results in a time-efficient manner (Creswell, 2009). Other methods, such as interviews or focus groups, would limit us too much in the scope of involving standards and respondents.

Our research addresses semantic standards in general, so that the survey had to cover a broad range of semantic standards. The intended respondents were standards developers from SSOs. Unfortunately, to the best of our knowledge, no up-to-date list of semantic standards existed upfront. Nevertheless our professional networks was activated and a list of possible respondents was set up, mainly from the Netherlands, Belgium, and Germany. In order to get additional and more international respondents, literature on semantic standards was assessed. Following other research (Zhao et al., 2005), xml.org was used. This list was enhanced with other semantic standards mentioned in literature (Hasselbring, 2000; Markus et al., 2006; Nelson, Shaw, & Qualls, 2005; Steinfield et al., 2007). The Internet was searched for standards developers involved in those standards. In cases where the standards developers could not be identified, we decided to send the invitation to the general e-mail address of the development organization.

The survey was designed using generally accepted survey principles (Creswell, 2009; Maxwell, 2005). The questionnaire consisted of a set of questions representing the propositions, and additional items to get background information (see Table 6). For comparability reasons, only closed questions are used, with the exception of three open (control) questions about the background of the respondent (see Appendix A). Invitations were limited to three people per semantic standard. All questions used the same five-point scale (Strongly disagree / Disagree / Partly disagree and Partly agree / Agree / Strongly agree). Several questions were deliberately formulated as a negation.

Subsequently, the survey results for each statement were analyzed and related to the propositions. A correlation analysis was done to find additional insights. Section 3.2 will relate the survey results to the propositions, and give possible explanations for the results.

Research	Concepts	Aspect/indicator	Statement
Is there, based on the current standards development processes and experienced interoperability and adoption problems, a need to elicit the quality of semantic standards?	Current standards development process	Quality as part of the current process	"Quality assurance is an explicit part of our current development process of the standard"
		Quality end-check implemented	"There is not a minimum quality check in place before the standard is released"
		Usage of tools for quality	"An instrument/tool is used to measure the quality of our standard"
		Perceived need for quality improvement	"The quality of the current standard can be improved"
	Interoperability problem	Avoidable errors	"New or updated releases cover avoidable corrections to the previous versions of our standard"
		Achieved interoperability	"The achieved interoperability is worse than expected"
		Influence of standard on interoperability	"Currently the achieved interoperability is affected by the limitations of our standard"
		Future interoperability	"Improvements to the quality of our standard will lead to improved implementations and ultimo lead to improved interoperability"
	Adoption problem	Current Adoption	"The current adoption is better than expected"
		Influence of standard on adoption	"Design choices of the standard have influenced the adoption process"
		Future adoption	"The adoption will be more successful when the quality of the standard is explicitly known to the users, and proven sufficient or improved"
	Desired assessment and visibility of quality	Minimum quality for interoperability	"A minimum quality level of our standard is needed to achieve interoperability"
		Minimum quality for adoption	"A minimum quality level of our standard is needed for high adoption rates"
		Potential usage of instrument	"I will not use an instrument/tool to measure the quality of the standard when it will be available"
Assessment of quality		"It would be helpful to have an instrument/tool/ knowledge to gain insight in the quality of the standard"	
Visibility of quality		"If the quality of the standard is not known then it is hard to improve the standard"	

Table 6 – Structure of survey statements

### 3.1.2 Limitations

Our research approach implies several limitations. First, it does not address the factors that influence quality. Although the characteristics of the development process of standards is expected to influence the quality of the standard, just as the management process of the standard and other context dependencies, this is out of scope for this problem survey. Another limitation is the broad definition and invitation of semantic standards, resulting in a heterogeneous group of semantic standards, making it hard to generalize our results to all individual or sub-groups of semantic standards. Finally, although we broadly defined quality as “fitness for use”, implying the relation with achieving interoperability, we are aware that in practice different views on quality exist. Some relate quality solely to the specification document, while others see quality as the adoption success of the standard. To align responses to some extent, the survey started with presenting a definition of quality, but still we are aware that some respondents will use a different view on quality.

### 3.1.3 Survey population

The survey was held from August 25<sup>th</sup> to September 25<sup>th</sup>, 2009. In total, 111 persons were invited, of which 48 responded, yielding a response rate of 43.6%. These 48 respondents represent 34 different semantic standards. Among these 34 standards were both international (e.g. HL7) and national (e.g. SETU), both governmental (e.g. StUF) and industry (e.g. Chem eStandards), and from different industry domains (e.g. healthcare, education, tourism, agriculture, finance, etc.).

Table 7 lists the semantic standards covered in the survey responses, classified by industry segmentation (based on SAP Industry Solution Maps (SAP, 2009) which has appropriate aggregation level and alignment with practice, in contrast to the official NACE classification), and with the number of respondents (N). Based on this initial list, a more comprehensive list of standards is published and maintained on [semanticstandards.org](http://semanticstandards.org).

#	Class	Name of Standard	Short description	N
1	Financial	ACORD	Exchange of insurance data	2
2	Public	Aquo	Information exchange about water management	1
3	Manufacturing	bcXML	Product modeling with integrated parametric geometry	1
4	Public	CEN/ISO EN 13606	Exchange between Electronic Health Records	2
5	Manufacturing/Trade	Chem eStandards	Chemical Industry B2B data exchange	2
6	Public	Content Packaging	Interoperability between digital learning related systems	1
7	Public	Content-ZoekProfiel	Standardization of the description of, mostly Dutch, learning objects	2
8	Manufacturing/Trade	Edibulb	Data exchange between bulb growers, intermediaries and bulb traders	1
9	Service / Trade	EDSN	Data exchange standard for the energy domain	1
10	Public	ELD	Exchange of student information	1
11	Service	ELSSI-EMD	Data for use in assessment procedures	1
12	Service / Trade	eTOUR	Interoperability in tourism (accommodation, events, activities, attractions, food & beverage)	1
13	Manufacturing/Trade	Floecom	Interoperability in the supply chain from grower to retail	2
14	Financial	FpML	Trade processing between firms for Over The Counter (OTC) Derivatives	1
15	Manufacturing	gbXML	Simplified and distributed data exchange for complex building resource analysis simulation engines	1
16	Trade	GS1 System	Identification of products, locations, documents, etc.	1
17	Public	HL7	Meaningful exchange of health information across different healthcare information systems	2
18	Public	ISO/IEC 19796-1	Quality Approaches for Learning, Education and Training	1
19	Manufacturing/Trade	JIS	Information exchange between jewelers and suppliers	1
20	Financial	MDDL/FIX	Financial services trading and market data	1
21	Public	MetaLex	Standard for legal sources	1
22	Public	NEN3610	Exchange of geo-information in different communities (spatial planning, cultural heritage, etc.)	2
23	Service	OTA	Business information and transaction standards for the travel industry	1
24	Public	OWMS	Metadata standard for Dutch government organizations	1
25	Manufacturing/Trade	papiNet	Different ways of electronic message exchange in the forest and paper industry	1
26	Service/Trade	SETU	Exchange of information in the staffing industry	3
27	Service	Shortsea XML	Interoperability between systems in short sea shipping	1
28	Service/Trade	SIDES	Sharing of recruiting and staffing data between suppliers, customers and third parties.	3
29	Public	SIF	Administrative and Instructional applications within a school, school district and State	2
30	Public	SNOMED CT	Connecting clinical domains and cross country borders	1
31	Public	StUF	Information exchange between Dutch municipal systems	2
32	Public	SuwiML	Structured information exchange in social security	1
33	Public/Financial	VEKTIS	The sending of declarations by health practitioners to several health insurers and also the receiving of declarations by a health insurer from several health practitioners	1
34	Financial	XBRL	Taxonomies developed locally/regionally in use globally	2

Table 7 – List of semantic standards covered by the responses

### 3.2 Results

The detailed survey results, including correlation analysis of the question results, are presented in Appendix B. Significant correlations and other survey results that contribute to the propositions are mentioned when we review the propositions in this section.

Proposition 1: The quality of standards can be increased by improving the current standards development process.

This proposition was agreed upon by 64.6% of the respondents, while 8.3% disagreed. Other results show that quality is embedded within development organizations as 77.1% of the respondents have quality assurance as an explicit part of their standards development process. And 81.3% already have some kind of minimum quality check in place before a standard is released. Although the survey did not ask for specifics, this quality check could be in the form of a final review before a new (version of the) standard is released.

It seems that current standards contain avoidable errors since 45.8% of the respondents stated that new or updated releases of their standard include corrections to avoidable mistakes in previous versions. This correlates to the statement about whether the quality of the standard can be improved ( $P=0.29$ ;  $p=0.05$ ). This result suggests there is considerable room for improvement within the current standards. But, is there any value in additional quality? This question is covered by the second proposition.

Proposition 2: Improved quality of standards leads to improved interoperability.

A substantial 66.7% of the respondents agreed with this proposition; only 8.4% disagreed. An even higher percentage (89.6%) viewed that a minimum quality level is a necessary requirement for interoperability. These figures lead us to conclude that the respondents correlate the quality of a standard with the achieved interoperability. At the same time, 64.6% disagreed with the statement that the achieved interoperability is affected by the limitations of the standard. Respondents seemed to anticipate and accept the interoperability level achieved. Only 10.4% said that interoperability is worse than expected. This satisfaction, or acceptance, of achieved interoperability may seem surprising in relation to interoperability problems in practice. However it might be explained by the population of the survey, consisting of standards developers. Responses from standards implementers or users might lead to different results, because they might have a different opinion of interoperability in practice.

We also see a positive correlation between a standard's achieved interoperability and whether there is quality assurance as part of the development process of the standard: Standards that have quality assurance as part of their current development process also have a minimum quality check in place ( $P=0.57$ ;  $p=0.00$ ), and rank high on the achieved interoperability ( $P=0.32$ ;  $p=0.03$ ). The data also show that where the quality of the standard could be improved the achieved interoperability is actually worse than expected ( $P=0.32$ ;  $p=0.03$ ). These results all confirm the positive correlation between quality and the achieved interoperability. This is in line with literature suggesting that the need for interoperability is one of the key drivers for the development of standards (Nelson et al., 2005). Will the same hold for adoption, that is, will quality improvement increase adoption rates? This is addressed by the third proposition.

Proposition 3: High-quality standards will have a better chance of being adopted.

Although 60.4% saw a relation between design choices of the standard and the adoption of the standard, this proposition is not completely supported by this survey. The results show more diverse opinions on this topic; 37.5% agreed and 37.5% disagreed with the question whether adoption will be more successful when the quality of the standard is known, proven sufficient, or improved.

Also, several respondents annotated their responses with factors other than quality yet more critical to the adoption of the standard:

"Don't forget the important role of communities and the community owner(s) (dominant players). Often these companies have a strong influence on the adoption and quality of standards."

"The degree of adoption depends on many other things than just the quality of the standard."

“So, although improving the standard itself is always a worthy goal, if the action to be transacted cannot be agreed upon, then adoption will always be limited.”

Even so, 83.3% thought a minimum quality level of the standard is needed for high adoption rates. A significant correlation is present for the following: In the cases where the adoption is rated better than expected, design choices of the standard have influenced adoption ( $P=0.48$ ;  $p=0.00$ ), and the adoption of the standard will be more successful if the quality is known, proven sufficient or improved ( $P=0.42$ ;  $p=0.00$ ). This suggests dependence between adoption and the quality of the standard which is also supported by case studies: Both the MISMO case within the mortgage industry (Markus et al., 2006) and the RosettaNet case within the electro technical industry (Boh, Soh, & Yeo, 2007) reported a similar relation between adoption and the content of the standard. Also in a recent innovation-centric adoption and diffusion framework for standards, standards characteristics are included as area that impacts adoption (Wapakabulo Thomas, 2010). Our conclusion is that a standard's quality is seen as necessary but is not really a sufficient condition for adoption. But, will knowledge about a standard's quality contribute to quality improvement? This is addressed by the fourth proposition.

Proposition 4: There is a need to make the quality of the standard visible by assessment.

The majority (54.2%) agreed with the statement that it is difficult to improve the standard if the quality is not known (14.6% disagreed). 85.5% considered it helpful to have some kind of instrument to make quality transparent, thereby supporting this proposition.

A notable 81.3% would use a quality measurement instrument, when available. Interesting to see is that standards which aim for quality are potential users of such an instrument: Standards that have quality assurance as part of their current development process also have a minimum quality check in place ( $P=0.57$ ;  $p=0.00$ ), and score high on the achieved interoperability ( $P=0.32$ ;  $p=0.03$ ), and will use an instrument to measure the quality when available ( $P=0.40$ ;  $p=0.00$ ). And those respondents who agreed to the statement that improved quality might lead to improved interoperability, thought that it is helpful to have some kind of instrument to gain insight into the quality of the standard ( $P=0.30$ ;  $p=0.04$ ). This goes even further than the proposition; not only is there a need to make the quality visible by assessment, but also some kind of instrument would be welcomed to assess quality.

Remarkably, the majority of respondents who already deploy a quality check before a standard's release also use some kind of instrument to measure the quality of the standard ( $P=0.39$ ;  $p=0.01$ ), but nevertheless would welcome a newly developed instrument and would use it when available ( $P=0.38$ ;  $p=0.01$ ).

Finally, there is no negative correlation between the current use of an instrument to measure quality of standards, and whether a new instrument would be welcomed. This suggests that respondents see room for enhancement or improvement of their quality assurance.

### **3.3 Discussion**

The results of the survey indicate that quality of standards is not properly addressed in current standardization practice, and this reduces standards' quality, and therefore interoperability. Possible explanations might be as follows.

First, developing standards, much like enterprise interoperability, is not considered to be a profession yet (Oude Luttighuis & Folmer, 2010). Even though considerable standardization experience and professionalism in formal standardization bodies is present, the semantic standards realm is characterized by a wide range of development processes. Most semantic standards are developed by a specific domain organization, and are outside of the traditional SDOs. This disperses standards development knowledge and experience and limits the proliferation of such experience and re-use of process and product components. This effect is increased by the fact that most standards developers are domain experts. Domain knowledge is crucial to standards development, but differs from general standardization expertise. Moreover, education and certification, so common for other professions, are hardly available or required in the standardization field. On the positive side, standards developers are intrinsically motivated (Teichmann, 2010) and eager to improve the quality of their standards when appropriate knowledge and tools are available, as shown by the

survey results. Adequate assessment of the impact of this factor would require further research and might yield opportunities for a more mature standardization profession.

Second, notwithstanding the role of standards in acquiring interoperability, and that topics such as enterprise integration, business process improvement and reducing enterprise costs have been priorities among many CIOs (McDonald, 2010; Park & Ram, 2004). This does not seem to result in standardization, as a topic on its own, being given high priority. The explanation might be that standards are seen as technical solution and not for both practitioners and scholars more interesting business topic.

A third possible explanation of the survey results is that, even though standards developers may think interoperability could be improved, current interoperability levels satisfy current business needs. Thus there could be a discrepancy between the supply side (standards developers) and the demand side (end users) regarding standards quality and the importance of interoperability. This research focuses on the supply side, and therefore cannot reflect the viewpoint of the demand side. An imperfect standard (from the viewpoint of standards developers) might be quite acceptable to the end user. Lower interoperability levels might satisfy current needs. This argument nevertheless deserves some nuancing. The respondents surveyed are standards developers, who have diverse backgrounds but include standards users (Zhao et al., 2005). Some are employed by software vendors. Others work for user organizations. So, at least some user perspective may be expected to have been included in the survey. In order to assess the impact of imperfect standards and interoperability on the demand side, we would need to extend our study to end users.

### 3.4 Conclusions

Our main research question in this chapter was: *Is there, based on the current standards development processes and experienced interoperability and adoption problems, a need to elicit the quality of semantic standards?*

We interpret the survey results as a positive answer to this question. The results of the survey show that basic procedures for quality are in place in the standardization process. Most standards developers see a need for further improvement of the quality of standards and for instruments and tools that can aid in the assessment and measurability of standard quality. Figure 12 summarizes our conclusions.

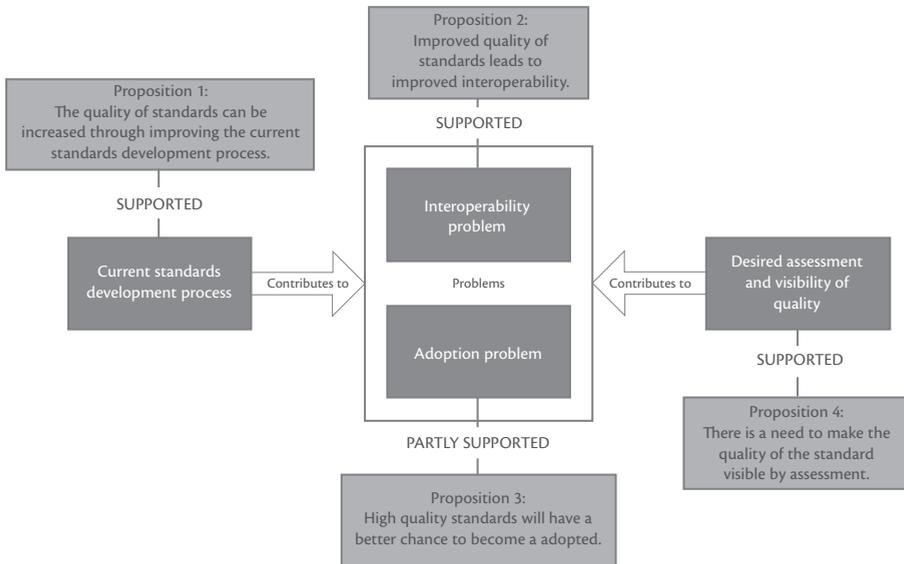


Figure 12 – Summary of propositions

# Identification of Research Gap

The previous chapter provided evidence for the practical relevance for quality measurement of semantic standards. This chapter will continue with describing the existing literature by analyzing an overview of studies based on a systematic literature review.

The goal of this part of our study is assessing the topic of quality of semantic standards as a possible research gap. A derived goal, and contribution to the knowledge area, is the analysis of coverage of this research subject within the most highly ranked (top) IS and management literature.

### **4.1 Research questions and method**

In order to get an overview of existing state-of-the-art in top journals regarding the topic of quality of semantic standards, the following research questions have been constructed:

1. What trend can be noted by looking at the amount of publications per year?
2. Are there any studies related to quality of semantic standards published?
3. Is there a strong research fundament for semantic standards, and specifically for certain domains (verticals)?
4. What is the maturity of the academic standardization discipline?

A systematic literature review (Petticrew & Roberts, 2006) has been set-up and performed to enable grounded answers to the research questions and to assure that no major publication will be missed. The search was constructed based on Rumsey's (2008) description of planning the campaign. The goal of identifying a research gap implies that the top 25 IS journals and top 25 management journals should be included (and restricted to) in the search phase. Search engines were selected based on their coverage of these journals. The selection of journals and search engines was based on previous work (DuBois & Reeb, 2000; Mylonopoulos & Theoharakis, 2001; Schwartz & Russo, 2004). More information on the journals and search engines selected is available in Appendix C.

From the domain of quality measurement of semantic standards, keywords were selected. To assure the quality of the keywords, the selection was done iteratively by testing the keywords in the search engine and by adding multiple synonyms. The selected keywords are visualized in Figure 13, while the synonyms and search strings are mentioned in Appendix C.

---

This chapter is based on:

(Folmer, Berends, Oude Luttighuis, & Van Hillegersberg, 2009): Top IS research on quality of transaction standards, a systematic literature review to identify a research gap. In: Jakobs, Kai (Ed.), Proceedings of the 6th International Conference on Standardization and Innovation in Information Technology SIIT 2009, The EURAS Board Series, ISBN: 978-3-86130-243-8, pp. 65-78, 2010, Verlagshaus Mainz GmbH Aachen.

(Folmer, Berends, Oude Luttighuis, Van Hillegersberg, & Lammers, 2010): Research on Quality of Transaction Standards: The Maturity of a Research Topic. In: Bernus, P., Doumeingts, G. and Fox, M. (Eds.), Enterprise Architecture, Integration and Interoperability (EAI2N), part of WCC2010, pp. 101-115, Brisbane: Springer.

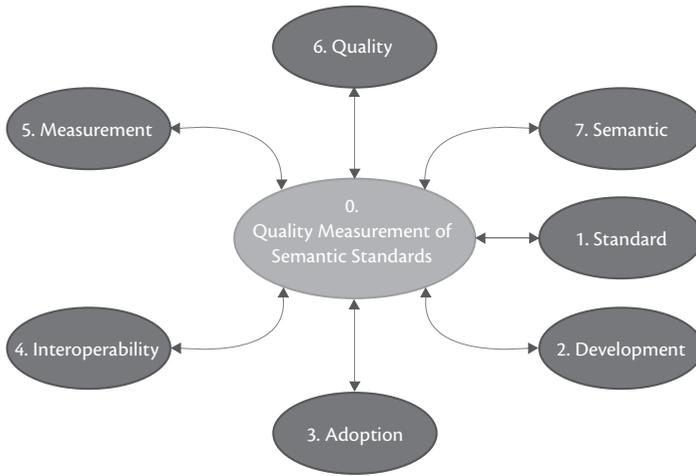


Figure 13 – Keywords

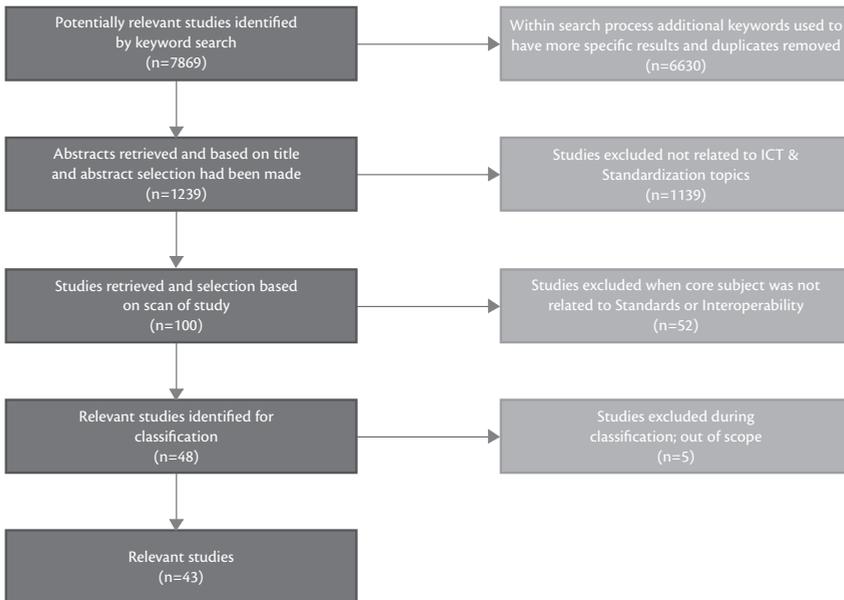


Figure 14 – Quorum flowchart of selected studies

The searches conducted with the search engines yielded several articles per query. Search queries were designed so that manageable amounts of publications were found. Then, an exclusion process has been initiated as described by Van der Linde et al. (2004). First, abstracts and keywords were assessed manually on relevance; in order to ensure that nothing was overlooked this process was done twice and by two individuals. This resulted in a list of 100 papers. A second screening on relevance took place by scanning each publication, again double-blinded. This resulted in a list of 48 publications; these were classified according to the developed framework. During this classification we found that an additional 5 papers were out of scope, which resulted in a final list of 43 publications (the complete list can be found in Appendix D). This selection process, visualized in Figure 14, is a weak spot in this methodology, because the selection criteria are subjective and difficult to trace. In the first step many papers related to software engineering, healthcare, multimedia and accountancy were removed. The second step removed publications with only marginal attention for standards.

Next, a classification framework is needed to arrange the studies found, in order to be able to answer the research questions. This framework was set up before the classification process itself started.

#### 4.2 Classification framework

Based on the research questions and other systematic literature review research (Wareham, 2005) several classifiers regarding the standardization subject were selected, as well as classifiers regarding the research rigor. These are:

- Topic: The topic (domain) of the research.
- Standard Lifecycle: The phase within the lifecycle of a standard.
- Standards View: The actor’s viewpoint on the subject.
- Type of Standard: What kind of standards is the paper about?
- Research Approach: The research approach (fundament) for the paper.
- Research Method: The applied research method of the paper.

Several other classification schemes were considered, such as the IS core theories used in many publications. The model of Benbasat and Zmud (2003) consisting of ICT-artifact, Usage, Impact, ICT managerial, methodological, and technological capabilities, and ICT managerial, methodological, and technological practices would have been applicable. And also, on a subset of the papers, the diffusion of innovation theory of Rogers (2003) is applicable. But the main reason for selecting the mentioned classifiers is the relevance to the research questions. Next, the six chosen classifiers will be further decomposed.

#### Topic

Based on the keywords (Figure 13) and brainstorming, five different topics have been identified.

Topic	Description
<b>Standards Lifecycle</b>	The publication discusses one or more steps from the standards life cycle, such as standards development or standards diffusion.
<b>Standards and Interoperability</b>	The publication concerns interoperability issues, or other higher-level aspects of standardization.
<b>Standards Quality</b>	The publication addresses the quality aspects of standards.
<b>Standards Policy/Strategy/Impact (PSI)</b>	The publication concerns economics of standardization, business cases, general advantages, the impact of usage of the standard, or the effectiveness of standards.
<b>Standards Organization</b>	The publication concerns standards setting organizations (SSO) and standards development organizations (SDO), National Standards Organizations, etc.

Table 8 – Topic

## Standards Lifecycle

We chose the earlier presented extended general lifecycle model (Söderström, 2004) as a start, because it takes most other lifecycle models into account. Although this model fits our purposes we condensed it for pragmatic reasons; it contains too many stages, which may result in fragmented results. We combined the Initiate and Standards Development phase (and kept the latter name), and did the same for Develop Product, Conformity Assessment, Educate and Implement. Also, Feedback is combined with Maintain.

In comparison with lifecycle models from other domains (e.g. software domain (Ambler, 2009)), the standardization lifecycle models found are open-ended: they lack an “end” phase. Based on the Enterprise Unified Process, we therefore decided to add a Retirement phase to the lifecycle model.

Standards Lifecycle	Description
<b>Develop</b>	The creation and development phase of a standard.
<b>Implement</b>	Implementation of the standard in products or systems, including implementation services.
<b>Use</b>	The usage of the standard, the adoption in the market (diffusion).
<b>Maintain</b>	The maintenance phase where standards (periodically) are improved to current needs.
<b>Retire</b>	The phase when a standard is withdrawn from maintenance.

Table 9 – Standards Lifecycle

## Standards View

Different roles take part in the stages identified in the lifecycle model. We however see no one-to-one correspondence between lifecycle stages and roles. For instance, it is possible to have a user view on the implementation of standards, but also the view of the creator of the standard on implementation phase. Krechmer (2006) identifies three main views on standards: user, implementer and creator. We added the policy maker role. One might argue that this constitutes a specific type of user, but for our goals it might become relevant to analyze the role of government related to quality and semantic standards.

Standards View	Description
<b>Creator</b>	The developer of the standard. (creates the standard)
<b>Implementer</b>	The implementer of the standard. (implements the created standard)
<b>User</b>	The (end) user of the standard. (uses the implementation of the standard)
<b>Policy Maker</b>	The policy maker about standards. (develops policy about the standard)

Table 10 – Standards View

## Type of Standard

As this study focuses on semantic standards we chose to use the classification used by Steinfield et al. (2007), as this is the closest fit to our third research question.

Type of Standard	Definition
<b>Syntactical</b>	The scope is related to technical standards like TCP, IP, SOAP
<b>Semantic – Horizontal</b>	The scope is related to cross industry standards like ebXML, UBL
<b>Semantic – Vertical</b>	The scope is related to industry standards like MISMO, HR-XML
<b>All</b>	Multiple types are covered

Table 11 – Type of Standard

## Research Approach

An often-used classification of the research approach is from Orlikowski and Baroudi (1991):

- Positivist
  - Descriptive
  - Theoretically grounded
- Critical
- Interpretive

Klein and Myers (1999) uses the same categories to classify IS research but without the shown subdivision of positivist research. For an analysis of e-commerce research, Wareham (2005) distinguishes between positivist, interpretivist, descriptive and design science. The critical approach has been left out, perhaps because of low expectations on finding articles that fit this category. Design science has been added as a more recent research approach (Wareham, 2005). Other options would be to distinguish between:

- Qualitative and quantitative approaches, and
- positivism, post-positivism, critical theory, and constructivism for qualitative research as described by Guba & Lincoln (1994).

For our purposes, we used the original list of Orlikowski and Baroudi (1991).

Research Approach	Description
<b>Positivist (Theoretically grounded)</b>	Propositions or hypothesis are formulated and tested, or analytical propositions are derived. Typically quantifiable measures on stated populations (Klein & Myers, 1999).
<b>Positivist (Descriptive)</b>	Describes current practices, without theoretical grounding or rigorous data collection and analysis. They describe issues to be shared with the community. Typically case studies (Orlikowski & Baroudi, 1991).
<b>Critical</b>	Critical perspective if the main task is being seen as being one of social critique, whereby the restrictive and alienating conditions of the status quo are brought to light (Klein & Myers, 1999).
<b>Interpretive</b>	A basic premise is that the perspective is fundamentally subjective, and thus, attempts to understand the phenomena through the meaning that participants assign to them (Orlikowski & Baroudi, 1991; Wareham, 2005). Typically orientated at social constructs, or the context of IS.

Table 12 – Research Approach

## Research Method

Research methodology is a vast and diverse field. For our research, the amount of methods should be limited in order to avoid fragmented results. Also, it should match our research questions. In our case, this means that a general, high-level classification of research methods will suffice.

Wareham (2005) uses for his e-commerce literature review: conceptual, survey, experiment, development, data analysis, case study, review, others. Orlikowski and Baroudi (1991) uses a somewhat different list: survey, laboratory experiment, case study, mixed method, field experiment, instrument development, protocol analysis, action research.

Our literature review parallels Wareham's, although the subject is different. Therefore we chose Wareham's list as a start. The following table is based on Wareham's (2005), but slightly adapted by combining survey, experiments and data analysis into one category.

Research Method	Description
<b>Conceptual</b>	Conceptual analysis, theoretical analysis, mathematical models, analysis or narration based upon author's experience, observation or thoughts. No strong empirical evidence to support author's conclusion. Descriptions of current practices, situations and imagined scenarios.
<b>Data Analysis/ Survey/ Experiments</b>	Mail survey, online survey, use of questionnaires to obtain quantitative or qualitative data. Lab experiment, field experiment, free simulation. Document analysis, content analysis, secondary data analysis, field data analysis, and other analysis based on data not from questionnaire instruments and/or experimentation.
<b>Review</b>	Literature review, historical rendition, commentaries, current status review, practice review.
<b>Development</b>	Techniques, methods, frameworks, instruments to develop some technical application, system, protocol, etc.
<b>Case Study</b>	Intensive analysis of cases based upon interviews, observations and analysis in some specific context.
<b>Other</b>	Ethnography, action research, other.

Table 13 – Research Method

### 4.3 Classification process and results

Like the selection process, the classification process has been carried out double blinded to improve the quality of the results. Differences in the classification have been solved by analyzing the differences and achieving consensus from both individuals and to make use of a third individual. The complete list of papers and their classification can be found in Appendix D.

Table 14 contains an overview of the distribution of papers across the journals, and over time.

Journal	<	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08	Sum	
Academy of Management Journal													1				1	
ACM Computing Surveys													1				1	
ACM SIGMIS Database												1					1	
Communications of the ACM (CACM)	1		1			1				2	1	2	1	1	1	2	12	
Decision Support Systems								1					1				2	
European Journal of Information Systems								1									1	
IEEE Transactions on Industrial Informatics (TII)														1			1	
IEEE Trans. on Information Technology in Biomedicine (TITB)						1	1									1	3	
Information and Management	1		1											1	1		4	
Information Systems Journal												1					1	
International Journal of Electronic Commerce (IJECE)				1								1					2	
Journal of Management Information Systems (JMIS)					1							1	1	1	1		5	
Journal of Strategic Information Systems (JSIS)		1															1	
Management Science			1														1	
MIS Quarterly: Management Information Systems (MISQ)														6			6	
Organization Science					1												1	
<b>Totals</b>		2	1	2	2	2	1	2	2	0	0	2	4	8	10	2	3	43

Table 14 – Distribution of relevant papers

The peak in 2005 and 2006 is remarkable, and is partly explained by the special issue on standardization in MISQ in 2006 (nr. 30). Communications of the ACM hosts by far the most relevant publications on this subject.

Table 15 contains the classification based on topic.

Topic	Count
Standards Lifecycle	16
Standards and Interoperability	13
Standards Quality	1
Standards Policy/Strategy/Impact (PSI)	11
Standards Organization	2

Table 15 – Results on Topic

Remarkable is the low number of studies in the third and fifth categories. The second category contains papers that are more high-level and standards are often not the main subject. This is also why especially these papers could not be scored on the Standards Lifecycle (Table 16).

Standards Lifecycle	Count
Develop	4
Implement	1
Use	23
Maintain	-
Retire	-
Not applicable	15

Table 16 – Results on Standards Lifecycle

Remarkable are the low scores for the maintain and retirement phases, and the high score for the use/adoption phase. Table 17 contains the results on the Standards View.

Standards View	Count
Creator	7
Implementer	15
User	20
Policy Maker	1

Table 17 – Results on Standards View

This shows, in combination with the results on standards lifecycle, that most of the papers are dealing with a user view on standards. Hardly any have a creator's view, or deal with the development life cycle phase of the standard.

Table 18 contains the results on the Type of Standard.

Type of Standard	Count
Syntactical	10
Semantic – Horizontal	11
Semantic – Vertical	14
All	8

Table 18 – Results on Type of Standard

The classification process for this category was somewhat difficult, because many papers did not completely focus on one type. Also, the emphasis was not always clear. It is remarkable that only 14 papers have been found that mainly deal with vertical standards, as the keywords were specifically aimed to find as many as possible.

Next are the results on the Research Approach and Research Method (see Table 19 and Table 20).

Research Approach	Count
Positivist (Theoretically grounded)	5
Positivist (Descriptive)	26
Critical	6
Interpretive	6

Table 19 – Results on Research Approach

Research Method	Count
Conceptual	11
Data Analysis / Survey / Experiments	5
Review	9
Development	7
Case Study	11
Other	-

Table 20 – Results on Research Method

Remarkable is the low amount of papers with a positivist approach, fundamentally grounded with thorough data analysis, and the high amount of descriptive research.

#### 4.4 Findings

This section revisits our research questions.

##### 1. What trend can be noted by looking at the amount of publications per year?

Based on this study, no upwards/downwards trend can be derived from the statistics. The publication peak in 2005 and 2006 can be seen as an incident with 42% of the selected publications is published in 2005 and 2006. We conclude that the area identified is currently not a continuous research area.

##### 2. Are there any papers related to quality of semantic standards?

Within these top journals hardly any (only 1 paper) research has been published about the quality of semantic standards. This clearly suggests that quality of semantic standards constitutes a research gap. With only two results, the subject of standardization organizations can be called a research gap as well.

##### 3. Is there a strong research fundament for semantic standards, and specifically for certain domains (verticals)?

The key indicator to answer this question is the amount of papers found. Although the keywords were specifically aimed at semantic standards, including search terms such as *e-business* and *vertical*, only fourteen papers have been found that deal with vertical industry standards. Much attention is paid to technical standards, but research regarding vertical standards seems not to reach major journals. The fourteen papers found moreover revisit the same vertical standards, which makes the unique number even lower.

##### 4. What is the maturity of the academic standardization discipline?

Given that only the Communications of the ACM regularly pays attention to this subject, this is no good sign for the maturity of the standardization discipline. Another negative sign is the lack of fundamentally

grounded positivist research, and the high amount of descriptive research approach without fundamental background. The case studies are almost all related to the fourteen papers identified as related to vertical standards. Empirical research is in the minority. Based on these observations, we may conclude that the standardization discipline is not mature. Yet, a more thorough benchmark with other disciplines is needed to make this conclusion more definite.

#### **4.5 Conclusions & discussion**

At least two research gaps have been identified, both standards quality and standards organization, which was the primary focus of this research part. Also achieved is an overview including some remarkable insights of the coverage of standardization research within the top IS and management journals.

It is important to notice though that the validity of these conclusions is limited to the set of journals we have investigated. There seems to be a major difference between the standardization research covered in the top journals and the research covered in the less known specific standardization literature (for instance the International Journal of Standardization Research). Some topics (like Standards Organization) that are hardly covered in top journals are often covered in those journals and other edited books by members of the EURAS community.

Another important remark is the delay caused by the academic publication process. Recent trends related to interoperability and standardization like for instance cloud computing, open data, and social media: the interoperability between Facebook and Google+, will take several years to be noticed in systematic literature reviews like ours.

The next step is to analyze in detail the 43 selected studies on its value related to quality and semantic standards, and to broaden the horizon with including studies beyond the top journals.



## Semantic Standards Literature Exploration

The following paragraphs elaborate on the content of the previously identified studies and go beyond by including a broader set of studies than that have been identified in the previous chapter, but will focus on studies that are particularly dealing with semantic standards. It was carried out according to the general principles of literature review (Silverman, 2010). Figure 15 presents the structure of this chapter: It starts with terminology, technology and examples of semantic standards. The second paragraph continues with discussing literature about development and adoption, the most prominent topic within standardization literature as shown in the previous chapter. It ends with presenting a set of tactics for development and adoption that to a certain extent is related to quality. However the third paragraph is completely focused on quality. This topic is first covered by looking at different domains (e.g. software engineering) before the quality in relation to standardization in general (5.3.2) is discussed and in particular related to semantic standards (5.3.3).

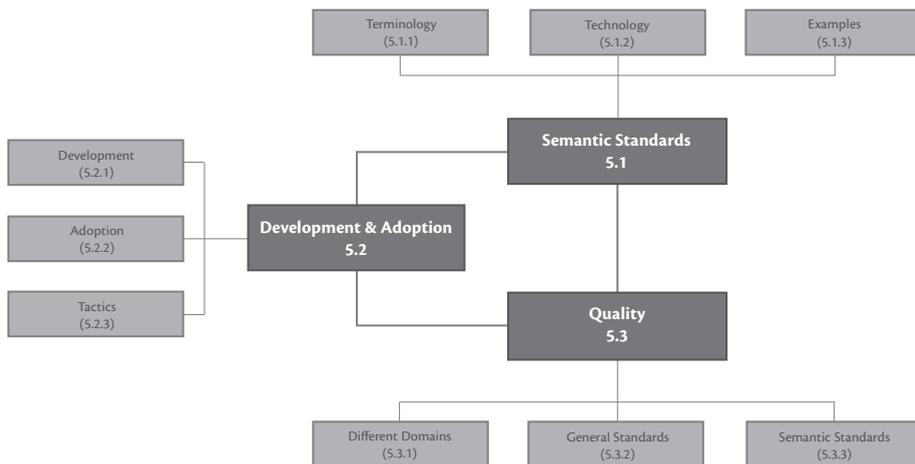


Figure 15 – Structure of this chapter

### 5.1 Semantic standard terminology

More specific terms used in literature are business transaction standards (Rukanova, 2005) and Vertical Industry Standards (VIS) (Steinfeld et al., 2007). Semantic standards can focus on a single vertical industry sector or purport to be applicable across sectors. An example of a cross-industry standard is electronic business XML (ebXML).

---

This chapter is based on:

(Folmer & Verhoosel, 2011): State of the Art on Semantic IS Standardization, Interoperability & Quality, Enschede: TNO, University of Twente, CTIT, NOiV.

We do not want to exclude cross sector semantic standards, hence we stick to the term semantic standards and by doing so we include both “vertical” and “horizontal” standards. But then we avoid the word “industry” as we do not want to exclude government oriented standards. Like our definition in Chapter 1 already captured, semantic standards may address product identification, data definitions, business document layout, and/or business process sequences (Steinfeld et al., 2007). Although intended for verticals, the following descriptions are also appropriate for describing semantic standards:

- “Vertical information systems (VIS) standards are technical specifications designed to promote coordination among the organizations within (or across) vertical industry sectors” (Markus & Gelinas Jr., 2008).
- “Trends are converging in new forms of cooperation among ICT-using organizations, for example, the user-led development of voluntary, open, industry-specific interorganizational coordination standards, here called vertical information systems (VIS) standards” (Steinfeld et al., 2007).
- “Vertical IS standards prescribe data structures and definitions, document formats, and business processes for particular industries” (Wigand, Steinfeld et al., 2005).

Löwer (2005) sums up the different terms used for what he calls inter-organizational standards, which to a large extent are synonyms: “Inter-organizational System Standards and Process Innovations”, “Open E-Business Standards”, “Standards for Domain-Specific Interoperability”, “Vertical Industry Languages”, “Vertical IS Standards”, “XML-Based E-Business Frameworks” and “XML-based E-Business Standards”.

### 5.1.1 Semantic approaches, languages and technology

Semantics deal with the meaning of signs, symbols, words and phrases (Brzezinski, 2010b) in the special sense of how these notifiers relate to reality, how they represent, designate and signify things (Rukanova, 2005). This statement is used to discuss how the world of semiotics can be used as a donor for developing theories for the standardization world which currently lacks theories (Brzezinski, 2010b). One example for that is for instance Stamper’s semiotic framework (as cited by (Rukanova, 2005) which is also applicable in the standardization world, either to identify interoperability levels or to classify standards.

Problems related to semantic mismatch and misunderstanding are common, while some think they will vanish over time whilst others think they won’t (Rebstock, 2009). If everyone were to use a single standard then semantic referencing would not be necessary, and although developments like core components are steps towards standards convergence, one universal standard would be an illusion. This means we have to cope with multiple e-business standards permanently, which will keep changing, resulting in a lasting situation of semantic variety, and will then be the source of mismatch and misunderstanding (Rebstock, 2009).

Besides specific horizontal and vertical semantic standards, there are also standards that can be used to describe (part of) the semantics that have to be defined by the standard. These include XML, UML, OWL, BPEL, BPMN and other similar types of standards.

The open standard-based Inter-Organizational System (IOS) uses semantic standards based on XML technology (Nelson et al., 2005). The XML (eXtensible Markup Language) 1.0 specification was introduced in 1998 by the World Wide Web Consortium (W3C) and was designed to improve the functionality of the Internet by providing flexible information structuring (Nurmilaakso & Kotinurmi, 2004). An XML document can be validated against an XML schema (XSD) that is included or referenced from the XML document. XML Schema Definition Language is an XML language for describing the valid structure of XML documents (Nurmilaakso & Kotinurmi, 2004). Alternatives for XML Schema are DTD (Document Type Definition) Schematron and RelaxNG. XML documents can be transformed by using another important XML standard called XSLT: eXtensible Stylesheet Language Transformations (Nurmilaakso, Kotinurmi, & Laesvuori, 2006) and the Naming and Design Rules (NDR).

Semantic Web technologies offer possibilities to express knowledge about the objects on the web. Standards in this area are RDF (Resource Description Framework), RDFS (Resource Description Framework Schema) and OWL (Web Ontology Language). Other core technology is UN/CEFACT CCTS (Core Components Technical Specification; ISO 15000-5) which presents a methodology for developing a common set of semantic building blocks that represent the general types of business data in use today and makes a provision for the creation of new business vocabularies and the restructuring of existing ones (Lampathaki, Mouzakitis, Gionis, Charalabidis, & Askounis, 2009).

Ontologies can also help by relating different semantic standards. For instance OWL is used to create an upper ontology of the CCTS specifications, to which different semantic (horizontal) standards can be linked like UBL 2.0, GS1 XML and OAGIS 9.1 (Dogac et al., 2010). If they do work, interoperability can be achieved among organizations that are using different standards.

The Web Services standards (SOAP, WSDL and UDDI) are used to create services based on XML. SOAP (Simple Object Access Protocol) defines the message, while WSDL (Web Services Description Language) defines the service itself. UDDI (Universal Description, Discovery and Integration) is used to search for trading partners. While on the one hand Web Services are dependent on standards (Kreger, 2003), on the other hand these standards are the fundament for the development of IOS and semantic standards.

Yet another related conceptual solution is related to business rules, which might be used to bridge the gap between business and technical people. Business rules can be business wise expressed in SBVR (or SBVR based equivalent) (Bridgeland & Zahavi, 2009) while the technical representation might be based on OCL (Warmer & Kleppe, 1998). Other rule based approaches are RuleML and SWRL (Horrocks et al., 2004).

Many more developments are taking place in this area at high speed, by standardization organization like W3C, OASIS, the Open Group (UDEF: Universal Data Element Framework (Schuldt, 2011)) and OMG. The latter develops the Model-Driven Message Interoperability standard (MDMI), focused on automatically adapting and translating multiple protocol syntaxes that share semantics.

### 5.1.2 Examples of semantic standards

As being mentioned, XML is one of the languages that provides a basis for defining the semantics of a term. Many authors have underlined the need for aligning semantics (Legner & Lebreton, 2007). There have been many XML based semantic standards, already since the early 21<sup>st</sup> century; in august 2001 XML.org contained 105 different standards spanning 25 vertical and 7 horizontal industries, while another list contains 450 submissions spanning 54 vertical and 9 horizontal industries (Nelson et al., 2002). Although EDI as technology is outdated, in several industries EDI-based standards are still *used* and *maintained*, in terms of the standards lifecycle (Söderström, 2004). EDI will not disappear that fast because often there is no positive business case for only a technology change from EDI to modern technology.

#### Horizontals

The use of the term vertical would imply that there are also horizontal standards. However, in the literature, a good definition of horizontal standards is hard to find. The main characteristic of horizontal semantic standards is that they can be used by various industries and sectors and is thus cross-sector oriented. Examples of horizontal, or cross-industry frameworks are for example cXML, OAGIS and xCBL (Nurmilaakso & Kotinurmi, 2004). Other important horizontal standards include UBL, GS1 XML and ebXML. The latter has specifically initiated the concept of core components, elements that can be used as the core and starting point of vertical semantic standards that make use of these core components (Folmer, Hinderer, & Otto, 2003; Van Blommestein, 2007). Since 2005 the ebXML Core Components Technical Specification (CCTS) has become an official ISO standard (ISO/TS 15000-5:2005).

A horizontal case study dealing with collaborative planning, forecasting, and replenishment (CPFR) based on a standards point of view (amongst others) is present in current literature (Markus & Gelinas Jr., 2008). The survey and analysis of horizontal standards (Kabak & Dogac, 2010) included EDI, UN/CEFACT CCL, UBL 2.0, OAGIS BOD 9.0 and GS1 XML and with the exception of EDI, they all use the CCTS in some (different) way. Other differences between these standards include the document artifacts, the use of code lists, the use of name spaces, and the naming and design rules used (Kabak & Dogac, 2010). Also important is the fact that there are major differences in how these standards do accommodate customization and extensibility. The horizontal OAGIS BODS are used in many vertical semantic standards, among others AiAG, ODETTE, STAR, AAIA (all automotive), but also in the human resources (HR-XML), chemical and aerospace industries (Kabak & Dogac, 2010).

## Verticals

With time many authors have included lists of semantic standards, including (Chari & Seshadri, 2004; Hasselbring, 2000; Lampathaki et al., 2009; Nelson et al., 2005; Steinfield et al., 2007; Von Westarp, Weitzel, Buxmann, & Köning, 2000). Since the list on xml.org has ceased, a new list is available on semanticstandards.org, containing nearly 100 standards and growing. Many of those are “industry specific” (vertical) for instance electronics (RosettaNet), chemicals (CIDX), assurance (ACORD), petroleum (PIDX), mortgage (MISMO), human resources (HR-XML), reporting (XBRL) and finance (ISO 20022). The following sections will mention several domains that have been studied by academics.

### Health care

Interoperability in the health care is well documented (Dogac et al., 2008; Eichelberg, Aden, Riesmeier, Dogac, & Laleci, 2005; Mori & Consorti, 1998). Several standards are available, and an overview is given by Eichelberg et al. (2005). Introducing an Electronic Health Record (EHR) is also seen as setting a standard (Hanseth et al., 2006), although it is a complex one and is not suitable for current standardization processes. With respect to standardization, the EHR is characterized by several problems (Sahay, Akhtar, & Fox, 2008):

- Most hospitals still use obsolete standards or protocols.
- Healthcare standards are not stable.
- ICT or Healthcare professionals may diverge from the use of the meaning that is defined by various healthcare standards (e.g. HL7, CEN 13606, openEHR, etc).
- Healthcare standards in XML solve the interoperability problem at syntactical level, but domain specific solutions are required to achieve semantic interoperability.

There are several competing standards approaches available which have been compared and show that achieving interoperability in the EHR domain has a long way to go (Blobel & Pharow, 2009).

### Education

There are many e-learning standards, in line with the Tanenbaum quote, for which overviews are available (Friesen, 2005; Hoel, Hollins, & Pawlowski, 2010). The IMS Global Learning Consortium Inc. (IMS) develops and promotes open specifications for facilitating online distributed learning activities (Friesen, 2005), but also ADL, IEEE, ISO, and other communities release standards for the e-learning domain. Often used standards are IEEE Learning Object Metadata (LOM), for the discovery of learning objects based on metadata. IMS Learning Design is a meta-language which can be used to model learning processes. ADL (Advanced Distributed Learning) SCORM (Sharable Content Object Reference Model) deals with real-time communication within the learning environment and deals also with the packaging of the learning material. SCORM aims at reusability, interoperability, durability and accessibility, and SCORM can be used in conjunction with LOM (Gonzalez-Barbone & Llamas-Nistal, 2007).

### Tourism

In the world's largest industry, tourism, many standardization efforts have failed because of their lack of flexibility (Fodor & Werthner, 2004). Given the heterogeneity of the market because of the web, the specific history of standards in the tourism domain, and the lack of a central authority that can impose such a standard, it seems unlikely that one global, all-embracing standard will be achieved. Instead, different standards for different market segments will co-exist (Fodor & Werthner, 2004).

### Building and construction

In the building and construction sector, a couple of XML based standards have been developed, such as bcXML and IFC (ISO12006-3 and eCognos) (Barresi, Rezgui, Lima, & Meziane, 2008). The EDI based standardization in the Dutch Building industry was used as a case study (Thissen & Stam, 1992). The main lessons learnt include:

- EDI *among* organizations is receiving increasing attention in the business community. The emphasis is on electronic communication of semantic standards in a standard format. It initially concentrated on technical protocols rather than on the content. Attention has shifted since the nineties towards higher-level layers of the OSI stack.
- Critical success factors for inter-organizational systems are:
  - Awareness of the strategic, long-range benefits.
  - High-level management support.
  - Support of industry leaders and/or the government.
  - Strong participation and membership in industry-wide organizations (needed for standardization).
- Standardization strategy was a lower-level result of the central issue of improving industry competitiveness!

### Automotive

The main standardization initiative in the automotive sector is taken by STAR (Standards in Automotive Retail) in which the AIAG (Automotive Industry Action Group) is participating (Anicic, Ivezic, & Jones, 2006). Other initiatives, also by AIAG, include the development of the MOSS (Materials Off-Shore Sourcing) standard for improving efficiency in global supply chains (Steinfeld et al., 2011a). The development of the Internet hub Covisint has been described by Gerst et al. (2005).

### **5.2 Semantic standards: Development and adoption**

The overview of studies within the domain of inter-organizational interoperability shows the focus on business aspects like adoption, governance and the organizational consequences (Robey, Im, & Wareham, 2008). Including the two main activities of standards organization: development and diffusion (Boh et al., 2007). It also shows the lack of studies that engage with the ICT artifact, going beyond the point of only describing the ICT artifact. One of the difficulties causing this is the short life cycle of ICT artifacts. Exemplary is the remarkable 56% of the studies that focus on EDI systems, while this is outdated technology and research outcomes might be not applicable to current generations of IOS based on open standards (Robey et al., 2008). Robey et al. proposes that researchers go beyond the superficial view of the ICT artifact, and engage with the ICT artifact, in our case semantic standards. Also remarkable is that the literature on semantic standards is often related to case studies regarding the adoption of the standard. For example, the adoption of STEP (Thomas, Proberts, Dawson, & King, 2008), MISO (Markus et al., 2006) and RosettaNet (Boh et al., 2007). This results in a rich knowledge base regarding development and adoption which will be described in this section.

Of note is the use of the wording of diffusion and adoption. Diffusion and adoption are slightly different concepts: Whereas adoption is normally used as the stage in which the standard is selected by an organization, diffusion is used to spread the standard for application. Adopted does not necessarily mean implemented: An organization may have chosen to adopt the standard but decided to wait with the implementation of (some of its) products or services. Here, we use the words adoption and diffusion as synonyms, because their slightly different meaning does not affect our research and it is often confused within the literature that we used for this chapter.

According to Zhao et al. (2007) development and adoption are interrelated since choices in development phases will influence adoption. Zhao defines a three-stage model of consortium based e-business standardization, simulating firms' strategic decisions:

1. First stage: Consortium participation
2. Second stage: Standard development
3. Third stage: Standard adoption

In addition, Zhao et al. (2007) notice that developers are adopters and most probably the early adopters. Moreover, the members' contribution is critical to the sustainability and success of a standards consortium and thus of the adoption of the standard. There are three ways to improve firms' involvement, as they will only contribute if the expected payoff is higher than otherwise:

1. Increase awareness of the potential benefits.
2. Improving inside benefits: Membership benefits like voting rights.
3. Reduce development costs.

### 5.2.1 Development

The state-of-the-art literature on the development of semantic standards is mainly concerned with the reasons for joining a standardization development trajectory and tactics for development. Zhao et al. (2007) presents various reasons for being involved in development. One reason is to contribute and orient the standard towards one's own business practices. The better the standard and the faster it is developed, the greater is the direct benefit for the developers. By being involved in the development of the standards, there is an increase in the understanding of the standard details which helps to reduce future implementation costs.

In addition to the work of Zhao et al. (2007), Boh et al. (2007) describe the paradox of participation in standards development. The greater the number of stakeholders, the more difficult it is to achieve consensus. It will slow down the process. On the other hand, involved stakeholders will be early adaptors. But why do organizations participate in standards development? Vendors are driven primarily by their perceived standard benefits, while users take only the perceived benefits from consortia participation activities into account (Zhao, Xia, & Shaw, 2011). This might be explained by the fact that users, in contrast to vendors, can easily act as free rider, implementing the publicly available standards. In general also the standards consortium effectiveness has a positive impact on the decision to participate in development activities. During development, users contribute more to the essential business domain knowledge (Zhao et al., 2011). Therefore SSOs should strive for increasing the perceived benefits from consortia participation activities to get the users into play.

There are various practical cases that show the different factors that play a role in the success of standard development. One of these examples is Rosettanet. The Rosettanet standards-setting process is not really open, and this might be one of the success factors (Boh et al., 2007). The strategies that have been used for standards development in RosettaNet are:

- Commitment of resources to the milestone program.
- Clear roles and restrictions.

- Validation beyond full implementation.
- Informal norms and social networks.

Another example of an open development process is the process of ebXML that has been studied (Choi, Raghu, & Vinze, 2004) and has led to the following propositions (adapted from (Choi et al., 2004)):

1. An open standardization process helps collaborators to create a functionally comprehensive standard.
2. An open standardization process promotes the convergence of technologies in the long run, paving the way to its domination over “proprietary” standards.
3. User participation is a moderating factor in an open standardization process for achieving a comprehensive and converged standard.
4. Interoperability, backward compatibility, feasibility and sponsor support (both SSO and technology providers) are critical factors that influence the creation of standards.

A comparison of multiple cases on semantic standards development in vertical industries is given by Nelson et al. (2005). Based on a comparison of nine different vertical standards, Nelson et al. identify key drivers, differences and similarities. Key drivers for vertical standards development are:

- Technological innovations (Internet, XML, etc.).
- Need for interoperability (to survive).
- Value proposition of the vertical standards consortium (pooling of R&D, time saving renegotiating with each new trading partner, etc.).

Differences between vertical standards include alignment with more established organizations, balance between vertical and horizontal focus, and adoption of the target domains including the use of tracking mechanisms for monitoring adoption. Similarities include non-profit status, vertical orientation, provision of standards freely, vendor neutral, platform independent, membership and fee structures. Another important contribution of Nelson et al. (2005) is the inter-organizational system (IOS) standards development cycle, containing the following phases:

- Choreography & Modularity (key cross-company business processes)
- Prioritize & Schedule (planning of business processes)
- Document & Standardize (develop specifications sets, including technology)
- Review & Test (permit user community to provide feedback)
- Implement & Deploy (provide implementation support and forecast adoption)
- Compliance & Certification (validate standards conformance to ensure interoperability)

More generally, Zhao et al. (2005) mention some unique characteristics of the vertical, semantic, e-business standards development process. They prove the uniqueness of e-business standards, in comparison with other standards (in particular ICT product standards). They describe challenges faced by the vertical e-business SSOs (in comparison to traditional SDOs like ISO) such as rapid technology development and divergent preferences of stakeholders. And most importantly a Participants - Technical content - Institutional structure framework is presented for studying vertical e-business standards. These three components are interrelated and determine the performance of the SSO, implying that the SSO should address all three components in an efficient and balanced way. The three components consist of the following features (Zhao et al., 2005):

- Participants (number, sector, bargaining power)
- Technical contents (maturity)
- Institutional structures (structure, procedures, openness)

Since semantic standards are being developed by many different SSOs, it might be expected that they will make a lot of (re)-use of each other's specifications. However the contrary seems true. There seems to be a lot of re-inventing of the wheel, based on a study of 33 SSOs (Löwer, 2005) (including horizontals like ebXML, cXML, W3C, etc. and verticals like ACORD, OTA, etc). Exceptions are RosettaNet, which makes significant use of the specifications of 8 other SSOs, and the specifications of UN/CEFACT are used by 10 other SSOs. The 33 SSOs that were studied only make marginal use of other specifications (Löwer, 2005).

The observation is made (Weitzel, Beimborn et al., 2006) that if standardization costs are too high we face the start-up problem and if standardization costs are too low we will face inefficient multi-standard equilibriums (for high and low standardization costs (as compared to network effects) monopoly outcome is quite rare). The implication of this observation is that with high standardization costs, standards development (and adoption) is less likely to take place in decentralized coordinated networks. With low standardization costs, the first mover advantage is limited and it should not be expected that partners simply follow, resulting in multiple standards. This shows already the link between development and adoption.

### 5.2.2 Adoption

Understanding standards adoption (and diffusion) stands out as an important research topic (Lyytinen and Rose 2003 as cited by (Zhu et al., 2006)) - probably because widespread standards adoption is critical. Simply explained by the fact that semantic standards, like other network technologies, are susceptible to network externalities (Boh et al., 2007; Cathomen & Klein, 1997; Katz & Shapiro, 1985).

There was some related research (empirical study) on adoption during the EDI-era (e.g. (Von Westarp et al., 2000)). And others like (Cathomen & Klein, 1997; Hart & Saunders, 1997; Kaefer & Bendoly, 2000; Kauffman & Mohtadi, 2004). A good overview containing even more studies is presented by Löwer (2005). Other comparisons have resulted in models to predict the adoption (Chwelos, Benbasat, & Dexter, 2001; Kaefer & Bendoly, 2000).

The research on adoption of IS standards continued in the XML standards-era according to Zhao et al. (2007), probably because of the low adoption of EDI-based solutions. Despite all promotional efforts, only 5% of the organizations that could benefit from the standard use it (Beck & Weitzel, 2005), or an estimated 2% of businesses worldwide (Wigand, Markus, & Steinfield, 2005). Just like within the EDI-era, several adoption models have been constructed primarily to predict and explain adoption (Chen, 2003; Kelly, Feller, & Finnegan, 2006; Mendoza & Ravichandran, 2007; Zhu et al., 2006).

Many case studies, such as STEP (Thomas et al., 2008), RosettaNet (Boh et al., 2007; Chong & Ooi, 2008; Löwer, 2005; Nelson et al., 2002), XBRL (Chang & Jarvenpaa, 2005) and MISMO (Markus et al., 2006; Steinfield et al., 2007; Wigand, Steinfield et al., 2005) focus on diffusion, leading to a strong research fundament. To explain adoption the following theories are often used:

- Diffusion of Innovation (DOI)
- Economics of standards (including network effects and switching costs)
- Game theory

DOI (Rogers, 2003) is often used, amongst others by Hovav et al. (2004) to analyze the adoption of IPv6, a technical standard. Some, like Weitzel et al. (2006), use both DOI and Network Effects. DOI lists five innovation attributes that influence the adoption decision, these include: relative advantage, compatibility, complexity, trialability and observability. Studies into the setting up of adoption models specifically for standards use complete DOI (Chen, 2003) or the DOI concepts complexity, compatibility and relative advantage (Kelly et al., 2006; Mendoza & Ravichandran, 2007), but they add other concepts that are, for instance, in the organizational and external context.

Hovav (2004) introduce two paths to standards adoption: Adoption through replacement and adoption through co-existence; XML and EDI is an example of the latter. Schwind et al. (2008) introduce “Determinants and parameters simulating diffusion dynamics in supply networks”. This is a model with factors, and each factor (determinant) is represented by one or more metric (parameters). Based on these metrics (including formulas), diffusion can be simulated.

For the migration to an IOS based on open standards, including XML-based horizontal and vertical standards, a conceptual model has been constructed. This open standard IOS adoption model indicates three variables influencing adoption of the standard (Zhu et al., 2006):

1. Network Effects (Trading community influence, Peer adoption)
2. Expected Benefits (influenced by Network Effects)
3. Adoption costs (Financial costs, Managerial complexity, Transactional risk, Legal barriers)

While adoption costs are a significant barrier there is a dependency based on the path taken. In this study non-EDI users were insensitive to adoption costs, in contrast to EDI users.

Adoption (Diffusion) strategies can be classified in four categories (Boh et al., 2007):

1. Market: Promote awareness among potential adopters about capabilities and benefits of the standard and how to implement.
2. Technology: Improve standard (lowering the costs of implementation and increasing the ease of implementation and use).
3. Policy: Change social and regulatory environment.
4. Relational: Co-opt key players to pressure their trading partners.

Diffusion of IOS has, just like a new product, a life cycle (Cathomen & Klein, 1997). The image of the life cycle depends on several factors (Cathomen & Klein, 1997):

- IOS: comparative advantage, compatibility, complexity, observability, cost, risk, availability
- Providers: strategy, structure, pressure, applicability, potentials
- Market: industry, tradable goods and services, competition
- Environment: economy, technology, law, society

ISO 10303, the Standard for the Exchange of Product Model Data (STEP), has been adopted worldwide and is often used in literature (Brunnermeier & Martin, 2002; Hardwick, Spooner, Rando, & Morris, 1996; Thomas et al., 2008; Wagner et al., 1995; Wapakabulo Thomas, 2010). Based on a case study of the adoption of STEP at the UK Ministry of Defense several barriers and facilitators of the adoptions have been identified (Thomas, Proberts, Dawson, & King, 2010):

Barriers	Facilitators
Difficulty understanding the standard	Other implementations (network effects)
Standards revision process	Pilots and demonstrations
Cost of the standard	Internal (economic) drivers

*Table 21 – 6 (out of 17) barriers and facilitators of the STEP standard (Thomas et al., 2010)*

With respect to adoption, implementation of ERP can be seen as a standardization of processes (intra-organizational interoperability). Many implementation related issues from ERP will be useful for standards as well. Benders et al. (2006) mention:

- Best practice (competitors will use the same best practices, and catch up quickly)
- Risks of non-conformance (ERP system does not fit)
- Power position of individuals
- Costs
- Implementation methodologies (SAP: ASAP, Oracle: FastTrack, Baan: DEM)

In comparison with standards, the first four are well known, although the power position of individuals is lacking attention. However, implementation methodologies are new to the standardization arena: Implementation methodologies offered for standards are very hard to find.

Another interesting case study is the adoption of RosettaNet which is well documented (Boh et al., 2007; Chong & Ooi, 2008). Rosetta has one of the biggest organizational memberships among supply chain standards consortia (Boh et al., 2007; Nelson et al., 2005). The case study of RosettaNet in China is described by Lu et al. (2006). Within the context of Malaysia the adoption factors trust, partner's power and product characteristics have influenced the adoption of RosettaNet positively, while the Malaysian Government's policy (financial incentives) seems not to have contributed (Chong & Ooi, 2008).

Standardization gains are often not operationalized due to lack of adoption. Possible cause is an asymmetry between individual and collective standardization gains and that there are thus multiple equilibriums between the two extremes (Weitzel, Beimborn et al., 2006). The standardization gap as a difference between the theoretical first-best and the realistic second-best standardization outcome, determines maximum possible coordination gains. Thus, depending on the situation, some available standardization gains can be internalized by communication (ballot problem: identify affected agents, arrange round-table talks). Others require an explicit redistribution of standardization costs and benefits (welfare problem: side payments). Consortia could provide institutional settings for binding agreements between agents (Weitzel, Beimborn et al., 2006).

### 5.2.3 Tactics for semantic standards

Semantic standardization differs from ICT product standardization. It is dangerous to generalize the research outcome to both groups. One of the differences for instance is the concept of standards war. In product ICT standardization this is a common phenomenon where various small groups can arise that want to standardize a certain ICT product in their specific way. Within semantic standardization, this is not likely because a semantic standard needs the support of all stakeholders. Semantic standardization is characterized by the heterogeneity of interests among participating user organizations.

Based on a case study, Steinfield et al. (2007) identifies the following maintenance characteristics and issues that are specific for semantic standards:

- Ongoing maintenance, since the user requirements of the vertical sector can change often in order to react to a flexible environment.
- An impertinent organization may not be adequate and a more formal institutional structure is needed for structure and the removal of uncertainty: Create a permanent organization.
- Early steps for legal challenges (intellectual property rights).
- Show how the standards can evolve as newer technology arrives.

Successful semantic standardization is characterized by jointly setting up tactics for development and diffusion (Boh et al., 2007; Markus et al., 2006; Zhao et al., 2007). In addition, this set of tactics should fit to the situation. Every domain is different, and requires its own specific set of tactics. What works in the MISMO case does not have to work in the HR-XML case, or any other case. Lessons learnt in the building and construction sector show that a plan of action for standardization must include a strategy for promotion, development, implementation and maintenance of vertical standards (Thissen & Stam, 1992). Several strategies have been introduced; including the do nothing approach (standardization will occur eventually). The other strategies fall into three categories (Thissen & Stam, 1992):

1. Stimulation of user consciousness of the need for standardization.
2. The introduction or use of power-related mechanisms as vehicles for speeding up the willingness for change and innovation, including standardization.
3. Coordinated theoretical development of standards, including the creation of a special organization to accomplish it.

In summary: Standardization is challenging, but in order to successfully achieve adoption a set of tactics is needed that jointly solve the standards development dilemma without jeopardizing the solution to the adoption dilemma and is fitted to the specific situation (Markus et al., 2006), and might also be locally adapted (Boh et al., 2007). Keep in mind there is no one right approach for the standards development process, not even a full open approach (Boh et al., 2007). Set of tactics for successful consortia to consider are:

Related to participation:

- Only involve the organizations that are committed to solving the problem (Boh et al., 2007).
- User-groups that have the greatest ability to influence adoption must be present in the development process, be committed, without having a disproportionate influence on the content of the standard (Markus et al., 2006).
- Involve all stakeholder groups (including ICT vendors), including key stakeholder and assure they do not drift apart during standardization (Markus et al., 2006; Steinfield et al., 2007).
- Create a social group (Steinfield et al., 2007).
- Active efforts for further participation (Steinfield et al., 2007).
- Increasing the perceived benefits from consortia participation activities (Zhao et al., 2011).

Related to the content:

- Focused: Promote a focus on solving real-world business problems (Boh et al., 2007).
- Limited scope (to keep intra-organizational conflicts out of the scope) (Steinfield et al., 2007) on maximum and visible benefits where already a certain degree of formalization and structuring of activities has occurred (Thissen & Stam, 1992).

Related to the process:

- Quick: Move the standard-setting process along quickly without negatively affecting the quality of the standards (Boh et al., 2007).
- Ensure open sharing of valuable knowledge across a range of stakeholders (Boh et al., 2007).

Related to organization:

- Structure governance (open memberships, voluntary participation in particular workgroups, transparency in decision making, fair voting rules, efforts to reduce costs of participation, separate governance committee) (Steinfield et al., 2007).
- Openness (participation, work accordance to fair procedural rules, decision making (consensus)) (Werle & Iversen, 2006).
- Free standards, and distribute standards through the Internet (Steinfield et al., 2007).
- Investing significantly in standards adoption (Boh et al., 2007).
- Organize effective management of the consortium (Zhao et al., 2011).

### **5.3 Quality**

This section will focus on quality starting with a broad view of quality in different domains, and via general standardization it will narrow the scope to quality in relation to semantic standards.

#### **5.3.1 Quality in different domains**

Quality has multiple meanings in different domains. Although our interest lies in quality related to standards it is worthwhile to study different domains where quality has a rich history. The different perspectives we included are:

1. Product engineering / manufacturing domain
2. Software engineering domain
3. Information system domain
4. Data / Information domain
5. Evaluation frameworks
6. Standards domain (sections 5.3.2 (general) and 5.3.3 (semantic))

Many more disciplines, like the management discipline (EFQM, Six Sigma, etc), might contain relevant studies relating to quality, but in this literature review we chose to limit the study to probably the most relevant disciplines related to standards. An elaborate overview of these domains is available (Folmer & Verhoosel, 2011), and therefore will be summarized briefly:

### 1. Product engineering / manufacturing domain

Quality has become a major topic since the reconstructions after the second world war until the eighties especially in the manufacturing industry, but later on it spread its wings beyond manufacturing to both private and public services (Ghobadian & Speller, 1994). Quality in product engineering is really associated with the philosophies of Guru's, like Juran (Juran & Gryna, 1988), Crosby (1979), Deming (1986) and many others like Feigenbaum, Groocock, Taguchi and Ishikawa. Many aim at the quality of physical products, including the presentation of quality attributes of which overviews are available (Chase & Aquilano, 1995; Garvin, 1984; Ghobadian & Speller, 1994). Since the nineties, ISO (9001) and other quality standards have become quite popular, while focusing on the processes, instead of the end product. These process-related standards have also become quite popular in software engineering.

### 2. Software engineering domain

The overall quality level of software is low (Davenport, 2005), which might explain research attention on quality within the software engineering domain. 30 years have passed since the up rise of this subject, but it has still not really penetrated into mainstream software engineering (Fenton & Neil, 2000). The APGAR score (for newborn babies) is also requested for software (Glass, 2008). A 2002 study from the U.S. National Institute for Standards and Technology estimated that software bugs cost the U.S. economy almost US \$60 billion a year (Davenport, 2005). The quality and cost problem of software development have led to the development of the Capability Maturity Model (CMM) by the Carnegie Mellon's Software Engineering Institute (SEI) in 1987. According to Davenport (2005) CMM has become such a huge success because of its simplicity, government support, its governance structure, and its flexibility in application within organizations. Current quality standards include CMMi and ISO 9126 family where the latter has been superseded by the ISO 25000 family, but many more standards that are not always coherent (Abran, Khelifi, & Suryn, 2003). Also many extensions to these standards for usage in specific situations exists (Lew, Olsina, & Zhang, 2010; Moraga, Moraga, Calero, & Caro, 2009; Van Zeist, Hendriks, Paulussen, & Trienekens, 1996; Van Zeist & Hendriks, 1996).

### 3. Information system domain

Software engineering has, in comparison with information systems (IS), a longer history in quality which makes it interesting to study both domains. Quality attributes for IS have been set up (Bailey & Pearson, 1983; Delen & Rijsenbrij, 1992; King & Epstein, 1982), just like models for IS success (DeLone & McLean, 1992), E-Commerce (DeLone & McLean, 2003), Enterprise System success (Sedera & Gable, 2004), and also a proposed IS quality model (Rodriguez & Casanovas, 2010).

Based on the quality aspects from mainly software engineering, IS, and data quality domain, a specific quality model has been constructed for knowledge management systems (Owlia, 2010). This work shows some valuable insights into how many quality aspects are available within existing literature with slightly different nomenclature and meanings. Other specific models include service oriented architectures (Joachim, Beimborn, & Weitzel, 2011; O'Brien, Bass, & Merson, 2005).

#### 4. Data / Information domain

Within the IS discipline, data quality is seen as a relevant area focusing on the quality of information inside an organization. Data or information quality is part of the IS success models presented in the previous section, and it is also part of the software engineering quality standard (ISO 25012). However it is an important area of research: 60% of the surveyed firms (500 medium-size corporations with annual sales of more than \$20 million) have problems with data quality (Wand & Wang, 1996; Wang & Strong, 1996). Within the domain of data quality, Juran's definition of fitness for use is commonly used (Wang & Strong, 1996; Zhu & Wu, 2010). Data quality has strong ties with standards, as data quality can be affected by protocols and standards (Madnick et al., 2009). Research in this area develops protocols and standards to improve the quality of data exchanged among multiple organizations or within a single organization (Madnick et al., 2009). Many models that include quality attributes are available within this domain (Byrd & Byrd, 2012; Kahn, Strong, & Wang, 2002; Knight & Burn, 2005; Stvilia, Gasser, Twidale, & Smith, 2007; Wand & Wang, 1996; Wang & Strong, 1996).

#### 5. Evaluation frameworks

There exist several frameworks for comparison and selection purposes of standards. We identified four as being relevant for standards:

1. Evaluation Taxonomy (Lampathaki et al., 2009): To classify business transaction standards based on taxonomy related information of a standard.
2. Comparative Analysis (Nelson et al., 2005): Analysis and comparison of multiple vertical standards.
3. Evaluation Framework (Mykkanen & Tuomainen, 2008): Evaluation of a specific (or relative comparison) interoperability standard(s).
4. Reference Model Analysis Grid (RMAG) (Pawlowski & Kozlov, 2010): Assessing, testing and validation of standards for achieving interoperability in the education domain.

The latter two are most exhaustive, where the RMAG is focusing on a specific domain (education) and the evaluation framework is generic and customizable for each domain. This framework consists of nine subjects, ranging from meta-topics, technical aspects, semantics, domain-specific, etc. An evaluation form has been constructed for each of them, consisting of, in total, 54 questions, excluding the many lower level questions. It includes a process model of how to perform the evaluation. More work should be done on validating the model and the forms. The RMAG, although intended for learning technology models, it is generally applicable. It includes different categories for standards classifications and a long list of analysis and assessment aspects and metrics for evaluation. Finally it consists of a structured survey to be used when evaluating the standard. Several categories deal with metadata like the objectives, domain, methodology and documentation of the standard. The category "In-depth analysis" deals specifically with interoperability on different levels: practical, semantic, and technical integration (Pawlowski & Kozlov, 2010). In comparison with the evaluation framework of Mykkanen & Tuomainen (2008) this RMAG evaluation contains fewer details and might be more practical (less time investment) to use, but the results will be less detailed as well.

This section showed that we can learn from other domains: e.g. CMM stresses the importance of configuration management and requirements management; both concepts are applicable to standards as well. Even the Software Quality Assurance is a concept that could be copied to a Standards Quality Assurance for developing standards. Furthermore, from the product engineering domain concepts are useful for the standards domain as well. For instance the quality grid by Crosby (which is also used in CMM) is applicable to standards as well, although most SSOs will be part of stage 1 - uncertainty, and not ready for the more sophisticated stages 2-5 (from Awakening to Certainty). Many of the quality attributes from the data quality domain might be applicable to standards as well, etc. This is very valuable as within the standards domain the quality subject is less mature. The next section will deal with quality from the standards domain itself.

### 5.3.2 General standardization quality

In the field of standardization, most research focuses on how standards develop, adoption decisions, types of existing standards, and those needing further development (Rukanova, 2005). Both Rukanova (2005) and Söderström (2004) found that there is little research in the area of standards implementation and even less on how to evaluate the fit between the requirements of a specific situation and a standard. This fitness for use in a specific situation is what we call *quality*.

In literature, quality is sometimes related to the adoption of the standard in practice. For instance Zhao et al. (2005) mention the penetration rate of a standard as a proxy for standards quality. Although adoption is important, this does not line up with a view on quality of “fitness for use”, for which adoption might be a proxy with too many limitations. A distinction is often made of a standards quality between the standardization process and its outcome, which we also see in the following theories.

#### Quality requirements and legitimacy model

Related to quality and the ICT domain, but not particularly semantic standards, we identified two theories that we will discuss: The EU quality requirements for recognition of ICT standards, and the legitimacy model from the learning technology domain.

Already mentioned in Chapter 1, the EU sets out a policy to include ICT standards maintained by others than the European Standards Organizations (EC, 2011b). To be able to only include high quality standards being developed outside traditional SSOs, it has defined requirements for the EU recognition of ICT standards. The first requirement is the acceptance of the market and the standards’ implementations may not hamper interoperability with the implementations of other standards. Furthermore it sets requirements for both process and product (the specification): These process requirements are *openness* to all stakeholders, *consensus* decision making and *transparency* related to information and participation. Requirements related to the specification are *maintenance*, *availability*, *intellectual property rights*, *relevance*, *neutrality and stability*, *quality* (EC, 2011b). Appendix O contains the full details on these requirements. Another source is the SEMIC.EU project of the European Government that provides a quality framework for so-called interoperability assets, and corresponding artifacts, that are useful within e-government interoperability projects (Nentwig et al., 2008).

Based on the work of Werle & Iversen (2006), an input and output legitimacy model has been introduced within the learning technology domain (Hoel & Hollins, 2008). Legitimacy is seen as essential for attracting necessary support from the stakeholders. Table 22 contains this model. The input legitimacy is focusing on the process side, while the output legitimacy is aiming at both the specification (the product) and its implementations. Other studies in this learning technology domain mention the “issue of quality” as problem (Hoel & Mason, 2011), but without searching for solutions.

Input legitimacy	Output legitimacy
All “interests” considered and ideally represented	Inscription of stakeholders’ interest
Open process	Enactment status (is the specification implemented and used in services?)
Balanced choice of SSO (either formal or community based)	Technical maturity of the specification

Table 22 – Input & Output legitimacy model (Hoel & Hollins, 2008)

#### Stakeholders

Different stakeholders have different views on quality, as they have different interests (Sherif, Jakobs, & Egyedi, 2007). Figure 16 shows the different angles of interest of the stakeholders. This stakeholder viewpoint is quite interesting because it is not the producer of standards but the end-user who bears the cost of change

(Egyedi & Blind, 2008). “In particular where lack of quality of an initial standard is the reason for a revision, the people responsible may not be the ones to pay” (Egyedi & Blind, 2008; Sherif et al., 2007). Or like Sherif et al. (2007) put it (Egyedi, 2008): “The diverse interests that affect standardization, the distributed nature of its management process and the time lag between a standard and its implementation in products and services mean that there is no clear accountability in terms of profit and loss responsibilities due to deficiencies in an ICT standard. In some cases, those who pay the cost of the lack of quality are not those who made the decisions. Thus, market mechanisms will rarely provide the driving incentive to carry out the intensive planning and coordination across organizational boundaries that are needed to produce a quality standard”.

The abundance of corrective market incentives to address lack of standards quality also applies to the corrupt use of standards, another issue regarding standards implementation (Egyedi, 2008).

The fact that different stakeholders will have different interests has been translated to a project management view on quality for the telecom domain. The core of the view is that within different aspects of project management like scope management, resource management, quality management, etc., quality needs to be addressed and symptoms of poor quality might be sighted (Sherif et al., 2007).

Stakeholder	Angle of interest	Quality Emphasis	Relevant Project
Owner (standards body)	Legitimacy Due process	P	Resource
Producer (technical committee)	Technical Due process	O, P	Quality, resource, time
Supplier (committee participant and standard developer)	Technical Due process	O	Resource, quality, documentation
Sponsor (Companies financing participants)	Marketing Financial (possibly technical)	O	Time, cost, resource
Consumers (implementers of standard)	Technical Ease of implementation	O	Quality, documentation
End-users (users of standard-compliant product)	Useability (interoperability and functionality) of standard-compliant product or service	O	Quality
Regulators	Legitimacy Due process	O, P	Quality, documentation

Figure 16 – Stakeholders' interest in standards quality (O=Outcome, P=Process) (Sherif et al., 2007)

### Standard development process

Egyedi (2000) calls for focus on the procedures of SSOs, because among others, there is a concern for the quality of standards. This concern is not new. Farrell already showed in 1996, based on a game theoretical model, how diverse interests of standard developers will cause delays in standardization and will influence a standard's quality (Zhao et al., 2005).

In another empirical study (Egyedi & Heijnen, 2008), the stability of standards is presented although limited to ISO/JTC1 ICT standards. The results show that 40% of the standards have changed over the years. Whether these changes are the result of a lack of quality is not known.

One approach to improve the quality of the telecom standards is to develop so called anti-products in parallel development (Brzezinski, 2010a). The antiproduct assesses the quality of the main product, because by parallel developing and sharing knowledge both the main product and the antiproduct will gain quality. For telecom standards this comes down to the development of four products (Brzezinski, 2010a): The base specification with an anti-product during early implementation. And includes a test specification (to validate the testability of the requirements from the base specification) with a test system as its antiproduct.

Remarkably a study by Aben (2002) shows that the number of user complaints concerning language equals the number of complaints about the technical content of the standard (Teichmann, Vries, & Feilzer, 2008).

From a more generic point of view, Morell and Stewart (1995) describe best practices for standards development based on a workshop method. The best practice consists of using Quality Function Deployment for the needs and requirements analysis of the standard. Total Quality Management (TQM) and Continuous Improvement (CI) can be used to keep the process ongoing and to assure progress is made. The best practice also includes two kinds of metrics:

1. To assess the progress of the process.
2. To measure the quality of the standards that are produced.

The best practice contains only some guidelines for metrics and the first attempt to suggest several metrics, including metrics like the number of redundant standards (process metric) and meeting the user needs of products (product metric). It stresses the importance of quality metrics even at early stages, because knowledge of those metrics can be used to set objectives and to install a sense of mission (Morell & Stewart, 1995).

### Openness

The openness of standards is seen more and more as a major selection criterion for standards by governments, software vendors and other users. Although openness is important it does not guarantee a high quality standard, and moreover it does not guarantee to be a good solution to the interoperability problem. Openness is just one quality attribute out of several. To achieve interoperability in an efficient manner, it is not enough to have openness as a single selection criterion. An overall view of quality is needed for selection purposes.

#### 5.3.3 Quality of semantic standards

Semantic standards development is different to the development of other standards. For instance, intrinsic motivation is particularly important in the context of a semantic committee (Teichmann, 2010). Intrinsic motivation can be compared to having a hobby in standards development, which impacts motivation and quality.

Two well-known case-studies regarding semantic standards are related to the MISMO and RosettaNet standards, and in both studies traces of the importance of quality can be found. In the search for critical success factors for a RosettaNet IOS project, quality was identified as a critical success factor, in the opinion of respondents: "Thanks to the high quality of RosettaNet standards, the implementation of IOS in Cisco and Xiao Tong was very efficiently carried out and at low cost" (Lu et al., 2006).

Based on the case study of the MISMO standard within the mortgage industry, a proposition (Figure 17) has been set up that supposes that the tactics that are used during standards development will influence the quality of the standard, which will, in turn, affect the success of standards diffusion (Markus et al., 2006).



Figure 17 – Proposition that relates the development process and adoption to the standards quality (Markus et al., 2006)

In addition, Markus et al. (2006) state that due to heterogeneous interests it is likely that the standard contains compromises that affect both nature and quality of the standard, and makes diffusion more challenging. MISMO shows that the “keep it simple, stupid” approach to promote diffusion is better than a perfectly designed technical standard. Based on the same case, quality related questions, currently unanswered, are raised (Markus et al., 2006):

1. Do semantic standards initiatives re-invent the wheel? Or are they borrowing from other initiatives?
2. Are there problems created in the area of cross-sector interoperability by sector initiatives?
3. What is the (diffusion) effect of the fact that semantic standards are developed by many different organizations?

The second question will become more important in the next few years, when vertical based standards become more and more adopted resulting in achieved interoperability within the vertical domain, and challenges in cross-sector interoperability. The first conflicts have been reported in literature, for instance competences that have been standardized within different domains (e.g. human resources domain and education domain) leading to conflicting standards and the need for models to deal with it (Grant & Young, 2010).

### Customizations/localizations/profiles

Information quality is an issue, and although semantic standards are intended to improve that quality issue, it is often not directly achieved (Stvilia et al., 2007). The results of a case study about exchange metadata about culture heritage based on the DublinCore standard (Stvilia et al., 2007), show that data quality is an issue (Table 23), and was not solved by applying the DublinCore standard, although it does not have to be caused by the standard.

Problem clusters	% of sample	Quality problem incidents counted
<b>Ambiguity</b>	56	Contradicting values of the same elements
<b>Inaccuracy</b>	25	Broken links to related objects
<b>Incompleteness</b>	100	Empty elements or element tags; less precision or completeness than expected for an elements; elements missing from a recommended set of elements
<b>Inconsistency</b>	82	Inconsistent formatting or representation of the same elements
<b>Redundancy</b>	54	Repeated elements containing the same values (duplicates)

Table 23 – Metadata quality problems based on a sample (Stvilia et al., 2007)

However, often these semantic standards seem far from being perfect since they are overlapping, incompatible, and not limited to their main scope (Mykkanen & Tuomainen, 2008). Even the gaps between the requirements and standards limit their usefulness (Mykkanen & Tuomainen, 2008), and thus quality issues. The lack of interoperability when using the same standard might be solved by customization and testing (Brutti et al., 2011). The lack of interoperability by using standards is caused by (Brutti et al., 2011):

1. Interoperability covers multiple layers (technical, semantic, organizational), while standards often do not cover all layers.
2. Generality and specificization; supporting as much as possible scenarios results in surpassing the acceptable level of complexity of the standard (Satisfying 20% of the requirements may lead to support of 80% of the cases).
3. The freedom within the standard that leads to uncertainty and redundancy.
4. Static and dynamic; in some case the XML schema structure is static and is always valid, while based on the changing context (dynamics) restrictions should be applied.

Based on analyzing 26 B2B standards, containing more than 3000 XSD files, and 170.000 tags, it was found 15% of the tags contained unmatched words with the dictionaries (Bedini et al., 2011), which might lead to non-interoperable implementations. It also proved that the standard could be much more simple: By using WordNet for identification of synonyms, the whole set of tags, called the e-business vocabulary is built with about 3300 different words (Bedini et al., 2011). In summary it showed that the quality of data dictionaries is not sufficient.

To be useful in real business, standards need semantic profiles that define restrictions for a specific context (e.g. specific domain, business processes, country, etc.) (Brutti, Cerminara, D’Agosta, Sabbata, & Gessa, 2010). Figure 18 shows this relation between horizontal (general), vertical standards and the needed customizations for local use. Other terms for profiles are localization and customization. This is especially needed for horizontal semantic standards, but often also for vertical ones. Otherwise, these standards have too much redundancy and uncertainty that limits interoperability in practice.

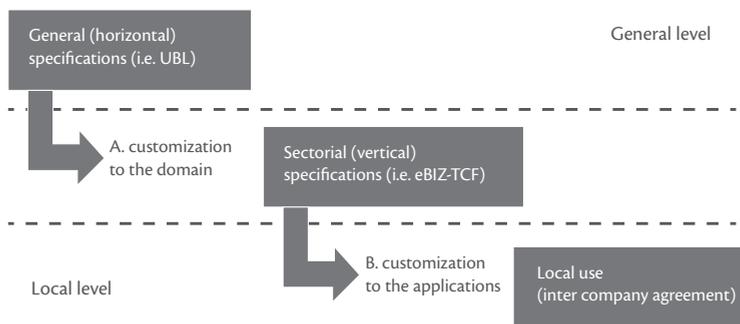


Figure 18 – The need for sectorial (vertical) standards (Brutti et al., 2010)

To point this out, Table 24 shows an example of uncertainty in practice within UBL (horizontal standard); two elements with the same semantics.

	<b>XPATH of element</b>	<b>Description</b>	<b>Occ</b>
1	OrderResponse/cbc:SalesOrderID	An identifier for the Order issued by the Seller	0..1
2	OrderResponse/cac:OrderReference/cbc:SalesOrderID	Identifies the referenced Order assigned by the Seller	0..1

Table 24 – Two different elements with the same semantics in UBL (Brutti et al., 2010)

A case study was set up to test the redundancy parameter. For several document templates (e.g. order, invoice) the number of required data objects was defined, and then tested to see how many options (redundancy) for storing this data object were available in different standards: in the horizontal standard (UBL 2.0), the vertical standard (Moda-ML XML) and a domain profile on UBL (eBIZ-TCF). Table 25 shows the results.

Document template	eBIZ-TCF Textile clothing scenario: data to be transferred	UBL 2.0 XML Schemas # of XPATHs containing data	Moda-ML XML Schemas for a fabric purchase proces # of XPATHs containing data	UBL Use Profile for a retail-side purchase process from eBiz-TCF # of XPATHs containing data
catalogue	55	38.630	99	60
order	22	2.893.732	163	36
order response	28	2.895.909	163	39
despatch	27	915.815	136	40
receipt advice	29	913.812	69	41
invoice	37	61.162	148	66

Table 25 – Case study results for testing redundancy (Brutti et al., 2010)

One might expect that a number closest to the number mentioned in the second column has the highest chance of achieving interoperability. And, on the other hand, a number much higher than in the second column might suggest low quality because this standard will be difficult to implement and will probably not lead to interoperability. Although just based on a single case, it shows that the risks of redundancy and uncertainty are much lower in vertical standards than in horizontal standards. This is because vertical standards are already much more tailored for a specific task within a more specialized context from the real world. It also shows the importance of profiles or localizations, which limit the redundancy and uncertainty of a specification. In contrast to horizontal standards, “vertical standards appear much more focused and effective to support real eBusiness” (Brutti et al., 2010).

Next to subjective methods like surveys, quality can be objectively assessed in multiple ways (Zhu & Wu, 2011):

- Manual inspection of a standard’s fitness.
- Direct measurement of quality metrics.
- Indirect assessment by measuring interoperability and other aspects of data instances created using data standards.

Within the financial domain, Bovee et al. (2002) evaluated the quality of the eXtensible Business Reporting Language (XBRL) standard to see if its vocabulary is comprehensive enough to support the needs of financial reporting (Madnick et al., 2009). This is the manual inspection of standards’ fitness. It studied the fit of XBRL with practice. Although on average the fit was good, significant differences across financial statements and industries were identified. Based on these results the need for industry-specific taxonomies is proposed (Bovee et al., 2002), in line with the study of eBiz-TCF.

Direct measurement can be automatically done by measuring two quantifiable metrics: completeness and relevance, based on counting the use of XML elements within the implementations (Zhu & Fu, 2009; Zhu & Wu, 2010). Standard quality can also be indirectly measured by assessing interoperability of data instances (Zhu & Wu, 2011). A set of data instances is interoperable if the instances use the same set of data elements defined in a data standard. Interoperability measures the extent to which the data instances have overlapping data elements. The interoperability between a pair of data instances is based on the common data elements used. The results of XBRL show that there is major room for improvement into the quality of XBRL, based on the three objective quality measures. It also shows that there is a lot to learn regarding the quality by collecting implementations of the standards in practice.

This ends our literature exploration. Although literature related to quality of semantic standards is scarce (Chapter 4), there is still relevant literature available, as we have shown within this chapter, that can be used to build upon.



# Design Approach

The previous chapters showed the problem relevance and the current state of the art on semantic standards and in particular the need for a quality measurement instrument. This chapter will set up a design research approach for this instrument.

### 6.1 Introduction

Chapter 1 already classified our quality instrument as typical artifact in terms of design science research, making it obvious to choose this approach as core methodology. Given this fundamental choice, many additional design approach decisions needs to be made to construct a design approach. This section will set up the design approach by applying appropriate research methodology. The research question for this chapter is: *How should the design process of an instrument to measure the quality of semantic standards be structured in order to achieve valuable outcome?*

We will start by presenting our research set up in section 6.2. After that we will classify our research in terms of research characteristics in section 6.3. This section will be followed by a validation of the research process on existing design principles for this type of research (section 6.4). Finally, conclusions are presented in section 6.5.

### 6.2 Research set up

Our main research question: *“What are the characteristics of an instrument to measure quality of semantic standards that will aid standard developers in improving their standards?”* can be interpreted in terms of design science as the building and evaluating of an instrument. In other words this research is focusing on developing (building) an instrument. In design science terms an ICT artifact. The process focus of design is related to two processes: build and evaluate.

In our research the build and evaluate phases are not separate, but more interwoven activities as described in action design research (Sein, Henfridsson, Pura, Rossi, & Lindgren, 2011). They are interwoven because of the interaction with the organizational context, which is activated by the main researcher who organizes the build and evaluate phases including research methodology (focus groups and case studies) in a way that utilizes the authors' context both by access to experts and access to cases.

The general structure of the build and evaluate phase is based on the work of Hevner et al. (2004) and on the process model for developing ICT maturity models, which is based on a design science research design (Becker, Knackstedt, & Pöppelbuß, 2009). As the quality instrument has similarities to a maturity model, this procedural model is to a large extent applicable especially while building. We use the sub-steps for iterative development: select design level (select parts of the instrument), select approach, design model section (design by executing approach) and test result.

The build-evaluate loop is typically iterated a number of times before the final design artifact is generated ((Markus, Majchrzak, & Gasser, 2002) as cited in (Hevner et al., 2004)). In our case the goal will be a maximum of three iterations by which the instrument will be accepted, declined or needs further revisions or re-evaluation (as suggested by Becker et al. (2009)) within the validation of our research.

Although it can be seen as part of the build process, we decided to explicitly mention the requirements specification phase in our process, in line with user centered design approaches such as ISO 13407. Requirements specification, including user involvement, is our essential starting point for both the build and evaluates processes.

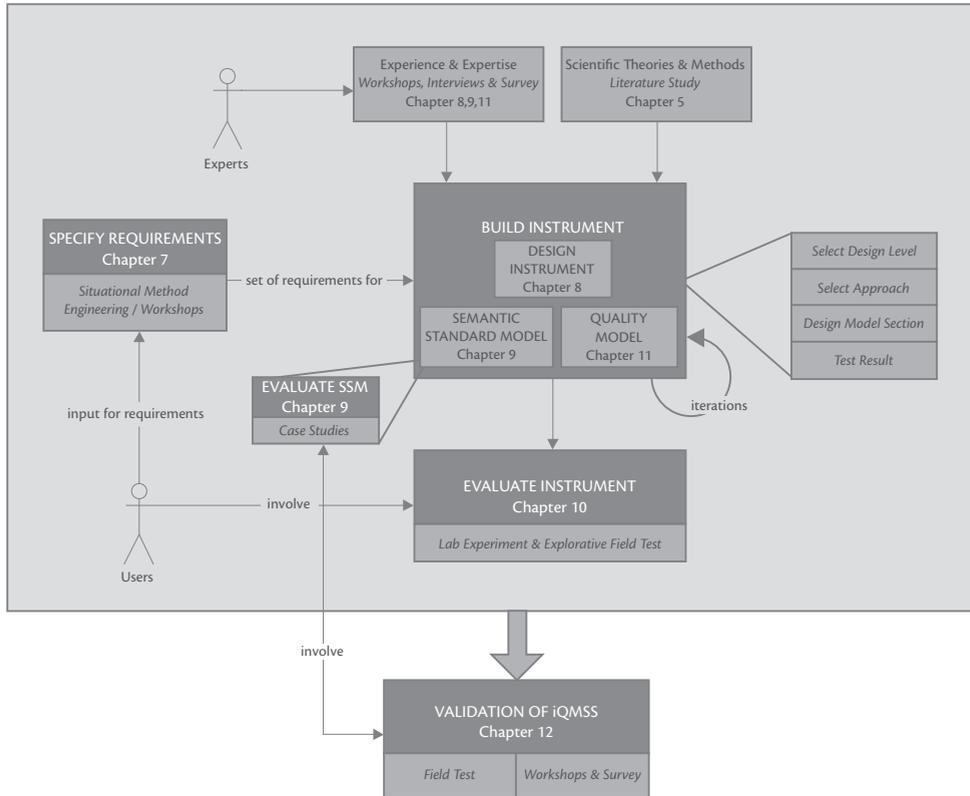


Figure 19 – Design research approach

The overall research method, including its iterative design and the chapter layout, is depicted in Figure 19. The four main phases will be explored in the next sections, followed by classification and validation of this approach.

### 6.2.1 Specify requirements

“Requirements are the things that you should discover before starting to build your project.” (Robertson & Robertson, 1999) The gathering of requirements, or requirements elicitation (Maciaszek, 2001), or requirements trawling (Robertson & Robertson, 1999), is the process of discovering and describing the desired product in terms of purpose, functionality, and other non-functional aspects like quality, easy to use, etc. Requirements specification is needed to gather the requirements from the users; it is the “what” (to build) of the instrument.

Many different requirements engineering approaches exist, many specifically created to engineer software products (Kotonya & Sommerville, 1998; Wieringa, 1996). In practice, the selection of the most suited approaches should be dependent on the project, for instance based on the value on the product (Sommerville & Sawyer, 1997). Many of the existing approaches use similar method components such as creating data and behavior models, or involving user groups during requirements elicitation. All components can be seen as method fragments, and when combined a new (situational) method is created. This situational dependency, and combining method fragments, is the core approach in situational method engineering (Brinkkemper, 1998).

For our research we decided to use situational method engineering to combine requirements engineering method fragments that are best suited for our situation. By doing this we can avoid a too cumbersome and formal process of requirements on one hand, while on the other hand we can select fragments that deliver the requirements at the level of detail and quality that we think fit our project best.

In our project two important high-level aspects are relevant for selecting method fragments:

- Involvement of end users by organizing a focus group.
- Requirement specification suited for evaluation purposes.

The latter will have impact on the level of detail and characteristics of a requirement.

### 6.2.2 Build instrument

Based on the requirements, the first iteration will start with designing and building of the first version of the ICT artifact, the quality instrument. Both practical and theoretical foundations are used to include existing expertise and experiences as well as scientific theories and methods. The instrument is an aggregation of different kind of artifacts. Although dependent on the requirements study it seems sense that an instrument to measure quality will need (Weber, 2010):

- Constructs: The vocabulary used within the instrument.
- Models: A quality model and a model of the measured object and its context (the semantic standard).
- Methods: The approach on how to use the instrument.
- Instantiation: A specific instance of construct, models and method to measure a specific semantic standard.

The complete set of artifacts will be determined based on the requirements study. Each artifact has its own build and evaluation process during the project, but in general the design steps will be similar: An approach will be selected, based on the requirements and foundations the specific artifact will be developed, and the result will be tested (Becker et al., 2009). In timing and dependency relation, there is a clear distinction between developing constructs and model artifacts and on the other hand developing instantiations and user method artifacts. To start working on the latter, the former needs be ready.

Rigor is achieved by appropriately applying existing foundations and methodologies (Hevner et al., 2004). The literature review (Chapter 5) has identified appropriate existing foundations, like existing ISO standards (ISO/IEC, 2001). Also theories from the IS field like DOI theory (Rogers, 2003) or TAM (Technology Acceptance Model) (Davis, Bagozzi, & Warshaw, 1989) might become useful, especially when the instrument will cover adoption related aspects of standards. On the other hand we want to include experiences from practice that are not part of the current literature. Therefore we need to include experts during the build iterations by including research methodology like focus group and survey.

This approach leads to multiple builds of the instrument that cannot be all presented in this thesis. We decided to include the final build as well as an intermediary build that was used during the explorative field test as evaluation.

### 6.2.3 Evaluate instrument

The other half of the loop is the evaluation phase, in which the artifacts are evaluated with respect to the utility they provide in solving the problems for which the artifacts have been built (Hevner et al., 2004). According to Hevner et al. (2004) rigorous evaluation of the utility, quality and beauty (style) of the artifacts are needed. In our case our primary goal is not to build the most beautiful and perfect instrument and therefore we choose to focus on evaluation of utility.

We need to evaluate the artifacts in a real-life situation, and get feedback on how the instrument performs within the problem situation. Methodologies, like case studies, experiments, field test, simulation, present guidelines on how to conduct the evaluation. Experiments are an appropriate strategy when “how or why” questions are being posed, just like case studies and field tests within a real-life situation (Yin, 2009).

A challenging task is how to set up the evaluation studies, for instance in the early build iterations we prefer to have explorative field test to be able to have quick iteration based on what we learn. Our evaluation studies combine case studies, lab experiments and field tests, but all with a specific semantic standard selected as “case”. For setting up our evaluation studies we combined the components of research design for case studies (Yin, 2009) with the process model on theory building from cases (Eisenhardt, 1989), and simplified it to an pragmatic approach suitable for this research:

1. Getting started (set up study's questions)
2. Selecting cases (define unit of analysis)
3. Crafting instruments and protocols
4. Entering the field
5. Analyzing data
6. Defining conclusions and implications (by applying the criteria for interpreting the findings).

Within the next section we will explain how we have implemented steps 1, 2 and 3 in our research.

#### Step 1. Getting started

For all evaluation studies the research question will be similar: How is the instrument performing on the standard? With two sub questions:

- Is the outcome satisfactory for the end user?
- What can be learned from this evaluation in order to improve the instrument?

#### Step 2. Selecting cases

First the unit of analysis needs to be determined: For instance will the complete instrument be used, or specific parts? Second the cases, the standard as subject, need to be selected. The selection of cases is an important aspect therefore some selection criteria have been used, to avoid random selection (Eisenhardt, 1989). As criteria we use maturity, complexity, functionality and domain. And finally the cases should be assessable to us as researchers. Our contribution to the knowledge base would be optimal if we included situations that are not accessible to other researchers. For generalization purposes it is important that the cases include a wide variety within the selection criteria. Which means they should differ in maturity, complexity and so on. Within the first build iterations the cases will be more explorative and participative by nature; ideally involving standards already familiar by the researchers. Later iterations, when the build of the instrument is more mature, cases should be selected that avoid researchers' bias.

In our research we have chosen for an explorative field test and lab experiment within the context of the researcher and one final field test outside this context. The first explorative field test will have action design research as main research approach (Sein et al., 2011), while the second study will be set up as an lab experiment involving groups of students (Creswell, 2009). The final evaluation should be focused on the

complete artifact. The cases itself will be selected based on arguments that will be explained in the chapter covering the evaluation studies.

### 3. Crafting instruments and protocols:

Within step three the data gathering methods are selected (interviews, documentation, archival records, etc.) based on (Blumberg, Cooper, & Schindler, 2011) and the open toolbox of techniques (Henderson-Sellers, Simons, & Younessi, 1998). Typically multiple data collection methods are used (Eisenhardt, 1989). Especially for the lab experiment an extensive preparation is needed.

Each evaluation phase may lead to either the rejection of the model or another iteration of the development phase or the reconceptualization of the transfer and evaluation methods (Becker et al., 2009).

#### *6.2.4 Validate iQMSS*

To perform credible research, validation is of utmost importance (Silverman, 2006). After the final field test has ended the evaluation phase of the final iteration, it is time for the validation of our research or more particular the iQMSS. When validating our research, we take a 3-step approach:

1. Check on requirements: Does the instrument implement all requirements?
2. Check on problem relevance and potential use: How is the instrument perceived by potential users?
3. Check on needed improvements and further research: What is needed to improve future use of the instrument?

First the requirements have to be checked if these are met. Second we have to get back to the problem survey, to revisit the conclusions and find out if our developed instrument contributes to solving the identified problems in practice. Lastly we want to know more about potential improvements, since the final build will probably not be perfect it is important to obtain suggestions for further research.

Both step 2 and 3 requires the involvement of a broad set of potential users to be able to answer these questions reliably. A focus group consisting of multiple potential users is most appropriate, and to achieve deeper understanding of individual viewpoints a questionnaire is added (Henderson-Sellers et al., 1998). Due to pragmatic reasons we combined the focus group for both steps.

Just like with the selection of the validation studies, the selection of people in focus group is of greatest importance. With focus groups the geographical representation is often an issue, just like other pragmatic reasons as time involvement of participants. To overcome this issue we decided to have two focus groups: the first one being held co-located with the OMG Technical Meeting (United States) and the second one being held in the Netherlands and organized by the government program Netherlands Open in Connection. This guarantees the participation of experts with different backgrounds and operating from different contexts. Participants of the problem survey (Chapter 3) that are geographically located in the Netherlands are invited for the second session in the Netherlands. By using partly the same group we are able to revisit the statements from the problem survey reliably.

#### *6.2.5 Summary of our research design*

To be able to answer the main research question our design approach has been decomposed in four main phases each having its own objectives, research questions and approaches as depicted in Table 26. This previous section presents only a brief high level design of the requirements specification, build and evaluate phases, and final validation. All parts are dedicated research studies that will all have a more detailed research design that are presented in upcoming sections according to the table.

Already early in the design process a distinction is made between two artifacts: The (model of) semantic standard is the subject of measurement, the quality model is the view to be used to analyze the subject. The

design of these two artifacts are together the object-design (van Aken, 2004). The instrumentation are the tools that can be used in which both models are implemented and by which the measurement can be done more effectively and efficiently; the realization design (van Aken, 2004). This separation is implemented in both the building and evaluating parts as some parts are related to one single artifact (for instance the Semantic Standard Model), while others are related to the complete set of artifacts.

The fourth part will focus on the validation of our research before Chapter 13 will discuss implications of this research and further research. This research design, or in van Aken (2004) terms “process design” completes the three design types every professional should use.

Phase	Section	Objective	Research Question	Research Approaches
<b>1. Requirements specification</b>	Ch. 7	To set up design criteria that can be used for both design and evaluation.	<i>What are the requirements for an instrument to measure the quality of semantic standards?</i>	Requirements engineering method, including elicitation techniques that involve end users (focus group).
<b>2. Building instrument</b>	Ch. 8	To build the conceptual model of the instrument.	<i>What constitutes a measuring instrument for the Quality Model of Semantic Standards (iQMSS)? What languages are available as quality language and semantic standard language? What are possibilities related to the development environment for the instrument?</i>	Design research, literature study, prototyping.
	Ch. 9.2	To build the semantic standard model.	<i>What are the characteristics of a semantic standard?</i>	Reverse engineering, literature review, expert session.
	Ch. 11	To build the quality model and its usage model.	<i>What are the characteristics of the quality model? What is an appropriate user method of the instrument?</i>	Literature study, expert session, survey.
<b>3. Evaluate instrument</b>	Ch. 10	To get better understanding how the instrument works in practice and receive input for improvements.	<i>The quality model: Is the model adequate? The quality assessment results: How useful are the results in practice?</i>	Participatory action research within explorative field test and lab experiment.
	Ch. 9.3-9.5	To evaluate the semantic standard model, in particular its applicability and usefulness in “real-life” context.	<i>Does the SSM work in practice?</i>	Case studies.
	Ch. 12.3	To evaluate the complete instrument.	<i>Does the process and outcome of iQMSS work in practice?</i>	Field test & interview.
<b>4. Validation</b>	Ch. 12	To validate the research outcome.	<i>Does the iQMSS fulfill its requirements? Do the process and outcome of iQMSS have value in practice? Does the iQMSS contribute to solve the identified problem in practice?</i>	Requirements validation, two focus groups and survey.

Table 26 – Research approach

### **6.3 Research classification**

Different components are involved in setting up a research approach (Creswell, 2009). We distinguish the research typology from the research epistemology, the research design and research methods/approaches.

#### **6.3.1 Research type**

Research can be divided between behavioral science and design science (Hevner et al., 2004). While behavioral science is aimed at prediction and explanation of behavior, the design science paradigm is aimed at the creation of new artifacts. The knowledge and understanding of a problem domain and its solution are achieved in the building and application of the designed artifact (Hevner et al., 2004), which emphasizes its problem solving paradigm. There are several characteristics of design science, which also apply to our research, including:

1. It involves the design of an artifact. This means that the result of the research is best described as a utility. Different types of artifacts exist: constructs, models, methods and instantiations (Hevner et al., 2004). In our research the intended instrument is an artifact, which inherits other artifacts as well (constructs, models, etc.).
2. It is new and innovative (no routine) design. In Chapter 4 we have seen that quality of semantic standards has been declared as research gap, and no instrument for quality measurement of semantic standards has yet been designed. Based on that, we conclude that our intended instrument is new and innovative.
3. It is a wicked problem. In our research there is complex interaction and relation between the problem and its solution, and there is critical dependence upon human cognitive abilities to produce effective solutions. In real practice it is also experienced as problem (Chapter 3).

Having said this, the applicability of design science to our research seems evident.

#### **6.3.2 Research epistemology**

Orlikowski & Baroudi (1991) state that several philosophical approaches are available for IS research, including positivist and interpretive as most used approaches. Positivist studies generally attempt to test theory, with the aim of increasing the predictive understanding of a phenomenon (Wapakabulo Thomas, 2010). The basic premise for interpretive research is that the perspective is fundamentally subjective, and thus, attempts to understand the phenomena through the meaning that participants assign to them (Orlikowski & Baroudi, 1991; Wareham, 2005). Both types require different research approaches and different roles for the researcher: an observer role within positivist research, a participating role in interpretive research (Wapakabulo Thomas, 2010).

Our study is classified as an interpretive approach, because it is assumed that our knowledge of reality is gained only through social constructions such as consciousness, shared meanings, documents and other artifacts (Klein & Myers, 1999).

#### **6.3.3 Research design**

Often three types of research design are used: qualitative, quantitative and mixed methods (Creswell, 2009). Mixed method (or multi-method (Mingers, 2001)) is an approach in which qualitative and quantitative approaches are combined, for instance the combination of a case study with a survey to draw a stronger claim on generalization (Jansen & Brinkkemper, 2009). Just like the research of Jansen & Brinkkemper, our research combines both in the sense that the focus is on qualitative research but will be supported by quantitative elements.

### 6.3.4 Research methods/approaches

Several lists of research methods/approaches exist. Chapter 4 already introduced the lists of Wareham (2005) and Orlikowski & Baroudi (1991). These contain the following approaches: conceptual, survey, (laboratory or field) experiment, (instrument) development, data (or protocol) analysis, case study, review, mixed method and action research. Not all approaches are suited for each research. Suitability depends on research epistemology and research design. For instance surveys and data analysis are not likely for qualitative research, while on the other hand (instrument) development, case studies or field experiments are more common for this type of research. In our mixed method research design the focus will be on approaches such as (instrument) development and field tests, and will be supported by quantitative approaches (e.g. survey).

Within the OPEN framework (Henderson-Sellers et al., 1998) and the Business Research Methods book (Blumberg et al., 2011) lots of techniques/methods have been identified and can be selected within our research phases. The choices for research approaches and their implementation will be more elaborated within each research phase.

### 6.3.5 Summary

The previous paragraphs have classified our research:

- Research type: Design science in IS.
- Research epistemology: Interpretive.
- Research design: Mixed methods.
- Research methods/approaches: Several, including focus groups and validation studies.

By classifying and selecting research characteristics, we gather applicable knowledge from the large academic knowledge base on this topic, which guides us designing and validating our research approach.

## 6.4 Validation of research approach

The study as being presented takes an enormous effort to perform. Dependent on the results it may be worthwhile. By having a sound research design we increase the chance of getting results that make all the effort worthwhile. Still it is hard to say if our presented research design will do the job effectively. Therefore we want to validate our research design against well known and accepted guidelines and principles for specific research methods. For this validation we chose the guidelines for design science research (Hevner et al., 2004) and checklist (Hevner & Chatterjee, 2010), and more specific in relation to the evaluation we chose the principles of conducting interpretive field studies (Klein & Myers, 1999) and design tests (Yin, 2009).

### 6.4.1 Guidelines for design science research

Design science emphasizes the role of rigor and relevance. Hevner et al. (2004) introduced 7 design guidelines, of which all should be applicable to design science research. More recently also a 8-question checklist has been developed (Hevner & Chatterjee, 2010). Both guidelines and checklist deal with the three main design science research cycle that every design science research should have, including ours: Relevance Cycle, Design Cycle and Rigor Cycle.

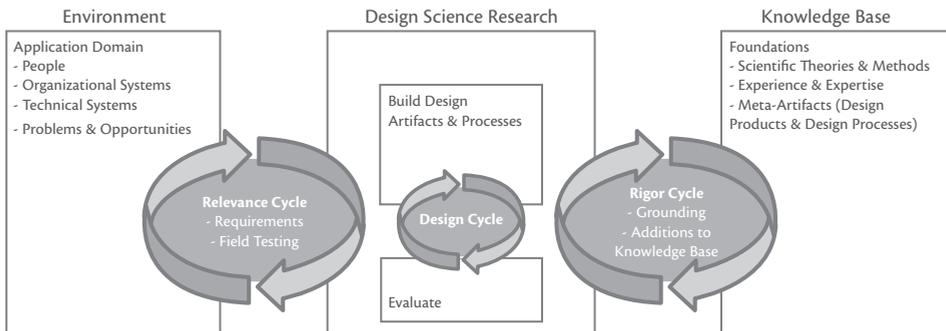


Figure 20 – The three cycles of design science research (Hevner et al., 2004)

We will explain how this is implemented in our research:

### Relevance Cycle: Environment

Within Chapter 1 we presented a research question that has been tested on relevance in our problem survey that is presented in Chapter 3. The artifacts build will fill in the needs identified in this survey. Our relevance cycle continues with the gathering of requirements from the potential users, and performing field testing by having our design tested in multiple evaluation studies. Finally we did a validation if our research contributes to solving the problem (Chapter 12), including a survey for testing the utility of the instrument.

In all, we have a strong focus on solving a real-life problem, by testing the problem statement, by deliberately adding a requirements specification phase in this research and evaluating the design in practice and validated our research with problem stakeholders to check if our research question has been satisfactorily addressed.

### Rigor Cycle: Knowledge Base

Our rigor cycle starts with identification of a research gap based on a systematic literature review (Chapter 4). It also identifies the foundations available in existing studies to build upon. The identification of the research gap (Chapter 4) and literature review (Chapter 5) describes the current state within scientific community. Next to the identification of a research gap, our contribution focuses on the developed artifacts that now have become available to the knowledge base. The artifacts build will fill in (part of) the gap, and are grounded on fundament of theories from existing literature, based on the extensive literature review (Chapter 5). The conclusion section (Chapter 13) will present the contribution of this research to both practice and scientific community.

Research rigor is implemented by designing a validated research method containing research techniques for both build and evaluate phases that builds upon existing knowledge from the design science arena. This chapter contains this research design including in this section the validation of our research approach.

Finally we put a lot of effort in communication of our research. Not only this book is available, but also many parts have been published at several conferences, journals, doctoral consortia, etc. But not only the scientific community; we also held sessions for the practice community to disseminate the results of our work, and will continue to do so.

### Design Cycle: Build & Evaluate

A sound detailed research method is needed to be able to develop a relevant and rigor instrument. The first step within the build phase is the identification of the artifacts. In our design of the instrument we recognize different artifacts, including constructs, models, methods and instantiations. Our design is a search process in which an optimal instrument may not be feasible. This research focuses on a satisfactory solution, and recognizes that an iterative strategy is needed for good design which is implemented in several design cycles.

The developed artifacts will be evaluated, and based on the results the artifacts will be improved. In total we have seven build versions of the instrument, of which three evaluation results are presented within this research. Another cyclic process is the use of requirements that in the validation are used to check if the built artifact meets the requirements. Although rigorous evaluation consists of the utility, quality and beauty (style) of the artifacts (Hevner et al., 2004), our research focuses on the utility. In our methods for design evaluation we combine observational, analytical and experimental approaches. The final phase will comprise an overall evaluation of the project and will answer the main research question, additions to the research community, and possibilities for further research.

#### *6.4.2 Principles for conducting and evaluating interpretive field studies*

In particular we wanted to validate our approach in how to use our evaluation studies. Next to the four design tests for case studies (Yin, 2009), there are 7 principles available for interpretive field research (Klein & Myers, 1999). We will use these principles in validating our research as our research is classified as interpretive research. Differences between case studies and field studies are minimal. There is no hard distinction between both (Klein & Myers, 1999), only that field studies require long period of time in the “field” (Yin, 2009). Field studies include in-depth case studies ((Walsham, 1993) as cited in (Klein & Myers, 1999)). Based on that conclusion both the design tests for case studies and the set of principles for conducting and evaluating interpretive field studies (Klein & Myers, 1999) will both hold for our evaluation studies.

Both design tests and seven principles deal with performing basic good research. Such as the reliability test that implicates that data will be carefully managed. Without going into detail on these general applicable good research practices we describe three fundamentally important aspects for our research based on the principles and design tests:

##### 1. Context

A critical reflection on the context of an evaluation study is required. We implemented this by including contextual descriptions of all evaluation studies in our research, and discuss potential consequences.

##### 2. Interaction between researchers and subjects

The interaction between researcher and subject is particularly relevant to our study since we earlier stated that in our design science approach we included action research concepts in line with action design research (Sein et al., 2011). We took several steps to positively deal with this interaction: First we described the involvement of the researchers with the subjects in our evaluation study approach. Second, we were sensible about the consequences like biases, and did not generalize these results. Third, we selected different research methods and different cases that have led to different level of obtrusiveness and interaction.

##### 3. Sensitivity & validity

Good research requires to be designed for validity by including sensitivity to possible differences in “interpretations”, “biases” and other “distortions”. Again we constructed validity by including multiple research methods for the collection of data within the evaluation study, including experiments, interview and observation. We report our sensitivity when relevant, and our sensitivity has led that we are careful about the results of our evaluation studies, and performed an additional survey to show the general value of our research.

#### **6.5 Conclusions**

The research within this chapter has been carried out to answer the following research question: *How should the design process of an instrument to measure the quality of semantic standards be structured in order to achieve valuable outcome?*

A sound detailed research approach is needed to be able to develop a relevant and founded instrument. Based on design science methodology, a research approach was constructed to build and evaluate the instrument. The presented research approach is carefully constructed in several iterations and based on good practice in design science research. Although presented as is, different options have been considered. The designed research approach was considered as optimal as it reflects many good practices and suits the context and goal of this research. Our research approach is characterized by a focus on an interplay between design idea of researchers and the forces in the environment; in line with the principle of authentic and concurrent evaluation of action design research (Sein et al., 2011). It can also be classified as a mixed method (qualitative and quantitative) research design for interpretive research in which multiple research methods will be used. The research design has been validated by utilizing theory of guidelines, principles, and tests related to both design science and case and field study research. The result of this section is a sound research design for developing an instrument to measure the quality of semantic standards. The next chapters will be the continuation of the execution of this research approach.



# Requirements

The previous chapter outlined the design approach for building the quality instrument for semantic standards. Before we can start building the instrument, it is necessary to determine its requirements, bringing us to the main research question in this chapter: *What are requirements for an instrument to measure the quality of semantic standards?*

The goal of this chapter is to answer this research question by performing a requirements study. We start by presenting the chosen research method, followed by details about the gathering process. We will present the resulting requirements, as well as our conclusions.

### 7.1 Research approach

For our purposes, we decided to embrace the notion of situational method engineering (Brinkkemper, 1996; Brinkkemper, 1998; Coulin, 2006), implying the configuration of an approach tuned to the project at hand. We assemble our requirements gathering process using method fragments from three well-known requirements engineering methods: QFD, KAOS, and Volere. Although many more methods exist and could have been selected, this selection combines requirements engineering methods with long-standing history (QFD) and formal (KAOS) and pragmatic (Volere) oriented methods.

QFD (Quality Function Deployment) is a method for requirements elicitation and transformation of requirements into product design. It has been developed by Akao (1990), based on the quality concepts of Deming. It is primarily used for designing physical products, but can also be used for ICT products. Its best known aspect is the so-called House of Quality. But, QFD also includes a team-based iterative method for understanding customer requirements. The House of Quality is a matrix with the “whats” and the “hows” plotted on each of its axes. The “whats” represents the customer requirements and the “hows” represent the functional requirements for the system. In consecutive steps, the “hows” from the previous step are the “whats” for the next step. This gives a leveled structure to requirements, while maintaining the link with the customer requirements at the highest level.

The KAOS (Knowledge Acquisition in automated Specification) methodology has been designed at the University of Leuven (Louvain) in the early 1990s, and continued to be improved (Al-Subaie & Maibaum, 2006). KAOS is a so called “Goal-Oriented Requirements Engineering (GORE)” (Van Lamsweerde, 2004) method. In a GORE method, a “goal-tree” is developed, in which all goals and requirements must contribute to a higher-level goal in the goal tree, and eventually leading to one pre-defined top-level goal. This property supports the requirements elicitation and selection process because one can find higher level goals by asking the “why” question, and lower level goals by asking the “how” question. It is a flexible method in the sense that it supports top-down, bottom-up, and middle-out requirements gathering. It also makes it possible to start by providing guidelines on how different requirements relate, and by relating requirements to a pre-defined, top-level goal.

---

This chapter is based on:

(Folmer, Krukkert, Oude Luttighuis, & Van Hillegersberg, 2010): Requirements for quality measurement instrument for semantic standards. In: Graz, Jean-Christophe, Jakobs, Kai (Eds), EURAS Proceedings 2010 Service Standardisation, The EURAS Board Series, ISBN: 978-3-86130-245-2, pp. 151-162, 2010, Verlagshaus Mainz GmbH Aachen.

The pragmatic oriented Volere Requirements Process Model (Robertson & Robertson, 1999) is a process for gathering and testing requirements. An important pragmatic element of Volere is the Volere shell: A template to make sure you gather all information about a requirement, such as the history of the requirement, customer satisfaction, its rationale and fit criteria.

Although each of these three methods might have done the job sufficiently for our requirements study, neither one perfectly matched our situation. Therefore, we chose a combination of elements from each of the methods, based on our constraints (Maiden & Rugg, 1996).

The goal-tree approach of KAOS was selected for its ability to structure requirements, just as the reasoning approach (asking how and why questions). KAOS' formal information modeling approach was not chosen, because it was too extensive for our purposes.

From QFD, we took the customer approach and the workshops, as efficient ways to involve stakeholders. Since domain expertise is essential in requirements elicitation, we involved potential end users of the instrument in our workshops. Participants have backgrounds in international standardization initiatives and compliance testing. A two-step approach was chosen to improve the results and also not to ask too much time of participants. The workshops were led by requirements gathering experts, who afterwards processed the outcome in a consistent and complete result. The House of Quality was not used, because it includes considerable amounts of physical product-related aspects, and the goal-tree from KAOS presents a viable alternative intended for use in the domain of IS.

Like QFD, Volere emphasizes end user involvement, the role of domain expertise and the use of workshops for elicitation. Our use of the KAOS goal tree ends with the identification of the requirements. From the Volere shell, we took additional attributes of the requirements, like "Fit Criterion", enabling us to express how compliance to a requirement can be tested, thus preventing requirements from vagueness. And "Priority", useful when requirements compete, either because they conflict, or because implementation resources are scarce. Another important attribute from Volere: "rationale" was not explicitly used, since it follows from the goal tree.

Our situated requirements specification method was constructed from a set of method fragments that were selected based on the research context, which is summarized within Table 27.

<b>Lightweight Requirements Specification Method for Quality Instrument</b>		
<b>Research context need:</b>	<b>Method fragment:</b>	<b>Usage:</b>
1. Linking goals and requirements	KAOS-goal tree	Setting up a goal tree.
2. Pragmatic feasible approach	QFD and Volere	Customer approach in workshops (QFD and Volere), and avoiding KAOS formal modeling approach, just as House of Quality.
3. Structured description of requirements	Volere	Adapted version of Volere template used for documenting requirements.

Table 27 – Situational method engineering for requirement gathering quality instrument.

## 7.2 The process

Preparations were carried out for two workshops involving potential end users. In our research we have access to a broad range of experts within TNO, the largest independent research institute in the Netherlands. To utilize this access we organized the workshops at two different TNO locations, Delft and Enschede. Five domain experts participated in the first workshop, which was held in June 2009. The experts brought experience from various semantic standards to the table: temporary staffing (HR-XML, SETU), finance/e-invoicing (UBL), disaster management, education (IMS, Edustandaard) and healthcare (HL7, CEN/EN 13606, Nictiz, Vektis).

In the second workshop, also had contributions from five experts. This time, experience from technical SSOs was involved: IEEE, 3GP, OMA, OPT, and ITU-T. Although these are not the main type of potential users, this session was extremely valuable. Experts involved in technical standards have many years of experience, while expertise within semantic standards is relatively new. Semantic interoperability does not have the same rich history as technical interoperability.

The exact form of the instrument (e.g. software tool, method or book) was not determined prior to the workshops. We wanted the participants not to feel restricted beforehand. Also, the meaning of concepts like “quality” and “semantic standard” was left implicit. This turned out to work quite well, since interesting discussions started on details of definitions. Figure 21 was used as the starting point of the workshops. It shows the instrument as a black box converting input (standard) to output (report). It also suggest possible forms of the instrument, such as a kind of handbook and/or tool. Different actors are shown as future users of the instrument. The distinction between the principal and tester shows a possible differentiation in the person who commissions the use of the instrument and is selecting the measurements and the persons who is carrying out the measurements (tester).

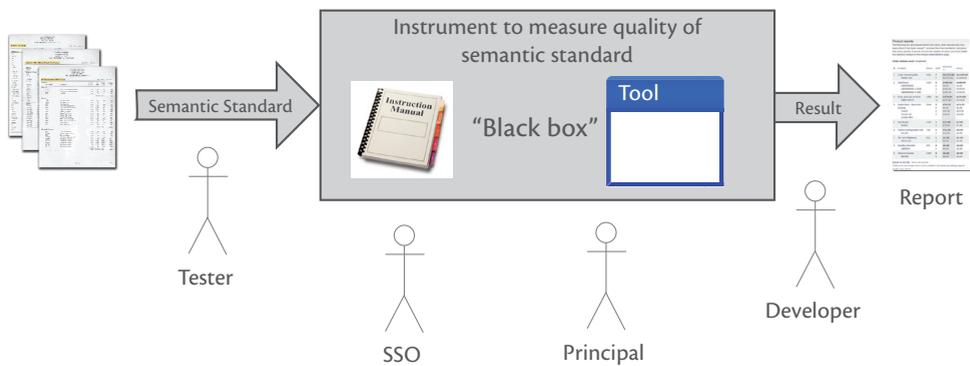


Figure 21 – Context diagram

During the workshop, participants were asked to think about, and write down the requirements and, after several minutes, present them to the other participants. With help of the requirements expert, the requirement was then added to the goal tree. This process was repeated several times during the workshop. This constitutes a bottom-up iterative approach, starting with a set of initial requirements and expanding it by asking how and why questions.

The result was a large amount of post-its, including redundant requirements, vague descriptions, general remarks, etc. Processing these involved selection and removal (redundant requirements, remarks), structuring within a tree, completing the goal-tree by adding requirements, and formulating the requirement. Then, the requirements were annotated with fit criteria and priorities.

### 7.3 The results

This section presents the goal tree gathered from the workshop sessions. It starts with the top-level goal, and the three level-two goals. In each of the following sections, one of the level-two goals is further decomposed. At the bottom of the goal tree (the leaves), requirements are specified.

#### 7.3.1 Overview

The top-level goal of the instrument is to support semantic SSO's in developing high quality standards. The rationale for this goal is the general believe that higher quality standards will lead to improved interoperability (Chapter 3). Figure 22 shows the top-level goal, and the three level-two goals.



Figure 22 – Top of the goal tree

These three sub-goals will be decomposed in the following sections. A detailed description of all requirements is presented in Appendix F.

### 7.3.2 Useful for semantic standards of different SSOs

Figure 23 gives an overview of all the sub-goals and requirements that need to be fulfilled for this level-two goal.

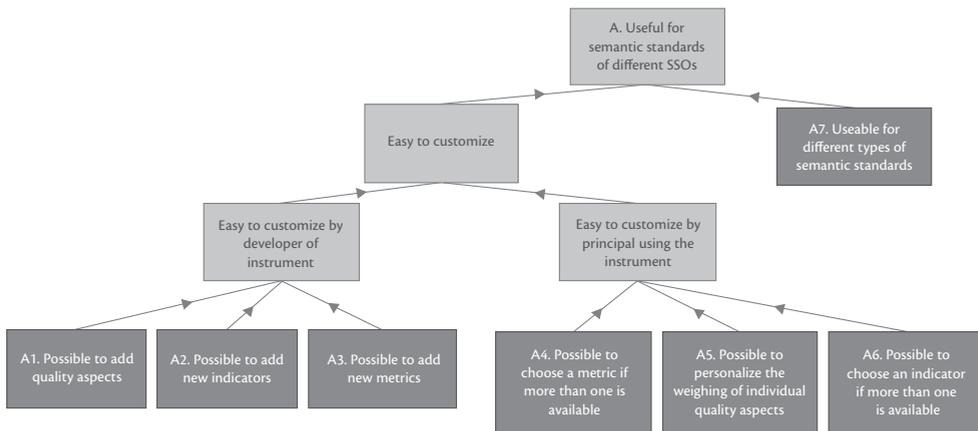


Figure 23 – Useful for semantic standards of different SSOs

First, the instrument should be easy to customize. This is because SSOs differ in their approach and in the quality aspects they find important. Also, the instrument should be useable for different types of semantic standards.

Regarding the customizability of the instrument, it is important that several roles involved in using the instrument can perform the customization. For the designer of the instrument, it is important that new elements (quality aspects, indicators, and metrics) can easily be added to the instrument. For the principal using the instrument, it is important that he can easily choose between different elements (e.g. measurements) of the instrument (if more are available), and that he can customize his “view” on quality by easily changing the weight factors.

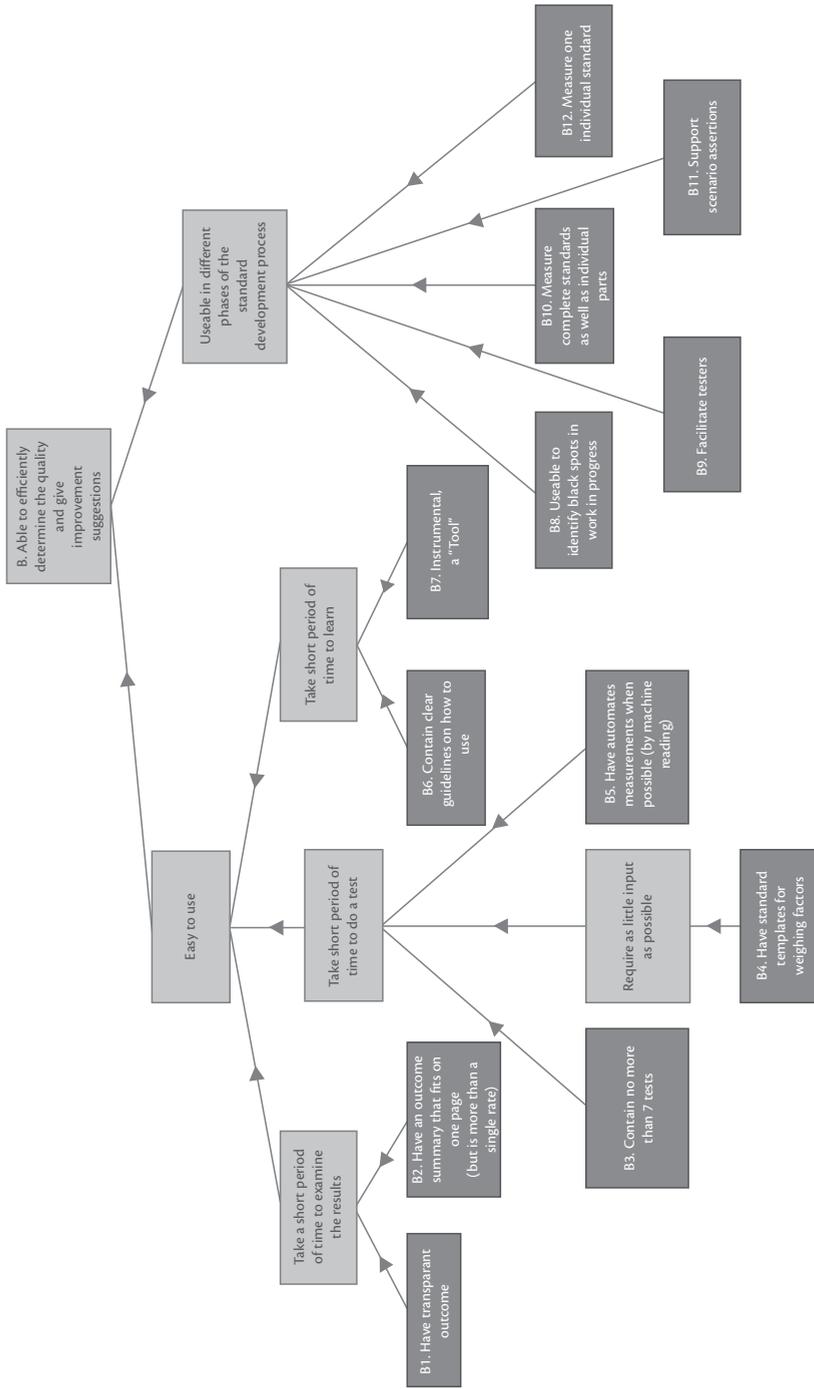


Figure 24 – Efficiently determine the quality and improvement suggestions

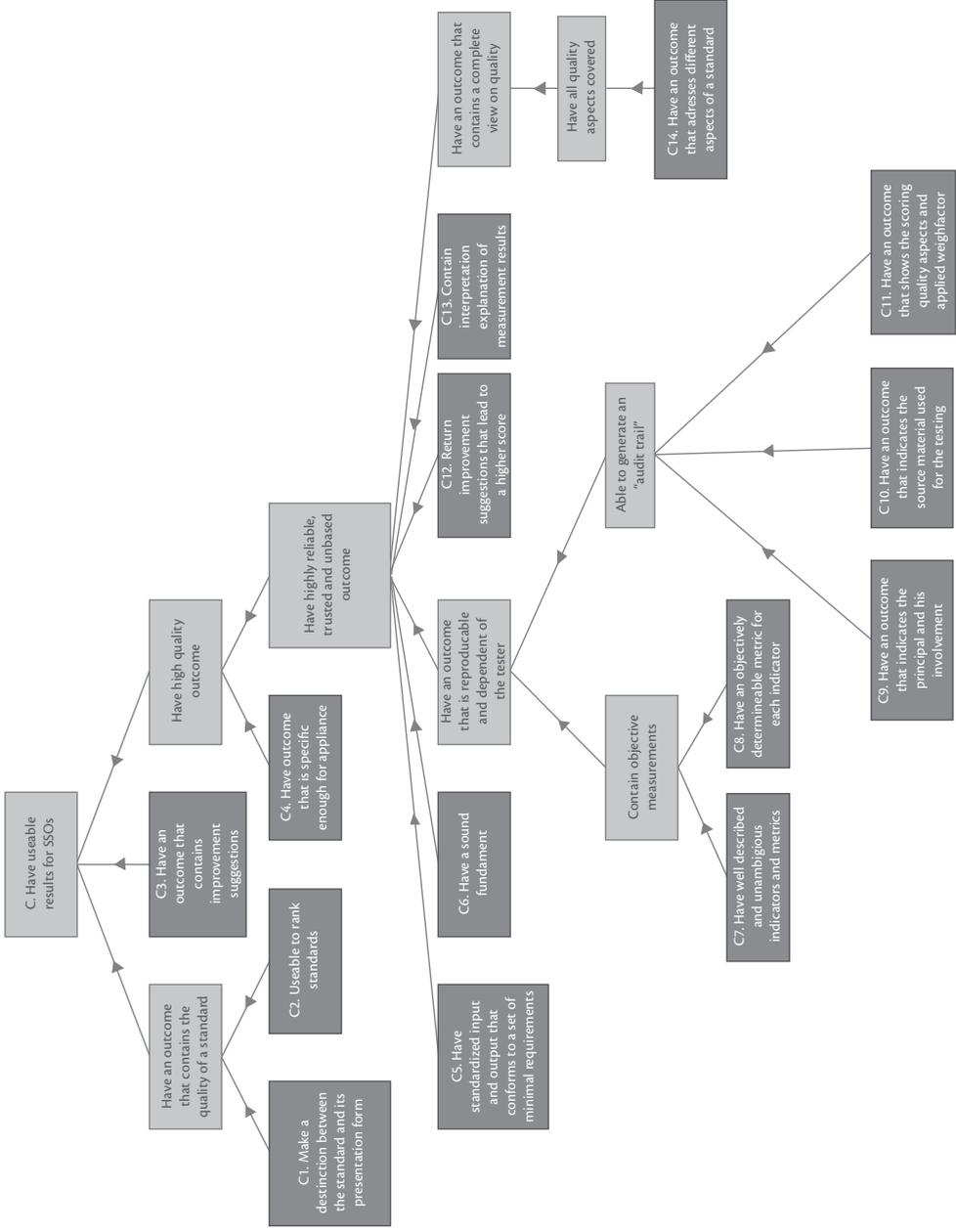


Figure 25 – Useable results for SSOs

### 7.3.3 Efficiently determine the quality and improvement suggestions

Figure 24 shows the decomposition of this level-two goal. During the workshops, ease of use mainly focused on the time required to execute certain activities. A distinction was made between the time for learning how to use the instrument (short learning curve), the time for executing a test, and the time for interpreting the results. Requirements were specified for all three aspects. In order to reduce the time taken for execution, the instrument should require as little input as possible.

In order to make the instrument useable in different phases of the standard development process, a number of requirements have to be met. These requirements focus mainly on the input that has to be provided to the instrument, as well as some functional aspects of the instrument.

### 7.3.4 Have useable results for SSOs

Besides providing quality scores for a standard, the instrument should also provide the user with suggestions for adjusting the standard so that a higher quality can be achieved. An instrument for determining the quality of standards should, of course, have a high quality output itself.

Figure 25 shows the goals and requirement that have to be met in order to have useable results.

In order for the outcome of the instrument to be of high quality, it should be reliable, trusted, and unbiased. Besides, the outcome has to be reproducible independent of the tester. This can be achieved by generating an audit trail, and having objective measures.

Also, the instrument should enable a complete view on quality, meaning that all quality aspects can be covered.

### 7.3.5 General observations and discussion

During the workshops, we focused on gathering requirements for the quality instrument. Nonetheless, we received several suggestions for specific quality aspects. These quality aspects were not used in this phase of our research, but are an interesting “by catch” for usage in a later stadium.

Another important notice is that quality is situational and time-dependent. This means that quality statements may change over time. It also implies that aspects of the problem environment should be part of quality.

Another valuable contribution was the suggestion of the following requirement: The instrument should indicate the value of the standard for: 1. Investment, 2. Solution/Cost reduction, 3. Commercial (Patents). Although interesting, we think it does not support the highest goal in our goal tree. The commercial value of a standard seems irrelevant, and might even contradict, for the highest goal related to achieving interoperability. This requirement might lead to an interesting but different instrument for example a kind of adoption measurement instrument that can be used by individual organizations to determine whether or not to invest the adoption or development of a standard.

Finally, people can hardly think of requirements without thinking of possible solutions. Requirements gathering processes gain focus as they proceed, but have to end as well avoiding designing solutions during requirements engineering. In our case, the scope was set by having a short presentation about the problem domain within each workshop. We stopped the requirements engineering process after two rounds of workshops and engineering the results of the workshops because by then the results were sufficient to start the design cycle and an additional requirements workshop would have limited value.

## 7.4 Reflection on requirements

The second workshop was held with experts having a technical background, who usually are involved in technical, telecom-related standards. Although the instrument is aimed to focus on semantic standards, the impression came to mind, based on the similarities in workshop discussions, that an instrument that focuses on technical standards might have similar requirements. This may imply that the requirements are not detailed enough, but might also imply that the instrument might be useful for other type of standards as well.

On the other hand, we also found that all requirements engineering methods we examined assume that a product (physical item or software) is going to be produced. In our case, we have a more abstract concept “instrument”, without having chosen the exact representation yet. Although this is not unique, this may have resulted in requirements that are abstract. One drawback of abstract requirements is that it is hard to determine whether we have a complete set of requirements. This makes it even more important to not only test whether the instrument fulfills the requirements, but also whether it presents a solution to its users.

The lightweight situated requirements specification method worked quite well and produced requirements that seem relevant. The result is a set of structured requirements presenting a rich set of information. We did notice however that a lot of functional requirements were identified, and only very few non-functional. A possible explanation is again the abstract notion of the instrument and possible abstract requirements.

In both workshops, the experts made a distinction between a standard (consisting of a set of agreements, but quite an abstract concept), and the representation of the standard, for example a paper document. One standard can have multiple representational forms, for example in different languages. Both the standard (as an abstract concept) itself as well as the representational form have quality aspects. It may even occur that the standard itself has a good quality, while one of the representational forms has a poor quality. This poses an interesting problem: how does one measure an abstract concept? Also, if the quality of a standard is measured using one of the representation forms, how can one distinguish between the quality of the standard itself, and the quality of the representation? This is quite similar to measuring the quality of software.

We already concluded that the instrument might be useful for multiple types of standards. It would have been interesting to compare our results with other studies regarding requirements for quality instruments. Unfortunately, to our knowledge, very little research has been done on requirements for instruments that can be used to measure quality, which makes comparisons hard.

## 7.5 Conclusion

Using a situational method combining fragments of QFD, KAOS, and Volere requirements engineering methods, we constructed a set of requirements for a quality measurement instrument for semantic standards, and structured them in a goal tree.

The top goal “Support semantic SSOs in developing high quality standards” has been decomposed into three level-two goals, which have been further decomposed and can best be summarized as:

- Usefulness for different semantic SSOs.
- Efficient to use.
- Usable end results.

Overall we can conclude that the presented set of requirements do contribute to our knowledge about the desires from standardization practitioners regarding an instrument for quality measurement. Within the next chapter a start will be made building the instrument based on these requirements.

# Design of the iQMSS

This chapter describes the design of a reference model for an instrument to measure quality of semantic standards, within the first section. Section 8.2 will introduce the meta elements of the reference model, while section 8.3 will discuss in more detail the implementation and tooling related elements of the reference model.

### 8.1 The reference model

This section continues the design process and presents an overall reference model for an instrument for quality measurement of semantic standards. It will not present the instrument itself, but rather the overall model and concepts needed for developing it. The following research question is the central theme for this chapter: *What constitutes a measuring instrument for the Quality Model of Semantic Standards (iQMSS)?*

The answer to this question lays out the core elements of the instrument. These elements and their relations form the model of the iQMSS. In terms of design science research (Hevner et al., 2004), the reference model of the iQMSS is a design artifact. Evaluation will be done in later stages. An iterative design approach, involving several experts, has been chosen to execute this stage of the design science research. The design of an iQMSS contains similarities with the design of an IS, more specifically, the design of the reference model is comparable to the development of a system architecture (Nunamaker, Chen, & Purdin, 1990).

As a starting point, design constraints have been set up based on existing knowledge and the requirements gathered for the iQMSS (Chapter 7). We use meta-modeling techniques and MDA (Model Driven Architecture) (Kleppe, Warmer, & Bast, 2003) for the clustering of design artifacts, and UML (Unified Modeling Language) (Fowler, 2004) for the method description.

#### 8.1.1 Design constraints of the iQMSS

Research issues related to the development of a system architecture within IS development, relevant for a quality instrument have been summarized by (Nunamaker et al., 1990):

- Develop a unique architecture design for extendibility, modularity, etc.
- Define functionalities of systems components and interrelationships among them.

Also useful are the elements from a quality framework for interoperability within the government domain defined by the SEMIC.EU project. It contains amongst others (Nentwig et al., 2008):

- Requirements for the definition of quality criteria for assets and their artifacts.
- An appropriate structure and representation of these quality criteria assigning these quality criteria to assets and their artifacts according to stakeholder requirements.
- Requirements for assessing, preferably measuring the quality of assets and their artifacts.

---

This chapter is based on:

(Folmer, Oude Luttighuis, & Van Hillegersberg, 2011c): A Reference Model of an Instrument for Quality Measurement of Semantic IS Standards, In: K. Blind, K. Jakobs (eds.), Proceedings of the 7<sup>th</sup> International Conference of Standardization and Innovation in Information Technology (SIIT 2011), ISBN: 978-1-4577-2020-8, pp. 69-78, Mainz Publishers, Aachen.

Although the SEMIC.EU requirements are useful, it is important to involve potential users as well. The main requirements for iQMSS result from a requirements specification study (Chapter 7), in which the top goal is defined as “support semantic SSOs in developing high-quality standards”. This requirement restricts the usability of the intended instrument to standard developers. This implies that, for instance, the selection of standards by standards users is taken to be out of scope for the iQMSS. This goal has been decomposed into three level-two goals, which all have been further decomposed, grouped and reformulated into five main requirements (REQ):

1. Usefulness for different semantic SSOs.  
REQ1: Accommodate the differences between semantic standards.
2. Efficiency of use.  
REQ2: Focusing on the quality needs/question of a specific SSO.  
REQ3: Implementation of the instrument in easy to use tools.
3. Usable results.  
REQ4: High-quality outcome based on scientific quality model.  
REQ5: Output report as fundament for improvement project.

### 8.1.2 iQMSS use case

This section will describe use cases for quality measurement with the iQMSS and identify the main actors and constructs of the iQMSS. The modeling notation used in the use cases and package diagram is conforming to the UML definition. We identify four actors:

- Initiator: Expert with in-depth knowledge of the iQMSS, knowing how to customize it for application for a specific standard.
- Principal: Expert with knowledge about applying the iQMSS and semantic standards in general. The principal has the lead in carrying out the measurement.
- Client: The client has an information need regarding the quality of a specific semantic standard and basic understanding about that semantic standard.
- Standard Developer: The standard developer has detailed knowledge regarding the semantic standard. This knowledge is needed during the process of quality measurement.

In practice, the initiator and the principal may be one person. The same holds for the client and standard developer. This use case diagram (Figure 26) shows the actors and their activities during and preceding the measurement.

Because semantic standards are diverse and will have different attributes, the measured concepts defined in the iQMSS have to be adapted before they can be used in the measurement process of a specific semantic standard. Together with the client, the initiator selects the measurable concepts to be used in the measurement, based on client’s information needs, and adapt these concepts to the specific situation for the semantic standard at hand. Based on the generic iQMSS, the initiator is responsible for the customization of the tools to the specific quality information needs and specific semantic standard under subject. After customization, the actual measurement of the standard can be executed.

The measurement process is performed by the principal. He or she performs the measurements and reports the results of the measurement to the client. These results are also evaluated by the standard developer. The latter should be able to interpret the results for the semantic standard, and to define quality improvement actions when needed.

The principal uses the customized iQMSS, in two basic steps, each presented as a lower-level use case. First, the principal gathers detailed information about the semantic standard at hand. The standard developer serves as the major source for this information. Second, the quality model needs to be initiated based on the gathered information resulting in an outcome that will be analyzed and documented.

Alternatively, a fifth actor may be imagined: an analyst who is equipped to analyze the overall results of several measurements of different semantic standards. After this analysis, the results of a measurement can be compared to other results and the quality measurement of a particular semantic standard can be positioned against other results.

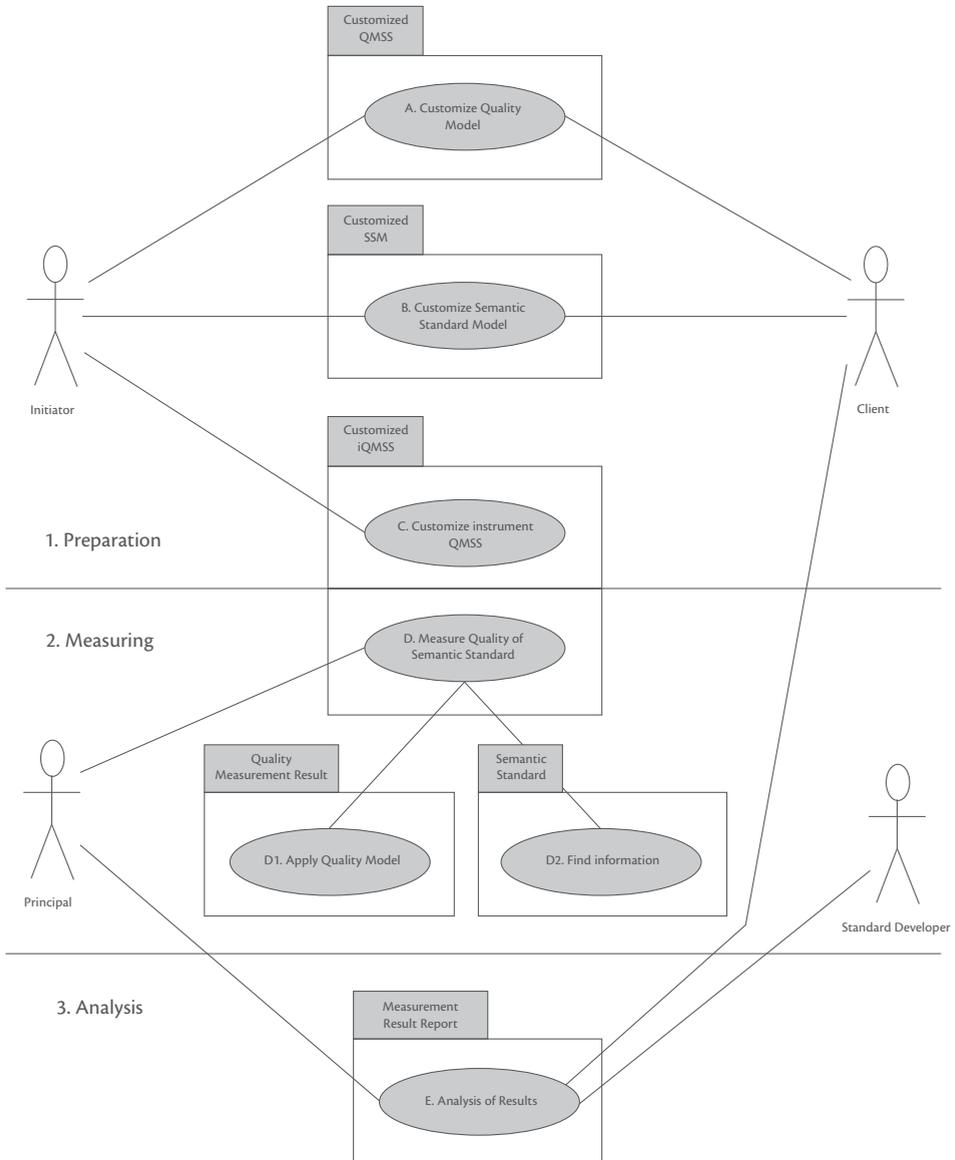


Figure 26 – Use case diagram iQMSS

8.1.3 Design of model iQMSS

When measuring, the initiator has to match the information need with what is measurable (the attributes of the measurement object). It makes no sense to measure things the client does not want to know, and if the client wants to know things that are not measurable makes no sense as well. This leads to the first design rule in which the measurement subject and “what we want to know” are separated as different perspectives:

- The first perspective is from what we want to know which in our case is related to quality. The information needs from the end users will be part of what we call the quality model for semantic standards (QMSS).
- The second perspective contains the semantic standard as measuring object, in a broad sense, covering more than just the specification of the semantic standard, and including the broad context (including implementations, etc.) of the semantic standard.
- A third perspective is added to visualize that in the end we need an implementation of the artifacts in some kind of tooling: the instrument itself.

Each perspective contains different artifacts on different abstraction levels. Our second design rule is the distinction between different levels: In terms of MDA (Kleppe et al., 2003) we have three levels of each artifact: The model (M1) is expressed in a language (M2), and the model is instantiated (M0).

Based on our requirements (REQ1, 2, 4) we have to make a generalization-customization separation of concerns. A general model needs to be customized based on the context. The relation between the generic and customized version is a subset relation type. The customized version is a subset of the generic version, customized to the specific context of the instance of the semantic standard as subject.

From another MDA viewpoint, our QMSS and SSM can be seen as PIM (platform-independent model), just as our usage models of the instrument. The instrument part is the translation of the PIMs into a PSM (platform specific model) and code. The PIMs can also be translated into different PSMs, meaning that different instruments based on the PIMs can be introduced.

This leads to the following reference model for iQMSS, based on the design constraints mentioned. The three perspectives are captured on the horizontal axis, the abstraction levels on the vertical axis. The iQMSS itself is the overall artifact. Emphasized in bold is the development core of the instrument: the generic elements of the iQMSS; the sub artifacts.

**instrument for Quality Model of Semantic Standards (iQMSS)**

		Specification Quality Model Semantic Standard (QMSS)	Specification Semantic Standard Model (SSM)	Implementation
M2	Model of	1. Quality Language (QL)	I. SS Language (SSL)	A. Development Environment
M1	Customization of	<b>2. Generic QMSS</b>	<b>II. Generic SSM</b>	<b>B. GMI Generic Model Implementation</b>
M1	Model of	3. Customized QMSS	III. Customized SSM	C. CMI Customized Model Implementation
M0	Model of	4. Measurement Result	IV. Semantic Standard	D. Measurement Result Report

Figure 27 – The reference model of iQMSS

This model contains a total of twelve constructs. The table below contains definitions of the three columns and the overall model of iQMSS.

Term	Definition
<b>Model of iQMSS</b>	The overall model of the measuring instrument for Quality Model of Measurement of Semantic Standards.
<b>Specification QMSS</b>	Specification on all levels related to the Quality Model for Semantic Standards (What we want to measure).
<b>Specification SSM</b>	Specification on all levels related to the Semantic Standard Model (What we are able to measure).
<b>Implementation</b>	The implementation is related to practical tooling; the output of the tools, the generic and customized version of the tools, and the development environment for the development of the tools. The tools are supporting the QMSS and SMM.

Table 28 – Definitions overall iQMSS

The twelve constructs from the iQMSS model are listed in the next table.

Term	Definition
1. Quality Language (QL)	The language in which the QMSS is expressed, consisting of defining constructs and their relations.
2. Generic QMSS	The generic version of the Quality Model of Semantic Standards using the QL to express the aspects within the model.
3. Customized QMSS	A specialized version of the QMSS adapted to the characteristics of the semantic standard.
4. Measurement Result	The result of a quality measurement on a specific standard with use of the QMSS.
I. SS Language (SSL)	The model of the SSM, consisting of a language in which the constructs are defined.
II. Generic SSM	The generic model of a semantic standard expressed in concepts from the SSL.
III. Customized SSM	A customized version of the generic model of the semantic standard, customized to be fit for a specific standard.
IV. Semantic Standard	The instance of a semantic standard as subject for measurement.
A. Development Environment	The components useful for building tools.
B. Generic Model Implementation (GMI) (generic iQMSS)	The generic version of tooling making the QMSS instrumental of nature. The tooling should support the measure process.
C. Customized Model Implementation (CMI) (customized iQMSS)	A customized version of the tooling based on the characteristics of the specific semantic standard.
D. Measurement Result Report	The output (report) of the tooling when performing the measurement.

Table 29 – Definitions iQMSS constructs

The implementation column applies to implementations in the broadest sense; this might consist of software tools, questionnaires, instruction document, or any other physical or digital tool for performing the measurement. It should guide and assist the initiator and principal of the iQMSS in effectively and efficiently making use of the iQMSS leading to useful results in the measurement result report.

There are more relations between the artifacts than Figure 27 shows. This section takes a closer look at the relations between the artifacts in the model. Most important is the relation between the QMSS (quality model) and the SSM (semantic standard model). The SSM defines the measure points that can be used by the QMSS (Figure 28).

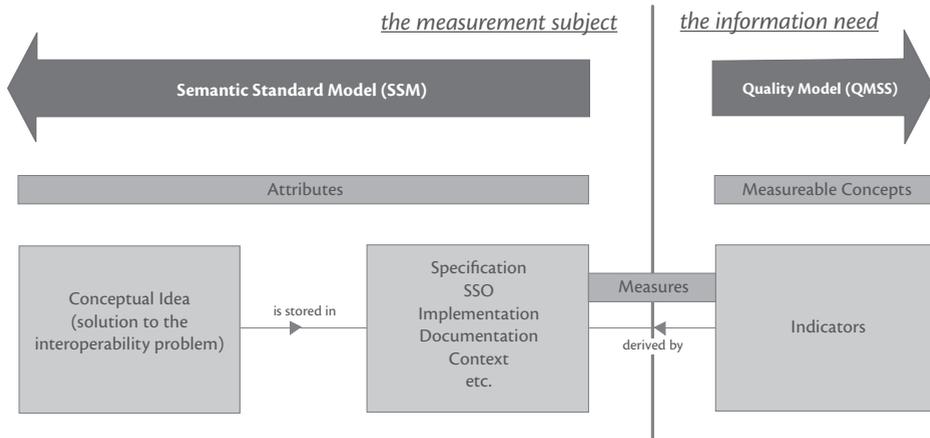


Figure 28 – The link between the measurement subject and the information need

The modeling notation used in the following package diagram is from the UML specification (Fowler, 2004). Three types of dependencies are used. Each of the dependency between the packages is labeled:

- *use*: Indicates that one package element in the client requires another package element from the supplier for its full implementation or operation.
- *specialize*: Relates two package elements, or sets of elements, that represent the same concept at different levels of abstraction, or from different viewpoints.
- *realize*: Indicates that the client model element is an implementation of the supplier model element, and the supplier model element is the specification.

Figure 29 shows the packages from the model and position these packages to each other. The type of dependency defines what type of relation is identified between the packages. Most of the associations are of the <<use>> type.

This UML representation (Figure 29) of Figure 27 shows also some refinements. The three horizontal and vertical layers are recognizable: from a generic instrument, to a customized instrument that is initiated and leads to the result, and the distinction between the quality model constructs (left side), the semantic standard constructs (right side) and the instrumentalization in the middle. On each of the three layers, the instrument is the central element and is built upon the notion of the quality model and the semantic standard.

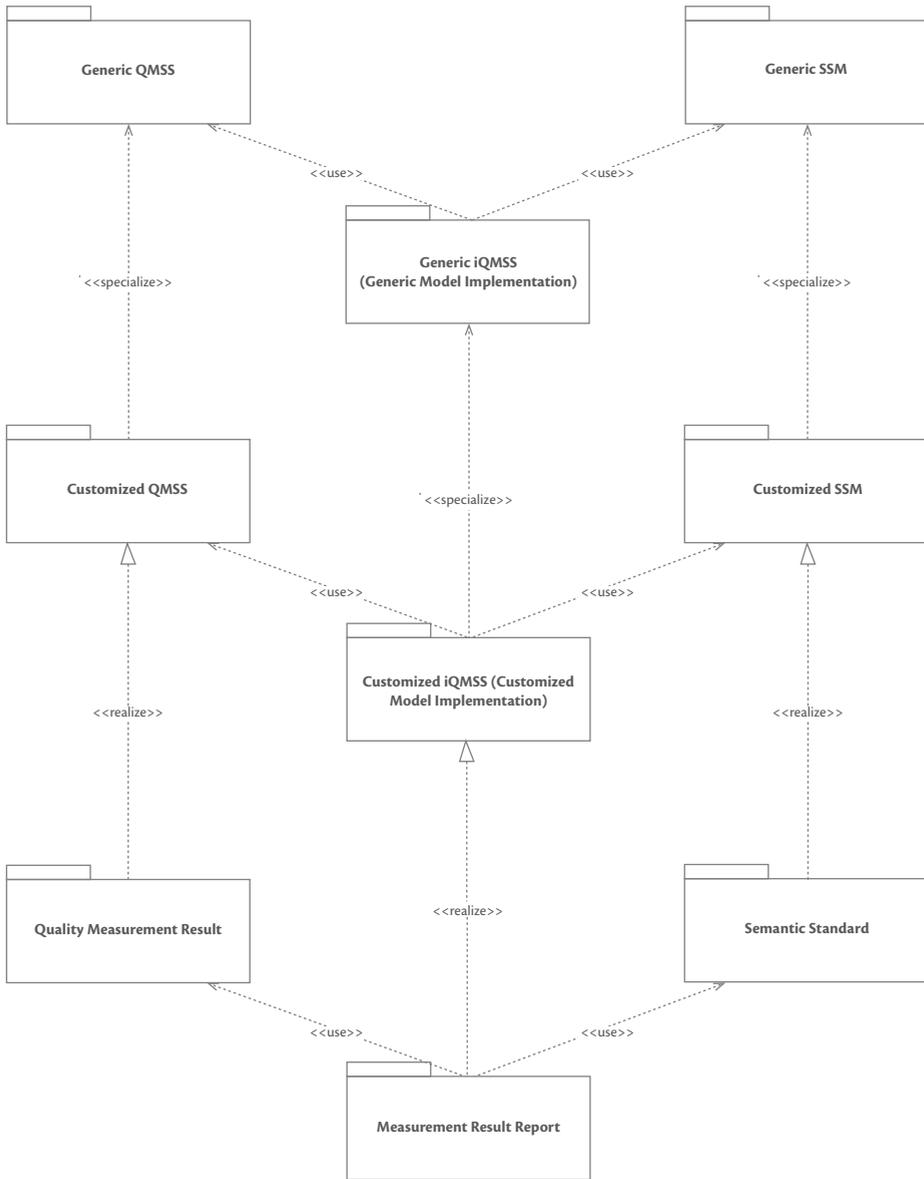


Figure 29 – Package model iQMSS

The previous figures show how the customized iQMSS links the semantic standard to the quality model. Figure 30 is an example of how this could work in practice; the crosses in the grid are the matches between what we want to know and what we are able to measure. A combination of crosses is an indicator for the value of the measurable concept for the semantic standard (based on a formula). The general information need for our research is the quality of semantic standards. However in practice users will be interested

in more detailed information needs like what is the quality of the functionality of the standard. These information needs are an aggregation of measurable concepts. Examples of measurable concepts might be functionality, usability, implementability, etc. Information needs might consist of multiple measurable concepts, and measurable concepts might be decomposed into other measurable concepts.

The (grouped) attributes of the semantic standard are measures for the measurable concept. This relation is described by the indicator which has become the linking pin between the semantic standard and the quality model. This potential relation is shown in the next figure.

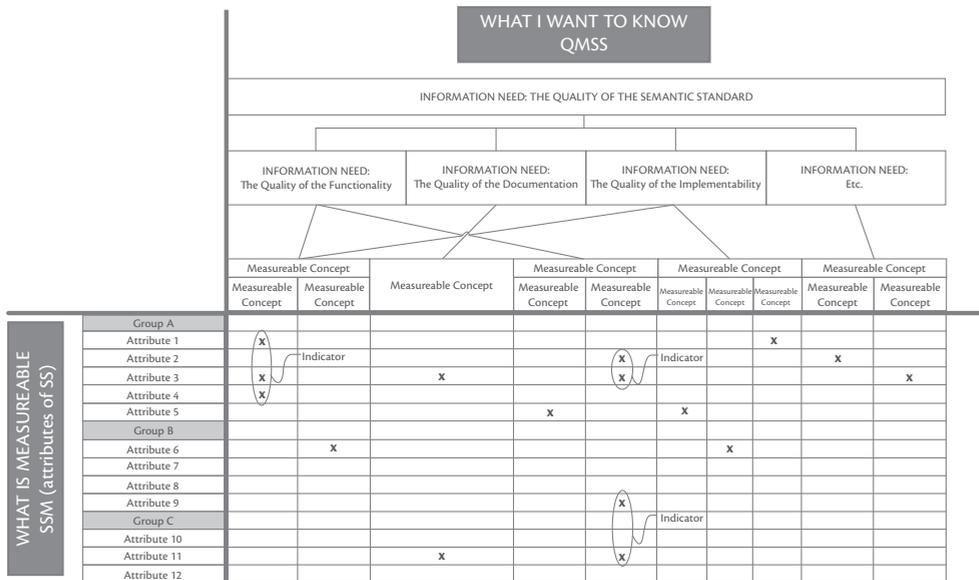


Figure 30 – The potential relation between QMSS and SSM in practice

## 8.2 The iQMSS languages

This second paragraph describes the design choices on the meta (M2) level of the Quality Model and SSM part of the iQMSS.

### 8.2.1 Research approach

The main goal is to develop these M2 level constructs by which they can be used in a later stage for designing the iQMSS. This leads to the research question: *What language is available as QL and SSL?*

To be able to answer this question we start by looking at existing quality literature, mainly from software engineering, and other areas. Starting point is the extensive literature review as described in Chapter 5. Based on the existing knowledge we will be able to distill an appropriate quality language, and will describe it in a way to be useful as quality language within the iQMSS framework.

### 8.2.2 The quality languages

This section will describe the content of, and the search for, the two highest level (M2) artifacts within the model of iQMSS: The quality language (QL) and the semantic standard language (SSL); in the remainder of this document these two are combined and called quality language. Existing knowledge is the main input for the design of this artifact. Many existing languages, including all domain specific languages (DSLs) are available, so starting point should be a state-of-the art review of existing languages/concepts.

## Overview of quality concepts

The intended quality language should be able to fit the semantic standards domain, and preferably based on existing knowledge. As shown in Chapter 4, there exists no literature on quality and semantic standards that introduces a quality language for this domain. However, much is written about quality in other contexts. Many definitions exist, especially from production management literature (like for example (Chase & Aquilano, 1995; Crosby, 1979; Deming, 1986; Garvin, 1984; Ghobadian & Speller, 1994; Juran & Gryna, 1988). Another domain with much literature on quality, but more related to semantic standards, is the domain of software engineering. Based on an extensive and broad literature review on quality, it might be expected that the domain of quality within software engineering will bring most relevant work to be useful within the domain of standardization. Concepts like quality and quality attributes are well described within this domain, including the introduction of Domain Specific Languages (DSLs) and several ISO and IEEE standards that deal with quality of software.

## Frameworks

One of the few existing well known process meta-models is SPEM (Software Process Engineering Meta-Model), a meta-model for the development of software ("Software Process Engineering Meta-Model, version 2.0," 2008). It distinguishes the method content and processes, which come together in the configuration phase, where configuration for the specific instantiation (project) takes place. Another interesting goal driven measurement approach is GQM: Goal Question Metric (Basili, Caldiera, & Rombach). Starting point within this approach are the goals set at the conceptual level. These goals are operationalized by one or more questions, and each question contains one or more measures.

Koziolok (2008) shows how GQM can be implemented as part of a larger framework within organizations. Goals should be taken from the project plan, and also the planning of the collection of data should be part of the project plan. Based on the collected data, interpretation should take place in which measurements are the interpretation of the collected data for the measures. Based on the measurements the answers to the questions can be arranged, just as determination of goal attainment.

Others use SPEM and GQM to build upon for more specific and detailed purposes (García, Ruiz, & Piattini, 2004). SEMIC.EU defines and uses the terms quality factor, criterion and indicator in a hierarchical way (Nentwig et al., 2008).

## ISO & IEEE standards

Both ISO and IEEE standards organizations have been active in relation to standardizing the quality domain in software engineering. Probably the most fundamental and generic is the ISO International Vocabulary of basic and general terms in metrology (ISO/IEC, 1984). Although this document seems quite outdated and focusing on physical measurements, still it might contain some valuable definition regarding measurement results, including indication, corrected result and uncorrected result. This document is superseded by the VIM (International Vocabulary of Metrology) (JCGM/WG2, 2008).

Specifically related to software quality there is an abundance of formal standards. However there is no single standard that embraces the whole software measurement area (García et al., 2009). This results in inconsistencies and discrepancies between different ISO standards which has been recognized by standardization bodies (including ISO/IEC and the IEEE) (García et al., 2009).

Several (Abran et al., 2003; Staron, Meding, & Nilsson, 2009) have studied differences in these standards, some (Bundschuh & Dekkers, 2008; García et al., 2006; García et al., 2009) contain comparisons of definitions of concepts like measures, metric, indicator and value, and the number of ISO standards differ. Standards included are ISO/IEC 25000:2005 version 4.32, ISO/IEC 25000:2005 version 4.33, 20926:2003, 14598-1:1999, 24765, 14102:1995, 15939:2002, 19500-2:2003, 19761:2003, and IEEE standards (Bundschuh & Dekkers, 2008).

Scholars of the university of Castilla La Mancha developed the Software Measurement Ontology (SMO) (García et al., 2006; García et al., 2009; García, Serrano, Cruz-Lemus, Ruiz, & Piattini, 2007), in which many concepts from measurement are introduced. It defines the concepts and compares definitions with other ISO and IEEE standards amongst others. Its greatest value lies in the combination of many existing definitions into one coherent model.

### 8.2.3 *The selection of SMO as language*

Terminology about quality is characterized by huge diversity and even conflicting definitions. The requirements study (Chapter 7) did not reveal specific requirements for the quality language (QL), nor the semantic standard language (SSL). Therefore it was decided to select and build upon the SMO, instead of developing our own language. SMO was chosen based on two main reasons:

1. The richest and most elaborate model found and applicable without known issues.
2. Integration of the most interesting ISO standards on quality.

Further on, our distinction between the measurement object and the quality model fits within SMO as well. In this section we will introduce the SMO and show its applicability to the world of standardization. In SMO terminology: The QMSS (quality model) is defined for semantic standards (entity class), standards (entities) like HR-XML, RosettaNet, XBRL, etc. belongs to this entity class. The quality model contains a set of measurable concepts that are associated to real life information needs. These measurable concepts relate to attributes of semantic standards. Measures, processed by a measurement approach are defined for the attributes. A measure has a scale and is expressed by a unit of measurement. There are several types of measures: base, derived and indicators. The latter are related to the information need. A base measure receives its value by performing a measurement method, while a derived measure is calculated based on a measurement function. An indicator is calculated based on an analysis model that uses decision criteria. A measure can be instantiated at a certain point in time: the measurement, including measurement result, performed on the attribute of the specific semantic standard under subject.

The relation between attributes and measurable concepts is complicated. In practice the value of measurable concepts is determined based on probability tables and the values of attributes. Still it is arbitrary to determine when a certain quality aspect should be labeled as measurable concept or as attribute. Therefore additional guidelines are needed, which we gathered by personal communication with Felix García: "one criterion to decide between attribute and measurable concept is that attributes are not composed of any other attributes. We consider attributes like the measurable physical or abstract property of an entity, a property which can be directly measured. For example, the size (attribute) of a program (entity). Related to that, a measurable concept could be maintainability, which can be divided in other properties."

Concepts like readability seems both attributes and measurable concepts; a design guideline for that would be: If you can define a measure to directly obtain the readability then it is considered to be an attribute. However concepts like readability are often abstract and divided in more specific properties that can be directly measured. The abstract concepts should be labeled as measurable concepts. The table contains a complete overview of the concepts, definitions, some example relations to other definitions, and the application within the domain of semantic standards.



Concept (García et al., 2009)	Superconcept (García et al., 2009)	Definition (García et al., 2006; García et al., 2009)	Relation with other ISO definitions (García et al., 2006)	Specific semantic standards design guidelines / examples
<b>Measure</b>	Concept	The defined measurement approach and the measurement scale. (a measure approach is either a measurement method, a measurement function or an analysis model). (14598-1 – adapted)	ISO 14598-1: [metric]: The defined measurement method and the measurement scale.	-
<b>Base Measure</b>	Measure	A measure of an attribute that does not depend upon any other measure, and whose measurement approach is a measurement method. (14598-1 –adapted)	ISO 15939: Measure defined in terms of an attribute and the method for quantifying it. (Note: a base measure is functionally independent of other measures) ISO 14598-1 [direct measure]: Measure of an attribute that does not depend upon a measure of any other attribute. ISO VIM [base quantity]: one of the quantities that, in a system of quantities, are conventionally accepted as functionally independent of one another.	Number of words.
<b>Derived Measure</b>	Measure	A measure that is derived from other base or derived measures, using a measurement function as measurement approach. (14598-1 – adapted)	ISO 15939: Measure that is defined as a function of two or more values of base measures. ISO 14598-1 [indirect measure]: A measure of an attribute that is derived from measures of one or more other attributes. ISO VIM [derived quantity]: quantity defined, in a system of quantities, as a function of base quantities of that system.	Average number of words per sentence.
<b>Indicator</b>	Measure	A measure that is derived from other measures using an analysis model as measurement approach.	ISO 15939: An estimate or evaluation of specified attributes derived from a model with respect to defined information needs ISO 14598-1: A measure that can be used to estimate or predict another measure.	Gunning Fog index.
<b>Information Need</b>	Concept	Insight necessary to manage objectives, goals, risks, and problems. (15939)	-	Insight in difficulty of implementing standard.
<b>Measurable Concept</b>	Concept	Abstract relationship between attributes of entities and information needs. (15939)	-	Abstract and decomposable concepts. E.g. Readability
<b>Entity</b>	Concept	Object that is to be characterized by measuring its attributes. (15939)	-	HR-XML.
<b>Entity Class</b>	Concept	The collection of all entities that satisfy a given predicate.	-	Semantic standards.
<b>Attribute</b>	Concept	A measurable physical or abstract property of an entity that is shared by all the entities of an entity class. (14598-1 –adapted)	ISO 14598-1: A measurable physical or abstract property of an entity. ISO 15939: Property or characteristic of an entity that can be distinguished quantitatively or qualitatively by human or automated means. ISO VIM [Measurable quantity]: attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively.	Attributes are not composed of any other attributes. We consider attributes like the measurable physical or abstract property of an entity, i.e, a property which can be directly measured.

<b>Quality Model</b>	Concept	The set of measurable concepts and the relationships between them, which provide the basis for specifying quality requirements and evaluating the quality of the entities of a given entity class.	ISO 14598-1: The set of characteristics and the relationships between them which provide the basis for specifying quality requirements and evaluating quality.	The QMSS.
<b>Scale</b>	Concept	A set of values with defined properties. (14598-1)	ISO 15939: Ordered set of values, continuous or discrete, or a set of categories to which the attribute is mapped.	Real between 0-10.
<b>Type of Scale</b>	Concept	The nature of the relationship between values on the scale. (15939)	-	The scales nominal, ordinal, interval and ratio. (Stevens, 1946)
<b>Unit of Measurement</b>	Concept	Particular quantity, defined and adopted by convention, with which other quantities of the same kind are compared in order to express their magnitude relative to that quantity. (15939, VIM)	ISO 14598-3 [unit]: A quantity adopted as a standard of measurement.	Number.
<b>Measurement Method</b>	Measurement approach	Logical sequence of operations, described generically, used in quantifying an attribute with respect to a specified scale. (A measurement method is the measurement approach that defines a base measure)	ISO VIM: Logical sequence of operations, described generically, used in the performance of measurements. ISO 15939: Logical sequence of operations, described generically, used in quantifying an attribute with respect to a specified scale.	Count Words.
<b>Measurement Function</b>	Measurement approach	An algorithm or calculation performed to combine two or more base or derived measures. (A measurement function is the measurement approach that defines a derived measure)	ISO 15939: An algorithm or calculation performed to combine two or more 'base measures'.	Divide words by sentences.
<b>Analysis Model</b>	Measurement approach	Algorithm or calculation combining one or more measures with associated decision criteria. (An analysis model is the measurement approach that defines an indicator)	ISO 15939: Algorithm or calculation combining one or more base and/or derived measures with associated decision criteria.	ReleaseIndicator = Releases divided by Specifications.
<b>Decision Criteria</b>	Concept	Thresholds, targets, or patterns used to determine the need for action or further investigation, or to describe the level of confidence in a given result. (15939)	ISO 14598-1 [Rating Level]: A scale point on an ordinal scale, which is used to categorise a measurement scale.	IF ReleaseIndicator = 1 then = Stable.
<b>Measurement Approach</b>	Concept	Sequence of operations aimed at determining the value of a measurement result. (A measurement approach is either a measurement method, a measurement function or an analysis model)	-	-
<b>Measurement</b>	Concept	A set of operations whose objective is to determine the value of a measurement result, for a given attribute of an entity, using a measurement approach.	ISO 15939, VIM: A set of operations having the object of determining a value of a measure. ISO 14598-1: The use of a metric to assign a value (which can be a number or category) from a scale to an attribute of an entity.	The set of operations (the measurement) carried out a certain point in time. E.g. Measurement of HR-XML on 2011-05-01.
<b>Measurement Result</b>	Concept	The number or category assigned to an attribute of an entity as a result of a measurement. (14598-1 [measure])	ISO 15939 [Measure]: variable to which a value is assigned as the result of a measurement.	Good.

Table 30 – The concepts of SMO applicable for semantic standards

This section has defined a quality language, making extensive reuse of existing work from the software domain, that can be used in the area of semantic standards.

### **8.3 The development environment of iQMSS**

We continue by looking at more detail at the instrumentalization part, based on the following research question: *What are possibilities related to the development environment for the instrument?*

This research question implies to investigate the possibilities of tooling as development environment for setting up the iQMSS. The chosen quality language should be implemented in such a way that it will be efficient and effective to use for the development of the iQMSS. This study will be explorative and will demonstrate possibilities for tooling as part of the development environment.

The design of the iQMSS envisioned tool based support for the standard quality analysis process in the terms of both modeling and visualization support. Prototypes of such tools have been realized based on the Software Measurement Modeling Language (Mora, García, Ruiz, & Piattini, 2008; Mora Rivas, García Rubio, Ruiz González, & Piattini), a language developed in conjunction of the SMO. These prototypes are intended to demonstrate the ideas, models and concepts as generated during the project. These prototypes are in line with the concepts of the Model Driven Architecture, where the distinction between levels of abstraction play an important role.

#### *8.3.1 Development environment for quality models*

Based on this meta-model, tooling can be built that are interoperable for exchanging models. A model *M* that conforms to the meta-model and has been developed in tool *A*, can be interpreted by tool *B* as long as both tools implement the information model as defined in the meta-model. Hence, model instances that conform to the same meta-model are regarded interoperable. In our work we used MagicDraw as modeling tool as this was selected and made available within TNO for information modeling.

#### **Modeling Tool – defining the meta-model**

By modeling SMO as a meta-model in MagicDraw, we are able to define a so-called “profile”. This profile contains all the concepts, relations and restrictions that are defined in the ontology. By applying this profile during modeling activities, we are able to create models that comply to the higher level meta-model.

#### **Create SMML compliant models**

To create an SMML compliant model, one can apply stereotypes as defined in the previously mentioned profile to common UML classes. A part of such a, rather bloated, diagram is depicted within Figure 33.

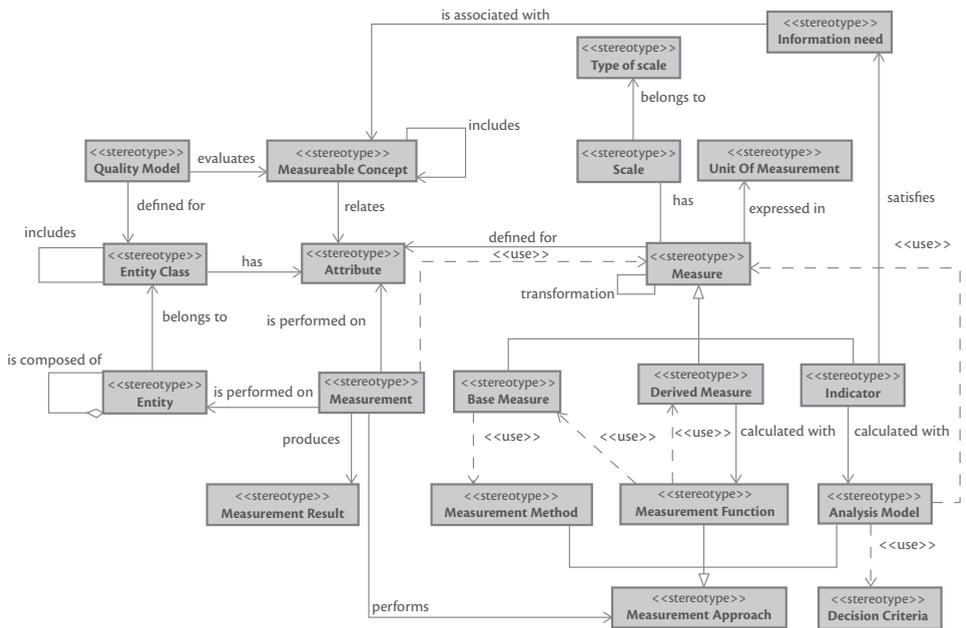


Figure 32 – The SMML profile, modeled as a class diagram

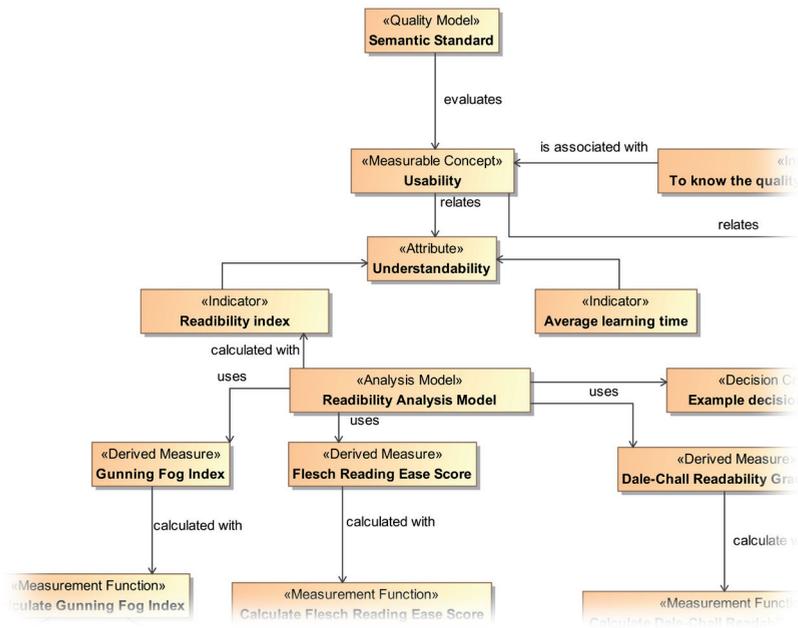


Figure 33 – Example of a class diagram (part) with SMML stereotypes applied

A more user friendly way is to define a custom diagram which is tailored to model SMML compliant models. The user is presented a user interface that contains all the building blocks he needs to define a model that complies with the profile behind the custom diagram, which on his turn contains the SMML meta-model.

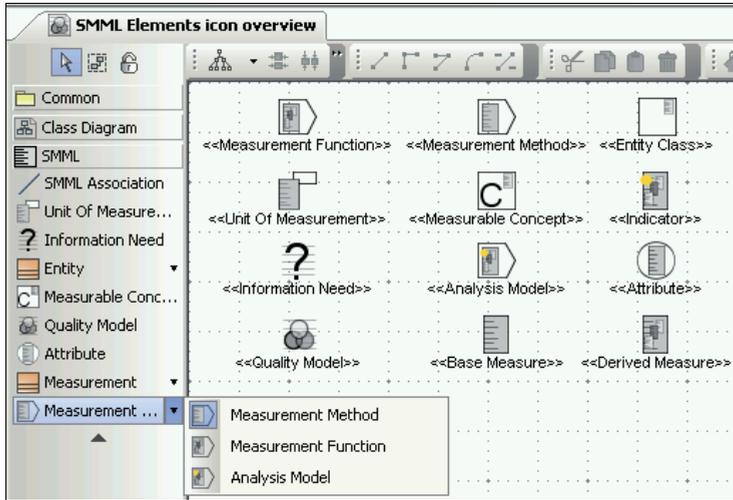


Figure 34 – User-friendly SMML diagram using the SMML icons

To create this customized diagram, the icons for concepts as used in SMML were defined in SVG and applied to the SMML profile as described previously.

The result of using the customized “Software Measurement Modeling Diagram”, is an easy to interpret diagram that is still fully compliant with the meta-model, a small example is depicted within Figure 35.

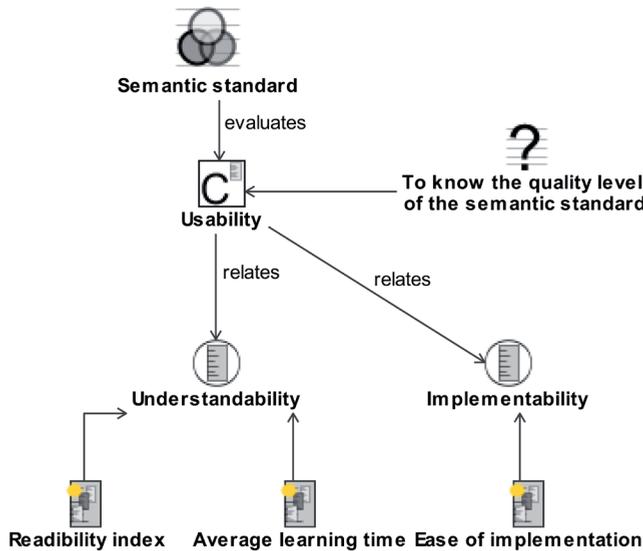


Figure 35 – Example of a SMML diagram which conforms to the SMML meta-model

### 8.3.2 Export and visualization

This tooling is not only essential for the development of comprehensive quality models, it also opens up new possibilities. It can be used to actually determine the quality of a semantic standard. To do so, we export the developed model to a file in the XML Interchangeable (XMI) format. This XMI file is an independent file which contains the serialized version of all the elements and associations as originally contained in the model. In essence, this is just a file-based “view” on the instance of the SMML meta-model. As this view is in XML, we can easily parse it with freely available components to create custom tailored tooling. For this project we have developed a web based visualization tool that can be used by humans to provide information concerning the quality of a semantic standard.

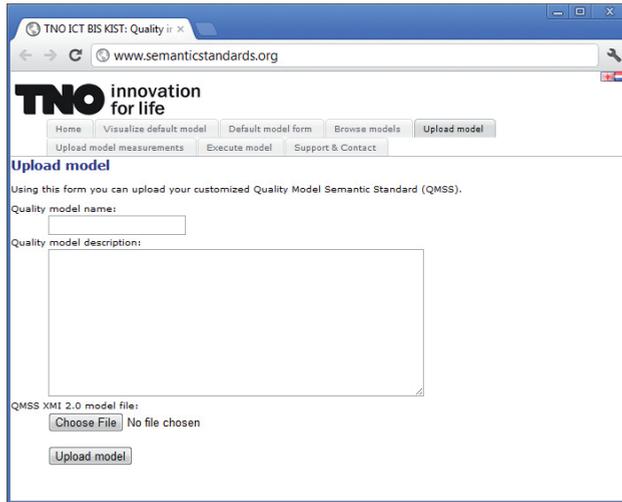


Figure 36 – Upload form where users can upload “their own” quality model

After uploading, parsing and interpreting the specific SMML elements contained within an XML file, a hierarchic visualization is presented to the user. This visualization can be navigated interactively using a common Internet Browser. As we have implemented the SMML meta-model in the visualization tool, it can determine what to do when a certain type of element (for instance an attribute) is clicked upon.

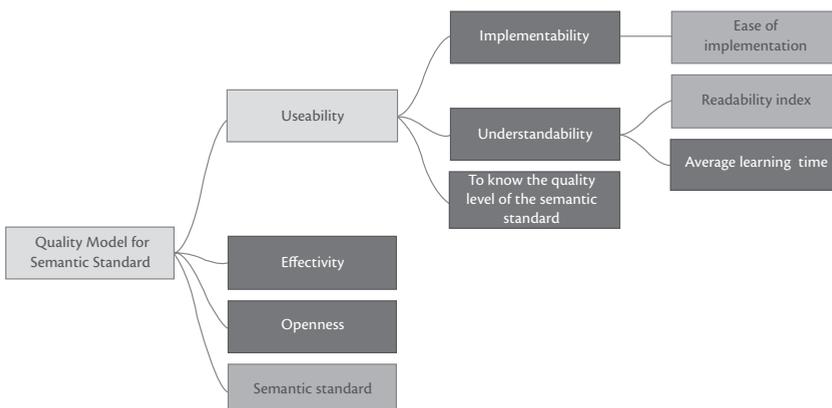


Figure 37 – Screenshot visualization of quality model

As the purpose is to develop an instrument to measure the quality of a semantic standard, we intend to use the aforementioned web based visualization to have users answer questions concerning a specific standard. Based on the provided information of a single or multiple stakeholders, the quality of the standard can be determined to a certain extent.

### 8.4 Conclusions

Our first research question was as follows: *What constitutes a measuring instrument for the Quality Model of Semantic Standards (iQMSS)?*

Based on the design constraints, the instrument should contain several artifacts, including a quality model, a model of a semantic standard, and an implementation as instrument. To accommodate the differences between semantic standards, a distinction is necessary between generic and customized artifacts. To be more specific, Table 31 summarizes how the requirements are met in the iQMSS reference model.

Nr.	Requirement	Accommodated
1	Accommodate the differences between semantic standards.	By initiating the model of a semantic standard knowledge is gathered about the measurement object. Only attributes of the specific semantic standard that have been included in the initiation of the model might be included in the measurement.
2	Focusing on the quality needs/question of a specific SSO.	This implies that dynamic selection of measures and partly measures need to be accommodated. In our model this requirement is satisfied by the introduction of a generic quality model that will be customized for every specific need from a client.
3	Implementation of the instrument in easy-to-use tools.	This is implemented by the "implementation" level in our model. This will focus on the development of appropriate tools that support the quality measurement.
4	High-quality outcome based on grounded quality model.	By having a generic QMSS based on existing theory.
5	Output report as fundament for improvement project.	Our model includes the measurement result report. However the content of the report will determine if it is useful for an improvement project.

Table 31 – Summary of requirements and how these have been accommodated within the design

	Specification Quality Model Semantic Standard (QMSS)	Specification Semantic Standard Model (SSM)	Implementation
M2	1. Quality Language (QL)	I. SS Language (SSL)	A. Development Environment
M1	2. Generic QMSS	II. Generic SSM	B. GMI Generic Model Implementation
M1	3. Customized QMSS	III. Customized SSM	C. CMI Customized Model Implementation
M0	4. Measurement Result	IV. Semantic Standard	D. Measurement Result Report

Figure 38 – iQMSS parts discussed in this chapter

We continued by studying the M2 layer of artifacts, including the implementation artifacts within the iQMSS. The research questions and corresponding answers are:

**What language is available as quality language for semantic standards?**

Different frameworks and languages from different domains exist. Especially the software engineering domain shows an abundance of quality related studies and definitions of quality concepts. The Software Measurement Ontology is an in-depth and consistent quality approach for software engineering, and has been selected within this study for application in the domain of standardization. The measurable concepts (Figure 39) stem from the quality model and are the concepts that we want to know (the information needs). Attributes of the semantic standard, the object of measurement, that are related to the measurable concepts can be measured by combining the use of indicators and base and derived measures.

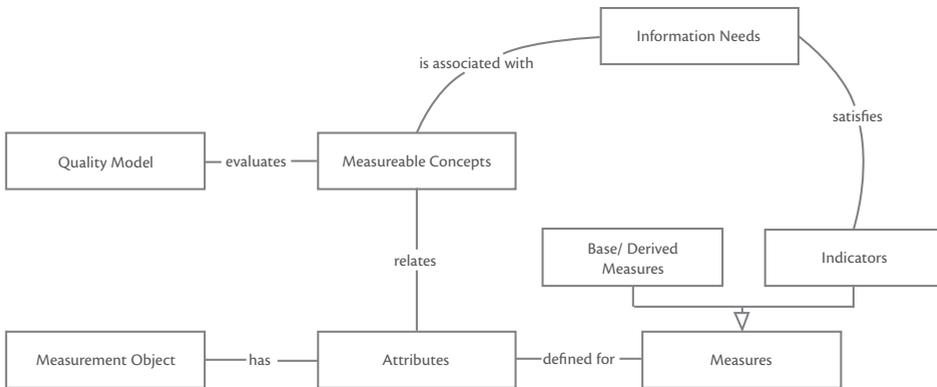


Figure 39 – Simplified version of terminology used in SMO (Adapted from (García et al., 2009))

### **What are possibilities related to the development environment for the instrument?**

By having a formal ontology incorporated in a tool like MagicDraw including graphical representation it becomes fairly easy to define and customize your own quality model, specific for a semantic standard. Within the current tooling we can create the models and visualize them accordingly in a web browser.

To fulfill the requirements like easy to use, tooling is essential. As the current visualization tool is designed in a modular design, it is relatively easy to develop plug-ins as the knowledge about the quality of standards evolves. Depending on what element a user has clicked, the tool 'knows' what to do, and can support the actual measurement. For example, when the 'number of words' of a document must be determined, a Word document can be uploaded and analyzed.

A measuring instrument without proper instrumentalization will be difficult to use in practice. Having tooling will definitely contribute in the usage of the instrument. However when there is one strict instrument/tool it is often difficult to handle differences in standards. It is probably impossible to have one tool to be used by all standards. Therefore there is a need for customization and flexible tooling. Only in some cases a dedicated tool for a specific standard (e.g. XBRL or HL7) might be worthwhile, especially when intended to be repeatedly used in many instances. This is a trade-off between a dedicated tool versus a meta tool. The choice depends on the business case, which for most standards will be in favor of the meta tool.

Optimal use of tooling set demands to the measures defined. Measures can be of different types, like absolute measures, relative measures, coefficients and index figures (Bundschuh & Dekkers, 2008), and have to be defined with scales (nominal, ordinal, interval and ratio (Stevens, 1946)).

This chapter has set the fundament of the iQMSS; the following chapters are aimed on developing parts of the iQMSS.

## The Semantic Standards Model

A comprehensive understanding about the characteristics of semantic standards is missing, which is needed for quality measurement. The chapter addresses this gap by developing a characteristics model for semantic standards (SSM). This SSM is part of the iQMSS as shown in Figure 40. Two case studies are used to check the applicability of the model in a “real-life” context.

	Specification Quality Model Semantic Standard (QMSS)	Specification Semantic Standard Model (SSM)	Implementation
M2	1. Quality Language (QL)	I. SS Language (SSL)	A. Development Environment
M1	<b>2. Generic QMSS</b>	<b>II. Generic SSM</b>	<b>B. GMI Generic Model Implementation</b>
M1	3. Customized QMSS	III. Customized SSM	C. CMI Customized Model Implementation
M0	4. Measurement Result	IV. Semantic Standard	D. Measurement Result Report

Figure 40 – The topic of this chapter (SSM) within the iQMSS

### 9.1 Research approach and concepts

This chapter introduces the notion of semantic standards to take up on the abovementioned need. It addresses the research question as to how semantic standards can be characterized, in particular in the context of other IS standards, and what characteristics are needed to describe, analyze, compare and design them. To adequately respond to this question, the previous chapter proposed a characteristics model for semantic standards, the SSM. The model is based on a deductive analysis of the state of the art in theory and practice and on multiple evaluation measures in the field. Among these evaluation measures are two case studies, one by the Dutch Ministry of Economic Affairs and one by Siemens Corporate Technology.

This chapter is based on:

(Folmer, Oude Luttighuis, & Van Hillegerberg, 2011b): A Model for Semantic standards. In: V. Folmin, K. Jakobs (Eds.), Proceedings of the 16<sup>th</sup> EURAS Annual Standardisation Conference “Standards for Development”, pp. 81-94, ISBN: 3861306298, Verlag Mainz.

(Otto, Folmer, & Ebner, 2011): A Characteristics Framework for Semantic Information Systems Standards. *Information Systems and E-Business Management* (Online First - 24 November 2011).

The remainder of this chapter begins with a description of the research process. The model design is presented before being applied in a case study setting (Yin, 2009). After that, the model is evaluated against its design objectives. The chapter discusses the research results with regard to their theoretical and practical contribution and concludes with a brief summary.

### 9.1.1 Research process

Overall, existing literature comprises valuable work regarding semantic standards (Chapter 5). A comprehensive model, however, which supports the analysis and evaluation of semantic standards is missing as of today.

The work presented in this chapter followed a two-step research process (see Figure 41) to fulfill this gap. Step 1 included the design of the characteristics model for semantic standards and consisted of three activities. Activity D1 used a reverse engineering (Chikofsky & Cross, 1990) approach to analyze existing semantic standards and to increase the understanding regarding their characteristics and constituents. The results were continuously combined with the results of a parallel literature review (Activity D2) regarding approaches for classification of semantic standards (see Appendix H). Preliminary versions of the model were reflected and discussed in multiple expert interviews (Activity D3). In general, expert interviews allow for explication of the knowledge of experts in the field, without taking the effort of conducting a comprehensive quantitative analysis (surveys, for example) (Meuser & Nagel, 1994). Two expert interview streams were used:

1. Four sessions comprising six experts from the Dutch research organization TNO, who have been involved in the development of semantic standards.
2. One expert session comprising the members of the BOMOS working group of the Dutch government, which created the Development and Management model of Open Standards (Folmer & Punter, 2011). The session included ten standards developers from different domains and was held in November 2010.

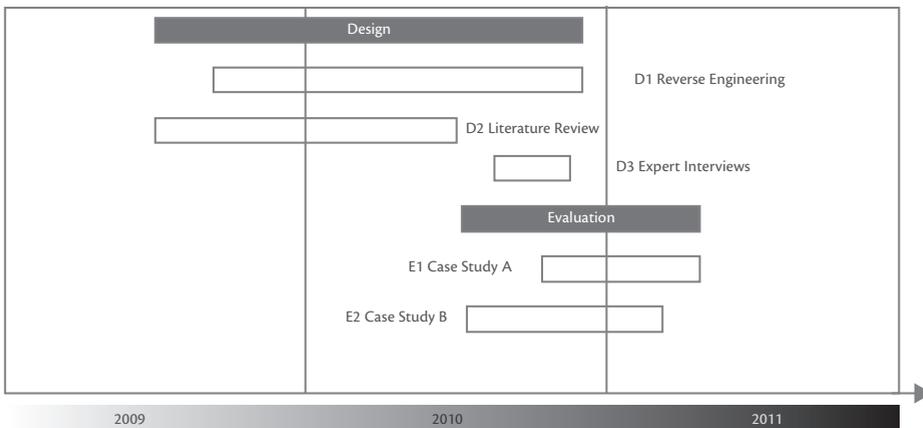


Figure 41 – Research process

The result of Step 1 is the SSM. Describing reality, the model forms an analytical theory according to the topology of theories in IS proposed by Gregor (2006).

Step 2 of the research process aimed at validating the model, i.e. in particular its applicability and usefulness in a “real-life” context. Case studies, in general, are well suited for this purpose, since they allow studying a contemporary phenomenon within its context (Benbasat & Zmud, 1999; Eisenhardt, 1989; Yin, 2009). Case A involved the Dutch Ministry of Economic Affairs and was conducted between December 2010 and February 2011. From a research perspective, Case A pursued two goals. First, it aimed at providing practitioner feedback during the model design phase. Second, it was used to check the practical applicability of the model. Case B was conducted in a collaborative research project with Siemens Corporate Technology and was conducted between August 2010 and January 2011. It aimed at demonstrating the applicability of the model in a private business environment. Both cases studies were of participatory nature (Baskerville, 1997). The researchers played an active role by moderating and facilitating the progress of the projects. However, they did not influence the decision of the practitioners with regard to analysis and evaluation of semantic standards.

## **9.2 SSM design**

### *9.2.1 Design objectives*

Gregor (2006) has stipulated a number of requirements an analytical theory needs to meet in order to contribute to the scientific body of knowledge. Among these requirements are usefulness, appropriateness, clarity, and completeness. The usefulness of the theory materializes in the use cases in which it is supposed to offer support. Regarding the characteristics model for semantic standards, three major use cases were identified:

- Analysis and comparison of semantic standards: Both researchers and practitioners as well as standardization bodies must be able to analyze and compare existing semantic standards in order to determine overlaps and “white spots”.
- Evaluation and selection of semantic standards: Organizations which are in the process of evaluating and selecting semantic standards must be able to base their assessment on a complete, comprehensive and unbiased foundation.
- Influence and further development of semantic standards: Assessments of the diffusion and acceptance of certain semantic standards as well as their quality must be based on a solid terminological foundation.

The appropriateness of a model is what Becker et al. (1995) in their “Guidelines for Orderly Modeling” (GOM) call “systematic structure”. One example to support a systematic structure would be to introduce different views. Apart from appropriateness, the requirement of clarity can also be found in the GOM. Becker et al. (1995) have stipulated that the information model must be understandable by its addressees. Finally, the completeness of a model ensures—according to Gregor (2006)—that no important concepts are omitted. The GOM see completeness constrained by the economic viability of the application of a model and by the focus on relevant concepts only. Economic viability also requires that the model must be adaptable with regard to the use case, because each case might require a different level of detail and only a subset of the concepts covered by the model.

### *9.2.2 SSM overview*

The characteristics model for semantic standards consists of three levels, of which the first two levels comprise categories and sub-categories, while the third level represents concepts. The concepts represent characteristics of semantic standards and can be assigned with values. Therefore, description, analysis, and comparison by means of the model are carried out on the basis of values on the conceptual level. All concepts are assigned to exactly one sub-category, and every sub-category is assigned at least to one concept. Both the introduction of a hierarchical structure and the grouping of concepts into sub-categories and of sub-categories into categories aim at ensuring the model’s comprehensibility (see GOM in the previous section).

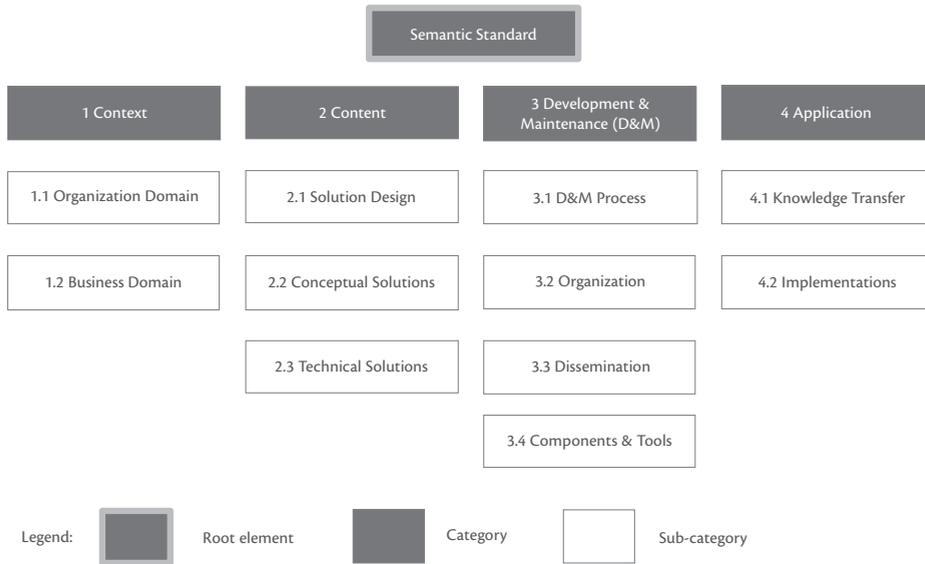


Figure 42 – SSM

Figure 42 shows the categories and sub-categories of the characteristics model for semantic standards. In total, 37 concepts are assigned to 10 sub-categories, which themselves are assigned to the four level-one categories. The model proposes the use of metadata, such as “Name”, “URL” etc. This kind of information is supposed to be attributed directly to the root element, i.e. the semantic standard. It is not included in the model itself, because it serves identification purposes rather than analysis and evaluation of different semantic standards.

The first level of the model includes four categories, namely “Context”, “Content”, “Development & Management”, and “Application” (Table 32).

ID	Category	Description
1	Context	Concepts related to the standard's environment
2	Content	Concepts related to the solution offered by the standard
3	Development & Maintenance	Concepts related to the standardization activities and their organization
4	Application	Concepts related to implementation and use of the standard

Table 32 – First-Level categories

The context (category 1) is the environment in which different stakeholders are facing a certain business problem for which a standard solution is required. The actual solution, i.e. the content of the standard, forms the second category (category 2) on the model's first level. This is what many researchers and practitioners refer to a standard in the narrower sense of an information model (see above). Moreover, each standard must be developed and maintained (category 3). All concepts related to the use of the semantic standard are included in category 4.

### 9.2.3 SSM in Detail

#### Sub-categories and concepts of the category “Context”

Table 33 shows the sub-categories and concepts related to the category “Context”. The column on the right indicates how the sub-categories and concepts are supported by literature. The codes refer to the categories and concepts in Appendix H. This approach recurs in all following three tables.

While the community (1.1.1 to 1.1.3) is principally confirmed as a concept in literature on semantic standardization (Lampathaki et al., 2009; Pawlowski & Kozlov, 2010), no further specification of the term is given. The expert interview sessions revealed the need for further distinction between target and adopting community. For example, the classification standard eCl@ss was initially designed to help match the needs of purchasing departments in large chemical companies, but has been adopted by many other sectors and functional departments eventually. The concept “active community” was included as a result of the literature analysis on IS standardization in general. Nickerson and zur Muehlen (2006), for example, stress the importance of the active community during the development of a standard.

ID	Sub-categories and concepts	Description	Literature Support
1.1	Organizational domain	Stakeholders related to the standard.	B.1., B.3., C.1., D.6.
1.1.1	Target audience	The addressees the standard is intended for.	
1.1.2	Adopting audience	The community using the standard.	
1.1.3	Active community	Stakeholders actively participating in the design, maintenance, dissemination etc. of the standard.	
1.2	Business domain	The business purpose for which the standard is designed.	B.2., C.2.a., D.2.
1.2.1	Business goals	The real-life problem the standard aims at overcoming and derived business goals.	
1.2.2	Application domain	Description of targeted domain of use, including rules and constraints like laws and regulations.	
1.2.3	Cost & Benefits	Benefits and costs related to achieving the business goals through use of the standard.	

Table 33 – Sub-categories and Concepts related to “Context”

Including the “business domain” as a sub-category in the model is backed by literature on semantic standardization (Lampathaki et al., 2009; Mykkanen & Tuomainen, 2008; Pawlowski & Kozlov, 2010). But again the discussion with subject matter experts resulted in the demand for a more detailed elaboration of this category.

### Sub-categories and concepts of the category “Content”

Table 34 shows the sub-categories and concepts of the category “Content”.

ID	Sub-categories and Concepts	Description	Literature Support
2.1	Solution design	Approaches and methods underlying the design of the standard.	A.1., A.4.,
2.1.1	Design paradigm	A high-level paradigm underlying the standard design.	A.8., B.4.,
2.1.2	Methods & Languages	Methods and languages used in the design of the standard.	B.5., B.6.,
2.1.3	Architecture	Architectural design choices for the standard, including functional and technical architecture, and relationships with other standards.	B.7., B.8., C.2.b., C.4.a., D.3.
2.2	Conceptual solutions	The design of the solution in concepts like descriptions and models.	A.2., A.3.
2.2.1	Domain model (requirements)	A description of the domain environment of the standard.	
2.2.2	Constraints	Constraints described as a solution, expressed like business rules, related to the standard. Such rules can express data dependencies based on the process status.	
2.2.3	Process	The design of the flow of activities encapsulated within the standard. This might include process diagrams, actors involved, timing, error handling, cancellation process, etc.	
2.2.4	Data, information	The design of data and information objects encapsulated within the standard. This might include messages/documents, ontologies, code lists, taxonomies, data dictionaries, sharable data components, etc.	
2.3	Technical solutions	The design of the solution in technical artifacts.	A.5.
2.3.1	Format	The format of the technical solutions, in which the conceptual solutions are represented.	
2.3.2	Medium (transport)	Solutions related to technical communication aspects.	

Table 34 – Sub-categories and concepts related to “Content”

The category “Content” consists of three sub-categories, namely “Solution design”, “Conceptual solutions”, and “Technical solutions”. While “Solution design” addresses underlying design principles and foundations (e.g. XML for many electronic business standards), “Conceptual solutions” and “Technical solutions” represent two different layers of abstraction. While the “Solution design” is relatively well supported by existing literature, hardly any contribution can be found in the scientific body of knowledge regarding the “Conceptual Solution” and the “Technical Solution”.

### Sub-categories and concepts of the category “Development and Management”

Many studies on standardization are focused on the development of standards (Chapter 4). This is closely related to the management of standards, involving standards development organizations. Table 35 contains the sub-categories and concepts of the category “Development and Maintenance”.

ID	Sub-categories and Concepts	Description	Literature Support
3.1	Development & Maintenance process	Activities related to the development & maintenance of the solutions the standard offers.	C.3., D.4., D.6., D.11.
3.1.1	Initiation	The initiation process of exploring new maintenance requests or requirements related to the standard.	
3.1.2	Design & Formalization	The design process of creating solutions for requirements and maintenance requests. And the transformation of the design of the solution in the requested formats, both conceptual and technical.	
3.1.3	Review & Testing	The review of the formalized solution by the stakeholders. If possible, the solution may be tested in practice.	
3.2	Organization	Organization of the development and maintenance of the standard.	A.7., B.1.,
3.2.1	Quality management	Quality assurance and benchmarking of the standard.	C.3.a.,
3.2.2	Rights policy	The description of the rights policy chosen for the standard.	C.4.b., C.5.,
3.2.3	Governance	Governance model for the organization of the standard. Including decision-making, release policy, and complaints handling.	C.6., D.4., D.6.
3.2.4	Finance model	The model chosen for financing the costs of the development and management processes.	
3.2.5	Vision & Strategy	The long-term vision for the standard, and its strategy for fulfilling the vision.	
3.2.6	Operational model	The operational approach for the development and maintenance of the standard, including meetings policy (location, frequency, openness of meetings), versioning policy, backwards compatibility policy, documentation policy etc.	
3.3	Dissemination	Activities related to disseminating the standard.	A.7., C.7.,
3.3.1	Communication & Adoption strategy	The overall strategy regarding communication with different stakeholders, using different communication channels. Including the strategy related to publication of the documents in which the standard is described, and the adoption strategy for achieving the desired adoption rate, including addressing the status of the standard by both the own organization and external organizations.	D.7.
3.3.2	Compliance strategy	The strategy to test and assure compliance of implementations to the standard by a certification program, for example.	
3.4	Components & Tools	Artifacts useful during implementation.	D.8.
3.4.1	Components & Tools	Components and tools used for implementation of the standard (a validation service, or open source component, for example).	

Table 35 – Sub-categories and concepts related to “Development and Maintenance”

In spite of the fact that many studies exist dealing with the concepts of the Category “Development and Maintenance”, the experts from the practitioners’ community considered the level of detail to be insufficient. Lampathaki et al. (2009), for example, identify “openness” as a relevant concept, but do not elaborate it further. According to the expert feedback, more detailed information is needed in practice about the development process, the governance structures of the standardization body, and about quality management aspects. This demand is supported by literature on IS standardization in general. Detailed “Development & Maintenance” concepts are necessary as they represent what Grindley (1995) refers to as “standards reinforcement mechanisms” (p. 27). These mechanisms are decisive for widespread adoption of a standard. And a detailed representation of “Organization” concepts is required to be able to analyze behavioral aspects such as described by Backhouse et al. (2006) and Nickerson and zur Muehlen (2006), for example.

Moreover, it turned out that openness is too broad a concept to be included as one characteristic. In fact, openness is multidimensional referring to open meetings, open intellectual property rights and open access to documents, for example (Krechmer, 2009).

The sub-category “Components & Tools” can hardly be found at all in existing literature. Only Lampathaki et al. (2009) briefly address this point under “Ease of use and implementation” (D.8.).

### Sub-categories and concepts of the category “Application”

Finally, the application of the standard is an important category with regard to the use cases described above as it influences potential further adoption, for example. Table 36 shows the sub-categories and concepts of the category “Application”.

ID	Sub-categories and Concepts	Description	Literature Support
4.1	Knowledge transfer	Concepts related to the dissemination of knowledge about the standard.	A.8., B.10.,
4.1.1	Helpdesk	Helpdesk availability to answer (implementation) questions about the standard.	B.12., C.5., C.6.
4.1.2	Events & Training	Availability of events and a training program to share knowledge about the standard.	
4.1.3	Consultants	Availability of consultants/implementers for the standard.	
4.1.4	Pilots (support)	Documentations about pilot implementations and availability of support for pilots.	
4.1.5	Representation forms	Knowledge about the standard is available in all kinds of representation forms, like specifications, implementation guidelines, examples, code lists, websites, flyers etc.	
4.2	Implementations	Concepts supporting the implementation of the standard.	B.9.
4.2.1	Implementations	Information about implementations, including reference or certified implementations. Reference implementation can be used as a template for further implementation while certified means approval (by certificate) by an appropriate authority such as the SSO.	

Table 36 – Sub-categories and concepts related to “Application”

Concepts of the category “Application” are addressed only to a limited extent in literature. Some contributions, though, supporting the sub-category “Knowledge Transfer” can be found. They remain, however, unspecific. And regarding the category “Implementations”, only Pawlowski and Kozlov (2010) address the point when referring to “Usage and validation” (B.9.). The general relevance of reference implementations in IS standardization is supported by the prominent Amaya case, for example. Amaya is a reference implementations for Internet standards which is maintained by W3C (Quint, 2010).

## 9.3 SSM application Case A: Dutch Ministry of Economic Affairs

### 9.3.1 Context

The Dutch government is leading in Europe when it comes to defining public strategies for adopting and promoting open standards. The policy named “Netherlands Open in Connection” (NOiV, 2007) is characterized by a stringent definition of openness and a “Comply or Explain” regime for the public sector. The latter implies that open standards that have been selected by the standardization board after a stringent procedure must be used within the public sector. Several standards that are included in the “Comply or Explain” list are the result of public-private partnership efforts (Lammers, Folmer, & Ehrenhard, 2010).

The Dutch Ministry of Economic Affairs wanted to know what other standards are available that have public-private partnership properties and that are aimed at solving economically and socially important goals. Standards identified should be allowed for future addition to the “Comply or Explain” list.

Three main criteria for the selection of semantic standards were defined:

1. Maturity of a standard: A standard may be adopted if there has been sufficient practical experience in using it.
2. Potential of a standard: A standard may be adopted if it has the potential to substantially contribute to the achievement of certain economic and/or social goals.
3. Maintenance and development of a standard: A standard may be adopted if the processes for maintaining and developing it are organized, open and structured.

The Ministry of Economic Affairs added an additional criterion for their specific purposes, namely (4.) public-private partnership. The criterion implies that stakeholders from both the public and the private domain should have an interest in the standard.

### *9.3.2 Standards selection process*

The characteristics model for semantic standards was used to support both the definition of the selection criteria and the process of selecting appropriate semantic standards to be potentially included in the “Comply or Explain” list.

The definition of selection criteria consisted of five steps. Step 1 aimed at the identification of the needs with regard to the standards’ selection (see above). Step 2 mainly included the formulation of questions the stakeholders wanted to have answered with regard to the standard. In Step 3 these questions were mapped to concepts in the model. Step 4 looked for potential gaps, before Step 5 aimed at adjusting the questions so that they still would reflect the stakeholders’ needs and in parallel match the concepts of the model. Table 37 shows the mapping of model concepts to the questions identified in Step 2.

Several of the questions are related to metadata of semantic standards, namely name, website, abbreviation, or start date (attributed on “root” level). The level of adoption can be determined by looking at the number of implementations in relation to the targeted audience. As this is, of course, not an easy task to do, in Case A multiple sources of evidence were used (for example, not only the big players but also small and medium-sized companies).

The standard selection process started with the creation of a long list of semantic standards based on desktop research and input from subject matter experts. Based on the assessment of experts ten standards were selected for further analysis. The analysis of the shortlisted standards was conducted using the characteristics model for semantic standards.

One example of the application of the model is included in Appendix I in order to give an idea as to what information was gathered from each standard by using these questions.

Item	Question addressed	Needs addressed	Model Concepts
1	Name of the standard	General	Semantic standard (root element)
2	Functional and organizational domain	General	Target audience (1.1.1), Application domain (1.2.2)
3	Website	General	Semantic standard (root element)
4	Usage on national, European, international level	General, Maturity	Adopting audience (1.1.2), Implementations (4.2)
5	Start date	Maturity	Semantic standard (root element)
6	Important stakeholders	Maturity, Public-private	Adopting audience (1.1.2)
7	Number of participants	Maturity	Active community (1.1.3)
8	Level of adoption	Maturity, Potential	Implementations (4.2), Target audience (1.1.1)
9	Contribution to economic and social goals	Potential	Business goals (1.2.1), Cost & Benefits (1.2.3)
10	Contribution to reducing the administrative burden or to improved inspection	Potential	Business goals (1.2.1), Cost & Benefits (1.2.3)
11	Profit/non-profit orientation of organization	Organization	Finance model (3.2.4)
12	Finance model	Organization	Finance model (3.2.4)
13	Participation model	Organization	Operational model (3.2.6)
14	Decision model / Governance	Organization	Governance (3.2.3)
15	Availability	Organization	Rights policy (3.2.2)

Table 37 – Application of the model in Case A

### 9.3.3 SSM applicability

Table 37 shows which parts of the model were used. First, it shows that some metadata aspects are covered. Second, the biggest part deals with the context (the Organizational and the Problem domain) and with the implementation of the standard. This part is needed for gathering information regarding the maturity and the potential of the standard. The third part that is covered comprises aspects from the development and maintenance organization. The table also shows that in this application case no question relates to the Content category of the standards.

Apart from that, the model helped identify gaps in the set of selection criteria:

- Content (category 2): By not looking at the content it remains uncertain what the quality of the solution will be in relation to the problem.
- Knowledge Transfer (sub-category 4.1): Knowledge transfer will have an impact on the ease and speed of adoption. By not looking at this aspect essential input for assessing the maturity and potential of the standard might be disregarded.

In this case the Ministry of Economic Affairs was very much aware of the gaps. For example, neglecting content related concepts was intended as the project scope did not include assessments on the quality of a semantic standard. Overall, the applicability of the model for Case A was considered high by all project stakeholders.

## 9.4 SSM application Case B: Siemens Corporate Technology

### 9.4.1 Context

Siemens Corporate Technology is a corporate organization supporting the Siemens divisions in all three sectors with expert knowledge on recent research and technology developments in areas of importance for the company.

With regard to IS standardization, a variety of specifications and standards exists across the group. Business and data objects, for example, typically have been defined on an individual basis for each organizational unit, business process, and application system. In order to increase transparency on business processes and reduce costs for data integration and transformation, Siemens Corporate Technology aimed at introducing a process oriented Enterprise Data Architecture with a common definition of business objects both from a functional and an ICT view. The Enterprise Data Architecture was supposed to make use of existing business data standards as much as possible. A project was started aiming at identifying business data standards available on the “market” and evaluating them for the use at Siemens. Recommended standards were supposed to be used by Siemens business units in order to provide a common understanding of business objects.

### 9.4.2 Standards evaluation process

In the beginning of the project, Siemens Corporate Technology identified five criteria against which existing standards were to be evaluated:

1. Business focus: The standard should have a focus on the modeling of business objects from the business view.
2. Popularity / diffusion / reach: The standard should be widespread in the user community.
3. Topicality: The standard should be still in development or maintenance and the latest version should be up-to-date (i.e. the latest update should not be older than three years).
4. Industry scope of Siemens: The standard should focus on one of the domains Siemens engages in (industry, energy, and healthcare).
5. Definition of a data model: The standard should support the modeling of business objects and their relationships.

A sixth criterion was support of the standard by software tools available in the market. It turned out, however, that this criterion could not be assessed for all standards.

As a result of the overall goal to develop an Enterprise Data Architecture, Siemens Corporate Technology focused only on those semantic standards which included specifications of business objects and/or catalogue data. Other standards with a focus only on messages, item identification or classification, and business processes were disregarded.

As a first step in the evaluation process, an inventory of standards comprising about two hundred entries was created based on an extensive Internet research. Included in the search were the following sources:

- Standardization organizations (e.g. ISO, UN/CEFACT);
- Research funding agencies (e.g. European Commission);
- Inter-trade organizations;
- Software vendors (e.g. IBM, SAP);
- Key market players (e.g. Chrysler, Toyota from the automotive industry).

In a second step, the inventory was narrowed down to a shortlist of about 40 standards using the criteria described above, except for the industry focus. A third step reduced the number of standards to twelve, which were then transferred to the so-called evaluation list. For this evaluation list the industry focus and the relevance for Siemens Corporate Technology were regarded as criteria.

Item	Evaluation Criteria	Model Concepts
1	Version	Semantic standard (root element)
2	Description	Semantic standard (root element)
3	Registered standard	Communication & Adoption strategy (3.3.1)
4	Standardization organization	All concepts in sub-category "Organization" (3.2)
5	Industry	Application domain (1.2.2)
6	Scope	Business goals (1.2.1), Architecture (2.1.3)
7	Developer	Active community (1.1.3)
8	Origin, contributions	Active community (1.1.3)
9	Link	Semantic standard (root element)
10	Content	All concepts in category "Content" (2)
11	Structural cardinality	Architecture (2.1.3)
12	Semantic cardinality	Architecture (2.1.3)
13	Representation	Concepts in sub-categories "Conceptual solutions" (2.2) and "Technical Solutions" (2.3)
14	Predefined content	Data, information (2.2.4)
15	Extensibility	Architecture (2.1.3)
16	Integration with other models	Architecture (2.1.3)
17	Industry acceptance	Adopting audience (1.1.2)
18	Tool support	Components & Tools (3.4)
19	Openness	Rights policy (3.2.2), Governance (3.2.3), Finance model (3.2.4), Operational model (3.2.6)
20	Availability	Rights policy (3.2.2)

Table 38 – Application of the model in Case B

Table 38 shows the list of evaluation criteria and their mapping to the model concepts. One example of the standards evaluation process is included in Appendix J to illustrate what information was gathered from each standard by using these criteria.

On the basis of these criteria, an evaluation of the standards selected in the evaluation list was conducted (Table 39). For this evaluation only the evaluation criteria 11 to 20 were relevant. A first value indicated the extent to which the criterion was met by a certain standard. The criteria allowed ranking semantic standards on an ordinal scale. For evaluation of the case of Siemens Corporate Technology the scale was chosen to range from 1 to 3, with "1" meaning the criterion is not met at all, "2" meaning the criterion is met to a certain extent, and "3" meaning the criterion is fully met.

Priority	Criterion	ACORD	Operational Data Model (ODM)	IEC 61968/70 CIM	ISO/IEC 15476 CDIF	ISO/HL7 21731:2006 HL7 RIM	IBM Industry Models	OAGis	Oracle Global Single Schema (GSS)	TM Forum SID	UN/CEFACT CCTS	openEHR Reference Model	JT File Format Specification
3	Structural cardinality	2	2	2	2	2	n/a	2	2	2	3	2	2
3	Semantic cardinality	3	2	2	n/a	3	2	n/a	2	2	2	2	2
2	Representation	3	2	3	n/a	2	2	1	n/a	2	2	2	n/a
2	Predefined content	2	2	2	2	2	2	2	2	2	2	2	n/a
0	Extensibility	2	2	2	2	2	n/a	2	2	2	2	2	n/a
1	Integration	3	n/a	3	n/a	3	3	2	2	3	2	2	n/a
2	Industry acceptance	2	2	3	n/a	3	2	2	2	3	3	n/a	3
3	Tool Support	3	n/a	3	n/a	3	2	n/a	2	3	3	2	3
1	Openness	1	2	2	2	2	1	3	1	2	3	3	2
2	Availability	1	2	3	1	1	1	3	1	2	3	3	1
	Evaluation Result	4.4	3.0	4.8	1.4	4.5	3.0	2.7	3.1	4.4	4.9	3.7	3.1

Legend: n/a - not applicable, i.e. could not be evaluated in the case.

Table 39 – Evaluation of semantic standards in Case B

In addition to that, priorities were assigned by Siemens Corporate Technology, ordering the criteria according to the relevance of the organization. The priorities 0 to 3 were chosen, with “1” representing lowest and “3” representing highest priority. The value “0” was introduced to exclude the criterion from the evaluation due to its irrelevance. The result values were computed by creating the average, weighted with the corresponding priority. The evaluation resulted in five semantic standards with a value higher than 4.0, what was considered a threshold for recommendation. These standards were ACORD, CIM, HL7 RIM, SID, and CCTS.

### 9.4.3 SSM applicability

The characteristics model for semantic standards in the case of Siemens Corporate Technology covered all required concepts. Most of the evaluation criteria referred to the categories Context and Content, followed by Development and Maintenance. For the Siemens case a general description of the standards was necessary, e.g. name, industry focus, developer, contributors, and a short description of the standard's content. This description was also used for the selection of the standards relevant for Siemens Corporate Technology. For further evaluation of the standards and for contrasting them, more details with regard to content were gathered. Here, also some aspects of the applicability were regarded.

## 9.5 SSM: Evaluation, discussion and conclusions

### 9.5.1 Evaluation

Both Gregor (2006) in her contribution on theories in IS and the GOM introduced by Becker et al. (1995) stipulate a multidimensional evaluation approach for theoretical constructs such as taxonomies. In the following, the SSM is evaluated against the combined set of criteria as described in the “Design Objectives” section.

- **Usefulness:** From the three use cases identified for the model (see section “Design Objectives”), two could be tested in case studies. Whereas Case A is about selecting semantic standards, Case B aimed at evaluating semantic standards. In both cases, the model was deemed useful by the project stakeholders with regard to supporting the project’s purpose. In Case A the model helped identify concepts which had not been considered before, but were then considered important to be addressed. In Case B, the model was used with a special focus on business object related standards. The general applicability for evaluation purposes was confirmed by the stakeholders in Case B. Moreover, one Siemens participant pointed out that for corporate-wide recommendation of the top-ranked standards, further analysis was necessary. In particular, the demand for reference implementation in software tools was articulated—what supports the inclusion of the sub-category “Implementation” in the model. The third use case, namely the influencing of existing standards, was not tested and should be part of future research.
- **Appropriateness and systematic structure:** The hierarchical structure and the grouping of concepts into categories were not issues of discussion in the case studies. Since the model was applied successfully in both cases, the silence on this point might be interpreted as a high parameter value.
- **Clarity:** The validation in the practitioners’ community (both through expert interviews and case studies) has shown that the description of concepts is of high relevance for the model to be considered “clear”. As standardization in general, and semantic standards in particular, are considered an “abstract” topic in the practitioners’ community, a clear definition of the meaning of concepts was necessary.
- **Completeness:** The case studies have shown that the model is considered complete with regard to the scope it was designed for. However, for evaluation purposes the concepts included must be accompanied by scales (which is not included in the current version of the model). In particular Case B has delivered evidence that the identification and documentation of a reference scales would be a reasonable area of future research.

### 9.5.2 Discussion

The discussion of the findings comprises both the design of the model itself and its application in the two case studies. The characteristics model for semantic standardization is based on the analysis of both the scientific and practical state-of-the-art. While theory supports the inclusion of individual concepts—as shown in Section “SSM in Detail”—the model in return also allows for some theoretical discussion. For example characteristic 1.2.3 (“Costs & Benefits”) which was used in Case A, but was not applied in Case B. The fact that a private business rated the concept apparently less important than a public authority corresponds with findings by Weitzel et al. (2006). They found that costs of standards adoption are relatively easy to quantify while the benefits might not be quantifiable at all. Consequently, the concept might be of limited value for standards end users.

Furthermore, Case A shows that the characteristics model is considered an useful instrument for policy makers in their ambition to support and guide the standard development process. This guidance might help to reduce the risk of “power games” and “standard wars” as described in literature (Nickerson & Zur Muehlen, 2006; Stango, 2004).

Case B revealed the need for reference implementations and software support in the practitioners’ community. Reference implementations do not (often) exist in the area of semantic standards. In Case B, “tool support” was rated a top-priority during the evaluation process. In general, the importance of software support for semantic standards confirms the path dependency theory according to which standards adoption at a certain point of time depends on previous adoption (Economides, 1996).

## Grounding

The grounding of the SSM based on theory within existing models (Appendix H) has led to interesting results about the diversification and gaps that have not been covered in the existing models. For instance Nelson et al. (2005) is the most detailed on the organization related characteristics of a standard, while Mykkanen & Tuomainen (2008) is the only model that looks at the content of the solution in detail. The model from Lampathaki et al. (2009) is mainly high-level and looks at what seems some random detailed aspects. Pawlowski & Kozlov (2010) focus on the solution and the use of the solution within implementation within the domain. None of these models cover everything from our SSM. To the contrary, several aspects from our model cannot be grounded within the existing models, and might be called a gap. We identified two groups of aspects that were only marginally addressed by the existing models:

1. Knowledge transfer, including concepts like a helpdesk, events & training, availability of consultants, support of pilots, availability of components & tools, and implementations.
2. Development & Management process: the process of how it is being initiated, the collection of requests, the design process, the formalization of the design, and the review and testing during development.

Second we identified some individual aspects that we did not find in the existing models:

- Costs & Benefits related to the standard.
- Vision/strategy of the standard development organization related to the standard.
- Quality management of the standard development organization related to the standard.

Based on this outcome we decided for an explorative additional grounding step, by searching for evidence of all characteristics within case study literature of semantic standards. We chose the studies from the systematic literature review (Chapter 4) that have been classified as case study research. The results are presented in Appendix K, and shows evidence for almost all characteristics with the exception of “Methods and Languages” and “Initiation”. Although without evidence we feel confident this belongs to the SSM.

## Other applications

The SSM is intended to be used as part of the iQMSS to measure the quality of standards, this chapter has shown that there are more use cases in which the SSM can be used. Further development might help increasing the usefulness of SSM. An example would be a method which outlines process steps for the application of the model in the three use cases introduced earlier. The design of such a method would help to improve and further develop the model itself while at the same time supporting its more wide-spread application.

### 9.5.3 Conclusions

This chapter reports on the design of a characteristics model for semantic standards (SSM). The model is grounded in both theory and practice and its applicability was checked in two case studies. The model contributes to the scientific body of knowledge in the field as it adds to a better understanding of the characteristics of semantic standards. Moreover, the model seems to be useful for practitioners. In the case studies it supports a policy maker in the process of public guidance with regard to the use and development of standards. And furthermore, it supports a private company in the process of evaluation of standards for internal use. Future research on the analysis and design of semantic standards, and in particular related to quality, can build on these results.



# Explorative Evaluation Studies

Chapter 6 described an iterative design approach containing multiple versions (builds) of the iQMSS. This chapter describes two explorative evaluation studies that have been performed during the early iterations of development to gather experiences and improvement suggestions related to the iQMSS. The SETU case (using build 0.3) is followed by the XCRI case (using build 0.4). The differences in these two builds are minimal; therefore only build 0.3 of the iQMSS will be introduced.

### **10.1 The iQMSS – build 0.3**

Starting point for the development of the first build of the quality model is the identification of measurable concepts. We have used multiple approaches to gather a list of measurable concepts. The foundation was laid out by the ISO 9126 standard (ISO/IEC, 2001). As it targets the evaluation of software quality, we had to eliminate the aspects that are irrelevant to standards. Secondly, we used a focus group for validation of the concepts identified and identification of other measurable concepts. With the same purpose, several other reports were used as well, and finally the result was tested in a case study within the educational domain. A full description of the process and its outcome is given by Krukkert & Punter (2008). In summary, the measurable concepts that have been identified, are visualized in Figure 43.

---

This chapter is based on:

(Folmer, Van Bekkum, Oude Luttighuis, & Van Hillegersberg, 2011): The Measurement of Quality of Semantic Standards: The Application of a Quality Model on the SETU standard for eGovernment. In: V. Folmin, K. Jakobs (Eds.), Proceedings of the 16<sup>th</sup> EURAS Annual Standardisation Conference “Standards for Development”, pp. 95-108, ISBN: 3861306298, Verlag Mainz.

(Folmer, Oude Luttighuis et al., 2011b): A Model for Semantic IS Standards. In: V. Folmin, K. Jakobs (Eds.), Proceedings of the 16<sup>th</sup> EURAS Annual Standardisation Conference “Standards for Development”, pp. 81-94, ISBN: 3861306298, Verlag Mainz.

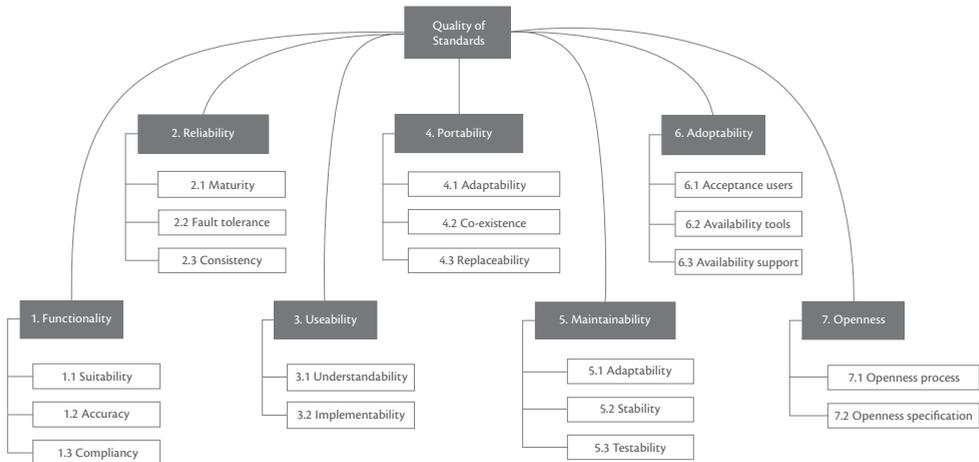


Figure 43 – Quality model for semantic standards (Krukkert & Punter, 2008)

Build 0.2 and 0.3 did not change the quality model as presented, but attempted to define related measures for the identified measurable concepts. These, including definitions of the measurable concepts are presented in Appendix L.

The quality model is intended to be used as self-evaluation tool by standards developers. In general this user group aims to achieve the highest quality in standards they develop, but is often unaware of the quality and possible improvements to their standards (Chapter 3). When the model is completed, it is foreseen that the use process starts with a preparation phase during which the appropriate measures will be chosen, combined with the measurement approaches, the weights of the measures and measurable concepts in relation to the overall score. The execution phase consists of performing the measurement, setting up the report by analyzing the results and will end by evaluating the measurement process.

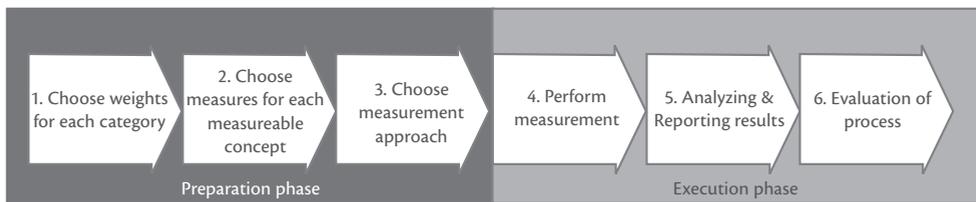


Figure 44 – How to use the quality model

Realizing that the quality model can be enhanced and further decomposed, we decided for an iterative design to make sure the developments will contribute in practice. Build 0.3 is suited to be tested in practice, although based on this build only step 4 and 5 from the execution phase can be tested. In order to be able to perform the preparation phase as described within Figure 44, a more advanced build is needed that includes measures, a scoring and weighting mechanism; because not every measure can be scored in the same way, and also the weights of a measure will differ based on the importance of the measure in relation to the quality.

The next sections will describe the two explorative case studies, SETU and XCRI, that have been performed based on build 0.3 (and build 0.4) of the iQMSS. All builds of the iQMSS contain the SSM that was presented in Chapter 9.

## 10.2 Explorative field test – SETU

We start in the next section by setting the scene, and explain why SETU has been chosen as explorative field test. The following sections will introduce the application of the quality model on the SETU standard. We report our findings and conclusions in the final section that will answer the research questions.

### 10.2.1 Research approach

Because of the explorative nature and the integration of theory and practice, a participatory action research approach is appropriate (Lau, 1999) for the first field test. This is implemented by having the same authors that have participated in the development of the quality model and that have participated in the development of the standard, perform the quality evaluation. This situation is present and accessible to the researchers for the SETU standard, a standard that is mandatory (“Comply or Explain”) in the public sector of the Netherlands in order to achieve eGovernment interoperability (NOiV, 2007). A second reason for selecting SETU is to assess the iQMSS applicability on rather mature standards. Finally this situation is not available to other researchers.

Currently within the SETU organization no knowledge about the quality is present. In this study we address this gap by applying the iQMSS (build 0.3) constructed mainly from the field of software engineering to examine the quality of a specific standard (SETU). By performing this research we gather knowledge about:

1. The quality model: Is the model adequate?
2. The quality assessment results: How useful are the results in practice?

In tradition of action research, application was done by two co-developers of the SETU standard, by going through the list of quality aspects. The results were validated based on review of two different co-developers.

### 10.2.2 SSM applied

The SSM is applied to receive the necessary knowledge about the standards and its context. Our approach was to describe all 33 aspects from the model for SETU, and make it a textual description without naming the 33 aspects. After that, we showed this textual description of what SETU constitutes to four other SETU experts.

During the creation of the textual description of the SETU standard the model performed like an easy to use guideline, by which the description of SETU was quickly written. The result however looks comprehensive and complete (see Appendix M for the full description). By giving more detail on every attribute the description could have become even more comprehensive. Other SETU experts did confirm that the description fitted the actual situation of SETU.

The application of the semantic standard model leads also to the identification of the sources needed for the application of the quality model (see appendix M). It also created awareness, that there is no one SETU standard, but at least four standards:

- SETU standard for ordering and selection version 1.1
- SETU standard for assignment version 1.1.
- SETU standard for reporting time & expenses version 1.1
- SETU standard for invoicing version 1.1

In the application of the quality model we aim to capture all four standards.

### 10.2.3 QMSS applied

Because of the extensive size, this section does not include the complete scoring of attributes from the QMSS, but Table 40 contains the overall scores for each measurable concept from the quality model. Recommendations for improvement of the SETU standard have been added by the researchers, based on the findings. The full set of results is available within Appendix N.

	Assessment	Recommendation
Functionality	<p>The standard provides the functions that support the stated needs and is highly suited to a small focused set of functions.</p> <p>Scoping however is not consistent between different documents, which may easily lead to confusion about the scope: Does the scope of SETU involve all electronic transactions or limited to the primary process? And does it only focus on temporary staffing through staffing agencies, or are consultancy organizations also part of the scope? There is also need for a broadening of scope.</p> <p>SETU standards contain options that have negative impact on interoperability.</p> <p>Although SETU is compliant with laws and regulations, this fact may be stated more explicitly.</p>	<p>Adjusting (broadening) and aligning scope is required.</p> <p>More strictness (less options) will lead to improved interoperability</p> <p>Compliance of SETU standards to laws and regulations can be more explicit.</p>
Reliability	<p>The SETU standards seem mature: A lot of information is available on the website. This however, includes outdated material. Users will not understand that, and may easily select the wrong document. The SETU organization has yet to prove its reliability in the future. The board could bring more balance to the representation of different types of stakeholders.</p> <p>The standards contain possibilities for corrections, but with options. The correction process needs to be standardized as well.</p> <p>The SETU data dictionary is an important addition to HR-XML, reducing ambiguity.</p>	<p>Keep deprecated material separated from current documentation.</p> <p>Update documents like the roadmap.</p> <p>Invite other types of organizations into the board of SETU.</p> <p>More attention on process standardization.</p>
Usability	<p>In general there are no major limitations to the usability of the SETU standard found in this study. The understandability of the highly structured specification documents seems straightforward. Especially the standard for reporting time &amp; expenses however, may benefit from improved readability, since it has a remarkably low score on readability. To shorten the time required to learn the standard, it may be useful to develop and share learning material (or courses). Implementability for staffing customers is affected by the lack of SETU support of procurement vendors.</p>	<p>Have a closer look at the readability of the standard for reporting time &amp; expenses.</p> <p>Availability of training material.</p> <p>A better connection with the infrastructure of the systems of the hiring company is advisable.</p>
Portability	<p>The portability of the SETU standard seems in line with expectations: It is adaptable to suit specific customer demands and the co-existence with HR-XML is perfectly logical. In the future, the alignment with invoice standards requires attention.</p>	<p>Alignment with invoice standards requires improvement.</p>
Maintainability	<p>Most importantly, the SETU standard is maintained, stable and adapts to needs in practice. It has some flexibility to adapt to the different needs in projects.</p> <p>Dependency on HR-XML is an issue, especially since HR-XML is changing its course lately. But also for instance the data definitions of HR-XML need improvement, just as version management.</p> <p>Implementations of the standard can be tested by using the SETU validation service.</p>	<p>Improvement of HR-XML data dictionary in line with the SETU data dictionary.</p> <p>Keep up with HR-XML versions.</p> <p>Next step of validation should be certification.</p>
Adoptability	<p>The adoption of the standard varies. The level of adoption is high on the supply side, but low on the demand side (staffing customer). The latter may improve now that SETU is on the “Comply or Explain” list of the Dutch government, which is specifically targeted at the demand side.</p> <p>In line with the above statement the adoption by the procurement software vendors needs improvement.</p>	<p>Promotion to staffing customers to enhance adoption.</p> <p>Promotion to software suppliers of staffing customers, especially the vendors of procurement software, like SAP and Oracle.</p>
Openness	<p>SETU is an open standard as confirmed by advice from an independent expert committee appointed by the Dutch government. Although SETU is not completely open on every detail, it apparently manages to deal with openness that befits the situation.</p>	<p>No recommendation.</p>

Table 40 – SETU high level results

## 10.2.4 Evaluation

In this section we look back at the three questions presented earlier and we will answer them subsequently:

### 1. The quality model: Is the model adequate?

The application of the instrument resulted in that the quality has become assessable, visible and relevant improvement suggestions are made, even for a well acclaimed standard such as the SETU standard. This leads us to answer this question positively, but there still is much to improve. In general, since the applied instrument was not finished in every detail, the following holds and was already established prior to this application:

- Attributes, measures, and measurements approaches needs to be further defined.
- The same holds for scoring mechanisms and guidelines on how to perform the tests.

We can summarize these as more details and guidelines are needed for improving the measurement. More detailed and based on the SETU application, we conclude:

- Future plans (like a roadmap) are not taken into account; the quality model is only looking at available results.
- The approach of modeling the processes and data is not part of the quality model.
- The content of a standard is captured in more items than just its specification. Especially for determining the quality, several other documents are important, apart from the specification.
- The openness of SETU is only marginally assessed by both the expert advice and the quality model. It would have been much more valuable to use the 10 requirements on openness (Krechmer, 2009). These 10 requirements are a broader and more balanced view on openness than a strict use of the definition of openness from the European Interoperability Framework.

### 2. The quality assessment results: How useful are the results in practice?

To be able to answer this question, the assessment results were validated by two other co-developers of SETU, not involved in the application of the quality model. Based on this study, it is impossible to state an explicit notion of semantic standard quality, like a certain number on a scale, or a value like perfect, sufficient, or not sufficient. Still assessment gives the impression that there are no major flaws in the quality, which supports the thought that the quality of SETU is rather good. Although this might be true, based on this study, it is an impression and feeling and not irrefutable proven. More importantly, this study does show some possibilities for improvement, exactly what the instrument is aimed for. In no particular order, the most important suggestions for improvement are:

1. Adjustment (broadening) and alignment (with practice) of scope is required.
2. More strictness (fewer options) will lead to improved interoperability.
3. Keep deprecated material separate from current documentation.

Worth noticing is that the SETU operating procedure is an important document for the quality assessment. This document can be further enhanced to be a “quality process” document. On the other hand, although the contrary was expected, the usability of the expert advice report was fairly limited in this assessment. It contains statements such as ‘the usability is good’, which may be sufficient for an assessment, but does not suffice as a fundament for starting quality improvements.

An unexpected eye-opener for the co-developers of SETU was the amount of outdated documentation on the website including deprecated versions of the standard. For the SETU organization the outcome is valuable, and will be a starting point for a quality boost.

## Research considerations

There are several considerations to be taken into account, when assessing the research presented in this section:

- A limitation of our research approach is that the assessors are also co-developers of SETU. Would the instrument also be of use to “independent” quality auditors that are new to a standard? To perform such a quality assessment, deep inside knowledge is required; only using documentation would probably not have given the same result.
- The self-evaluation has consequences on the outcome, especially the risk of bias.
- One application case is not enough to generalize: more studies are required in order to generalize our results. The same study can be performed on standards within other domains as well.
- The quality model needs to be further extended to make usage possible by others than the developers. The results so far are an instigator for further development of the model.

### 10.3 Lab experiment – XCRI

Following the SETU explorative field test, a complementary lab experiment was executed: the XCRI case. We chose for a different research method as every research method has its own strengths and weaknesses, and will lead to different results. In particular lab experiments are methodological complementary to field experiments (Harrison & List, 2004). We used this lab experiment, as part of our multi-method approach, to achieve additional insights in comparison with the SETU explorative field test.

#### 10.3.1 Experimental set up

Because of our work at the University of Twente we had access to master students enrolled in a graduate MSc course on Business Process Integration. This gave us the opportunity to work with multiple groups and compare results. The advantage of working with students is the homogeneity of the group that allow comparisons of subgroups (Harrison & List, 2004). An obvious limitation is that students are not the intended users of the instrument, and lack knowledge about the standard itself. We tried to overcome this problem by selecting a relatively simple standard close to the domain of the students. The standard XCRI (for exchanging course-related information) was chosen, because the education domain is familiar to students and the XCRI is relatively simple, and a quite complete overview is presented on the Internet ([www.xcri.org](http://www.xcri.org)). The goal of this experiment is to gather knowledge about the usage of the instrument to gather improvement suggestions. Lab experiments are a useful research method for generating these kind of qualitative insights (Levitt & List, 2006). The experiment is performed as assignment given to students that consists of three parts:

1. Describe the XCRI standard.
2. Describe the quality characteristics of XCRI.
3. Present top 5 improvement suggestions for XCRI.

On October 6<sup>th</sup> 2010, a class of students (master Business & ICT at the University of Twente) was involved in an experiment to measure the quality of the XCRI standard. Build version 0.4 of the instrument was used by half of the groups during class. The other groups had to find out how to perform a quality measurement without iQMSS. The class lasted for 1 hour and 45 minutes. As preparation, the students received a week in advance, two articles about semantic standards (Nelson et al., 2005; Zhao et al., 2005). They had no a priori knowledge of neither the XCRI standard, nor any part of the iQMSS.

The assignment to describe the XCRI standard was limited to approximately 20 minutes. Two simulated experts were present and responded to questions asked by the students by e-mail. They answered the questions based on a Q&A with one of the actual XCRI developers. The students were randomly divided in six groups of two or three students each, in line with the characteristics of a *true* experiment (Creswell, 2009).

Three groups (1,2,3) used the instrument, three (4,5,6) did not. The two experts ranked the descriptions of both groups. Based on a questionnaire each group was evaluated afterwards.

This approach led to 6 descriptions of the XCRI standard, 6 applications of the quality model, many improvement suggestions, e-mail exchanges between the students and the simulated experts, and notes taken. The complete set is available by contacting the author of this book.

### 10.3.2 Results of the use of SSM

This section will present the results of the groups when applying the semantic standard model on XCRI. The results of each group can be summarized as:

- Group 1: Broad view of different aspects of the standard.
- Group 2: Looks only at adoption and tests in practice: an “implementers view”.
- Group 3: Broad view of different aspects of the standard.
- Group 4: Small, technical view.
- Group 5: Some random characteristics.
- Group 6: Used the table from the paper of Nelson et al. to structure their description.

The results of the groups without iQMSS were generally narrow, technical, and seemingly random descriptions of XCRI, and did not contain information about the development and maintenance organization, or market adoption. One group did actually use the framework of Nelson et al. (2005) to describe some aspects of XCRI. They mainly used the easily accessible information on the front web page of XCRI, which is a limited view.

The user groups with iQMSS succeeded all in delivering a more comprehensive and broad description of XCRI. Yet, because they were focused on using the model, they had to rush to finish in time. In contrast to the other groups, they were looking for specific information about the XCRI standard. The results of these groups also included information about the development & management organization and information about implementations of the XCRI standard, in contrast with the other groups.

During evaluation, the groups responded that they welcomed the model as a guideline. The groups without the model requested a guideline. However, according to the user groups with iQMSS, the descriptions and definitions within the model needed clarification. Also guidelines on how to use the model were requested.

### 10.3.3 Results of the use of QMSS

The second and third part of the assignment relates to the quality model. First we present some remarkable results of the groups, related to the description of the quality characteristics:

- Group 1: Instrument was only used for 50% (time issue). It seems they did not have many problems using the instrument with the exception that they did not find much evidence within the standard for several quality attributes.
- Group 2: Only a small part of the instrument was used, and several quality attributes were misinterpreted.
- Group 3: Again, part of the instrument was used. The quality attributes were sometimes answered as yes/no questions, which is not the aim of a quality assessment.
- Group 4: This group did something completely different; they set up some characteristics of XCRI readiness (which is available on the website), but is not related to quality.
- Group 5: This group developed their own quality characteristics (which all are part of the instrument as well).
- Group 6: This group started by the goal of interoperability, and linked it to several quality characteristics like openness, ease of use, etc. and tried to set up some measures like amount of request for changes.

The outcome of group 1 to 3 may seem somewhat incomplete, the good thing is that they stayed on track contrary to group 4. The value of the results of groups 1 to 3 only became noticeable during assignment 3. Groups 5 and 6 showed a very limited view on quality, but they succeeded to relate the goal of interoperability to a couple of quality characteristics and defined measures for that. By doing this they were actually constructing a quality instrument.

Characteristics of the improvement suggestions for each group are:

- Group 1: 5 suggestions of which two are practical (like update the blog), and three suggestions seems valuable. Especially the observation of a missing roadmap seems in place.
- Group 2: This group created valuable observations, without suggestions for improvements related to quality. But other improvements such as improving the number of implementations and to include information about the cost of an implementation. The most valuable observation is that the maintenance of the standard is not structurally organized, which might become a risk.
- Group 3: Probably because of time limitations this group delivered only three short suggestions, including more implementations, and cooperation with international standards bodies.
- Group 4: Although 5 suggestions were presented they seem to be somewhat vague. This might be related to the misinterpretation of assignment 2 by this group as well. However valuable suggestions are: more information for adoption, more examples of interoperability with the Forum supporting implementations. And, more information about estimations for implementing projects might help.
- Group 5: Somewhat vague (and short) suggestions. Even one suggestion might be contra productive: limit the amount of contributors. Certification and training seem to be valuable suggestion, just like to plan ahead for future technologies.
- Group 6: This group has included some practical suggestions (use spam filter on the wiki), and some higher level suggestions: start sharing vision about quality. Another interesting suggestion is the sharing of usage statistics.

All groups delivered some valuable suggestions. However looking at the results in more depth, than group 1 & 2 have identified the most valuable suggestions: Set up a roadmap for future directions and organize structural maintenance (arguably the most important one).

#### *10.3.4 Evaluation*

Four out of six groups did have lively discussion while two other groups were very quiet; which might be explained by cultural differences (these two groups were the only groups with students having an Asian background). The results of the groups without instrument were handed in much faster than the results of the groups with instrument. The groups that used the instrument used it very rigorously, which caused time problems. Table 41 contains the evaluation results based on a questionnaire.

Group	iQMSS used	Satisfied with Process	Satisfied with Result	Enjoyed participation	Comments
1	Yes	Negative	Very Negative	Acceptable	Instrument is useful to organize your thoughts. Needed quite some time to understand the attributes and to map them to something operational in the XCRI-standard. More knowledge about standard is required.
2	Yes	Acceptable	Acceptable	Positive -Acceptable	Get to know the standard first. More explanation of the attributes is needed. Overall, though, it is much helpful than without one. An instrument improves the process. However not the complete model can used for every standard. It needs to be fitted to the standard.
3	Yes	Negative -Acceptable	Acceptable- Negative	Negative -Acceptable	In principle the instrument is really helpful. However the current definitions are not sufficient. The instrument helped us to structure the process. We did have some difficulties to find the necessary information.
4	No	Acceptable	Positive	Positive	Thinks that instrument will improve results.
5	No	Positive- Acceptable	Positive -Acceptable	Positive	A quality model might have helped.
6	No	Negative	Positive	Positive	An instrument with quality measurements would be preferred. Otherwise at least a quality definition is needed.

Table 41 – Evaluation results

Comments were made by groups 1, 2 and 3 about the limited time and that they were not satisfied with both the process and the result. They needed more time to get to know the standard, and also for understanding the instrument. In summary the results are:

- More knowledge about the standard is needed for using the instrument
- Much more time is needed for good use.

More remarkable is:

- The groups that did not use the iQMSS had enough time and enjoyed participation in the experiment. The groups with iQMSS all ran out of time, made comments about rushing and running out of time, and probably because of that enjoyed the experiment less.
- All groups say that an instrument will improve the process and result of the quality measurement.
- The groups that used the instrument already think that the current version is helpful mainly because it is structuring the analysis.

### Research considerations & discussion

A general drawback of a lab experiment is the low external validity which means that generalization is limited (Roe & Just, 2009). Other limitations include the fact that the users were students with no experience in this area, which is different compared to the intended user group that is involved in standardization. The intended users have more experience and knowledge in general but also regarding standardization. This will likely have influenced the results, especially since the students had some difficulties understanding the XCRI standard. Students declared they had mainly difficulties in mapping the quality attributes on characteristics

of the standard. It might be expected that for standards developers this will be much easier. However on a high abstraction level, and supported by the evaluation, an instrument helps to structure the process of quality measurement, even for students with hardly any knowledge about the standard itself.

Based on this experiment a conclusion might be drawn that the instrument (build 0.4) is already supporting the process of quality measurement. A second conclusion can be drawn that improvements are needed:

- Include more detailed measurements in addition to the quality attributes
- Improved descriptions of the quality attributes
- Flexibility of model usage since not all quality attributes are applicable to every standard.
- Be aware of too rigid usage of the iQMSS, by which full appliance of the instrument becomes the goal, instead of finding quality improvements. Strict, too rigid, usage should be avoided.

For follow-up experiments involving students it is recommend to have much more time available (for the users of the instrument) and to arrange a lecture being held upfront about the standard as subject.

#### **10.4 Conclusions**

Upfront we already knew that generalization of the explorative evaluation study results is not applicable, and that care is needed when stating conclusions about these evaluation studies. The selected research approach was particularly suited for this explorative work, and resulted in many improvement suggestions that were included in further builds of the instrument.

Both standards were relatively simple, although SETU actually consists of multiple standards. In our explorative field test it did not lead to problems. However it might be easier and more dedicated to select one standard that relates to one interoperability problem for quality measurement.

It is remarkable that even for a mature standard with “Comply or Explain” status the quality measurement resulted in valuable improvement suggestions. For achieving this status, the SETU standard passed a thorough procedure that included assessment of its quality, just as other (political) factors. SETU is expected to be of high quality, otherwise it should not have achieved the comply-or-explain status in the Netherlands.

The XCRI lab experiment led to some interesting improvement suggestions. Also the confirmation that an instrument is a must-have when performing quality measurement, is an important result of the lab experiment.

Amongst others based on the experiences gathered within these two explorative evaluation studies further developments on builds have taken place. These further developments and the outcome will be presented in the next chapter.

## Chapter 11

# The Quality Model of Semantic Standards

This chapter will describe how the QMSS was developed, its final version, and how it should be used. The focus is on the development of the generic QMSS, based on earlier defined fundamentals. Figure 45 shows the topic of this chapter within the reference model of the iQMSS.

	Specification Quality Model Semantic Standard (QMSS)	Specification Semantic Standard Model (SSM)	Implementation
M2	1. Quality Language (QL)	I. SS Language (SSL)	A. Development Environment
M1	<b>2. Generic QMSS</b>	<b>II. Generic SSM</b>	<b>B. GMI Generic Model Implementation</b>
M1	3. Customized QMSS	III. Customized SSM	C. CMI Customized Model Implementation
M0	4. Measurement Result	IV. Semantic Standard	D. Measurement Result Report

Figure 45 – The topic of this chapter (QMSS) within the iQMSS

This chapter starts with a description of the research approach used to construct the QMSS, and the output of this process is described in sections 11.2, 11.3 and 11.4. Section 11.1.3 will describe the overall QMSS structure and its measurable concepts will be presented in section 11.2. Section 11.3 will continue by adding more detail about the measures, with respect to the QMSS product quality. Finally section 11.4 will describe how the quality models should be used in practice. In summary:

### 11.1 Research methodology

### 11.2 The QMSS – The measurable concepts

### 11.3 The measures for product quality

### 11.4 The iQMSS usage process

Thus part 11.2 and 11.3 will present the generic quality model applicable in practice, while in 11.4 we will focus on the guidelines of how to use the QMSS in practice. These three sections together, are the most essential (i)QMSS parts, with the exception of tooling. The potential QMSS user should therefore start with reading section 11.4 and thereafter sections 11.2 and 11.3.

This chapter is based on:

(Folmer, 2011): The Quality Model of Semantic IS Standards, Paper presented at the 8th International Conference “Standardization, Prototypes and Quality: A means of Balkan Countries’ Collaboration”.

## **11.1 Research methodology and structure**

Our research methodology has been split logically between the overall approach of how the different builds of the instrument come together, and the research methodology applied to the final build. The overall approach will be described in the first section, while the detailed design approach of the final build phase is described in the second section. The final section will set the structure of the QMSS.

### *11.1.1 Overall approach for build iterations*

The design process of the final QMSS build was experimental in character whereby different builds of the QMSS were applied in explorative case studies. These different builds were used on different sources of literature, and yielded different results. Chapter 5, the literature review, has already shown that although a quality instrument for semantic standards does not exist, very many studies have been done and can be used to set up such a quality instrument. Although Chapter 5 describes many of these studies we take this one step further and searched for studies that particularly mention measurable concepts or measures related to quality for different kinds of artifacts that might be valid for semantic standards as well. A complete overview of these studies, including the mentioned quality related artifacts, has been developed and is presented in Appendix O.

Our development started with the design of the quality model within the Integrate project (Krukkert & Punter, 2008), which can be seen as the pre-successor of the QMSS, be it is our first build. During this first build some scientific studies were included, mainly from the software domain, but in the end it was generally more practical oriented.

In multiple iterations several builds were constructed accounting for a growing amount of practical experiences and theoretical studies in each build. The builds 0.3 and 0.4 have been used for evaluation studies as described in Chapter 10. Build 0.5 focused on indentifying applicable measurable concepts from the data quality domain by surveying the relevance of the concepts to semantic standards (Folmer & Van Soest, 2011). The first five (from 0.1 up to 0.5) builds of the instrument were all explorative in nature and did not have strict version management. As a result, build 0.5 is not a continuation of build 0.4, but instead was based on build 0.1. Therefore we created a new build, 0.6, which is an integration of all the previous builds and is the fundament for further development. In our final build, 0.7, all known sources have been included for completeness.

Figure 46 shows all the information sources used in the steps taken during the development of the instrument. A number of sources have not been specifically mentioned in the figure, but nevertheless were used in this research, as part of the software quality domain and were encapsulated within ISO 9126x and ISO 250xx. These are: (Boehm, 1973; Cavano & McCall, 1978a, 1978b; Humphrey, 1989; Larrucea, 2008; McCall, Richards, & Walters, 1977; Milicic, 2005). Also in the area of data quality other sources were used (Alexander & Tate, 1999; Dedek, 2000; Katerattanakul & Siau, 1999; Naumann & Rolker, 2000; Shanks & Corbitt, 1999; Zhu & Gauch, 2000) and were encapsulated by Knight & Burn (2005).

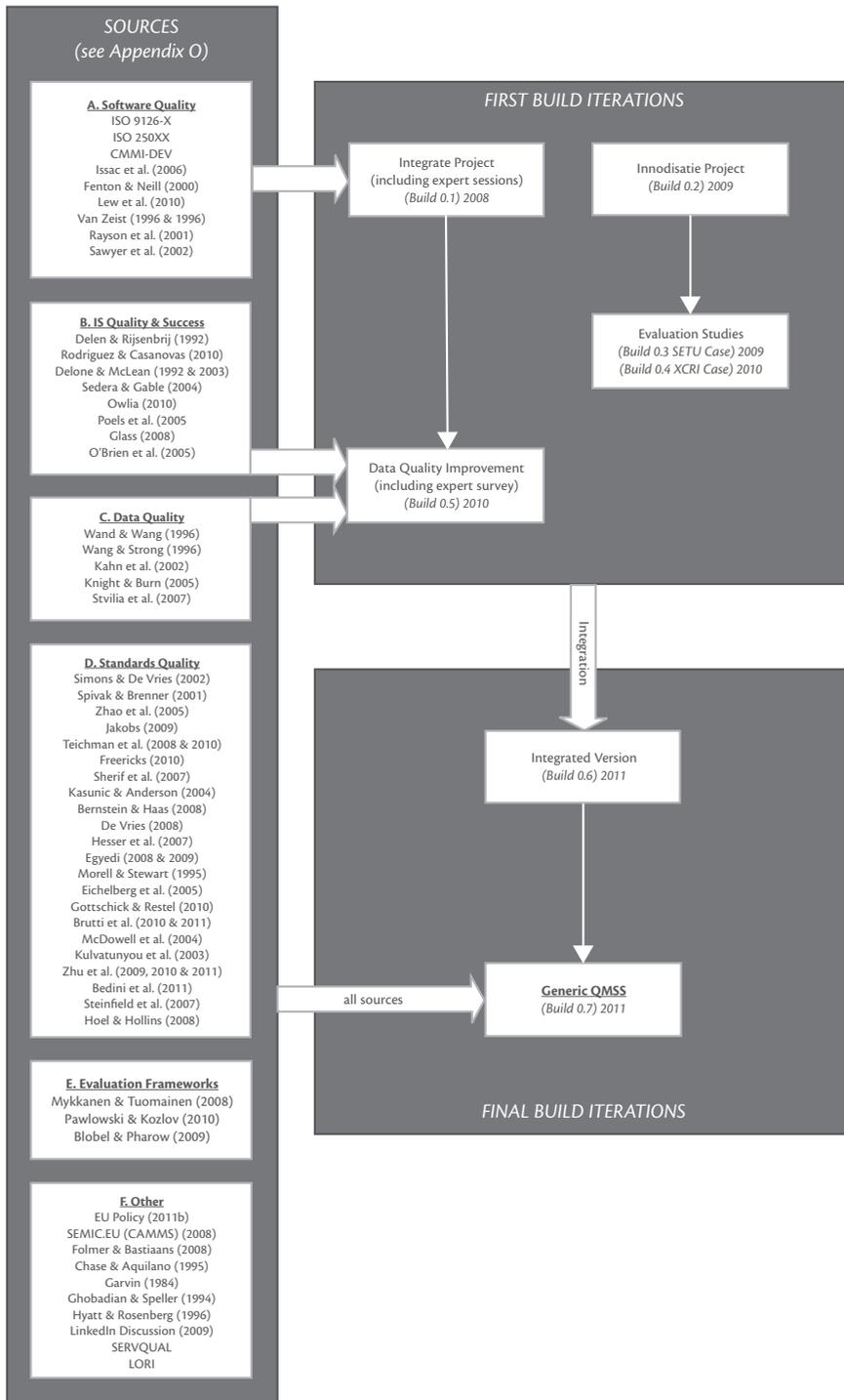


Figure 46 – Overview of QMSS sources and builds

11.1.2 Final build research approach (build 0.7)

This section describes the research approach of the final iteration within the scope of this thesis that leads to build 0.7.

**Research approach a priori**

The starting point for the final build of 0.7, was the previous integrated build (version 0.6). This bottom-up approach was continued by following four main steps: A. define the high level structure (section 11.1.3), B. define the quality model (measurable concepts) (section 11.2), C. define the measures (section 11.3), D. define usage process (section 11.4). These four steps are a work breakdown approach focusing on specific parts of the QMSS.

The same approach was used for each of the four main steps: carrying out a circle of sub steps, ensuring all the requirements were checked, literature sources were included, design rules were followed, etc.

Additionally, one sub step was added, when it became apparent that there was a lack of sufficient measures in the literature sources (step C); we set up an expert workgroup to gather additional measures. This workgroup session was also used as a means for reviewing the measurable concepts of the quality model.

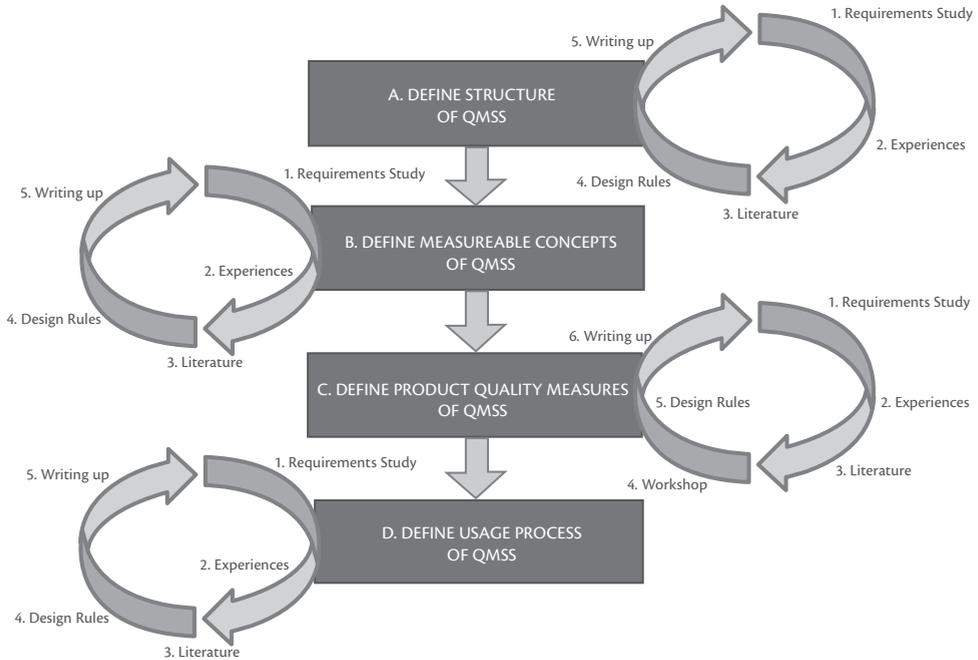


Figure 47 – Research approach for the final build

**The sub steps**

Within the final build phase it is important to: (1) meet the iQMSS requirements, (2) apply the knowledge gathered by explorative case studies, (3) apply existing knowledge, (4) apply the design rules, (5) and write them up according to the quality language. Each of these sub steps will be explained in more detail:

1. The design of the QMSS has to meet the requirements as presented within Chapter 7. This is implemented by checking the requirements actively during every design phase.
2. The explorative case studies, as described within Chapter 10, resulted in lessons learned that are included in every design phase.
3. Many studies have been performed for setting up quality artifacts for different domains. These are presented in Appendix O, and are used when applicable.
4. Apply the design rules on how to use and transform the knowledge gathered into solutions for the QMSS. These design rules will be explained in the next section.
5. The solution has to be written down according to the chosen terminology (e.g. measurable concepts, information needs, attributes, measures) of the quality language as described within Chapter 8.

The combination of these sub steps ensures the quality attributes of the software domain (as an example) are checked for their relevance to semantic standards (the requirements), and are aggregated and described according to the design rules and the quality language that have been selected for the QMSS.

### The design rules

Design rules give guidance on how knowledge can be transformed and used for the solution, in this case the QMSS. We start by describing the general design rules for IS theory, and will continue with more specific design rules for measuring instruments. A key lesson is simplicity (Glass, 2008), based on the parallel between the Apgar score for newborn babies, and an IS score is proposed. A model with “ilities,” such as reliability, maintainability and efficiency (like ISO 9126), with a simple measurement is enough and is a key to success (Fenton & Neil, 2000). Philip Theden distinguishes three applicable metric characteristics (Bundschuh & Dekkers, 2008):

- Information character: Where metrics permit one to make judgments about important subjects and relationships in organizations.
- Quantifiability: Where subjects and relationships between them are measured on a standard scale.
- Specific form of information: Where complicated structures and processes can be presented in relatively simple ways through a specific form of metrics.

The following design guidelines from the IS field are applicable (Gregor, 2006):

- A classification system is useful in aiding analysis in some way.
- Category labels and groupings are meaningful and natural.
- Hierarchies of classification are appropriate (most important divisions are shown at the highest level).
- The logic of the phenomena into categories should be clear, as should the characteristics of each category.
- Important categories should not be omitted; it should be complete and exhaustive.
- A previous classification system could be revised as new entities come to light, or some preferable way of grouping or naming categories is identified.

“A judgment as to the degree to which the theory satisfies these criteria allows one to assess the contribution to knowledge” (Gregor, 2006). More specifically, Cavano & McCall (1978a) propose evaluation characteristics related to the utility of the model. The intention is to use these evaluation criteria after the design (of software), but in our case they will be used as design guidelines. These characteristics are (Cavano & McCall, 1978a):

- Definition
  - What is the model measuring?
  - Is it detailed enough?

- Fidelity
  - Will different quality assurance personnel get similar results?
  - Are the actualities close to the predictions?
- Constructiveness
  - Does it help in understanding (software) quality?
  - Are the derived measures explainable?
- Stability
  - Can the model be manipulated to obtain desired results?
- Usability
  - Can the methodology be cost-effectively implemented in a quality assurance program?

Some guidance rules are available from both practice and theory for selecting measurable concepts and measures. The URU is a concept used in practice for instance by SNOMED (medical standard) and stands for: Understandable, Reproducible, Useful. It is applied to the terms within the SNOMED CT standard, but is also applicable to many other artifacts.

We can learn from CMM (Humphrey, 1989) as to how one should inspect the characteristics of quality measures, such as objective, timely, available, representative and controllable. It does not necessarily mean that all of them will be completely true for every measure. Other criteria for measures are (Morell & Stewart, 1995):

- The metric must be:
  - Easily observable (does not require elaborate or special means of data collection)
  - Reliable (different observers should get approximately the same result under similar circumstances)
  - Well defined (a general agreement on what the indicator means)
  - Useful (knowledge of the indicator must lead to practical actions in reasonable time)
- Data collection: it must be possible to systematically integrate the data into an ongoing activity, thus removing the possibility of measures being used too infrequently.
- Data must be in a form and location where it is easily available to those who can make use of it.
- The data must be useful at reasonably gross levels of approximation. Otherwise, too much effort will be expended on data quality assurance.
- Metrics must make sense when used jointly rather than individually. The problem is that any process can be distorted to meet any given metric, but joint metrics tend to preserve the intent behind the measure.

In practice not all criteria can be met (e.g. easily observable, no special data collection), but it is still valuable to keep them in mind during the selection of measures.

### Performing the research

During the research process it became apparent that the selection of quality measures was problematic. Most available studies within the different domains have put a lot of effort in setting up models including measurable concepts, but leave the question on how to measure, open. The process of how measures can be selected and validated has been described but it also shows the complexity (Poels, Maes, Gailly, & Paemeleire, 2005). In this research it was not feasible to set up a complete list of validated measures as it could not be constructed based on the current studies.

Since the available studies did not satisfy the needs of our QMSS, it was decided to gather indicative measures from practice based on a workgroup session. This workgroup session took place during June 2010 at the University of Twente. Ten experts (from TNO, Novay and the University of Twente) participated, of which 5 are actively involved in standards development of at least one specific standard (group 1), and 5 are experts

in the area of standardization but they are not actively involved in a specific standard (group 2). By creating 5 groups of 2 people (always combinations of both groups) in sessions of half an hour each, measures were constructed. Every half hour each group was assigned another part of the quality model which required the measures. Each group had to build upon the work of the previous group. In total 5 sessions were held, before a short evaluation was done. All groups also commented on several definitions of the quality model that needed improvements and were updated afterwards accordingly.

Other results from the evaluation suggest the experts did not miss elements in the model. Too complete was mentioned once, and several experts suggested situation dependent use of the quality model to avoid the need for all concepts. Another valuable remark was that by focusing on measures, a bias might be introduced, since concepts that are hard to measure are neglected.

### Observations

During the build process some measurable concepts appeared and disappeared. For example fault tolerance is part of ISO 9126, and was included within the first builds of the instrument. However based on survey results, it was excluded in build 0.5, and then included again in build 0.7 with an adapted definition.

Some specific elements from the model have been extensively researched, while others are new or hard to capture. For example completeness and relevance have been extensively researched and a measurement instrument has been developed specifically for these two measurable concepts (Zhu & Fu, 2009; Zhu & Wu, 2010, 2011).

#### *11.1.3 The structure of QMSS*

A flexible structure is part of the requirements, while the design rules talk about a logical structure. Within the explorative case studies, the different QMSS builds grew in terms of the number of quality measures and often the added measures were not strictly related to internal quality. Based on these three findings, we had to come up with a logical structure to make the instrument flexible to use.

The original information prerequisite for our research scope was related to the intrinsic quality of the standard. During the explorative case studies and on the basis of the requirements study, other information needs became apparent, asking amongst others:

1. The internal quality of the standard? – The original information need.
2. The implementability of the standard?
3. The durability (future-proofness) of the standard?
4. Should I select the standard?
5. Is the standard a good solution for the interoperability problem?

Looking at it from a broader perspective, we see that distinctions should be made. For instance the distinction between the product and the process, as has been proposed by many (e.g. (Hoel & Hollins, 2008; Morell & Stewart, 1995; Stvilia et al., 2007)). According to them two types of metrics are important (Morell & Stewart, 1995):

- Monitor the progress of the process = process metrics
- Quality of the standard (outcome) = product metrics

Our research has already shown that the relevant concepts for a semantic standard include its context, content, development organization, and its application (Chapter 9). This also reflects the ISO 9126 and 25000 family of standards for software engineering, that includes separation of concerns based on the product (internal and external), the process and its use (Figure 48). Also the legitimacy model of Hoel & Hollins (2008) make explicit notion of the specification (the product), the process, and the implementations of the specification (its use).

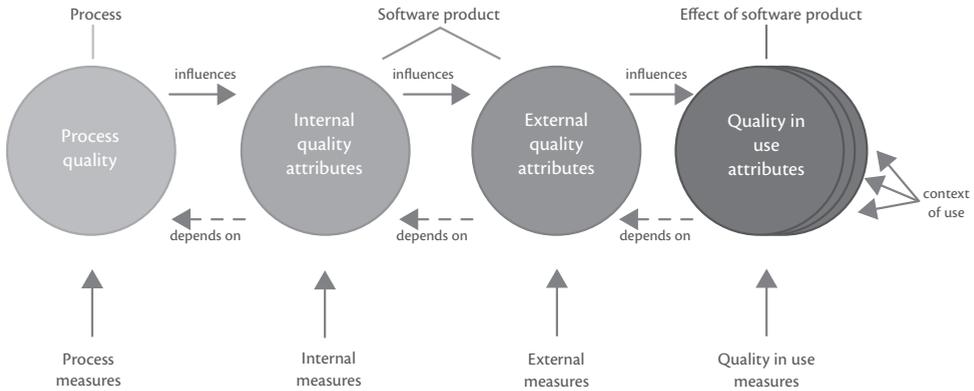


Figure 48 – ISO quality model for software (ISO/IEC, 2001)

The result of applying this separation of concerns to the quality model is a separation of the quality model into three parts: product quality, process quality, and the quality in practice. This maps the conceptual model of a semantic standard since product quality deals with the content (the specification), the process quality relates to the development & maintenance processes as carried out by the development organization, whereas quality in practice deals with the application environment, the performance of the implementations of the standard.

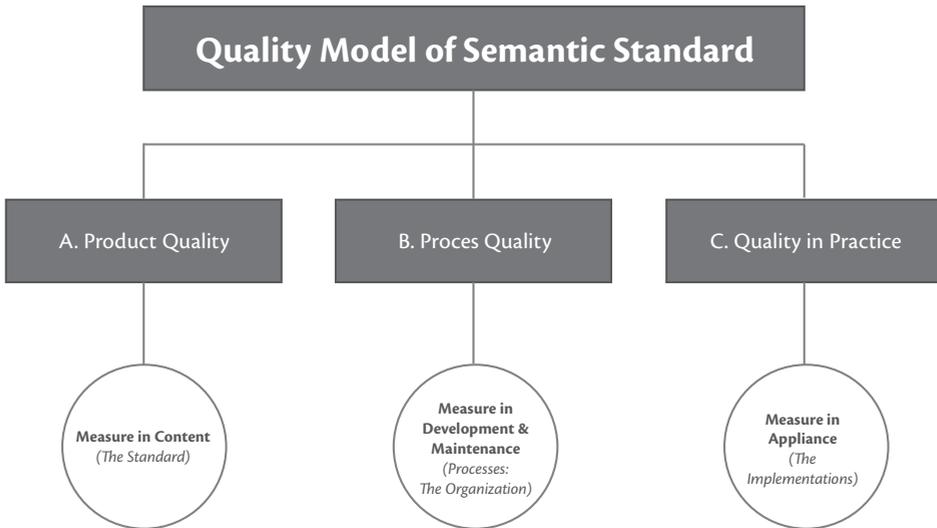


Figure 49 – Structure of QMSS

This structure makes the use of the quality model more flexible. Dependent on the information needs, only parts of the quality model have to be used. Our information needs can be mapped onto the three parts accordingly:

1. The internal quality of the standard? – Part A
2. The implementability of the standard? – Part A+ B
3. The durability (future-proofness) of the standard? – Part B + A (partly)
4. Should I select the standard? – Mainly part C
5. Is the standard a good solution for the interoperability problem? – All parts

Our focus throughout this research project has been on the internal, product quality of the standard. This model shows the boundaries and context of product quality, and during the next step we will present the first versions of the other parts as well. However the measurable concepts of part B and C might be less mature since our build-evaluate iterations were focused on product quality. Only the product quality measures (section 11.4) will be presented.

### **11.2 The QMSS (final build 0.7)**

This section contains the quality model and introduces all the measurable concepts for all three parts, including definitions.

#### *11.2.1 Product quality*

Based on the research approach the model for product quality was constructed. The product quality basically consists of three information needs:

1. Is the functionality of the standard appropriate? - Does it have the features to solve the interoperability problem?
2. Is the standard usable? - Can the standard be implemented and used without burden?
3. Is the standard durable? - Will the standard be future-proof?

These three information needs define the structure within the model.

With regard to technical complexity our measurable concepts, and later on the measures, focus on XML technology. If other technology is used, the model should be changed accordingly, including the measures. The latter might be quite difficult, because XML metrics are studied because of their commodity.

The model for product quality, as an output of the research approach described in section 11.1, is depicted within Figure 50.

# A. Product Quality

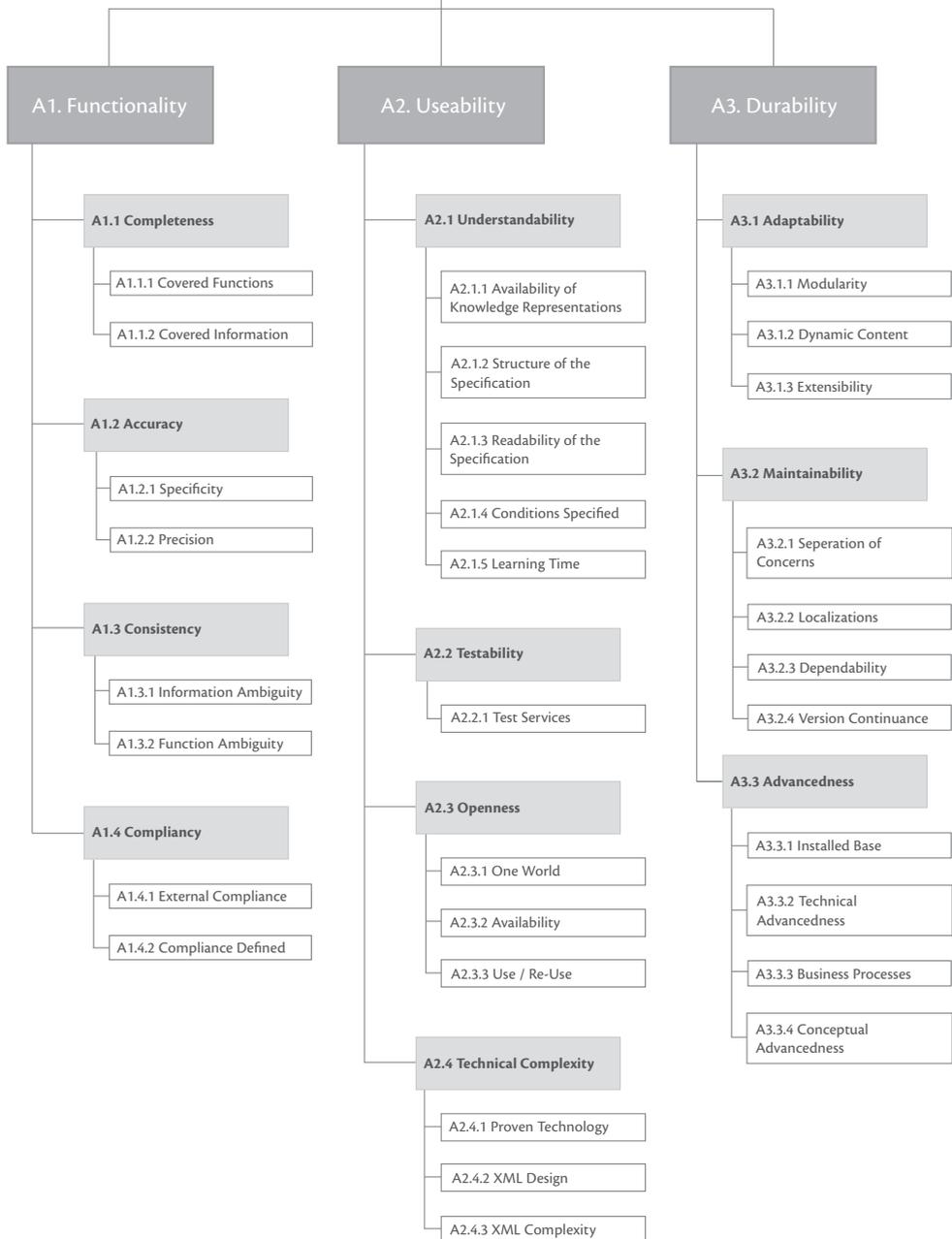


Figure 50 – Model for product quality

The definitions and some further explanation/remarks are presented in the following table. If the source of the definition is mentioned, it should be read as “originated from”, but the actual definition might be deferred.

Measurable Concept	Definition	Remarks
<b>A. Product Quality</b>	The total attributes of a standard that determine its ability to satisfy stated and implied needs when used under specified conditions. (ISO 9126)	This includes both internal and external quality in ISO terms.
<b>A1. Functionality</b>	The capability of the standard to provide functions which meet stated and implied needs when the standard is used under specified conditions. (ISO 9126)	The specification fulfills the functional needs of the intended job.
<b>A1.1 Completeness</b>	The extent to which a standard is of sufficient breadth, depth, and scope for the task at hand. (Wand & Wang, 1996)	This includes other terms like relevancy and suitability, and is the functional view on the content of the specification. The task at hand is aimed at solving an interoperability problem.
<b>A1.1.1 Covered Functions</b>	The level of functions specified in the specification in relation to the interoperability problem.	Indicates if the standard covers all functionality required to solve the interoperability problem.
<b>A1.1.2 Covered Information</b>	The specified level of information elements to support the interoperability problem.	When information elements are missing or when too many information elements have been added, it will have a negative impact on interoperability.
<b>A1.2 Accuracy</b>	The capability of the standard to provide true data with the needed degree of precision. (ISO 9126 & ISO 25012)	The level of needed specificity and precision in both semantic meaning and technical syntax. (This does not cover, but relates to, the quality of the content: consistency (A1.3))
<b>A1.2.1 Specificity</b>	The level of detail and in-depth of the scope.	Does the standard address a specific problem or a generic problem?
<b>A1.2.2 Precision</b>	The match between the unambiguously requested and provided precision. (ISO 25012)	Syntactic and semantic accuracy. (For example <i>surname</i> (instead of name, and not limited to 10 digits))
<b>A1.3 Consistency</b>	The extent of consistency in using the same values (vocabulary control) and elements to convey similar concepts and meaning in a standard. (Stvilia et al., 2007)	The degree of coherence and freedom of contradiction within the standard (ISO 25012). The quality of the content of the different models.
<b>A1.3.1 Information ambiguity</b>	The level of ambiguity of the information elements, and consistency of use.	The quality of the structuring and definition of the information elements.
<b>A1.3.2 Function ambiguity</b>	The level of ambiguity of the function elements and consistency of use.	The quality of the structuring and definition of the functions, processes and business rules.
<b>A1.4 Compliancy</b>	The capability of the standard to adhere to other standards, conventions or regulations in laws, but also to define what compliancy implies for this standard. (ISO 9126 & ISO 25012)	The compliancy when other standards are implemented, and how the conformance to this standard can be assured.
<b>A1.4.1 External compliance</b>	The compliance level to other standards, conventions, or regulations in laws and similar prescriptions.	Compliance with other standards at two levels: 1. Standards used to create this standard (e.g. UML). 2. Standards on different levels of interoperability (e.g. laws, or technical standards).

Measurable Concept	Definition	Remarks
<b>A1.4.2 Compliance defined</b>	The availability of a strict set of testable rules that define compliancy with the standard.	Is there a strict formulation if the implementation is to be conformant to the standard? This supports strict implementations.
<b>A2. Usability</b>	The capability of the standard has to be understood, learnt, used and attractive to the user, when used under specified conditions. (ISO 9126)	Also contains terms like implementability and readability, and is needed so as to estimate the efforts required for implementation.
<b>A2.1 Understandability</b>	The capability of the standard to enable the user to understand the standard for usage for particular tasks and conditions of use. (ISO 9126)	To enable it to be read and interpreted by users. (ISO 25012)
<b>A2.1.1 Availability of knowledge representations</b>	The level of available knowledge of the standard in different represented forms.	The availability of representations other than the specification, such as implementation guides, "how to's", "for dummies", training, etc.
<b>A2.1.2 Structure of the Specification</b>	The structure of the specification contains all needed and expected subjects in a logical manner. (Hyatt & Rosenberg, 1996)	Understandable and complete structure of the specification document.
<b>A2.1.3 Readability of the Specification</b>	To enable the standard to be read and interpreted by users. (ISO 25012)	The level of readability to the intended audience.
<b>A2.1.4 Conditions specified</b>	Both the required knowledge and abilities of the target audience are specified.	The tuning of the specification to its intended/target audience.
<b>A2.1.5 Learning Time</b>	The average time needed to understand the standard appropriately.	It is differentiated based on the actor. For an implementer it is the learning time necessary for implementation.
<b>A2.2 Testability</b>	The capability of the standard to be validated. (ISO 9126)	Intended to avoid faulty implementations.
<b>A2.2.1 Test services</b>	The availability of different kinds of test services.	E.g. validation service, helpdesk, test documentation, test procedure, etc.
<b>A2.3 Openness</b>	The implementation of open characteristics within the standard.	Includes "open specification".
<b>A2.3.1 One World</b>	One standard for one interoperability problem.	Are there (partly) competitive standards that may lead to variations of implementation for the same interoperability problem?
<b>A2.3.2 Availability</b>	The degree to which the standard is available for every user when and where he or she needs it. (Delen & Rijsenbrij, 1992)	Accessibility.
<b>A2.3.3 Use / Re-Use</b>	The possibilities to use and re-use the standard by implementers and end users, or others re-using it for standards development.	Intellectual Property Rights (e.g. patents) affect the use and re-use negatively.
<b>A2.4 Complexity</b>	The complexity of the technical solution proposed by the standard. (Stvilia et al, 2007)	The complexity of the technical representation of the standard.
<b>A2.4.1 Proven Technology</b>	The complexity of the technology used.	Has the most appropriate technology been selected?
<b>A2.4.2 XML Design</b>	The complexity of the XML Design.	Compliance to standards for XML design.
<b>A2.4.3 XML Complexity</b>	The complexity of the XML Schema.	The XML Schema characteristics.
<b>A3. Durability</b>	The capability of the standard to have a long lasting life cycle; Connecting to current situations and future proofs. (Garvin, 1984)	Includes concepts like changeability and flexibility.

<b>A3.1 Adaptability</b>	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 considered standard. (ISO 9126)	Includes customizability.
<b>A3.1.1 Modularity</b>	The logical structured design of the standard.	Re-use within the (content part of the) standard.
<b>A3.1.2 Dynamic Content</b>	The capability of the standard to deal with dynamic, often changing content, without resulting in version explosion.	Keeping a stable version of the standard, while implementing flexible structures to deal with dynamic content. (Product identification codes is an example of dynamic content)
<b>A3.1.3 Extensibility</b>	The extent to which a standard provides possibilities to extend the capabilities without affecting other parts of the implementation. (O'Brien et al., 2005)	Includes scalability, and forward compatibility, to support innovation.
<b>A3.2 Maintainability</b>	The capability of the standard to be modified in an efficient manner. (ISO 9126)	Modifications may include corrections, improvements or adaptation to changes in the environment, requirements or functional specifications. This also covers re-usability, replaceability and co-existence.
<b>A3.2.1 Separation of Concerns</b>	The level of separation of parts that are different in nature.	Separation of technology dependent and technology independent parts. But also separation between content addressing business people or technical people.
<b>A3.2.2 Localizations</b>	The availability and use of localizations and extensions in the main standard.	Several standards need localization for specific business/country use.
<b>A3.2.3 Dependability</b>	The dependences on other standards.	Co-existence with other standards, and optimal re-use of existing standards. (e.g. when the standard is a profile on top of other standards)
<b>A3.2.4 Version continuance</b>	New versions of the standard show functional and data continuance. (ISO 9126)	New versions support the same functions. Backwards compatible when possible.
<b>A3.3 Advancedness</b>	The state of the art notion of the standard, in relation to current practice and future innovation.	Advancedness in relation to the outside environment in stakeholder organizations. Currentness (ISO 25012): Fits the timeframe.
<b>A3.3.1 Installed Base</b>	The connection of the standard to the current ICT landscape in stakeholder organizations.	In line with the installed base.
<b>A3.3.2 Technical Advancedness</b>	The position of the chosen technology in its life cycle.	The maturity of the technology used.
<b>A3.3.3 Business Processes</b>	The connection of the standard with the current business processes of the stakeholders organizations.	Commodity level of business processes. (e.g. the standard supports only self-billing, while within the domain this is highly uncommon)
<b>A3.3.4 Conceptual Advancedness</b>	The expected life span of the conceptual solution.	The complexity of the solution fits the complexity of the interoperability problem on a conceptual level. And does it fit the foreseen developments within the domain, or is it way ahead of its time (e.g. automated plug and play e-business in 2001).

Table 42 – Measurable concepts defined for product quality

11.2.2 Process quality

The second branch of the overall quality model tree is process quality. This part has been added in the final build and is therefore less mature and has not been extensively researched. The results are the first step for further development.

The main information needs related to process quality are:

1. Is the Development & Maintenance process professionally organized? (Development & Maintenance)
2. How is the standard presented to the outside world? (Communication)
3. How capable is the organizational structure in relation to the standard? (Organization)

Figure 51 captures the model for process quality.

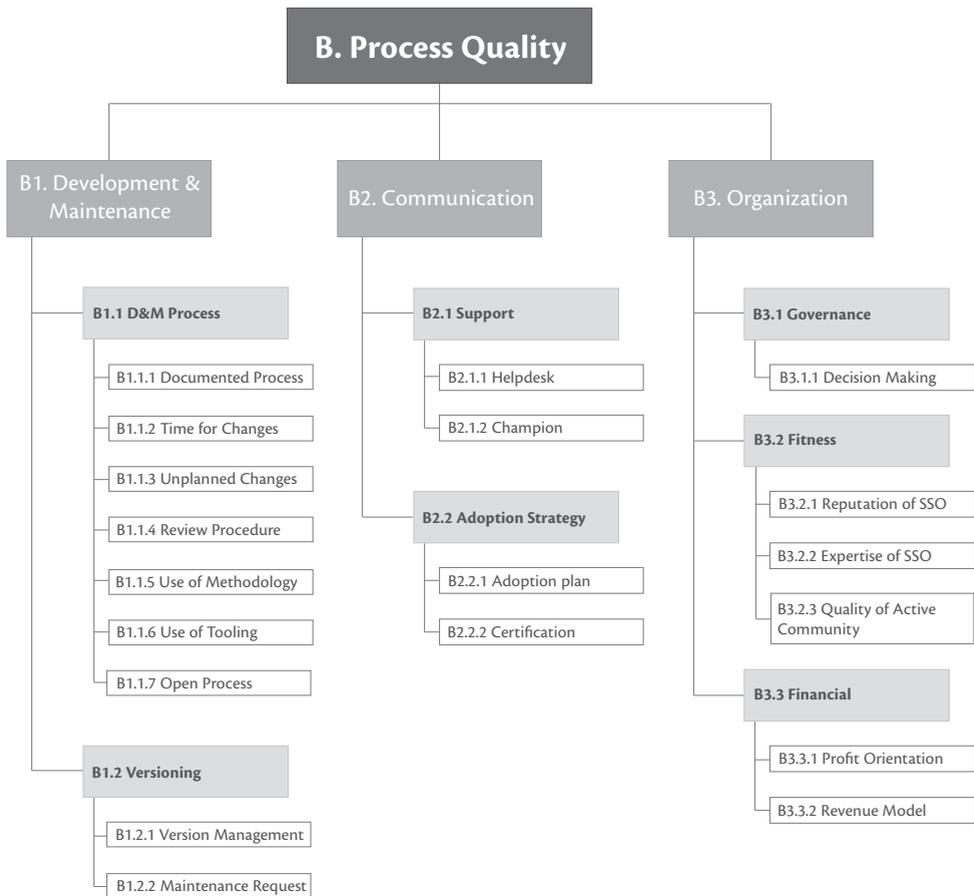


Figure 51 – Model for process quality

All measurable concepts have been defined in the following table, including remarks and directions as an aid when looking for potential measures.

Measurable Concept	Definition	Remarks/Synonyms/Direction for measures
<b>B. Process Quality</b>	The quality of the sequence of interdependent and linked procedures that resulted in the standard.	Organizational quality.
<b>B1. Development &amp; Maintenance</b>	The professionalism of how the standard development and the maintenance are organized.	This concept is based on BOMOS. (Folmer & Punter, 2011)
<b>B1.1 D&amp;M Process</b>	The capability of the D&M process to suit a standardization environment.	The D&M should guide developers in achieving quality.
<b>B1.1.1 Documented Process</b>	The availability of a document describing the development and maintenance process.	Including both the initial development approach and change procedure. A patent check should be part of the process, as well as a signed IPR statement of workgroup participants.
<b>B1.1.2 Time for Changes</b>	The time needed for changes to take place, beginning with the status of the new maintenance request until the release of the standard.	Check the version history of the standard, or the overview of the maintenance request history.
<b>B1.1.3 Unplanned Changes</b>	The ability to release bug fixes that solves major errors within the standard.	In what way is the process agile enough to enable a quick next version of the standard in which the bug is fixed? Are bug fixes released? How long after the bug was identified?
<b>B1.1.4 Review Procedure</b>	The presence of a thorough and documented review process.	Aspects that need to be measured are the number of review cycles and the number of reviewers. Different types of stakeholders as reviewers? Public review round? Passive reviews or active (like testing) reviews?
<b>B1.1.5 Use of Methodology</b>	The usage of a methodology, including language, within the process.	A proper methodology will lead to models that are maintainable. Is there an explicit choice made for the methodology?
<b>B1.1.6 Use of Tooling</b>	The usage of advanced tooling within the process.	Automation reduces errors. List the tools that are normally used in the development process.
<b>B1.1.7 Open Process</b>	The extent to which the D&M process is organized in openness, i.e. it is accessible for all.	The absence of stakeholder groups will have an impact. Open access means that no stakeholder group is excluded. Different rates for different stakeholders? (e.g. academia and SME's have lower participation fees than large industry players). Open meeting: Are the meeting locations accessible? Telco's and e-mail are used to limit meetings/costs? Is the calendar published? (Lammers et al., 2010)
<b>B1.2 Versioning</b>	The capability of the standard to have versioning in place that combines stability and the possibility to make changes.	Explicit version management is required to raise the appropriate expectations.
<b>B1.2.1 Version Management</b>	The organization and procedures of version numbering and version management, also in relation to backwards compatibility.	Documented and published policy on version management including an approach for version numbering, major and minor releases, for example with the numbering. What is the trigger to start a new version? Based on the number of maintenance requests? Or releases are limited in number on time-basis? Is there tool support to deal with version management and maintenance requests? Is backwards compatibility guaranteed for a certain period?
<b>B1.2.2 Maintenance Requests</b>	The ability within the D&M process to deal with maintenance requests.	Is an overview of maintenance requests present on the website? Are the MRs traceable, including history? How many Maintenance Requests, and how often are these processed, and how often do these lead to changes? Are all stakeholders invited to submit MRs?

Measurable Concept	Definition	Remarks/Synonyms/Direction for measures
<b>B2. Communication</b>	The totality of communication activities related to the standard.	The presentation of the standard to the outside world.
<b>B2.1 Support</b>	The availability of knowledgeable support.	Support aimed at helping stakeholders with (the choice of) implementing the standard.
<b>B2.1.1 Helpdesk</b>	The availability of knowledgeable support from the organization maintaining the standard.	It will improve implementations, and valuable feedback on maintenance requests. Is there a helpdesk, forum, or any other means for asking questions and receiving appropriate answers? How many channels (telephone, e-mail, face-to-face, Internet) are available?
<b>B2.1.2 Champion</b>	The availability of the ambassador of the standard, who enthusiastically promotes the standards and helps when issues arise.	This aspect might be related to the success of the standard. Is there a person or group that acts as a focal point for the standard?
<b>B2.2 Adoption Strategy</b>	The availability of a promotion strategy to strive for successful adoption in practice.	Adoption will not go automatically but needs to be striven for.
<b>B2.2.1 Adoption Plan</b>	The availability of a plan to raise the number of implementations.	Is there a plan that contains a broad range of adoption activities that are carried out accordingly? Has it been set up based on a structured approach on adoption, by using, for example an adoption instrument? Measure the number of communication means, from financial to legal means, intended to stimulate the use of the standard.
<b>B2.2.2 Certification</b>	The availability of a certification program aimed at adoption.	Is there a certification program? For which stakeholders? What is being certified? Is it aimed at adoption? How does it deal with new versions?
<b>B3. Organization</b>	The capability of the organizational structure to effectively set standards.	Perhaps there is a network of organizations involved.
<b>B3.1 Governance</b>	The organization of decision making within the governance.	Relates to the openness of the standard.
<b>B3.1.1 Decision Making</b>	The organizational structure of decision making, including the way decisions are made.	Which groups decide what and how? (e.g. What decisions are made by the board? What decisions by the workgroup?) How is the decision made? (e.g. consensus/majority/different weights) Does it discriminate groups of stakeholders for instance based on payments? Minimum number of votes, etc. Does it reflect the stakes of the actors involved?
<b>B3.2 Fitness</b>	The suitability of the development organization for the job.	The capability of the organization to support the standard appropriately.
<b>B3.2.1 Reputation of SSO</b>	The reputation of the development organization in the field.	Is it independent, does it reflect the stakeholders? Is it committed? Is it known? Is it trusted? How many years has it existed? Are stakeholders or members participating? Manages more standards?
<b>B3.2.2 Expertise of SSO</b>	The knowledge of the experts of the standard development organization in the area of standardization.	Standardization knowledge is needed, not only on standardization procedures, but also on content (how to develop information models, data definitions, transformation to technology). Is standardization knowledge available both on procedures as well as content aspects?

<b>B3.2.3 Quality of Active Community</b>	The knowledge that is present within the active community should be representative of the problem domain.	Domain knowledge is essential; is it presented within the active community? And are they committed?
<b>B3.3 Financial</b>	The capabilities to be financially neutral and stable for years to come.	Relates to the openness of the standard, and the adoption potential.
<b>B3.3.1 Profit Orientation</b>	The profit-orientation of the standard development organization.	Not for profit fits best.
<b>B3.3.2 Revenue Model</b>	The presence of an appropriate revenue model for acquiring a budget for development and maintenance.	An inappropriate model will temper adoption and might limit maintenance. Is it stable? What kind of income is gathered? (government funding, project funding from participants, membership fees, provision of services, licenses, donations)

*Table 43 – Measurable concepts defined for process quality*

### 11.2.3 Quality in practice

Just like process quality, this area has not been extensively researched: the results are therefore based on a first build iteration, and should be developed further in future iterations. Quality in practice deals with the environment of the standard, its potential and actual use by the stakeholders. Two main information needs are present:

1. Is the standard accepted as solution in practice?
2. Will the standard lead to interoperability in practice?

Based on these two information needs, and the research approach as described in section 11.1, the quality model was constructed and is presented in Figure 52.

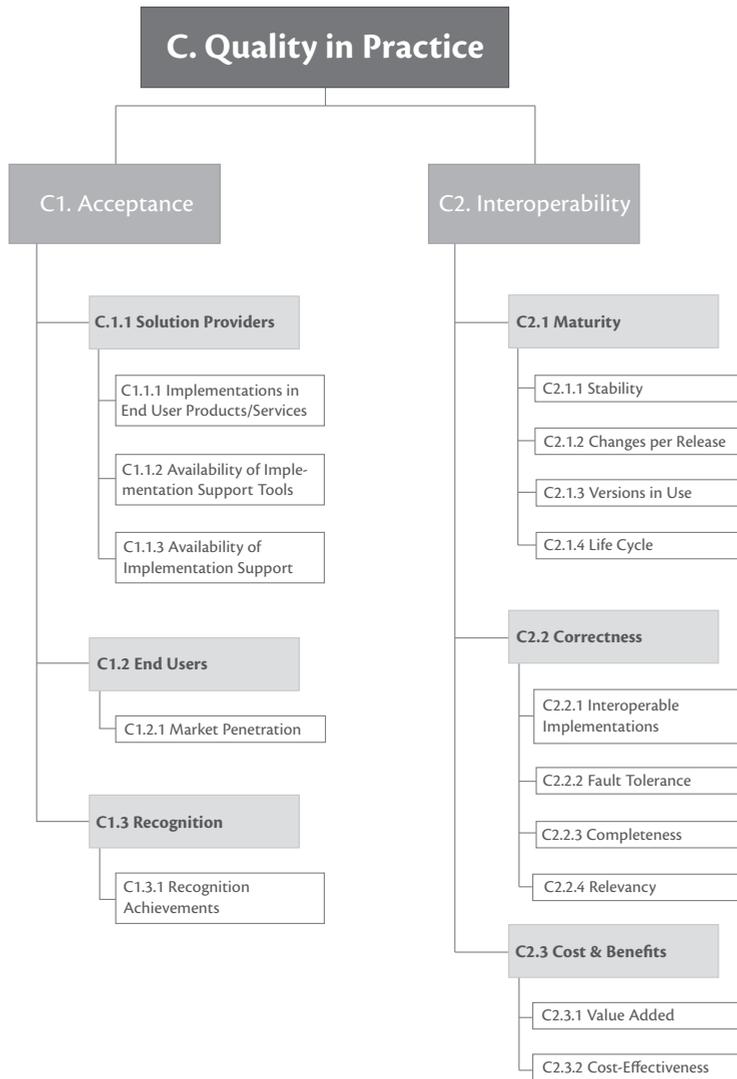


Figure 52 – Model for quality in practice

Again all the measurable concepts have been defined in the following table, including remarks and directions as an aid when looking for potential measures.

Measurable Concept	Definition	Remarks/Synonyms
<b>C. Quality in Practice</b>	The extent to which a standard can be used by specified users to meet their needs to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. (ISO 14598)	Quality in use.
<b>C1. Acceptance</b>	The adoption of the standard within the domain.	Adoption/acceptance in practice.
<b>C1.1 Solution Providers</b>	The extent to which solutions providers have adopted the standard.	Solution providers provide products and service that are used by the end users. The adoption by solution providers is a multiplier for adoption.
<b>C1.1.1 Implementations in End User Products/ Services</b>	The level of implementations in the products and services offered by solution providers and to be used by end users.	Is it possible for end users to use the standards by using products off the shelf provided by solution providers? Check the top 5 of products and service in the market on their standard usage.
<b>C1.1.2 Availability of Implementation Support Tools</b>	The availability of tools and components that can be used to simplify implementations of the standard.	Check if tools or components are available, for example open source components.
<b>C1.1.3 Availability of Implementation Support</b>	The availability of consultants, and implementation partners as a support for implementation.	The expertise available outside the SSO. How many organizations provide consultancy on the standard?
<b>C1.2 End Users</b>	The extent to which the end users have adopted the standard.	The actual use of the standard.
<b>C1.2.1 Market Penetration</b>	The level of usage within the intended end user audience.	What is the percentage of use? Of total organizations? Of total transactions? Differentiated by different user groups.
<b>C1.3 Recognition</b>	The extent to which the standards receive external recognition.	The credibility.
<b>C1.3.1 Recognition Achievements</b>	The external formal recognition of the standard.	Both external status and reputation in the domain. Is the standard formally acclaimed (eg. ISO status) Is the standard acclaimed by governments? (e.g. comply or explain list in the Netherlands) But does it also measure the fame/reputation of the standard? Is it above or under par? (Chase, 1995)
<b>C2. Interoperability</b>	The ability of two or more systems or components to exchange information and to use the information that has been exchanged. (Legner & Lebreton, 2007)	The capability of the standard to achieve meaningful communication between systems.
<b>C2.1 Maturity</b>	The capability of the standard to be a stable and proven solution.	A mature standard will positively influence interoperability.
<b>C2.1.1 Stability</b>	A stable release schema means ample time in between releases.	Count the number of versions within several years. Too many versions within a short time will temper both adoption and interoperability. A maximum of one major release per year is recommended. Is there a fixed release process (eg. a new version will become available yearly on Jan. 1 <sup>st</sup> )?
<b>C2.1.2 Changes per Release</b>	The number of changes that have been made to the standard.	Count the number of changes per release. Too many changes might indicate that current quality is low and might have an impact on interoperability. Calculate the adaptations needed in software implementations based on function points for a new version. How old is the oldest version of the standard in use? And how many versions in between the current version?

Measurable Concept	Definition	Remarks/Synonyms
<b>C2.1.3 Versions in Use</b>	The number of versions that are concurrently in use.	Too many versions will tamper interoperability. Count the number of versions that are currently used in practice.
<b>C2.1.4 Life Cycle</b>	The stage of the standard in the life cycle model of standards, but also inclusion of the timing in the market.	Both the life cycle of the standard as well its timing in the market (it might arrive too early/late). What is the life cycle status? (According to a model like the Gartner Hype Cycle, or a specific model for standards' life cycle: creation, fixes (changes), maintenance (changes), availability (no changes), rescission. (Krechmer, 2006) or the extended life cycle model (Söderström, 2004)) And what is the timing in the market? Is another solution already present?
<b>C2.2 Correctness</b>	Extent to which an implementation of a standard satisfies its specifications and fulfills the user's mission objectives. (McCall et al., 1977)	Also called "Free of Error".
<b>C2.2.1 Interoperable Implementation</b>	The level of interoperability that is achieved by the current implementations.	The level at which information exchange between organization is improved in terms of fewer faults, better correct interaction and understanding, or easier achieved. Also can exchange various implementations of the standard and check for correctness, by, for instance having plugfests.
<b>C2.2.2 Fault Tolerance</b>	The degree to which the standard supports free of disturbances when minor deviations occur. (Delen & Rijsenbrij, 1992)	Also robustness, performance, reliability or recoverability. If the standard is too sensitive for gold plated perfect use, it will lead to problems in practice. Check the complaints of other implementers.
<b>C2.2.3 Completeness Elements</b>	The extent to which the data standard specifies all the data elements needed by the standard user. (Zhu & Wu, 2011)	If many custom added elements are used it might indicate that the standard is not complete, especially when multiple implementations are using the same custom added elements. Count the number of custom added elements in relation to the used elements.
<b>C2.2.4 Relevancy Elements</b>	The extent to which the data of the standard only specifies the data elements needed by users of the standard. (Zhu & Wu, 2011)	If many elements are not used, then these are also not relevant, especially if the same elements are not used in multiple implementations. Count the number of used information elements in relation to the number of available elements.
<b>C2.3 Cost &amp; Benefits</b>	The extent to which the benefits cover the costs of standardization.	Although not necessary for all stakeholders, the total use of the standard should have a positive business case.
<b>C2.3.1 Value Added</b>	The extent to which a standard is beneficial and provides advantages when used. (Kahn, Strong & Wang, 2002)	Are expectations matched in practice? Are net benefits achieved? Are there: Cost savings? Expanded Markets? Incremental additional sales? Reduced Search Costs? Time Savings? (Sedera & Gable, 2004)
<b>C2.3.2 Cost-Effectiveness</b>	The extent to which the cost of collecting appropriate knowledge and implementing the standard is reasonable. (Wand & Wang, 1996)	What, if any, is the multiplier on the investment of implementing the standard? What is the effect (cost, loss of profit) of NOT implementing the standard?

Table 44 – Measurable concepts defined for quality in practice

### 11.3 The set of measures for product quality

To have a set of validated measures for the complete quality model would be the ideal setting, but it is currently not available. However during a workshop with experts many potential measures were gathered. These are highly valuable, but should be used with care since they have not been extensively validated.

The indicative measures are presented in the following table.

Measurable Concept	Measures	Function
<b>A1.1.1 Covered Functions</b>	<ol style="list-style-type: none"> <li>1. Functions required and specified</li> <li>2. Requirements specification</li> <li>3. User requests</li> <li>4. Not covered</li> </ol>	<ol style="list-style-type: none"> <li>1. Measure the number of functions requested in the interoperability problem description (the requirements specification), and compare this to the number of functions covered within the standard. (If more functions are covered than requested, this indicates that compactness is low.)</li> <li>2. Measure if the standard requirements have been fully specified. (Not specified at all, only at a high level, not completely, the specification covers more than the requirements.)</li> <li>3. Measure the number of maintenance requests (MR) related to added functionality (e.g. Number of MRs in certain periods, or the number of MRs related to functionality divided by the total number of MRs.)</li> <li>4. Measure if the functionality that is NOT covered, is explicitly mentioned.</li> </ol>
<b>A1.1.2 Covered Information</b>	<ol style="list-style-type: none"> <li>1. Necessity of elements</li> <li>2. Elements specified and needed</li> <li>3. Mandatory elements available to the actor</li> <li>4. Free elements</li> </ol>	<ol style="list-style-type: none"> <li>1. Assess the number of elements that are not mandatory and are just nice to have. (The optional features.)</li> <li>2. The number of elements specified in the standard divided by the number of essential elements. (A score below 1 is problematic, while a score much higher than 1 is not positive either.)</li> <li>3. Measure if there are mandatory data elements that might be difficult to capture by the responsible actor.</li> <li>4. Assess the necessity to use user defined elements. (e.g. There are no user defined elements available; or the standard specifies the use of user defined elements.)</li> </ol>
<b>A1.2.1 Specificity</b>	<ol style="list-style-type: none"> <li>1. SMART</li> <li>2. Specific requirements</li> <li>3. Problem specificity</li> <li>4. Distinctive naming</li> </ol>	<ol style="list-style-type: none"> <li>1. Is the problem stated in a SMART way?</li> <li>2. Assess the availability of specific requirements formulated for testing the specification. (To check the specification against the requirements.)</li> <li>3. Assess the problem: Is it one problem or multiple?</li> <li>4. Assess the naming of functions and elements on distinctive, specific names and the absence of terms that are open to interpretation.</li> </ol>
<b>A1.2.2 Precision</b>	<ol style="list-style-type: none"> <li>1. Semantics precision</li> <li>2. Syntax precision</li> </ol>	<ol style="list-style-type: none"> <li>1. Assess the naming of semantic elements on self-contained precision. (e.g. Surname, first name, name.)</li> <li>2. Assess the syntax on precision, field lengths, appropriate data types. (Integer, string, etc.)</li> </ol>
<b>A1.3.1 Information Ambiguity</b>	<ol style="list-style-type: none"> <li>1. Vocabularies</li> <li>2. Definition ambiguity</li> <li>3. Business language</li> <li>4. Coherence and Contradiction</li> </ol>	<ol style="list-style-type: none"> <li>1. The extent to which vocabularies / code lists are used.</li> <li>2. Check definitions on ambiguity, whether they are self-explanatory, and the use of variation in wording.</li> <li>3. Assess if terms and definitions are in line with business vocabulary used in practice, for instance by validating them with business people.</li> <li>4. Assess the specification on overlap and contradictions in element definitions. (e.g. Are elements being reused to overcome overlaps?)</li> </ol>

Measurable Concept	Measures	Function
<b>A1.3.2 Function Ambiguity</b>	<ol style="list-style-type: none"> <li>1. Function coherence</li> <li>2. Process input/output</li> <li>3. Business rules</li> </ol>	<ol style="list-style-type: none"> <li>1. Assess the functions, processes and business rules on logical structuring, strict definitions in recognizable business language.</li> <li>2. Assess if the process models have clearly defined input, output, triggers, pre and post conditions.</li> <li>3. Assess the quality of the business rules on compliance to a formal business rule language, e.g. SBVR.</li> </ol>
<b>A1.4.1 External Compliance</b>	<ol style="list-style-type: none"> <li>1. Defined compliancy</li> <li>2. Logical set</li> <li>3. Missing standards</li> </ol>	<ol style="list-style-type: none"> <li>1. Is there a definition of compliancy with other standards? Including regulations and laws?</li> <li>2. Is it a logical and coherent set of standards? (e.g. If XML, SOAP, etc. the complete set of web services standards)</li> <li>3. Are there other standards or laws for compliancy that are expected but are not mentioned?</li> </ol>
<b>A1.4.2 Compliance Defined</b>	<ol style="list-style-type: none"> <li>1. Conformance rules</li> <li>2. Testable</li> </ol>	<ol style="list-style-type: none"> <li>1. Assess if strict conformance rules are specified, like a minimum set that needs to be implemented. Assess if they are part of the standard?</li> <li>2. Assess the conformance rules on reliable testability.</li> </ol>
<b>A2.1.1 Availability of Knowledge Representations</b>	<ol style="list-style-type: none"> <li>1. Number of representations</li> <li>2. Reference implementation</li> </ol>	<ol style="list-style-type: none"> <li>1. Count the number of representation forms from a predefined list. (e.g. Specification, how to guidelines, online training, helpdesk, FAQ, reference examples, etc.)</li> <li>2. Assess the availability of reference implementations supporting the compliancy rules.</li> </ol>
<b>A2.1.2 Structure of Specification</b>	<ol style="list-style-type: none"> <li>1. Template based</li> <li>2. Essential parts</li> <li>3. MOSCOW</li> <li>4. Layered structure</li> <li>5. Links</li> </ol>	<ol style="list-style-type: none"> <li>1. Is a template used for structure?</li> <li>2. Assess the contents on the availability of essential parts like a conformance section, problem description, scope, etc.</li> <li>3. Assess if formal logic and structured language, e.g. "must have" (MOSCOW), etc., is used.</li> <li>4. Assess the contents on the use of a logical layered structure (like separation of business processes/semantics/technology).</li> <li>5. Assess if the specification uses links for easy navigation through the document.</li> </ol>
<b>A2.1.3 Readability Specification</b>	<ol style="list-style-type: none"> <li>1. Easy reading index</li> <li>2. Carelessness</li> <li>3. Linguistic indicator</li> <li>4. Consistent language</li> <li>5. Domain knowledge indicator</li> </ol>	<ol style="list-style-type: none"> <li>1. Apply well known indexes to the specification document (and exclude the XML parts). (e.g. Gunning Fog Index or Flesh Kincaid Reading Ease.)</li> <li>2. Assess the carelessness: no spell checking, open ends, etc.</li> <li>3. The linguistic quality of the English language used in the standard. (e.g. Non-native English writers, translations from other languages.)</li> <li>4. Is the structured language (e.g. MOSCOW), and the naming of actors/roles, consistently used?</li> <li>5. Assess the number of words that are not understandable for people without domain knowledge.</li> </ol>
<b>A2.1.4 Conditions Specified</b>	<ol style="list-style-type: none"> <li>1. Target group</li> <li>2. Knowledge &amp; abilities</li> <li>3. Reader specific parts</li> <li>4. Referencing</li> </ol>	<ol style="list-style-type: none"> <li>1. Is the target group specified?</li> <li>2. Assess if the pre-requisites on knowledge and abilities are defined.</li> <li>3. Assess if selective parts are aimed at selective target groups and whether this is clearly marked.</li> <li>4. Assess if the standard is sufficiently self-contained, and references are made to other documents and standards.</li> </ol>
<b>A2.1.5 Learning Time</b>	<ol style="list-style-type: none"> <li>1. Learning time per actor</li> <li>2. Search time</li> <li>3. Total number of pages</li> </ol>	<ol style="list-style-type: none"> <li>1. Assess the learning time by asking different actors about their learning time.</li> <li>2. Sample test the search time for some random aspects related to the standard. (Or assess the difference in search time between an expert and beginner: small differences reflect a short learning time.)</li> <li>3. Count the total number of pages of the main content of the specification, excluding white pages or irrelevant appendices.</li> </ol>

<b>A2.2.1 Test Services</b>	1. Test services availability	1. Assess the number and quality of available options for testing, including validation services, on amongst others syntactic, semantic, process levels, helpdesk, test documentation, test procedure, certification, etc.
<b>A2.3.1 One World</b>	1. Competitive standards	1. Assess if there are competitive standards, also on parts. The impact depends on the lifecycle status of the standard.
<b>A2.3.2 Availability</b>	1. Freely available 2. Burdens 3. Free information	1. Is it possible to download the specification without burden? (No costs, no registration?) 2. Assess the burdens (membership, registration, costs, etc.) necessary to receive all information (not only the specification) required to be able to use the standard. 3. Assess the amount of freely available information apart from the specification, like FAQ, tooling, examples, guidelines, course material, etc.
<b>A2.3.3. Use / Re-Use</b>	1. License	1. Assess if the license is specified and what restrictions for use and re-use are applicable. Anything less than royalty-free affects negatively.
<b>A2.4.1 Proven Technology</b>	1. Proven 2. Market acceptance	1. Assess if the chosen technology is seen as an appropriate solution by experts. 2. Assess if the technology is accepted by the market and won't lead to additional costs because of a limited choice in products and vendors.
<b>A2.4.2 XML Design</b>	1. NDR 2. Conform W3C 3. Core components	1. Assess if Naming and Design Rules are correctly implemented. 2. Assess conformance to W3C specifications. (XML Schema, Formatting and Namespace rules.) 3. Assess if core components, or any other implementation of re-usable components, are used for data elements.
<b>A2.4.3 XML Complexity</b>	1. XML complexity  2. Tags defined	1. Assess: a. the number of: Complex Type Declarations, Simple Type Declarations, Annotations, Derived Complex Types, Global Type Declarations, Global Type References and Unbounded Elements. b. the average: Number of Attributes per Complex Type Declaration, Bounded Element Multiplicity Size, Number of Restrictions per Simple Type Declaration and Element Fanning. 2. Check if all tags are defined within a data dictionary.
<b>A3.1.1 Modularity</b>	1. Data dictionary 2. Re-use of components 3. Independence of modules	1. Assess if a data dictionary is available, and if it can be used in different standard modules. 2. Assess if there is re-use of components, building blocks, within the standard. 3. Assess the possibilities of using certain modules of a standard independently.
<b>A3.1.2 Dynamic Content</b>	1. Dynamic content	1. Assess if code lists or other solutions are used for dynamic content resulting in a separation between dynamic content that is prone to changes and a standard core that proposes a stable representation of the business domain.
<b>A3.1.3 Extensibility</b>	1. Extension mechanism	1. Assess if there is an extension mechanism, such as additional data elements, with rules on how to use them.
<b>A3.2.1 Separation of Concerns</b>	1. Separation of concerns	1. Is separation of concerns applied within the standard (according to a defined policy)?
<b>A3.2.2 Localizations</b>	1. Option of localization 2. Localization rules	1. Are localization / profiles for specific purposes intended and are guidelines provided as to when to use localizations? 2. Is there a strict policy of defining the contents of localizations?

Measurable Concept	Measures	Function
<b>A3.2.3 Dependability</b>	1. Dependencies	1. What kind of strong dependencies exist, or are they expected, regarding other standards. A strong dependence means that a version update of the other standard has an impact on this standard. Especially focus on generalization-specialization relations; Is the dependency logical and managed?
<b>A3.2.4 Version Continuance</b>	1. Version check	1. Only assess the content of different versions, based on version history, and assess whether there is logic in the content changes, including backwards compatibility.
<b>A3.3.1 Installed Base</b>	1. Current ICT	1. How difficult is integration with existing systems (e.g. SAP) in stakeholders' organizations.
<b>A3.3.2 Technical Advancedness</b>	1. Lifecycle position 2. Competing technology	1. Assess the position of the used technology on hype cycles. (e.g. The Gartner Hype Cycle on Technology.) 2. Assess the competing technology. Is the choice of technology justifiable?
<b>A3.3.3 Business Processes</b>	1. Commodity	1. Assess proposed business processes and compare them to the standardized business processes commonly used in practice.
<b>A3.3.4 Conceptual Advancedness</b>	1. Conceptual complexity 2. Future fit	1. Assess the fit between the complexity of the solution and the complexity of the interoperability problem. Is it overkill? 2. What is the fit with future views on the interoperability problem domain?

Table 45 – Measures for product quality

## 11.4 The iQMSS usage process

This section describes how to use the iQMSS; the essential steps are explained. The use case diagram of the intended use of the instrument is presented in Chapter 8. This model shows that four different actors (defined within Chapter 8) are involved. These four actors are involved in three main phases: 1. Preparation, 2. Measuring and 3. Analysis. This section will describe each of those phases in more detail.

### 11.4.1 Preparation

First of all, when starting a project, all roles should be covered by people: in practice the client and the standard developer might be one person and the same holds for the initiator and principal actors. The initiator and principal should be familiar with the iQMSS and should have read the documentation (this document). The client should have a real information need regarding the quality of the standard, while the standard developer should have access to detailed information related to the standard. These are the pre-conditions to start the following steps.

#### A. Customize quality measurement

*Goal:* Define the information need of the client and select applicable parts from the quality model.

*Input:* The quality model.

*Output:* Description of the information needs of the Client, and the corresponding selection of measurable concepts from the quality model: the customized QMSS

*Activities:* This part contains two steps:

##### A1. Plan a meeting with the client to discuss the information needs (initiator & client)

The elicitation of the information needs. It might be helpful to use this list of possible choices, although the information needs should not be limited to these options:

- A. The client wants to know the total product quality?
- B. The client wants to know the appropriateness of functionality?
- C. The client wants to know the usability of the standard? (Is the specification implementable and can it be used without burden?)
- D. The client wants to know the durability of the standard? (Will it hold in future?)

As well as the determination of information needs some other background related questions are relevant as well:

1. What is the reason to perform the quality assessment? (e.g. There are complaints, or an improvement project has been started, etc.)
2. Are quality issues expected: in what areas?
3. Are there areas that need special attention?
4. Are there any other content related requirements for the analysis output?

The meeting can also be used to discuss process issues like the actual time needed, throughput time, when it should be ready, the budget, to whom the results should be presented, etc.

#### A2. Map the measurable concepts to the information needs (initiator)

Based on the results of the meeting, the initiator can map the information needs on the quality model and select the relevant parts needed for measurement, and this customizes the quality model.

When needed or requested the initiator might also set up a project plan. It is advisable to that the client approves the results of this phase.

### **B. Customize SSM**

This second part will change the focus of the subject from the requirements to that of the semantic standard.

*Goal:* Define the model of the semantic standard as the subject, to get a better understanding of the subject.

*Input:* The SSM.

*Output:* Customized SSM for the specific standard that is the subject of the quality measurement. And the main sources of information necessary for the measurement.

*Activities:* This part has two steps:

#### B1. Gather and analyze information about the standard (initiator)

Normally the initiator will not be familiar with the standard, and should start gathering information about the standard. All the gathered information should be documented and analyzed on relevance.

#### B2. Customize the model of the semantic standards (initiator & client)

To get a better understanding of the standard and to know what will be measurable or not, the model of the semantic standard should be used as guidance. Some parts might not be relevant because of the chosen information need or the standard does not cover certain parts of the model. The customization of the semantic standard model is therefore a subset of the basis of the selection which in turn is based on the selected information need and coverage of the standard. However the context part of the model should always be selected as a part of the customization.

### **C. Customize the instrument QMSS**

Based on the output of both previous steps we know what the measurable concepts are, as they are related to the information needs and we have a basic understanding of the measurable semantic standard. This knowledge should be combined by means of the following steps.

*Goal:* To define the measures as part of the QMSS and when applicable define the instruments to be used in the measurement.

*Input:* Customized QMSS and customized SSM.

*Output:* Set of measures (complete customized QMSS) and instruments (complete customized iQMSS).

*Activities:* This part has two steps:

#### C1. Set up the measures (initiator)

The initiator has to select the measures for each measurable concept. The mentioned measures which are part of the QMSS product quality can be used as an inspiration. However stringent use of these measures is not advisable: select the ones that are applicable and measurable for this standard, and look for other measures, for instance domain specific measures (Mykkanen & Tuomainen, 2008).

#### C2. Set up instrumentalization (initiator)

Instruments might be rather useful for measuring efficiently. However this build does not contain recommended instruments like, for instance, templates or software/Internet tooling. But it is worthwhile to have a broader look for instruments. As an example: the availability of several websites (instruments) to measure reading ease (Flesh-Kincaid, Gunning Fog, etc.) might be particularly useful. For other measures it might be worthwhile to set up a web-based survey to deal with multiple opinions.

### **The end of the preparation**

The preparation phase should end with a validation of the results so far, involving the client. Applying the instrument without any kind of specific case validation might be dangerous (Nyeck, Morales, Ladhari, & Pons, 2002). An enhanced validation might involve the use of a focus group discussion.

#### *11.4.2 Measuring*

After the preparation it is time to start with the actual measurement, which is captured in one main step:

### **D. Measure the quality of a semantic standard**

*Goal:* To collect the measurement results.

*Input:* Complete customized iQMSS.

*Output:* Measuring results and observation notes.

*Activities:* First start with an exploration of the requirements that were used to develop the standard. Then two steps will be carried out many times in iterations for each measurable concept:

#### D1. Apply the quality model: what is to be measured? (principal)

#### D2. Find the information: to score the measure (standard developer)

It is preferable that the work is carried out in sessions by the principal and the standard developer. Write down the score and other remarks. It is expected that some scores might not be easily retrievable during the work session, these will be homework.

### **The end of measuring**

It is advisable to end the measuring phase by letting the standard developer validate the measuring results.

### 11.4.3 Analysis

The project will end with the delivery of the final product, in which the previous results are used to answer the client's information needs. It consists of one main step:

#### **E. Analysis of the results**

*Goal:* To present the report that fulfills the need of the client.

*Input:* The measuring of the results and observation notes.

*Output:* Report and end evaluation.

*Activities:* Three steps need to be carried out:

##### E1. Set up a report (principal)

Start with an interpretation of the results of the previous phase. Based on the interpreted results, define any actions that might be taken, they will be part of the recommendations. Then continue with setting up the report, and start by clearly presenting the client's information needs in the report. Continue with describing the approach taken, and present issues encountered and how these have been dealt with. Then present the most remarkable results related to the information need, and end with recommendations on how the quality of the standard might be improved: as in the suggested next steps. The latter requires special attention as the client has often started this quality measurement as a starting point for an improvement project.

##### E2. Review the report and finalize (principal & others)

A good habit is to review the report, which can be done for instance by the standard developer, but might also involve the initiator, client or other experts. Based on the reviews the report can be finalized. When there are doubts about the results then an additional step might be to let another expert in standardization do the analysis as well, or try to compare the results with other standards.

##### E3. Discuss improvement suggestions (principal, client & standard developer)

The improvement suggestions are intended to be used by the client and standard developers. Based on that assumption it is recommended to discuss the improvement suggestions with the client and standard developer, but even a broader group, for instance the workgroup responsible for standard development, might be interested.

#### **The end of analysis**

The project might end with an evaluation to assert the satisfaction of the client. When it is expected to repeat the quality measurement in time, which is recommended, then it is also advisable to extend the evaluation to the standards developer, focusing on improvement of the iQMSS.

### 11.4.4 Recipe

In summary, we present the recipe for the initiator and principal in Table 46.

Phase	Steps
<b>Preparation</b>	0. Prepare the project by reading this book (initiator & principal) A1. Plan a meeting with the client to discuss the information needs (initiator & client) A2. Map the measurable concepts to the information needs (initiator) B1. Gather and analyze information about the standard (initiator) B2. Customize the model of semantic standards (initiator & client) C1. Set up the measures (initiator) C2. Set up instrumentalization (initiator) Validate
<b>Measuring</b>	D1. Apply the quality model: what is to be measured? (principal) D2. Find the information: to score the measure (standard developer) Validate
<b>Analysis</b>	E1. Set up a report (principal) E2. Review the report and finalize (principal & others) E3. Discuss improvement suggestions (principal, client & standard developer) Evaluate

Table 46 – iQMSS recipe

### 11.5 Conclusions

To set up this final QMSS build we followed an extensive research approach, including the integration of all the previous iQMSS builds. Next we improved and validated the model based on the sources that were either recently published or were previously not included. Also a detailed layer, the measures for product quality, was added to the quality model, just as a usage model of how to use the quality model. Together, the result of this step is the generic QMSS (build 0.7).

# Validation of the iQMSS

In Chapter 10, we reported on our explorative case studies aimed at acquiring knowledge to improve the instrument. Based on their outcome a final version of the instrument has been developed (Chapter 11). This chapter will present its validation for use in practice.

### **12.1 Research approach**

The goal of this step is to validate the practical usability of the iQMSS according to the approach, as presented in Chapter 6. This goal can be broken down into three research questions (RQ), and a research approach has been set-up for each of them:

RQ1. Does the iQMSS fulfill its requirements?

RQ2. Do the process and outcome of iQMSS have value in practice?

RQ3. Does the iQMSS contribute to solving the identified problem in practice?

#### *12.1.1 RQ1: Re-assessment of requirements*

Although the defined set of requirements (Chapter 7) have been used during the development of the iQMSS, it is still important to re-assess the requirements. Not all requirements will have been fully met, and it is important to know which part of the requirements have been met and why others are excluded. This re-assessment was carried out by taking the table of requirements and writing down whether the requirement had been met, and providing argumentation. Based on this outcome we can answer the research question; does the iQMSS fulfill its requirements?

#### *12.1.2 RQ2: Field test*

To answer this research question, we have chosen a field test because of its applicability for testing in practice. Because we need to gather in-depth knowledge and need to know how the instrument works in practice, a field test is applicable (Yin, 2009). Another option is to perform multiple field tests or case studies. Their advantage might be the generalizability of their results, although generalization of a single field test is possible as well (Lee, 1989). However, the time and costs required and our research limitations prevented us from choosing this option.

Hence, the single field test fits our research goals and available time and resources, best. Another rationale is that we as researchers have access to a particular situation that might be difficult to access by other researchers (Yin, 2009). In our research approach we tried to reduce the effects of the limitations of a single field test in contrast to a multiple field test or case study, for instance by triangulation. Implying that data would be collected through multiple sources including interviews, observations, and document analysis (Creswell, 2009; Maxwell, 2005). Other validation threats like “researchers bias” and “reactivity” (Maxwell, 2005), although important, do not have a major impact since we were not explicitly aiming for generalizability of the results.

IMS LRM (Learning Resource Meta-data) is used as an *instrumental* field test (Stake as cited by (Silverman, 2010)), in which a case is examined mainly to provide insight into an issue. The IMS LRM standard has been selected for the case, for several reasons:

- The context: International – national.
- The content: A broad semantic standard.
- The development and maintenance: The experienced IMS organization within the education domain.
- Application: The standard is widely adopted.
- Practical reasons: Experts available who are willing to participate.

One limitation of this choice is that IMS LRM might not be representative of all semantic standards, since IMS LRM is a document (metadata) standard, in contrast to many vertical message exchange standards. Albeit that there is probably no single semantic standard that is representative of all semantic standards. In summary, IMS LRM is suited for this field test and the three analytic features of our research (Silverman, 2010) can be summarized as:

1. The boundaries: Specifications that are the core of the IMS LRM, and the experts of Kennisnet.
2. Unit of analysis: Both the process and outcome of the iQMSS.
3. Limited research problem: Limitations on generalizability, while preserving the wholeness of applying iQMSS as intended.

The iQMSS has been fully tested, which implies that the process was followed as described in Chapter 11. During and after the application, additional steps have to be taken to answer the research question. To get more knowledge regarding the process we made use of observations during the process. We avoided too much involvement (action research) because we had already taken that approach during the explorative evaluation studies and wanted to minimize our possible influence in this field test. For a better understanding of the value of the outcome an interview with the client was done afterwards. Both will be described in more detail in the following sections.

### 1. Observations & Reflection

To study actual behavior, we need to be unobtrusive. Simple observation with written techniques is an appropriate method (Kellehear, 1993). Our first approach was to meticulously document the steps and observations during the process, including timing of steps, discussions between the actors, etc. (Miles & Huberman, 1994). Based on that knowledge outcome, we were able to consider these findings together with both earlier defined requirements for the instrument and general and common knowledge of performing measurements. The advantages of the observations is that we gathered first-hand experiences of actors, we could record information as it occurred and unusual aspects could be noted during the observation (Creswell, 2009).

In the case of observations, the researcher can opt for the full observer role (no participation) or can act as a participant as well. We chose to have one observer who supported the initiator and principal when needed, but his role was primarily the observer. We also chose an unstructured way of observing, in contrast to semi-structured, because we did not want to limit ourselves, and get open results.

In this study, observation creates opportunities to see through the eyes of the actors involved, as well as viewing the process (Silverman, 2006). Our observation notes focused on the following questions (Emerson as cited by (Silverman, 2010)):

1. What are the actors doing? What are they trying to accomplish?
2. How exactly do they do this? What specific means and/or strategies do they use?
3. What problems do they encounter? And how do they solve them?

## 2. Interview

As well as our reflection on the field test based on observation, we interviewed the client (standard developer), about his experience with the process and his satisfaction with the outcome. An interview gave us the opportunity to gather more specific details and opinions. And although it is obtrusive we did it afterwards so that the iQMSS results were not affected.

We knew in advance that the outcome of the quality of the iQMSS would be highly affected by the knowledge of the standard developer. Therefore we decided not to interview other standardization experts about the outcome, but only interviewed the standard developer who had participated in the measurement. By doing that we focused on the relative validation of the outcome in relation to the input of the standard developer.

The interview questions used were:

- What is your opinion of the results in general? Is it what you expected?
- Does the result contain improvement suggestions that you do not agree with?
- What is your opinion about the practical relevance? Can you, based on the results, start an improvement project for the standard?
- Was the process acceptable and flexible to address your requirements?

Based on the outcomes of the observation and the interview, we are able to answer the question: whether the process and outcome is acceptable in practice (RQ2).

### 12.1.3 RQ3: Expert workshops

Early in our research we identified a problem in practice (Chapter 3) for which we built the iQMSS, as an intended solution. To test if we had succeeded, the third research question was set up to assess the contribution of iQMSS to the identified problems in practice.

It is not recommended to set up a new survey that addresses the survey respondent of the problem survey (Chapter 3), since the respondent will not have enough knowledge about the iQMSS to be able to assess the merits of the iQMSS. To overcome this issue we chose to organize expert workshops of about two hours, where the first hour would be a presentation of our research results focusing on the iQMSS, and the second hour a discussion of the results. These discussions were structured by using propositions that deal with completeness of the iQMSS and its potential use in practice. Although not related to a research question, but more as a side effect, the expert workshops could also lead to new ideas related to the practical use of the iQMSS. The gathered ideas were used in our further research section in Chapter 13.

Due to the global scope of standards, we decided to do two workshops, one international workshop, as a special event in conjunction with the OMG (Object Management Group) Technical Meeting (United States), and one national-oriented workshop organized by the Netherlands Open in Connection. The audience of OMG is appropriate because OMG is involved in semantic standards, in OMG words domain standards (for instance for the financial domain), and are highly interested in developing high quality standards. Netherlands Open in Connection offered the inclusion of many different kinds of semantic standards, in the Dutch context. Both workshops were identical in program and approach.

To gather more in-depth understanding, we asked the participants of the workshop to participate in an additional short survey. This also adds the opportunity to receive more critical feedback as the survey was set up anonymously in comparison with the discussion within the workshop.

During the survey we first tested if the outcome of our problem survey is also shared by the participants in the workshops:

- More focus on quality of standards is needed.
- A quality instrument is useful to improve the quality of a standard.
- Improving the quality of the standard will lead to improved interoperability in practice.

Our expectations were that the results should be similar to the problem survey, meaning a high acceptance of all three propositions.

Secondly we aimed to learn the workshop participant's opinion of the contribution of the iQMSS to solving the identified problems, by setting the following propositions:

- The result, the iQMSS, contributes to the knowledge base regarding semantic standards.
- By using the iQMSS, the quality of a semantic standard will become explicitly known.
- By using the iQMSS, quality improvements for the standard can be identified.

Thirdly the aim was to assess the potential usage of the iQMSS, by means of the two propositions:

- I will use the iQMSS, when it becomes available.
- If agreed, to which standard? (open question)

Finally, we wanted to assess the participant's background (control questions) as well as to give the participants the opportunity to present (further research) ideas:

- What kind of standard development are you involved in (semantics (e.g. Finance/ISO 20222), technical (e.g. SBVR, MDA, etc), or none)?
- What more needs to be done? What is your suggestion for further research? (open question)
- Any other remarks/suggestions?

With the exception of the open questions, all the other propositions were presented based on a 5 point Likert scale. The following sections contain the results of the validation steps.

## **12.2 Validation based on requirements**

In Chapter 7, a situational requirement engineering method was used, resulting in a goal tree in which the requirements are structured. This structure shows the requirements related to the input of the instrument; stating that the instrument should be useful for a set of different semantic standards. It also shows that the instrument should be efficient and especially easy to use. Finally there is a set of requirements related to the outcome of the instrument, stressing that a high quality outcome is important, including improvement suggestions.

In our validation we assessed all the fit criteria for the individual requirements. Most of the requirements were met. We will present a summary in this section of each of the three top level goals within the requirements goal tree (Appendix F):

- A. Useful for semantic standards of different SSOs. The iQMSS is intended to be used for different kinds of semantic standards and it is also tested on different standards: SETU is a message exchange semantic standard, while IMS LRM is a metadata standard. Part of the solution to make it applicable to a wide range of semantic standard is the ease of customization. Part of the usage model is to deselect the irrelevant parts of the iQMSS and to select and add appropriate measures. It is also possible to adapt the iQMSS to the specific needs of the standards setting organization, or others.
- B. Able to efficiently determine the quality and to give improvement suggestions. The ease of use is essential for achieving this goal. The customization process makes the use more complicated because an additional (customization) step is used, in contrast to the out-of-the-box usage. However there is return-on-investment because the customization of the actual measuring is much easier and quicker. The iQMSS contains a usage model on how to use, and focus on the transparent outcome.
- C. Have useful results for SSOs. Many requirements are related to the usability and quality of the results, including its complete view on quality, objective independent reproducible test results and so on. So far the test results have been positive. The validation sessions (section 12.4) were particularly addressed to test if these requirements were met.

Although the above shows the achievement of the top level goals, when searching into the detailed requirements and fit criteria we found some that had not been met, because of a different reason. Below we mention those specific requirements including the numbering as used within Chapter 7 and Appendix E and F):

1. Ranking of standards (REQ C2): Currently not implemented because requirements C2 (ranking) and C3 (improving) are difficult to combine and they demand two somewhat distinct instruments. We, based on our problem statement and survey, chose to implement requirement C3: The iQMSS delivers an outcome that contains improvement suggestions. Currently it is too arbitrary to rank standards based on the outcome. However comparative analysis is possible based on comparing the results of the individual measures.
2. The weighing factors (REQ A5): As a consequence of not implementing the ranking of standards (C2) requirements, the value of applying weights and scoring becomes less important and has currently not been implemented.
3. One page results (B2): In practice one page was achievable by applying formatting, but was not met: the results of the IMS LRM case fitted onto two pages, of which the improvement suggestions were about 1 page. This seems fair.
4. Short period to learn (B5 and B6); the automation has not been fully implemented yet, and the customizability and completeness of the learning time of the iQMSS will take some time. Although the dependence on the iQMSS developer is not really a problem, it is recommended that the principal has a rich history in standards development.

### **12.3 Field test IMS LRM**

The field test was carried out during August and September 2011. Jos van der Arend from Kennisnet, volunteered as a client and standard developer, since he is actually a member of IMS and is involved in the IMS LRM standard. Jasper Roes, standardization expert from TNO, took the role of initiator and principal. Support was given by the author, the lead developer of the iQMSS, who also took the role of observer during the measuring meeting.

This section will describe the process remarks as well as the outputs for each of the three main phases of the usage model (Section 11.4).

#### *12.3.1 Preparation phase*

##### **The process**

The steps from the recipe:

##### 0. Prepare the project by reading this book (initiator & principal)

Not all the parts of this book were at the final stage when the field test was performed, but the description of the instrument was ready and read by the initiator and principal. However the initiator/principal was already involved in the development of the iQMSS and therefore quite familiar with the ideas and quality model.

##### A1. Plan a meeting with the client to discuss the information needs (initiator & client)

The client is mainly interested in the (internal) product quality of the IMS LRM standard, because the standard might be used to localize for the Dutch education situation. This localization task is performed by the client's organization. Most relevant in this context is the durability of the IMS LRM standard: will the standard be future-proof? In the context of localizations, it is further important to know how localizations are dealt with: The adaptability of the standard. But also the current advancedness and the ease of maintenance.

#### A2. Map the measurable concepts to the information needs (initiator)

The information needs reflect the product quality – durability part of the iQMSS.

#### B1. Gather and analyze information about the standard (initiator)

The IMS website contains a webpage that aggregates all official information regarding the standard. We also performed a Google search to find additional sources.

#### B2. Customize the model of semantic standards (initiator & client)

We applied the Semantic Standard Model to the IMS LRM standard. Based on the information needs the application and organization parts were removed during the customization. The information need justifies that we only take the context and content parts of the semantic standard model into account.

#### C1. Set up the measures (initiator)

We chose to select all the measures to be a part of the durability branch (A3) in the customized QMSS. Some have been slightly adapted to be more applicable to the context of this measurement.

#### C2. Set up instrumentalization (initiator)

Based on the project requirements, and the current lack of software instrumentation, it was decided not to use any software instruments, but to create a table with all the forms of the measures used during the measuring meeting.

#### Validate

The initiator checked this outcome and choices with the client, and based on review comments the implementation of the semantic standard model was updated, but the intended measures were unaltered.

### **The outputs**

#### From A:

1. A description of the client's information needs:

- The durability, longevity expectations, for IMS LRM, in order to build localizations upon it.

Other project requirements:

- A short report that focuses on showing the potential of a full quality measure on LRM.
- Minimum amount of resourcing by the client and standard developer.

2. Selection of measurable concepts; the customized QMSS (Figure 53). This is part of the iQMSS product quality.

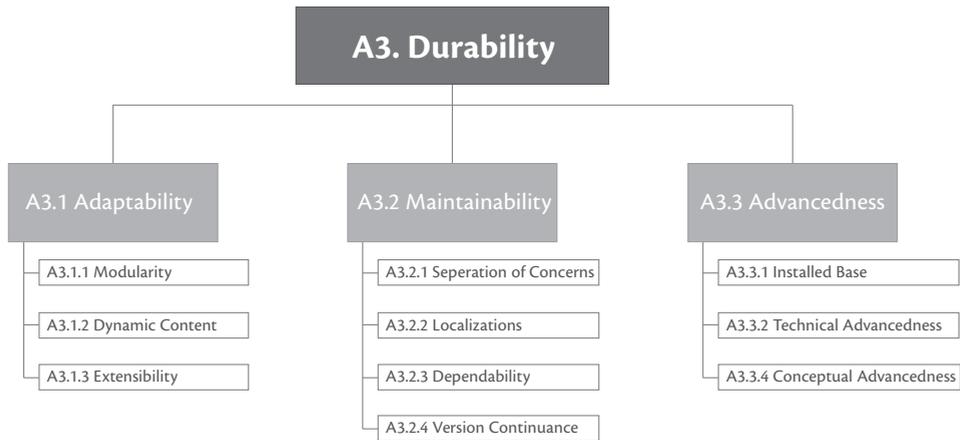


Figure 53 – Selected IMS LRM measurable concepts

From B:

1. Customized SSM and main sources of information: The customized and instantiated SSM is described within Appendix P. It contains an overview of the IMS LRM standard.

From C:

1. The set of measures to be used during measurement (Table in Appendix Q).

2. Set of instruments. The table is the instrument to use. In the table we created space to be able to answer each question in three ways, based on facts, feelings and a judgment/scoring. The Gartner Hype cycle was included as an instrument to express the lifecycle status of the technology parts. Furthermore no intended use of surveys, questionnaires, Internet tools or whatsoever was included.

Measurable Concept	Measures	Function
<b>A3.1.1 Modularity</b>	1. Data dictionary	1. Assess if a data dictionary is available, and if it can be used in different standard modules.
	2. Re-use of components	2. Assess if there is re-use of components, building blocks, within the standard.
	3. Independence of modules	3. Assess the possibilities of using certain modules of a standard independently.
<b>A3.1.2 Dynamic Content</b>	1. Dynamic content	1. Assess if code lists or other solutions are used for dynamic content resulting in a separation between dynamic content that is prone to changes and a standard core that proposes a stable representation of the business domain.
<b>A3.1.3 Extensibility</b>	1. Extension mechanism	1. Assess if there is an extension mechanism, such as additional data elements, with rules on how to use them.
<b>A3.2.1 Separation of Concerns</b>	1. Separation of concerns	1. Is separation of concerns applied within the standard (according to a defined policy)?
<b>A3.2.2 Localizations</b>	1. Option of localization	1. Are localization / profiles for specific purposes intended and are guidelines provided as to when to use localizations?
	2. Localization rules	2. Is there a strict policy of defining the contents of localizations?
<b>A3.2.3 Dependability</b>	1. Dependencies	1. What kind of strong dependencies exist, or are they expected, regarding other standards. A strong dependence means that a version update of the other standard has an impact on this standard. Especially focus on generalization-specialization relations; Is the dependency logical and managed?
<b>A3.2.4 Version Continuance</b>	1. Version check	1. Only assess the content of different versions, based on version history, and assess whether there is logic in the content changes, including backwards compatibility.
<b>A3.3.1 Installed Base</b>	1. Current ICT	1. How difficult is integration with existing systems (e.g. SAP) in stakeholders' organizations.
<b>A3.3.2 Technical Advancedness</b>	1. Lifecycle position	1. Assess the position of the used technology on hype cycles. (e.g. The Gartner Hype Cycle on Technology.)
	2. Competing technology	2. Assess the competing technology. Is the choice of technology justifiable?
<b>A3.3.3 Business Processes</b>	1. Commodity	1. Assess proposed business processes and compare them to the standardized business processes commonly used in practice.
<b>A3.3.4 Conceptual Advancedness</b>	1. Conceptual complexity	1. Assess the fit between the complexity of the solution and the complexity of the interoperability problem. Is it overkill?
	2. Future fit	2. What is the fit with future views on the interoperability problem domain?

Table 47 – Measures for IMS LRM

### 12.3.2 Measuring

#### The process

The steps from the recipe:

D1. Apply the quality model: what is to be measured? (principal)

D2. Find the information: to score the measure (standard developer)

The iterative circle of these two activities began with an interview between the principal and the standard developer (and the observer). After a brief introduction the instrument (table) was used.

Three questions were asked for each measure:

- What are the facts that support the answer to the question?
- What is your feeling about the answer to the question?
- What do you think of it? Is it good/acceptable, or not?

Based on the observation notes we can conclude that it is absolutely essential to be prepared on two aspects:

1. It is essential to have all documents at hand, and take the time to search for facts supporting the statement.
2. The principal should prepare the measures very well. It is quite important to have a good idea as to which questions will be asked, because some are only slightly different. Without knowing the follow-up questions, the question at hand might be wrongly interpreted and confused with the follow-up question.

Other lessons learned from the observation:

- The more directive the measures, the better they are for an easy measuring process. For instance the measurement using the Gartner Hype Cycle is very straightforward and easily done.
- But there is a downside to easy measurement; the fear of performing an autopilot measurement without critical thinking becomes apparent, and avoids the critical attitude that is needed during measurement. So, probably, a mix of straightforward and somewhat more complex measures is needed to keep an open critical mind.
- Timing; in this case there was a good balance of timing, implying that all measures took roughly the same time in measurement. During measurement the principal and standard developer checked the time several times to see if they were still on track, within the time they had reserved. In this case there was enough time in relation to the number of measures. One could imagine that if the set of measures becomes too large and the time for measurement is too limited the balance in time per measurement will be gone. Typically in such cases, the first measures receive more attention than the final measures. This will have an impact on the results.

### Validate

Afterwards, the notes were processed and the table with measurement results was sent to a standard developer to be reviewed. Based on the review some minor updates were made.

### **The outputs**

The output of this phase is the table with the measuring results (Appendix Q).

#### *12.3.3 Analysis*

### **The process**

The steps from the recipe:

#### E1. Set up a report (principal)

Based on the measuring result the analysis report was an easy process. By going through the remarkable issues from the measuring results, the relevance of each issue was questioned, and when known some improvements were presented.

E2. Review the report and finalize (principal & others)

E3. Discuss improvement suggestions (principal, client & standard developer)

### Evaluate

The analysis report was sent to the client (standard developer) for a review. The review continued with a follow-up session in which also the improvement suggestions were discussed, as well as an interview/discussion on evaluation. Remarkably, no changes were made to the analysis report. The results of the evaluation will be presented in the following section.

### **The outputs**

The analysis report: Available in Appendix R.

#### *12.3.4 Field test discussion*

We evaluated the field test in an interview with the client/standard developer on October 12, 2011, at Kennisnet. We present the results in two parts, one related to the process, the other to the result.

### **The process**

The process was felt to be flexible and acceptable for use by Kennisnet. The tailoring to the needs of Kennisnet made the measurement very efficient. However detailed preparation of the measurement phase is essential, and even during measurement a structured approach is really important. Otherwise misinterpretation of measures will occur resulting in less valuable results.

During the measurement we created a structured approach by always looking at the quality tree before reading out the definition of the measure. For each measure we looked for supporting facts, followed by an expression of the feelings, and finally a judgment. Although the facts or judgments were not always present, the structuring led to a satisfactory process for all. More structure, for instance by having stricter answering options for the measures, or more guidance/support in the measuring process would have been appreciated.

One potential issue is the dependence on the principal. Kennisnet would also be interested to use the iQMSS themselves. Currently, there are presumptions that the iQMSS will not be easy to use without having the developer's knowledge on board. It will require some additional time. The availability of tooling might also lead to easier usage. This becomes particularly apparent when the measuring results are transformed into the analysis report. More transparency and guidance on this process is needed to loosen the dependency on the iQMSS developers.

### **The end result**

The correctness or validity of the improvement suggestions in the analysis report were not discussed that much; they had been expected by the client. To a large extent, about 60-70% of the improvement suggestions were already known and expected to be in the report. However, it was really appreciated that although already known, they are now in an official report. Especially the independent characteristics of the analysis/report give the report an additional status, which makes it easier to use for improvements. The other 30-40% did not lead to disagreements in the discussion, although several suggestions are observations that need to be investigated further.

Although the practical relevance is highly estimated, the real value has to be shown during a standard improvement project. The feeling is that the analysis report is an ideal starting point for an improvement project. A suggestion was made to present the top 3 more prominently.

Currently the quality of the result highly depends on the quality of the people involved in the application of the instrument. Again, sophisticated tools might lower this dependency, especially with regard to the principal.

## 12.4 Validation sessions at OMG and NOiV

The first workshop as part of the OMG Technical Meeting was held on September 22, 2011, Orlando (FL), USA. Ten people participated in the session; most of them had longstanding experience in standardization practice. The second workshop took place on September 27, 2011, The Hague, The Netherlands. Twenty people participated of which about two thirds were actually involved in the development and maintenance of semantic standards while one third had other roles, such as involvement in government policies. Four survey forms were collected during the US session and fourteen during the Dutch session. Both sessions took approximately two hours and both were recorded. In the end, both discussion and survey outcomes were similar, without noticeable differences. Therefore we will discuss the combined results.

### 12.4.1 Discussion results

In general the discussion resulted in a lot of positive feedback related to the research in general and the iQMSS in particular. During the discussion the completeness of the quality models was acknowledged by participants. Although invited and challenged to do so, the participants were not able to identify potential gaps within the quality models. Thus all the aspects mentioned by the participants were already part of the quality model. The discussion even touched upon the possible over-completeness of the quality model. However the iQMSS usage model is focused on selecting only the most appropriate part of the quality model.

Another interesting discussion item was that the iQMSS should not only be applied incidentally; it should be a part of the SSO quality assurance processes. The current focus of the iQMSS is however on identifying improvement suggestions which might not be handy if it is applied within the quality assurance process as a final step before release. Early application in the quality assurance process or using an adaptation of the iQMSS might solve that issue.

The important parts of the discussion are related to the ideas for further research. This will be presented in Chapter 13.

### 12.4.2 Survey results

The respondents within the survey are involved in a wide range of standards, from education, health care, geographical, and agriculture. Similar to the discussion the survey results are coherent and positive. We used the five point Likert scale in which “3” corresponds to “neutral”, “4” to “agree slightly” and “5” to “agree strongly”. The results are presented in Table 48.

Propositions	N	Minimum	Maximum	Mean	Std. Deviation
More focus on quality of standards is needed	18	3.5	5.0	4.52	0.55
A quality instrument is useful to improve the quality of a standard	18	3.0	5.0	4.27	0.75
Improving the quality of the standard will lead to improved interoperability in practice	18	3.0	5.0	4.05	0.80
The result, the iQMSS, contributes to the knowledge base regarding semantic standards	18	3.0	5.0	4.25	0.80
By using the iQMSS, the quality of a semantic standard will become explicitly known	18	3.0	5.0	4.05	0.72
By using the iQMSS, quality improvements for the standard will be identified	16	3.0	5.0	4.37	0.71
I will use the iQMSS, when it becomes available	16	3.0	5.0	3.81	0.75
Valid N (listwise)	16				

Table 48 – Survey results

First of all, we can conclude that, in line with the problem survey, the workshop participants support the call for more attention on the quality of standards. They are also convinced that the iQMSS is actually contributing to the knowledge base and is useful for the identification of improvement suggestions. Although it is not relevant for all participants, a very positive result was the potential use of the iQMSS. This is also supported by the fact that, based on these two workshops, four people asked for investigating the opportunities to use the iQMSS for four respective semantic standards.

The open questions revealed some other interesting research directions; such as the relation of quality with adoption and development of standards and preserving correct implementations of the standard.

### **12.5 Conclusions**

This chapter validated the iQMSS by applying different research methods. If we re-assess our research questions, then we can conclude the following:

- RQ1. Does the iQMSS fulfill its requirements? - Yes, to a large extent the requirements and the fit criteria have been met. The most prominent requirements that have not been met are automated tool support and the ranking of standards.
- RQ2. Does the process and outcome of iQMSS have value in practice? - Yes, our IMS LRM field test shows that within a flexible and relatively simple process an output is created that is of value to the client.
- RQ3. Does the iQMSS contribute to solving the identified problem in practice? - Yes, our survey shows that the iQMSS contributes to the knowledge base and the identification of improvement suggestions. A viewpoint on standards quality has become available by means of the iQMSS, which was previously seen as a problem in practice.

The evaluation phase may lead to either the rejection of the model or another iteration of the development phase or the reconceptualization of the transfer and evaluation methods (Chapter 6 (Becker et al., 2009)). Based on the validation we conclude that more iteration will improve the iQMSS further. However applicability is not an issue: the current version is useful in practice.

# Conclusions and Future Research

This final chapter will revisit the main research question, summarize and draw conclusions and finally will present suggestions for further research.

### **13.1 General conclusions**

The main research question for this research was: *What are the characteristics of an instrument to measure the quality of semantic standards that will aid standard developers in improving their standards?*

The answer to this question is formed by the characteristics of the iQMSS, and is presented in Chapters 8, 9 and 11. In summary, the design model of the iQMSS involves three types of artifacts: artifacts related to the quality model, those related to the semantic standard model and those related to the instrumentalization. At the highest abstract level, a language was needed to create both a quality model and a semantic standard model with consistent concepts and definitions. For this purpose we selected the SMO from the software engineering domain, and used this terminology throughout this thesis.

The QMSS consists of three parts; product quality (intrinsic), process quality (the organization), and quality in practice (application). Each of these contains a hierarchy of measurable concepts, including definitions. Measurable concepts define what we want to measure, given some information need. The product quality tree has been enhanced with a set of measures for each measurable concept. For example, the instrument proposes the use of the Gunning Fog index as a measure for readability, which is in the understandability and usability part of the product quality tree. Another example is the proposed measure to assess the correct implementation of Naming and Design Rules (NDR), which is part of the technical complexity, which is part of the usability of the standard.

The SSM describes the measurement subject and supports the understanding of the standard. At the highest level, a semantic standard is seen to consist of its context, its content, its development and maintenance processes, and its application. The context involves the problem characteristics, while the content is related to the characteristics of the solution. The development and maintenance processes concern the organization of the standard, and the application deals with both application support and applications in practice. The semantic standards model is applied in the iQMSS to gather knowledge about the standard, to create an overview and an understanding about what we are able to measure.

The instrumentalization relates to the practical tools that accompany the models to make the iQMSS easy to use. In our research we have shown the possibilities of using a tool. Our current instrumentalization is limited to fill-in tables on paper.

Based on the requirements, we decided that, for every specific use case, the generic artifacts need to be adaptable to specific needs: Customization. The quality model, the semantic standard model, as well as the instrumentalization, need to be customizable for every measurement instance, according to the usage model. Although this is an additional step for the user, it reduces the efforts needed to use the iQMSS, and provides a better fit in practice.

The glue between the artifacts is the usage model, which describes how the iQMSS should be applied. Four roles are required. First the client role has the information need about the quality of a standard. The initiator role prepares the measuring process by customizing the iQMSS to the needs of the client. Actual measurement is performed by the principal and the standard developer. The result includes a description of the standard based on the SSM, the customization of the iQMSS, the measurement results, and the analysis report. The latter provide a basis for an improvement project for the standard.

As well as our main research question, we defined the main research objective as the design of an instrument to measure the quality of semantic standards, and that will aid standard developers in improving their standards. The previous section showed the characteristics of the instrument, but does it aid standard developers in improving their standards? Our validation studies suggest that the iQMSS identifies improvement suggestions. Thus it will help standard developers. However our limitation is that we do not have evidence that the iQMSS identifies the best or a complete set of improvement suggestions, and that the use of the iQMSS actually leads to an improved version of the standard. However, based on the validation workshops we believe that the iQMSS will be used in practice as an aid by standards developers.

We can therefore answer the main research question positively. Based on this research the developed iQMSS is able to assess the quality of semantic standards and is able to present improvement suggestions.

### **13.2 Conclusions: the contribution to practice and science**

This section will describe the contribution of this research to both practice and science.

#### *13.2.1 Contribution to practice*

One of the key questions when evaluating research is the importance of the topics and issues to the field of inquiry (Silverman, 2006). Practice needs more knowledge and guidance on how to improve the quality of standards. This statement is grounded by our survey of 34 semantic standards, which shows that the quality of semantic standards for inter-organizational interoperability can be improved. Based on an extensive design approach the iQMSS was developed, and it gives an insight into the quality and improvement suggestions. This is a much needed tool, and the need will continue to rise, as more and more new semantic standards are introduced.

Improvement of semantic standards requires transparency of their quality. A wealth of semantic standards has been developed for various industry sectors. Although many semantic standard development organizations already have quality assurance in place, this research shows that they could benefit from a quality measuring instrument. The introduction and use of the instrument may lead to an improvement in semantic standard quality. Improved standards may advance interoperability in networked business. Achieving interoperability may lead to significant cost savings, performance improvements and efficiency gains.

In summary, the main conclusions and contributions to practice are:

- According to standard developers, the quality of semantic standards is improvable and desired for achieving interoperability.
- Practice needs more knowledge and guidance on how to improve the quality of standards.
- The availability of the iQMSS tool to assess the quality of semantic standards and to suggest improvements.
- The flexible use, allowing for multiple usage scenarios and application to a wide range of semantic standards in an effective and efficient way.

Our final validation shows that the iQMSS received a warm welcome in practice, an indication of the practical relevance of this research.

### 13.2.2 Contribution to science

A key question when evaluating research is its scientific contribution to existing research and theoretical debates (Silverman, 2006). Despite the importance of standards in the evolution of ICT (Lyytinen & King, 2006), the issue of semantic standard quality is relevant in practice (Chapter 3), yet it is not often addressed (Chapter 4). Our research, based on a systematic literature review, identified a research gap in the topic of quality and semantic standards, which is already a contribution to science.

Our research further adds to science by providing evidence that in practice there is a great need for knowledge on standards' quality (Chapter 3) but also for the development of further knowledge about semantic standardization, which is under-addressed in science (Chapter 4). Finally, the QMSS and the SSM are significant new scientific results.

Two other key questions related to the scientific contribution are (Silverman, 2006):

- Their conceptual rigor through explicit specification of concepts and theoretical perspectives, clarity of objectives, appropriate treatment of relevant literature, logical reasoning, etc.
- Their methodological rigor through the use of appropriate methods, appropriate and sufficient data, rigorous and innovative analysis.

These relate to the validation and credibility of our research. We used triangulation (combination of multiple theories, methods, observers, empirical materials) (Silverman, 2006) as triangulation may improve the reliability (Silverman, 2010). In the end, our research approach has been extensive, particularly because we included experience from practice, given that the amount of relevant literature was limited. Several research methods were used to include the practical experiences effectively and efficiently. Even when validating the results, we took several approaches, including field test, workshops and survey.

Innovation is achieved by combining knowledge from different fields into a new area (Levina & Vaast, 2005). In our research this is applied in the iQMSS as it builds on knowledge from different domains, including the software domain, data quality domain and product engineering domain. It has led to two new scientific artifacts, the SSM and QMSS, which already have proven valuable in practice. Innovation can continue by improving and validating these artifacts by the scientific community.

In summary, the main conclusions and new contributions to science are:

- Evidence of the need for knowledge on the quality of semantic standards in practice.
- Despite its importance in practice, the topic on the quality of semantic standards is a proven research gap based on a systematic literature review.
- The presented state of the art is a fundament to build upon and can direct scientific research agendas.
- SSM & QMSS: Innovative artifacts with highly acclaimed practical value that can be built upon in future research.

### 13.3 Future research

Like any other research, this research has bias. The author stems from the standardization community, has many years of experience in developing semantic standards, and based on this experience strongly believes that the quality of semantic standards needs improvement. If we look back and reflect on our research then we conclude that this background and bias has led to a focus on inclusion of practitioners' experiences and the focus on an outcome that is primarily useful in the real life situation of semantic standards.

Other reflections on our research, together with the validation workshops as described in Chapter 12, have helped us to set up four directions for further research. Additional ideas were gathered during discussions with colleagues and in presentation and discussion sessions where the results of this work were presented. The first three directions are directly related to the iQMSS, while the fourth direction extends beyond that.

## 1. Improving content

Although our research involved an extensive process to build the iQMSS, improvements to the content are still achievable. This might apply to the quality model, the semantic standard model, and most likely to the usage model. Performing case studies and evaluating the use of the iQMSS in practice, will both be suitable to identify these improvements.

Moreover, content could be added as well, particularly by adding more measures to all the parts of the quality model. For “process quality” and “quality in practice” in particular more build and evaluation iterations needs to be done, and finally the measures need to be constructed. For now, we advise the use of BOMOS (Folmer & Punter, 2011) as a guidance on process quality and the work of Zhu et al. (Zhu & Fu, 2009; Zhu & Wu, 2010, 2011) for a guidance on (partly) measuring quality in practice.

The most interesting will be the search for existing and new research that leads to measures that can be incorporated in the iQMSS. Another example is the search for information on the development of tools to test the quality of data definitions, which has unfortunately not been published yet. This might be used as a measure for “A1.3.1 Information Ambiguity”. By adding more of these scientifically developed and practically validated measures, the outcome of the iQMSS will likely increase. Another interesting source might be to search for measures that are already used in practice by SSOs. For instance the OGC (the geospatial domain) developed the “Specification Model Standard for Modular specifications” (OGC, 2009), and uses this as a quality tool for their own specifications. Such documents might help to enrich the set of iQMSS measures.

From the perspective of the iQMSS user, there is a strong request for more directions about right and wrong or good and bad measurement values. The iQMSS can then be used to assess their conformance to their standardization approach, when possible, on every quality aspect. This requires not only that every measure includes a scale, but also that values for good and wrong are assigned. This becomes quite challenging since this is both context dependent and hardly any research is available as a foundation for assigning good and bad values. It also leads to a shift in the focus of the iQMSS, which might be conflicting. It could also lead to the development of a different and distinctive iQMSS.

## 2. Improving application

A second direction for future research is related to improving the application possibilities, which will have positive effects on the usage of iQMSS in practice. The ease of application can be influenced in three ways:

*A. Develop templates:* The current iQMSS model contains steps for its customization and for selecting parts to fine tune the instrument to the specific information needs of the client. The reason for that is to find a balance between the support of the instrument to many different information needs in practice and to keep the application lean. However, although the information needs will differ in practice, it is also likely that several “information needs” will be common and will be repeatedly used by different clients. For such commonly used information needs, it will be worthwhile to develop templates. Templates might also be developed for certain kinds of usages. For instance, a template on how to integrate the iQMSS into a quality assurance process, and another template on how to use iQMSS in workgroup sessions, etc. could be developed. By using such a template the usage of the iQMSS will be simplified, because the customization aspect during the preparation phase can be eliminated.

*B. Develop tooling:* In the requirements and the design of the iQMSS sections, the instrumentalization by means of tooling has been stressed. When developing the iQMSS we focused on the quality model and the semantic standard model, and only showed the potential of tooling. Further developments leading to the availability of an iQMSS toolset are needed and can be focused on simplifying the customization or the measurement itself. It might even involve toolsets for specific templates or information needs. This might lead to a relative simple Internet-based tool that lowers the burden when starting to use the iQMSS.

As well as the ease of application, application can be improved by broadening the application in two ways: by broadening the purposes and scope of standards:

*A. Broadening the purposes:* The iQMSS is directed at improving the standard from the standard developers' perspective. However we identified other potential interesting purposes for other users already during the start-up phase of this research, such as gathering knowledge related to the implementation project of a standard from an implementer's perspective. Such knowledge can be used for forecasting the effort needed for the implementation project and to assess potential problems. Another potential additional purpose concerns the selection of standards. Selection, or even comparison, can be done for multiple goals:

- Implementation in projects or products (standard implementers and standard users).
- Intake or release (SSOs).
- Re-use in other standard (SSOs).
- Inclusion in a (policy) list, or compliancy to certain criteria (policy maker).
- Benchmarking.

Users and implementers might be interested in selecting standards for their use based on certain criteria. In an organization running multiple projects or products, this is also called standards portfolio management. In the future, when dealing with the continuous growth of standards, the importance of standards portfolio management will grow significantly.

SSOs have further issues, for instance about the timing of releasing a standard, as part of a quality assurance process, or related to a decision regarding the intake of standards for maintenance that are developed by others. Also, selection of standards for re-use within the standards development process might be an issue.

Currently standards are also of particular interest to policy makers. Governments select standards to fit an interoperability framework. In the Netherlands for instance, the government has created lists with approved standards, based on certain selection criteria. Finally benchmarking standards might be interesting to many, such as standardization researchers who want to study the differences of the standards.

The iQMSS, or parts of it, might qualify for such additional purposes as well. Yet, there are important differences in requirements between improving quality on the one hand and these other purposes on the other. For example a selection instrument might have a Boolean outcome, in contrast to the open outcome of an instrument aimed at improving the quality. It is likely that these different purposes cannot be addressed by one single instrument. However, based on the iQMSS, a distinctive new instrument can be developed that complies with the requirements related to other purposes.

*B. Broadening scope:* Another way is to broaden the current scope of semantic standards to other standards, such as ICT standards in general. Part of the quality model, such as process quality, might apply to other types of standards, while other parts of the model might require adaptations.

Finally, application can be improved by becoming less dependent on the iQMSS principal. Currently, the role of the principal requires a lot of knowledge about both the iQMSS and scientific and practical knowledge about semantic standards. This makes it difficult for this role to be taken up by others, other than iQMSS insiders. By reducing the requirements of the principal role, others can take this role up. However the current knowledge requirements of the principal are particularly important for the transformation between the measurement results and the writing of the analysis report. Currently there is little guidance for this step; it relies on the experience of the principal. Future research should be aimed at automating or giving more guidance for this step.

### 3. Validating values

In the future, extensive work should be done on validation. This not only applies to the correctness of the outcome of the iQMSS on a large scale, but specifically also to the direct and indirect value of the outcome. The outcome draws its value from the actual implementation of improvement suggestions in updated versions of the standards. Follow-up research can then be aimed at studying the impact of those improved standards: Did it lead to improved interoperability? And what is the business or social impact of that? In other words, are the assumptions we made in this research correct. We advise the following for any further research on this topic:

*A. Validation of the correctness of the outcome and its practical use:* This can be done in multiple ways; we suggest a practical and a more sophisticated approach. The practical approach deals with performing many more studies whereby the instrument is applied as intended, and the outcome is evaluated by means of interviewing and surveying experts who are related to the standard. What are the experts' opinions about the results?

A more sophisticated approach is studying the effect of the iQMSS outcome. How many of the improvement suggestions have actually been addressed in updated versions of the standard? A longitudinal study can be done in which new versions of standards are appraised after the iQMSS has been applied. Statistical analyses on how many improvement suggestions have been implemented in updated versions, and the number of revisions that were not identified by the iQMSS, do indicate the correctness of the iQMSS result. As well as the statistical approach which addresses multiple iQMSS applications for different standards, more can be learned by studying a standards improvement project in-depth, by using, for example, action research techniques. We can then study the project in more detail and analyze it to see if the iQMSS results have been used optimally.

*B. Validation of the business case, or in other words the effect of improved standards:* Our assumption is that improved standards lead to improved interoperability. But if we invest 10,000 Euros in quality improvement, will that lead to improved interoperability that is worth more than 10,000 Euros? Or is every cent spent on quality improvements always worth the money spent and will it have a multiplier effect on the value of improved interoperability? Or is there a threshold? It would be very interesting to develop a decision support model for quality investments. This model should encapsulate the knowledge on how to calculate the value of an interoperability improvement of for example 10%, and deal with the fact that the benefits of interoperability are not equally divided between stakeholders. It should also deal with the different perspectives of an individual stakeholder versus the whole domain. To our knowledge this is complex and has not been studied before, which makes it difficult to solve. A starting point might be to apply scientifically validated business case approaches to interoperability problems.

### 4. Broader view

If we take a broader helicopter view of the subject, then many more suggestions for future research become apparent. First, the iQMSS needs to be maintained to be sure that lessons are learned, or that new measures will be added to the iQMSS. Actually a suggestion was made during the validation workshops, to make sure the iQMSS, or in particular the models, become a (European) standard itself. By becoming a standard, future maintenance is covered, and it will gain additional status that might have a positive effect on the application.

More generally, the science of semantic standards (and quality) needs further maturation. This would result in the availability of more measures that could be used within the iQMSS. The science of semantic standards should cover studies, using multiple research epistemology, that address the relationship between adoption, development and quality of semantic standards. It should include viewpoints from different domains, such as marketing, economics and software. And it should continue innovating by applying that knowledge to the domain of semantic standards.

Specifically more research needs to be done on standards as a solution to the interoperability problem. Both existing studies (e.g. (Brutti et al., 2010)) and our IMS LRM field test show that the relation between the main standard and the needed localizations or profiles needs more exploration.

The standards world is evolving; future directions include the use of ontologies in standards, the use of business rules to capture and maintain business knowledge and the automatic transformation between the business and technical content of semantic standards.

Also negativist studies should be encouraged that question standards as appropriate solutions for achieving interoperability. Other solutions, other than standards, should be studied. On that track, many developments are expected to take place in the near future. An example is the OMG MDMI (Model Driven Message Interoperability) initiative (OMG, 2010) that defines transaction data transformations or “maps”. These maps are computer readable and unambiguously define and preserve the business payload of any financial message regardless of its original protocol. One of the MDMI pillars is the separation of data structures from business meaning. Usage of MDMI might cause a switch in priority related to quality, focusing on the quality of the MDMI transformation map of a certain standard. Other SSOs have initiatives in the same area, such as The Open Group with UDEF; a framework for enterprise-centric global interoperability (Schuldt, 2011; TOG, 2011). Also more fundamental solutions, such as integration based on Artificial Intelligence (AI), might become available in the future as well.

Currently the whole idea of semantic standardization is that we are able to capture information and business processes of the “human” world into “machine documents” like information models, XML Schema’s, and the current textual descriptions in SBVR or OCL. But what if this assumption is not correct, and there is a mismatch between the “human world” and that what we are able to capture in “machine-related documents”? Our research actually tries to improve the “machine-related documents” to be a better fit to the “human world”. If there is proof of a mismatch between the human world and machine documents, then every quality aspect (even improved) of current standards will still not suffice. Innovative thinking that searches for new solutions might then be more appropriate. New kinds of solutions that do a better job related to the human world are a direction for further research. As for now it is still worthwhile to improve the quality of standards given the potential limitations. Changes in this area are moving slowly. Many industries are still relying on EDI-based standards developed in the eighties. It is unlikely that within the next few decades the current standards will have disappeared from the arena. A focus on the quality of semantic standards will therefore be a safe investment. Without quality enhancement, standardization may become a failing paradigm, as argued by Cargill & Bolin (2007).



## Appendix A

# Problem Survey Design (chapter 3)

The survey consists of a number of general questions, not specifically related to the research question, and are mentioned within Table 49.

General question	Example answer	Reason for question
What standard are you mainly involved in?	SIDES.	Open control question. Respondent is invited to relate to a specific standard.
A short description of the interoperability problem	Exchange of data in the temporary staffing domain.	Context for interpretation of the results.
Would you like to participate in further research regarding the quality of the standard?	Yes/No.	Possibility to validate the results of the next phase of the research.
Remarks	Nice questionnaire.	Other feedback, such as problems with the questionnaire, aspects that are missing, etc.

Table 49 – Overview of general questions within the survey

In order to disambiguate the terms *quality* and *semantic standards* and to usher the respondent into the details of the survey, a general introduction text was added to the survey:

The main research question for this survey has been: Is there, based on the current standards development processes and experienced interoperability and adoption problems, a reason to develop more knowledge about quality of standards?

The scope of the survey has been limited to semantic standards, which include business transaction standards, ontologies, vocabularies, messaging standards, and vertical industry standards. In most cases, these semantic standards are based on XML syntax, but the core of such standards is their description of the meaning of data and process.

Respondents were expected to be involved in the development or maintenance of a semantic standard. The name of the semantic standard is asked for in question 1, as well as the intended purpose (interoperability problem) of the standard.

In all questions the term standard should be read as the standard that you are involved in. Another term often used in the questionnaire is quality. It is defined as a standard's ability to achieve its purpose, in other words, its fitness for achieving semantic interoperability. This implies that quality deals with both intrinsic aspects (the specification) and situational aspects (external environment) of the standard.



# Survey Results (chapter 3)

Table 50, Table 51, Table 52 & Table 53 show the survey results for each of the four research issues, respectively.

<b>1. Statements about the current standards development process</b>						
	Strongly disagree	Disagree	Partly disagree partly agree	Agree	Strongly agree	Rating Average
Quality assurance is an explicit part of our current development process of the standard.	0.0% (0)	4.2% (2)	18.8% (9)	35.4% (17)	41.7% (20)	4.15
There is not a minimum quality check in place before the standard is released.	52.1% (25)	29.2% (14)	8.3% (4)	8.3% (4)	2.1% (1)	1.79
An instrument/tool is used to measure the quality of our standard.	10.4% (5)	29.2% (14)	18.8% (9)	20.8% (10)	20.8% (10)	3.13
The quality of the current standard can be improved.	0.0% (0)	8.3% (4)	27.1% (13)	50.0% (24)	14.6% (7)	3.71

Table 50 – Survey results - current situation

<b>2. Statements about the interoperability problem</b>						
	Strongly disagree	Disagree	Partly disagree partly agree	Agree	Strongly agree	Rating Average
New or updated releases cover avoidable corrections to the previous versions of our standard.	10.4% (5)	14.6% (7)	29.2% (14)	35.4% (17)	10.4% (5)	3.21
The achieved interoperability is worse than expected.	18.8% (9)	52.1% (25)	18.8% (9)	10.4% (5)	0.0% (0)	2.21
Currently the achieved interoperability is affected by the limitations of our standard.	12.5% (6)	39.6% (19)	29.2% (14)	14.6% (7)	4.2% (2)	2.58
Improvements to the quality of our standard will lead to improved implementations and ultimo lead to improved interoperability.	2.1% (1)	6.3% (3)	25.0% (12)	54.2% (26)	12.5% (6)	3.69

Table 51 – Survey results - interoperability problem

<b>3. Statements about the adoption problem</b>						
	Strongly disagree	Disagree	Partly disagree partly agree	Agree	Strongly agree	Rating Average
The current adoption is better than expected.	2.1% (1)	25.0% (12)	33.3% (16)	35.4% (17)	4.2% (2)	3.15
Design choices of the standard have influenced the adoption process.	4.2% (2)	20.8% (10)	14.6% (7)	50.0% (24)	10.4% (5)	3.42
The adoption will be more successful when the quality of our standard is explicitly known to the user, or proven sufficient or improved.	10.4% (5)	27.1% (13)	25.0% (12)	35.4% (17)	2.1% (1)	2.92

Table 52 – Survey results - adoption problem

<b>4. Statements about the desired quality situation</b>						
	Strongly disagree	Disagree	Partly disagree partly agree	Agree	Strongly agree	Rating Average
A minimum quality level of our standard is needed to achieve interoperability.	0.0% (0)	6.3% (3)	4.2% (2)	56.3% (27)	33.3% (16)	4.17
A minimum quality level of our standard is needed for high adoption rates.	0.0% (0)	8.3% (4)	8.3% (4)	60.4% (29)	22.9% (11)	3.98
I will not use an instrument/tool to measure the quality of our standard when it will be available.	27.1% (13)	54.2% (26)	14.6% (7)	4.2% (2)	0.0% (0)	1.96
It would be helpful to have an instrument/tool/knowledge to gain insight in the quality of our standard.	0.0% (0)	8.3% (4)	6.3% (3)	68.8% (33)	16.7% (8)	3.94
If the quality of the standard is not known than it is hard to improve the standard.	2.1% (1)	12.5% (6)	31.3% (15)	37.5% (18)	16.7% (8)	3.54

Table 53 – Survey results - desired solution

Table 54 shows the results of a correlation analysis. On both axes the sixteen statements are projected. The table shows the Pearson correlation coefficient (P) for each pair of questions. This correlation coefficient measures the strength of a linear relation between two variables. A correlation between two variables is significant when the significance probability (p-value) equals or is smaller than the level of significant  $\alpha$  (Bhattacharyya & Johnson, 1977). The correlation analysis has been done with an  $\alpha$  of .05 and .01. Correlations which were found highly significant at the .01 level are indicated with \*\*, and correlations which were found significant at the .05-level are indicated with \*. Where significance is proven by the p-value, the correlation coefficient (P) represents the strength of the linear relation between the two variables. When the correlation coefficient (P) is negative (-), a negative relation is found, meaning that a higher score of one of the two variables is connected to a lower score of one of the other the two variables.

Correlations		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Quality assurance is an explicit part of our current development process of the standard.	P	-0.57**	0.57**	-0.21	0.08	-0.32*	-0.12	0.15	0.26	0.14	-0.03	0.15	0.06	-0.40**	-0.02
2	There is not a minimum quality check in place before the standard is released.	P	0.00	0.16	0.61	0.03	0.42	0.32	0.07	0.35	0.83	0.31	0.66	0.00	0.90	0.90
3	An instrument/tool is used to measure the quality of our standard.	P	1	-0.39**	-0.19	0.23	0.25	0.00	-0.14	-0.03	0.08	-0.09	-0.01	0.38**	0.09	-0.11
4	The quality of the current standard can be improved.	P	0.01	0.24	0.19	0.11	0.08	0.98	0.33	0.81	0.59	0.56	0.97	0.01	0.54	0.44
5	New or updated releases cover avoidable corrections to the previous versions of our standard.	P	1	0.03	0.29*	-0.17	-0.16	-0.25	0.33*	0.14	0.14	0.29*	0.18	-0.26	-0.22	0.06
6	The achieved interoperability is worse than expected.	P	0.82	0.05	0.25	0.27	0.09	0.02	0.34	0.34	0.34	0.05	0.22	0.07	0.12	0.68
7	Currently the achieved interoperability is affected by the limitations of our standard.	P	1	0.29*	0.32*	0.18	0.08	0.11	0.08	0.11	0.04	0.19	0.04	0.09	0.11	0.04
8	Improvements to the quality of our standard will lead to improved implementations and ultimo lead to improved interoperability.	P	0.05	0.03	0.22	0.59	0.44	0.76	0.20	0.77	0.56	0.44	0.80	0.80	0.80	0.80
9	The current adoption is better than expected.	P	1	0.19	0.19	-0.12	0.20	0.07	-0.16	0.03	-0.09	-0.06	0.06	0.06	0.02	0.09
10	Design choices of the standard have influenced the adoption process.	P	0.20	0.40	0.18	0.63	0.28	0.83	0.56	0.67	0.69	0.92	0.56	0.56	0.56	0.56
11	The adoption will be more successful when the quality of our standard is explicitly known to the user or proven sufficient or improved.	P	1	0.24	-0.11	-0.04	-0.09	0.25	-0.08	-0.02	0.17	0.05	0.09	0.09	0.09	0.09
12	A minimum quality level of our standard is needed to achieve interoperability.	P	0.10	0.46	0.80	0.80	0.52	0.09	0.57	0.87	0.25	0.72	0.55	0.55	0.55	0.55
13	A minimum quality level of our standard is needed for high adoption rates.	P	1	0.12	-0.09	-0.17	0.14	-0.10	-0.06	0.22	-0.09	0.08	0.08	0.13	0.55	0.59
14	I will not use an instrument/tool to measure the quality of our standard when it will be available.	P	0.44	0.54	0.25	0.34	0.51	0.68	0.13	0.55	0.59	0.59	0.59	0.59	0.59	0.59
15	It would be helpful to have an instrument/tool/knowledge to gain insight in the quality of our standard.	P	1	0.11	0.05	-0.03	0.14	0.14	0.01	0.30*	0.20	0.20	0.20	0.20	0.20	0.20
16	If the quality of the standard is not known then it is hard to improve the standard.	P	0.44	0.72	0.84	0.33	0.33	0.33	0.93	0.04	0.16	0.16	0.16	0.16	0.16	0.16
		P	1	0.48**	0.42**	0.23	0.20	-0.32*	0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
		P	0.00	0.00	0.11	0.17	0.03	0.77	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78
		P	1	0.31*	0.27	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26
		P	0.03	0.06	0.08	0.15	0.69	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
		P	1	0.27	0.22	-0.03	-0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
		P	0.06	0.14	0.84	0.35	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
		P	1	0.54**	-0.31*	0.09	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21
		P	0.00	0.03	0.54	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
		P	1	0.10	0.17	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
		P	0.48	0.24	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78
		P	1	-0.11	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22
		P	0.44	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
		P	1	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37

Table 54 – Correlation analysis



# Literature Search Strategy (chapter 4)

### The top journals and search engines

The nature of the research aimed at identification of research gap dictates that all top 25 IS journals and management journals must be included; therefore it is important to determine which search engines cover these top 25 journals. Schwartz & Russo (2004) produced such an overview, although the search engines significant changed since then. So, we re-examined the coverage of these search engines across the top fifty CS/IS journals (Mylonopoulos & Theoharakis, 2001) and the top thirty of International Business Journals (DuBois & Reeb, 2000). The search engines INSPEC, ACM DB and Ei Compendex as mentioned by Schwartz have been excluded because no published list of accessible journals was available. The results of this comparison, limited to the first 25 journals of both categories, are shown in Table 58.

Based on the coverage and availability, the decision was made to use Scopus and EBSCO as search engine. Three of the journals are not covered by either of these two search engines. These journals have been manually searched from the year 2000 until April 2009. These journals are:

- Database
- International journal of management
- Advances in international banking

Scopus and EBSCO contain partly the same journals. In practice only the journals not covered by Scopus were searched using EBSCO. These are:

- Communications of the AIS
- International studies of management and organization
- Multinational business review
- The international journal of accounting

### Using keywords

The keywords for performing the research are of crucial value. We selected seven keywords, determined by means of consults between peers, brain storming and by examining known literature regarding the subject. Then, we searched for synonyms, in its broadest sense, including different words in the same vein. Table 55 summarizes these results.

Keywords	Standardization	Development	Adoption	Interoperability	Measurement	Quality	Semantic
Synonyms	Standardisation	Process	Compatibility	Interoperable	Measuring		Vertical
	Standard	Developing	Comply	Interorganiz(s)ational	Measur(e)ment		Transaction
		Organiz(s)ation	Compliance	Compatibility	Measure		Exchange
		Consortium		Compatible	Metric		Domain
		Consortia		Integration	Criteria		e-Business
						Industry	

Table 55 – Overview keywords and other words in the same vein

The next step was to create regular expressions for capturing different forms and spelling of words, by using wildcards. For instance, *interoperab\** will yield results for *interoperability* but also for *interoperable*. Table 57 shows the expression used as query for each base term within Scopus and EBSCO.

### The combination of keywords

The keywords will be combined during the search process. The keyword “Standardization” is so important that it was decided to include it in every query. Table 56 contains the two-, three- and four-word combinations we have used. The table should be read so that every keyword stands for its corresponding query listed in Table 57.

Standardization Development	Standardization Interoperability Adoption Development
Standardization Adoption	Standardization Measurement Adoption Development
Standardization Interoperability	Standardization Quality Adoption Development
Standardization Measurement	Standardization Semantic Adoption Development
Standardization Quality	Standardization Measurement Interoperability Development
Standardization Semantic	Standardization Quality Interoperability Development
Standardization Adoption Development	Standardization Semantic Interoperability Development
Standardization Interoperability Development	Standardization Quality Measurement Development
Standardization Measurement Development	Standardization Semantic Measurement Development
Standardization Quality Development	Standardization Semantic Quality Development
Standardization Semantic Development	Standardization Measurement Interoperability Adoption
Standardization Interoperability Adoption	Standardization Quality Interoperability Adoption
Standardization Measurement Adoption	Standardization Semantic Interoperability Adoption
Standardization Quality Adoption	Standardization Quality Measurement Adoption
Standardization Semantic Adoption	Standardization Semantic Measurement Adoption
Standardization Measurement Interoperability	Standardization Semantic Quality Adoption
Standardization Quality Interoperability	Standardization Quality Measurement Interoperability
Standardization Semantic Interoperability	Standardization Semantic Measurement Interoperability
Standardization Quality Measurement	Standardization Semantic Quality Interoperability
Standardization Semantic Measurement	Standardization Semantic Quality Measurement
Standardization Semantic Quality	

Table 56 – The 41 combinations of keywords

	Used query expression
Standardization	Standard*
Development	Develop* OR Process OR Organi?ation OR Consorti*
Adoption	Adoption OR Compatibility OR Comply OR Compliance
Interoperability	Interoperab* OR Interorgani?ational OR Compatib* OR Integration
Measurement	Measur* OR Metric OR Criteri*
Quality	Quality
Semantic	Semantic OR Vertical OR Transaction OR Exchange OR Domain OR e-Business OR Industry

Table 57 – The query expression for each keyword

<i>Top 25 CS/IS journals</i>	ISSN	Scopus	Web of Science	Ingenta	ABI/Inform	EBSCO Business
1 MIS Quarterly: Management Information Systems	0276-7783	x	x		x	x
2 Communications of the ACM	0001-0782	x	x		x	x
3 Information Systems Research	1047-7047	x	x		x	x
4 Journal of Management Information Systems	0742-1222	x	x		x	x
5 Management Science	0025-1909	x	x		x	x
6 IEEE Transactions on Computer sciences ( <i>various</i> )	e.g. 0018-9340	x	x			x
7 Harvard Business Review	0017-8012	x	x		x	x
8 Decision Sciences	0011-7315	x	x	x	x	x
9 Decision Support Systems	0167-9236	x	x	x	x	x
10 Information and Management	0378-7206	x	x	x	x	x
11 European Journal of Information Systems	0960-085X	x	x	x	x	x
12 MIT Sloan Management Review	1532-9194	x	x		x	x
13 ACM Transactions on Database Systems ( <i>various</i> )	e.g. 0362-5915	x	x		x	x
14 Data Base	0095-0033					
15 Organization Science	1047-7039	x	x		x	x
16 Information Systems Journal	1350-1917	x	x	x	x	x
17 Academy of Management Journal	0001-4273	x	x		x	x
18 Communications of the AIS	1529-3181				x	x
19 IEEE Computer (Graphics and Applications) ( <i>various</i> )	e.g. 0272-1716	x	x			x
20 Journal of Strategic Information Systems	0963-8687	x	x	x		x
21 Administrative Science Quarterly	0001-8392	x	x		x	x
22 Academy of Management Review	0363-7425	x	x		x	x
23 International Journal of E-Commerce	1086-4415	x	x		x	x
24 ACM Computing Surveys	0360-0300	x	x		x	x
25 Accounting, Management and Information Technologies	0959-8022	x		x		
25 Information and Organization	1471-7727	x		x		x
<b>Total</b>		<b>23</b>	<b>22</b>	<b>7</b>	<b>20</b>	<b>23</b>

<i>Top 25 International Business Journals</i>		ISSN	Scopus	Web of Science	Ingenta	ABI/Inform	EBSCO Business
1	Journal of international business studies	0047-2506	x	x	x	x	x
2	Management international review	0938-8249	x	x		x	x
3	Journal of world business	1090-9516	x	x	x	x	x
4	International marketing review	0265-1335	x	x	x	x	x
5	Journal of international marketing	1069-031X	x	x		x	x
6	International business review	0969-5931	x	x	x		
7	International studies of management and organization	0020-8825				x	x
8	Journal of global marketing	0891-1762	x	x		x	x
9	International journal of research in marketing	0167-8116	x	x	x	x	x
10	Advances in international comparative management	0747-7929	x			x	
11	Advances in international marketing	1474-7979	x			x	x
12	Journal of international financial management and accounting	0954-1314	x		x	x	x
13	Multinational business review	1525-383X				x	x
14	Advances in international accounting	0897-3660	x			x	
15	International trade journal	0885-3908	x			x	x
16	International management	0020-7888	x			x	
17	The international journal of accounting	0020-7063			x	x	x
18	International journal of management	0813-0183				x	
19	Global finance journal	1044-0283	x		x	x	x
20	Journal of international management	1075-4253	x	x	x		
21	Thunderbird international business review	1096-4762	x			x	x
22	Journal of international consumer marketing	0896-1530	x			x	x
23	Advances in international banking and finance						
24	International journal of conflict management	1044-4068	x	x	x	x	x
25	International journal of finance	1076-9307	x	x		x	x
		Total	20	11	10	22	18

Table 58 – Overview Top journals and search engines

### Search process

The search has been performed on title, abstract and keywords (only SCOPUS). Searches within the top journals are conducted by means of the ISSN numbers of those journals and is performed during March and April of 2009. The combination of three and four keywords created a large, but manageable number of studies. Only three keyword searches that yield more than two hundred results were refined by adding a fourth keyword. To make sure to include the core studies a two keyword search has been performed, with the following additional rules:

- Studies from year 2000 until April 2009 that have been cited more than five times are included.
- Studies before year 2000 cited more than fifty times are included.

## Appendix D

# Overview of Selected Studies (chapter 4)

Authors	Title	Journal	Year	Topic	Lifecycle	View	Type	Approach	Method
Albrecht, C. C., D. L. Dean, et al.	Marketplace and technology standards for B2B e-commerce: Progress, challenges, and the state of the art.	Information and Management	2005	Interoperability	-	User	Syntactical	Critical	Review
Backhouse, J., C. Hsu, et al.	A question of trust.	CACM	2005	PSI	Use	User	Syntactical	Critical	Case Study
Bernstein, P. A. and L. M. Haas	Information integration in the enterprise.	CACM	2008	Interoperability	-	Implementer	Syntactical	Descriptive	Review
Boh, W. F. and D. Yellin	Using enterprise architecture standards in managing information technology.	JMIS	2006	PSI	Use	User	All	Positivist	Data Analysis
Boh, W. F., C. Soh, et al.	Standards development and diffusion: A case study of RosettaNET.	CACM	2007	Lifecycle	Use	Creator	Vertical	Descriptive	Case Study
Cathomen, I. and S. Klein	The development of FEDi in Switzerland: A life-cycle approach.	IJEC	1996	Lifecycle	Use	User	Vertical	Descriptive	Case Study
Chari, K. and S. Seshadri	Demystifying integration.	CACM	2004	Interoperability	-	Implementer	All	Descriptive	Development
Chen, H. M. and P. J. Sheldon	Destination Information Systems: Design Issues and Directions.	JMIS	1997	Interoperability	-	Implementer	All	Descriptive	Development
Chen, P. Y. and C. Forman	Can vendors influence switching costs and compatibility in an environment with open standards?	MISQ	2006	PSI	Use	Implementer	Syntactical	Descriptive	Data Analysis
Damsgaard, J. and D. Truex	Binary trading relations and the limits of EDI standards: The Procrustean bed of standards.	EJIS	2000	PSI	Use	Implementer	Horizontal	Critical	Conceptual
De Bruijn, J., D. Fensel, et al.	Using the web service modeling ontology to enable semantic e-business.	CACM	2005	Interoperability	Use	Implementer	Syntactical	Descriptive	Conceptual
Dogac, A., Y. Kabak, et al.	Collaborative business process support in eHealth: Integrating IHE profiles through ebXML business process specification language.	IEEE TITB	2008	PSI	-	Implementer	Syntactical	Descriptive	Conceptual
Eichelberg, M., T. Aden, et al.	A survey and analysis of electronic healthcare record standards.	ACM Computing Surveys	2005	Quality	-	Implementer	Vertical	Descriptive	Review
Elgarah, W., N. Falaleeva, et al.	Data exchange in interorganizational relationships: review through multiple conceptual lenses.	ACM SIGMIS	2005	PSI	Use	User	Horizontal	Descriptive	Review
Fodor, O. and H. Werthner	Harmonise: A step toward an interoperable e-tourism marketplace.	IJEC	2004	Interoperability	-	User	All	Interpretive	Development
Frenkel, K. A.	Politics of standards and the EC.	CACM	1990	PSI	-	Policy Maker	Syntactical	Interpretive	Review
García, R. G. and E. Gelle	Applying and adapting the IEC 613-46 standard to industrial automation applications.	IEEE TII	2006	Lifecycle	Implement	Implementer	Vertical	Descriptive	Development



Authors	Title	Journal	Year	Topic	Lifecycle	View	Type	Approach	Method
Thissen, W. A. H. and W. J. Stam	Electronic data interchange in an industrial sector: The case of The Netherlands' building industry.	Information and Management	1992	Lifecycle	Develop	Creator	Vertical	Descriptive	Conceptual
Venkatraman, S., H. Bala, et al.	Six strategies for electronic medical records systems.	CACM	2008	Interoperability	-	User	All	Interpretive	Review
Wang, E. T. G. and A. Seidmann	Electronic data interchange: competitive externalities and strategic implementation policies.	Management Science	1995	Lifecycle	Use	User	Horizontal	Interpretive	Conceptual
Weitzel, T., D. Beimborn, et al.	A unified economic model of standard diffusion: The impact of standardization cost, network effects, and network topology.	MISQ	2006	Lifecycle	Use	User	Horizontal	Positivist	Data Analysis
Wigand, R. T., C. W. Steinfield, et al.	Information technology standards choices and industry structure outcomes: The case of the U.S. home mortgage industry.	JMIS	2005	PSI	Use	User	Vertical	Descriptive	Case Study
Wybo, M. D. and D. L. Goodhue	Using interdependence as a predictor of data standards. Theoretical and measurement issues.	Information and Management	1995	PSI	Use	User	Vertical	Positivist	Data Analysis
Zhao, K., M. Xia, et al.	An integrated model of consortium-based e-business standardization: Collaborative development and adoption with network externalities.	JMIS	2007	Lifecycle	Use	User	Vertical	Descriptive	Conceptual
Zhu, K., K. L. Kraemer, et al.	Migration to open-standard interorganizational systems: Network effects, switching costs, and path dependency.	MISQ	2006	Lifecycle	Use	User	Horizontal	Positivist	Data Analysis
zur Muehlen, M., J. V. Nickerson, et al.	Developing web services choreography standards - The case of REST vs. SOAP.	Decision Support Systems	2005	Lifecycle	Develop	Creator	Syntactical	Descriptive	Case Study

Table 59 – Overview of selected studies

## Appendix E

# Set of Requirements for Instrument (chapter 7)

The Volere method emphasizes the rationale and fit criterion of the requirement (Robertson & Robertson, 1999). All of the requirements found in the requirement specification phase are described in the table below. For each requirement the following information, based on Volere, is provided:

- Number: the number of the requirement matches the numbers used in the figures in Chapter 7 and Appendix F.
- Short name: a short description of the requirement, matches the names used in the figures in Chapter 7 and Appendix F.
- Description: a detailed description of the requirement.
- Fit criterion: the criteria to determine whether the requirement is fulfilled.
- Priority: the priority of the requirement, used for choosing between conflicting requirements and when time limitations prevent implementing all the requirements.

#	Short name	Description (what?)	Fit criterion	Priority
A1	Possible to add quality aspects	The instrument should be flexible to add new quality aspects.	The end user should be able to add aspects without dependency on the instrument designer.	Medium
A2	Possible to add new indicators	The instrument should be flexible to add new indicators for existing quality aspects.	The end user should be able to add indicators without dependency on the instrument designer.	Medium
A3	Possible to add new metrics	The instrument should be flexible to add new metrics to measure existing indicators.	The end user should be able to add metrics without dependency on the instrument designer.	Medium
A4	Possible to choose a metric if more than one is available	The user should have the possibility to choose a metric if more than one is available for measuring an indicator. Depending on the preferences of the user, he could select a rigid but time-consuming metric or a less rigid but ease to determine metric.	The instrument should present the user a choice if more than one metric is available.	High
A5	Possible to personalize the weighing of individual quality aspects	The overall quality of a standard is determined by combining all the individual quality criteria. However, different users may have different opinions to which quality criteria are important. The instrument should allow users to personalize the weighing for all the individual criteria.	The user must be able to personalize the weighing factors himself, without the help of the designer of the instrument.	High
A6	Possible to choose an indicator if more than one is available	The user should have the possibility to choose an indicator if more than one is available for a given quality attribute. Depending on the preferences of the user, he could select a better but time-consuming indicator, or a lesser but easy to determine indicator.	The instrument should present the user a choice if more than one indicator is available.	High

#	Short name	Description (what?)	Fit criterion	Priority
A7	Useable for different types of semantic standards	Semantic standard may vary in content en format, but the instrument should be useable for all semantic standards.	The instrument should be useable for all standards presented on semanticstandards.org.	High
B1	Have transparent outcome	The outcome of the instrument should provide insight on how the results are determined. To do this, the instrument must relate quality aspects to attributes of the standard.	The outcome of the instrument should contain all applied metrics and weighing factors. For all metrics that require human interpretation, an explanation must be provided.	Medium
B2	Have an outcome summary that fits on one page (but is more than a single rate)	In order to be useable by the user of the instrument, the outcome summary should contain no more than one page.	Summary of outcome maximum of one page A4 size using font size 10.	High
B3	Contain no more than 7 tests	Contain no more than 7 tests (compare with car testing: city, snow, dessert, test track, long ride, etc).	The number of tests performed by the instrument should be no more than 7 for one single standard.	Low
B4	Have standard templates for weighing factors	The instrument should have "standard" templates for users who do not wish to tailor the weighing factors to their own need.	It must be able to use the instrument without spending any time on determining weighing factors.	Medium
B5	Have automated measurements when possible (by machine reading)	To make the instrument as easy as possible to use, the instrument should perform automated measurements when possible.	All metrics that can be determined by machine reading should require no human interaction.	High
B6	Contain clear guidelines on how to use	The instrument should be easy to use, and therefore contain clear guidelines.	A guidelines document should be available.	High
B7	Instrumental, a "Tool"	The instrument should be practical useful by being implemented as tool.	All parts of the instrument should be covered by physical or software products.	High
B8	Useable to identify blank spots in work in progress	The instrument should not only be useable for determining the quality of finished standards, but also give improvement suggestions when used on work in progress.	In the results the blank spots are presented.	Low
B9	Facilitate testers	The instrument should be useable by testers that are implementing a (draft) standard.	The instrument should make it possible to test parts of the standard.	Low
B10	Measure complete standards as well as individual parts	The user should be able to use the instrument not only on a complete standard, but also on parts of the standard.	When combining test of individual parts of a standard, 90% of the standards should have less than 10% deviation from the testing of the complete standard.	Medium
B11	Support scenario assertions	The instrument should support scenario assertions, "what if...".	It should be possible to use at least two scenarios (minimum and maximum).	Low
B12	Measure one individual standard	The instrument should take one individual standard (or a part) as input.	The instrument should never require a second standard to be used.	Medium

C1	Make a distinction between the standard and its presentation form	One can distinguish a standard (set of agreements) and the presentation (usually a document). Some standards are presented in different forms (e.g. different languages). The instrument should give insight in whether a quality measurement is done on the representation (document) or the standard.	For each measurement and attribute it must be clear whether the standard or the presentation was subject of investigation.	Low
C2	Useable to rank standards	The outcome of the instrument should be useable to compare and rank two or more standards (this standard is better than that standard).	The outcome of the instrument should also include one score on the scale of 1 to 10 (latter is better).	Medium
C3	Have an outcome that contains improvement suggestions	The instrument should not only return the quality of the standard, but also suggestions to improve the standard.	For each standard that has a quality score less than 10, the instrument should return at least one improvement suggestion.	Medium
C4	Have outcome that is specific enough for appliance	The improvement suggestions should be specific enough for the user to apply on the standard, without having to consult an experienced user.	When improvement suggestions are processed by 5 independent users, 4 out of 5 should make the same changes to the standard.	Medium
C5	Have standardized input and output that conforms to a set of minimal requirements	In order to process standards in a comparable way, the input should conform to a minimum set of requirements. When conforming to the minimum set of requirements, the output should also conform to a set of requirements.	For 5 standards that comply to the input requirements, at least 4 of the outcomes comply to the minimum set of requirements.	Medium
C6	Have a sound fundament	The result of the instrument should not be easy to devaluate, therefore the instrument should have a sound, theoretical fundament.	The model behind the instrument should be supported by at least one scientific theory.	High
C7	Have well described and unambiguous indicators and metrics	The indicators and metrics shall be well described and unambiguous.	When asking users to explain the indicators and metrics, 4 out of 5 users give the same explanation for at least 90 percent of the indicators and metrics.	High
C8	Have an objectively determinable metric for each indicator	Each indicator has at least one metric that can be determined objectively.	When 5 independent users test a standard, at least 90 percent of the metrics shall score within a 10 percent margin.	High
C9	Have an outcome that indicates the principal and his involvement	In order to determine the objectivity of the outcome of the instrument, the principal and his role in the standard (development process) should be known.	The outcome should always include the principal and his involvement.	Medium
C10	Have an outcome that indicates the source material used for the testing	The outcome of the instrument should always indicate all the source material (documents) that is used for the testing.	The outcome should always indicate the source material.	Medium
C11	Have an outcome that shows the scoring quality aspects and applied weight factor	The outcome of the instrument should provide insight on how the results are determined. To do this, the instrument must relate quality aspects to attributes of the standard.	The outcome of the instrument should contain all applied metrics and weighing factors. For all metrics that require human interpretation, an explanation must be provided. (Similar to B1)	Medium

#	Short name	Description (what?)	Fit criterion	Priority
C12	Return improvement suggestions that lead to a higher score	After processing the improvement suggestions given by the instrument, testing the standard should lead to a higher score.	When using the instrument on a standard that is not yet finished, a second test after applying the improvement suggestions should return a higher score.	High
C13	Contain interpretation explanation of measurement results	The outcome of the instrument should be easy to interpret, and therefore contain an explanation of the results.	Each of the score of an quality attribute should contain an explanation.	Medium
C14	Have an outcome that addresses different aspects of a standard	In order to give a complete view on the quality of a standard, all quality aspects that are important to the user should be measured.	Each attribute that has a weighing factor that is larger than 0, should be assessed during the testing.	Medium

Table 60 – Requirements specified

## Appendix F

# Requirements Goal Tree (chapter 7)

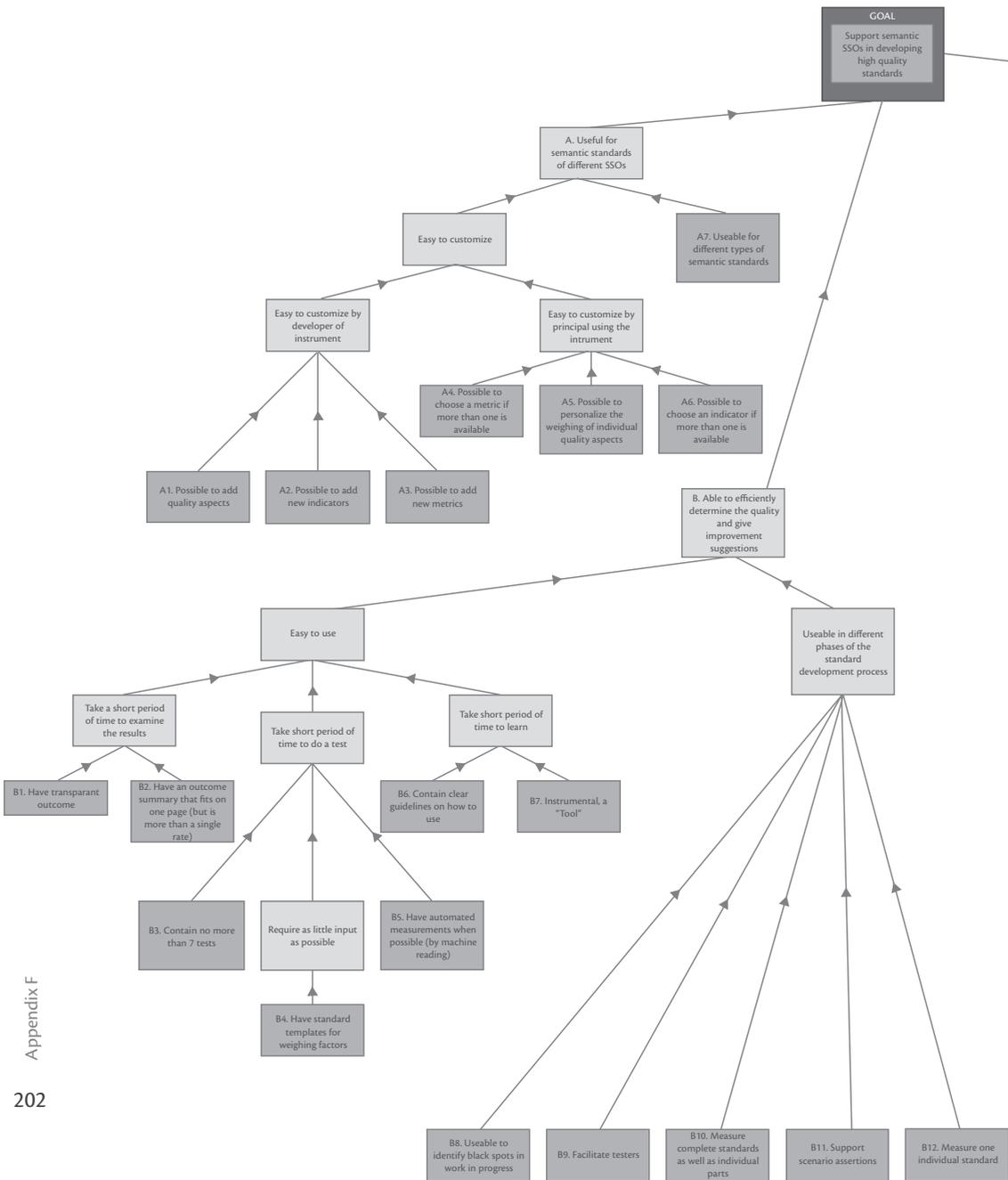
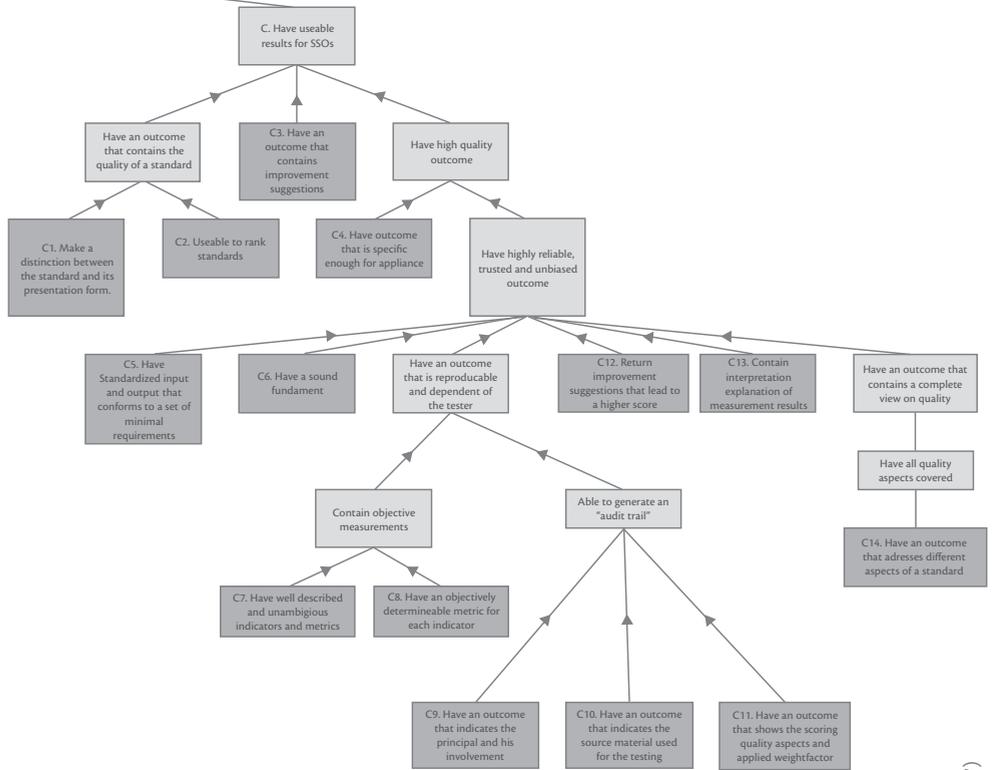


Figure 54 – Complete requirements goal tree





## Appendix G

# Graphical Presentation of SSM (chapter 9)

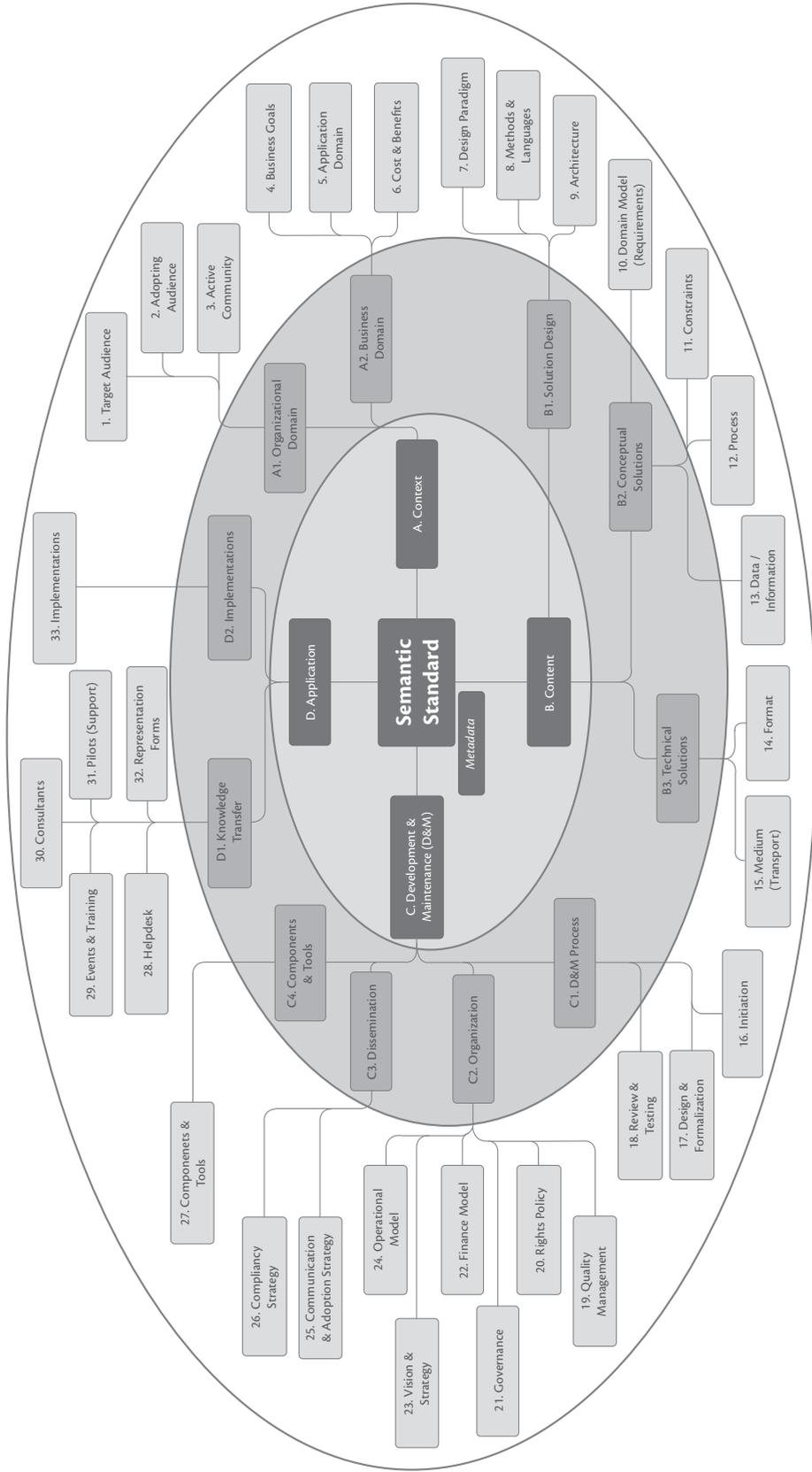


Figure 55 – Graphical presentation of SSM

## Appendix H

# Existing Models for Semantic Standards (chapter 9)

Approach/ Model	A. Evaluation Framework (Mykkanen & Tuomainen, 2008)	B. Reference Model Analysis Grid (RMAG) (Pawlowski & Kozlov, 2010)	C. Comparative Analysis (Nelson et al., 2005)	D. Evaluation Taxonomy (Lampathaki et al., 2009)
<b>Purpose</b>	Evaluation of a specific (or relative comparison) interoperability standard(s).	Assessing, testing and validation of standards for achieving interoperability in the education domain.	Analysis and comparison of multiple vertical standards.	To classify business transaction standards based on taxonomy related information of a standard.
<b>Categories and concepts included</b>	<ul style="list-style-type: none"> <li>A.1. Overview</li> <li>A.2. Information and semantics</li> <li>A.3. Functionality and interactions</li> <li>A.4. Application infrastructure</li> <li>A.5. Technical aspects</li> <li>A.6. Flexibility, accuracy, extensibility</li> <li>A.7. Maturity, usage, official status</li> <li>A.8. System life-cycle</li> <li>A.9. Domain-specific features</li> </ul>	<ul style="list-style-type: none"> <li>B.1. General data</li> <li>B.2. Objectives</li> <li>B.3. Domain</li> <li>B.4. Classification</li> <li>B.5. Methodology</li> <li>B.6. Evaluation</li> <li>B.7. Processes</li> <li>B.8. Levels</li> <li>B.9. Usage and validation</li> <li>B.10. Documentation</li> <li>B.11. In-depth analysis:               <ul style="list-style-type: none"> <li>a. Practical interoperability and credibility</li> <li>b. Semantic interoperability</li> <li>c. Integration: communication</li> <li>d. Bindings and conformance</li> <li>e. Integrity</li> </ul> </li> <li>B.12. Quality attributes:               <ul style="list-style-type: none"> <li>a. Seamlessness</li> <li>b. Adaptation</li> <li>c. Document guidance</li> <li>d. Compatibility</li> <li>e. Extensibility, adaptation</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>C.1. Context (Industrial Group, URL, Profit Orientation / Partnerships, Membership fee structure, Industry Participation, Decision Making, Standards Availability, Members, Year Incepted)</li> <li>C.2. Choreograph and modularity:               <ul style="list-style-type: none"> <li>a. High-level processes</li> <li>b. Specification sets</li> </ul> </li> <li>C.3. Prioritize and schedule:               <ul style="list-style-type: none"> <li>a. Decision-making</li> <li>b. Update</li> </ul> </li> <li>C.4. Standardize and document               <ul style="list-style-type: none"> <li>a. Typical specification sets</li> <li>b. Sub-committee structure</li> </ul> </li> <li>C.5. Reviews and tests</li> <li>C.6. Implement and deploy</li> <li>C.7. Compliance and certification</li> </ul>	<ul style="list-style-type: none"> <li>D.1. Scope</li> <li>D.2. Completeness</li> <li>D.3. Compatibility (with other standards)</li> <li>D.4. Openness</li> <li>D.5. Customization capabilities (modularity, expandability, composability)</li> <li>D.6. Maturity</li> <li>D.7. Standard support</li> <li>D.8. Ease of use and of implementation</li> <li>D.9. Modeling of messages</li> <li>D.10. Integrated management of enterprise and data models</li> <li>D.11. Configuration Management (versioning, backwards compatibility)</li> <li>D.12. Additional features (support for rules modeling, workflow capabilities, incorporated into the documents)</li> </ul>

Table 61 – Existing models for semantic standards

# Example of a Semantic Standard Classification in Case A (chapter 9)

### cfiXML

General characteristics	
1. Name of the standard	cfiXML (capital facilities industry XML)
2. Functional and Organizational domain	Worldwide standard for data exchange for the capital facilities industry: building and maintenance of large utility buildings (offices, shops, hospitals), industry plants and (technical) facilities.
3. Website	<a href="http://www.cfixml.org/">http://www.cfixml.org/</a>
The standard supports and aligns the information exchange between the stakeholders in the capital facilities industry, thereby considerably saving costs in the long run, according to NIST research.	
Current use ("usability" in terms of Dutch "Comply or Explain" policy)	
4. Usage on national/European/international level	International, primarily US.
5. Start date	2004; First public release.
6. Important stakeholders	The so-called sponsors; consortiums (FIATECH) and associations (DIPPR) that are representing owners, engineering procurement and construction contractors, technology suppliers, equipment suppliers, universities and research organizations.
7. Number of participants	Broad application of the standard in the United States, including several international companies (among them Royal Shell).
The standard is particularly used by American companies. Moreover, some sectors use it more intensively than others, for example the oil and gas industry.	
Potential ("potential" in terms of Dutch "Comply or Explain" policy)	
8. Level of adoption	Adoption is limited to the US and to some internationally operating companies.
9. Contribution to economic and social goals	The standard contributes to economic goals by lowering costs within the value chain.
10. Contribution to reducing the administrative burden or to improved inspection	This is an industry standard with hardly any relation to the government. It does reduce the administrative burden, but on a B2B level rather than on a B2G level.
The standard has economic potential for the Netherlands and might result in a more competitive building and maintenance sector by reducing the administrative work.	

**Development and Maintenance (“openness” in terms of Dutch “Comply or Explain” policy)**

11. Profit/non-profit orientation of organization	Non-profit.
12. Finance model	Payments of both sponsors and participating software companies.
13. Participation model	Everybody can participate.
14. Decision model / Governance	Not known (probably sponsors).
15. Availability	Open-source model is used, including minimal support. Information is freely available.

Although the governance structure is not known, the standard builds on volunteers for its development, and the sponsors seem to heavily influence and direct the standard.

**Conclusions:**  
Usage of the standard by the Dutch government might have a flywheel effect for usage within the Dutch building industry. The standard brings in a solution to B2B interoperability problems related to complex projects, involving different stakeholders, including designers and providers of equipment and maintenance. For the big players in the utility construction field (including the government) the potential savings are huge. On the other hand, the potential regarding B2G, and in particular specific government tasks, seems limited.  
Further exploration involving the different stakeholders is recommended, especially focusing on the benefits of this standard for the stakeholders.

## Appendix J

# Example of a Semantic Standard Evaluation in Case B (chapter 9)

## ACORD Framework

Evaluation criteria	Value
Version	Dictionary – Published, 2008. Capability Model – Published, 2007; ver. 2.0 published November, 2009. Information Model – Initial release, August, 2009, update 2.0 in 2001. Data Model 1.x – Incremental releases in 2010. Component Model – Start in 2010. AIM Harmonization with IBM BOM – 2010.
Description	The ACORD Framework represents a single streamlined business model for standards creation that is flexible enough to cross lines of business and geographic borders.
Registered standard	-
Standardization Organization	-
Industry	Insurance industry.
Scope	Clinical data, person data, product data.
Developer	ACORD (Association for Cooperative Operations Research and Development).
Origin / Contributions	IBM donated Insurance Application Architecture (IAA) Business Object Model (BOM) November 2009.
Link	<a href="http://www.acord.org/">http://www.acord.org/</a>
Content	<ul style="list-style-type: none"> <li>• Business Dictionary: Single business glossary to bridge communication gaps</li> <li>• Component Model with Services Maps</li> <li>• Capability Model with Process Maps: Reflects the industry's usual way of doing business. The Model offers an organizational baseline, a preferred approach</li> <li>• Information Model: provides relationships among insurance concepts, such as Policy, Product, Party, and Claims. It currently contains more than 1000 classes and 2000 attributes. It can be used to jump start application development, consume ACORD XML messages, and as a semantic model for integration, among other uses.</li> <li>• Data Model: logical level entity-relationship model. "Logical level" implies that it can be used in any database implementation. Some of the many uses of the ACORD Data Model include creating physical data models, data warehouses, or to validate your own data models.</li> <li>• Differences Data Model / Information Model <ul style="list-style-type: none"> <li>○ Same content, different formats: <ul style="list-style-type: none"> <li>▪ Information Model –Unified Modeling Language (UML)</li> <li>▪ Data Model –IBM InfoSphereData Architect, Computer Associates Erwin</li> </ul> </li> <li>○ Different naming conventions</li> <li>○ Added keys (big) to Data Model</li> <li>○ Discriminators added to resolve inheritance structures</li> <li>○ Associative classes added to resolve M:M relationships</li> </ul> </li> </ul>
Structural cardinality	Classes (~1000), Attributes (~2000), Relations, Cardinalities.
Semantic cardinality	Business Dictionary.
Representation	MagicDraw –for users of MagicDraw (that is the tool ACORD uses). UML (XML Metadata Interchange –XMI) for import into UML tools. HTML – for anyone who doesn't have a UML tool.
Predefined content	Component Model: Party, Contract, Product/Agreement, Physical Object, Claim
Extensibility	'plug in' of new data requirements of the standards-setting efforts to the data hierarchy, Extension of the model to support unique requirements is provided.
Integration to other models	ACORD XML for Life, Annuity & Health, Business Message Specification (TXLife), Object Model Specification (XMLife), Tabular Data Specification (XTbML) .
Industry acceptance	Supported / used by IBM.
Tool Support	Mapping to: XML standards (all versions), Forms eLabels, AL3, Other standards and models. MagicDraw, DataXtend Browsers for ACORD (PCS XML v1.15.0 and LAH Standards v2.20.00).
Openness	Working Groups & members, changes may be proposed.
Availability	Membership necessary, usage fee occurs.

## Appendix K

# SSM Grounded by Case Study Research (chapter 9)

#	Literature Reference	#	Literature Reference	#	Literature Reference
1	"All sectors, including .. in their interests." (p.20) [3]	14	"In January 2001 .. mortgage value chain." (p.13) [3]	27	"The eMortgage Implementation Guide .. electronic loan document." (p.13) [3]
2	"As in many .. declined to participate." (p.10) [3]	15	"Swiss banks try .. for customer transactions." (p.7) [2]	28	"Companies participating in .. support and services." (p.2) [1]
3	"Various mortgage industry .. the mortgage industry." (p.11) [3]	16	Not found	29	"We also attended .. conferences and workshops." (p.2) [1]
4	"The second MISMO .. entirely without paper." (p.13) [3]	17	"In this new .. for all departments." (p.8) [4]	30	"For example, Rosettanet .. inexpensive e-hub solutions." (p.5) [1]
5	"Mortgage banker similarly .. in their interests." (p.14) [3]	18	"The validation process .. knowledge sharing." (p.3) [1]	31	"Some standards consortia .. of organizational contexts." (p.3) [1]
6	"The cost of .. of participants' interests." (p.2) [1]	19	"Only council members .. become milestone programs." (p.3) [1]	32	"The eMortgage Implementation Guide .. electronic loan document." (p.13) [3]
7	"A second key .. ensuring valid transactions." (p.19) [3]	20	"MISMO considered and .. approach to IPR." (p.17) [3]	33	"Starting with only .. electronic loan document." (p.13) [3]
8	Not found	21	"MISMO, now a .. Committee at a time." (p.17) (whole paragraph dedicated to <i>governance</i> ) [3]		
9	"MISMO can be .. entirely without paper." (p.11-13) [3]	22	"Each participating organization .. at no charge." (p.17) [3]		
10	"DTA procedures enable .. to be processed." (p.4) [2]	23	"RosettaNet provides a .. of many consortia." (p.6) [1]		
11	"The DTI file .. the SWIFT network." [2]	24	"Non-subscribers can .. in MISMO meetings." (p.17) [3]		
12	"Although much data .. in this industry." (p.10) [3]	25	"These strategies are .. message about RosettaNet." (p.4-5) [1]		
13	"What we have .. that same definition" (p.19) [3]	26	"Some standards consortia .. of organizational contexts." (p.3) [1]		

Table 62 – Attributes from SSM grounded in case study research

The numbering of attributes is the numbering used within Appendix G. The numbers of literature reference corresponds with:

[1] = Boh, W. F., Soh, C., & Yeo, S. (2007). Standards development and diffusion: A case study of Rosettanet. *Communications of the ACM*, 50(12), 57-62.

[2] = Cathomen, I., & Klein, S. (1997). The development of FEDI in Switzerland: A life-cycle approach. *International Journal of Electronic Commerce*, 1(4), 129-145.

[3] = Markus, M. L., Steinfeld, C. W., Wigand, R. T., & Minton, G. (2006). Industry-wide Information Systems standardization as collective action: The case of U.S. residential mortgage industry, *MIS Quarterly*, 30, 439-465.

[4] = Hanseth, O., Jacucci, E., Grisot, M., & Aanestad, M. (2006). Reflexive standardization: Side effects and complexity in standard making. *MIS Quarterly: Management Information Systems*, 30, 563-581.

## Appendix L

# QMSS Build 0.3 (chapter 10)

The table below consists of definitions of all measurable concepts from the quality model, including the lowest level measurable concepts called attributes. This table has been used within the SETU and XCRI cases.

<b>Measurable Concept / Attribute</b>		<b>Definition</b>
1.	Functionality	The capability of the standard to provide functions which meet stated and implied needs when the standard is used under specified conditions.
1.1	Suitability	The capability to provide an appropriate set of functions for specified tasks.
1.1.1	Completeness	The level of covered functions in relation to the interoperability problem.
1.1.2	Compactness	The level in which unnecessary, in relation to the stated and implied needs, functions are not part of the standard.
1.1.3	Specificity	The level in which the standard does solve a specific real life problem in relation to a generic issue.
1.2	Accuracy	The capability of the standard to provide interoperability, with the needed degree of precision.
1.2.1	Match expectations in practice	On functionality the level of implementations that do satisfy the functional needs for which the standard is intended.
1.2.2	Match specification	The level of functions for which a solution is provided in relation to the described needed functionality. Internal consistency.
1.2.3	Level of interoperable implementations	The level of successful implementations, meaning resulting in interoperability.
1.3	Compliance	The capability of the standard to adhere to other standards, conventions, or regulations in laws and similar prescriptions.
1.3.1	Compliance	The compliance level to other standards, conventions, or regulations in laws and similar prescriptions.
2.	Reliability	The capability of the standard to maintain a specified level of performance when used under specified conditions.
2.1	Maturity	The capability of the standard to avoid failure as a result of faults in the standard.
2.1.1	Weak sentences	The level of weak sentences within the specifications.
2.1.2	Release schema	Stability of release schema.
2.1.3	Changes per release	The amount of changes in the releases within a specified time period.
2.1.4	Corrections	The level of number of reported request in relation to the amount of requests that has been processes in a new release.
2.1.5	Availability of documentation	The level of availability of documentation.
2.1.6	Documentation	The quality level of the documentation.
2.1.7	Development Organization	The level of reliability of the Standard Development Organization that maintains the standard.
2.2	Fault tolerance	The capability of the standard to maintain a specified level of performance when a fault occurs.
2.2.1	Manual interference	The level of amount of manual interference in the implementation of the standard.
2.2.2	Correction	The inclusion of correction processes, to automate the process of corrections and to avoid manual interference.
2.3	Consistency	The level of ambiguity within the standard.
2.3.1	Information ambiguity	The level of ambiguity within the information elements.
2.3.2	Function ambiguity	The level of ambiguity within the functions.

3.	Usability	The capability of the standard to be understood, learned, used and attractive to the user, when used under specified conditions.
3.1	Understandability	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.
3.1.1	Structure document	The level of structure of the specification documents.
3.1.2	Completeness descriptions	The level of readers that sufficiently understand the standard by reading the specifications.
3.1.3	Readability	The readability index according to a formula.
3.1.4	Learnability	The learning time needed to understand the standard.
3.1.5	Function recognition	The recognition of experts in the field of the functions and terms used by the standard.
3.2	Implementability	The effort needed to implement the standard.
3.2.1	Implementation experience	Experiences about implementations; the amount of implementations and how successful or not.
3.2.2	Reference examples	The availability of reference examples that can be used for implementation purposes.
3.2.3	Installed base	The connection of the standard with the ICT landscape at stakeholders organizations.
3.2.4	Faulty implementations	What can be learned from faulty implementations, if any.
3.2.5	Business Processes	Does an implementation of the standard require a major change in business processes.
4.	Portability	The capability of the standard to be transferred from one environment to the other.
4.1	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.
4.2	Co-existence	The capability of the standard to co-exist with other independent standards in a common environment.
4.2.1	Other standards	The relationships with other standards.
4.2.2	Modularity	The modular design of the standard.
4.3	Replaceability	The capability of the standard to be used in place of another specified standard for the same purpose in the same environment.
4.3.1	Functional continuance	New version of the standard show functional continuance by supporting the same functions.
4.3.2	Data continuance	New version of the standard show data continuance by supporting the same data definitions.
4.3.3	Replaces standards	Makes other standards superfluous.
5.	Maintainability	The capability of the standard to be modified. Modifications may include corrections, improvements or adaptation to changes in environment, and in requirements and functional specifications.
5.1	Adaptability	The standards possibilities to adapt.
5.1.1	Unplanned changes	Releases of bug fixes that solves major errors within the standard.
5.1.2	Time to change	The time needed for changes to take place, from filing the maintenance request till the release of the standard.
5.1.3	Other standards	Dependability on other standards.
5.1.4	Dynamic content	A design of a standard that put dynamic contents like code lists outside the main specification.
5.1.5	Localizations	The availability and use of localizations and extensions in the main standard.

5.2	Stability	The capability of the standard to avoid unexpected updates of the standard.
5.2.1	Version management	The organization and procedures of version numbering and version management, also in relation with backwards compatibility.
5.2.2	Versions in practice	The actual practice of versioning of the standard.
5.3	Testability	The capability of the standard to be validated.
5.3.1	Test services	The availability of different kind of test services, like a validation service, or helpdesk.
5.3.2	Autonomous testing	The availability of autonomous testing.
5.3.3	Reference implementation	The availability of reference implementations.
5.3.4	Certification	The availability of a certification program.
5.3.5	Test documentation	The description of how to test compliancy to the standard, including a minimal set of requirements for a compliant implementation.
6.	<b>Adoptability</b>	<b>The adoption of the standard within the domain.</b>
6.1	Acceptance of users	The level of adoption of the standard by the users.
6.1.1	Market penetration	Market penetration differentiated per stakeholder, including big fish involvement.
6.1.2	Lifecycle status	The maturity of the standard in terms of lifecycle status.
6.2	Availability tools	The availability of tools that support the implementation of the standard.
6.2.1	Products/Services	Implementations of the standards within products and services to be used by the end user.
6.2.2	Components	The availability of components and tools useful to support implementations.
6.3	Availability support	The availability of knowledge and implementation support about the standard in the market.
6.3.1	Focal point	The focal point for all related issues about the standard.
6.3.2	Support organizations	Organization that have knowledge about the standard or support implementations.
6.3.3	Community	The availability of an online community, Internet fora for discussions about the standard.
7.	<b>Openness</b>	<b>The openness of the standard.</b>
7.1	Openness Process	The level of openness of the development and decision-making processes.
7.1.1	Not-for-profit	The development and maintenance of the standard by a not-for-profit organization.
7.1.2	Meetings	The meetings are organized in a way that do not rule out certain stakeholders.
7.1.3	Decision making	The way decision making is organized. For instance consensus based.
7.1.4	Change procedures	The description of the change procedure of a standard.
7.2	Openness Specification	The level of openness of the specification.
7.2.1	Availability	The availability to download the standard.
7.2.2	Use	The possibilities to use the standard. (costs, IPR, etc)
7.2.3	Re-Use	The possibilities to re-use the standard. (costs, IPR, etc)
7.2.4	Open Source	The availability of implementations of the standard available as open source software.

Table 63 – Definitions of measurable concepts and attributes

# SETU Case Result of Applying SSM (chapter 10)

This section will describe the SETU standard by using the SSM. The SETU standard is characterized by an active community of representatives of the largest temporary staffers (Randstad, USG, Adecco, Manpower amongst others) together with the standardization experts of TNO and support of the sector organization ABU. The adopted audience includes the same temporary staffers but with also several middle-size temporary staffers, a list of staffing customers and software vendors offering solutions mainly for temporary staffing companies. The targeted audience is broader, since it includes every staffing companies and customers, and does also include software vendors from the procurement domain (e.g. SAP) which are currently not part of the active community or adopted audience.

SETU standards are a solution for effective and efficient inter-organizational communication about temporary workers. The application domain is within the Human Resources domain, including temporary staffers, staffing customers and service providers. Privacy laws are applicable, just as standard laws for e-invoicing. The domain is characterized by a handful of large dominant staffing companies, and thousands of niche suppliers regarding their ICT often dependent on software suppliers. The scope is limited to the primary process of selecting & ordering, the assignments, time card reporting and invoicing.

The SETU business case is quite obvious, although difficult to estimate. The main savings are present in the time card process and invoicing. Tools are available to calculate potential savings, but since everybody understands that for each stakeholder the business case is positive there is not much need for further research on the business case.

SETU has chosen for a model based approach, which is documented as development method for SETU standards. The paradigm can be called “message based”, in contrary to “service based”. As much as possible SETU makes use of existing methods and languages, like UML, SBVR, Schematron, XML & XML Schema. Architecture is very important to SETU on different levels. First the relation between the different SETU standards is important and documented. Second the relation with the foundational standard HR-XML is described and continuously monitored. SETU representatives are active in HR-XML workgroups.

The core of the SETU standards is its models. Starting with the domain model to sketch the problem situation. Although SETU does not standardize the processes, they are captured in process models as reference examples and include options for corrections as well. SETU standardizes the data in messages (for instance the timecard message), and includes a data dictionary and code lists (for instance expense types). The SETU technical format is XML and XML Schema. To support the SETU adopters SETU does also include a “transport” guidelines on which protocols to use for exchanging messages.

The SETU uses a development process, initiated by the demands of the stakeholders (and approved by the SETU board), within workgroups the topic is explored and solutions are designed and formalized. Finally a review process is started before the workgroup hands over its work to the board for the release decision. The maintenance process is based on the filed maintenance requests, which after a threshold has passed and after approval of the board, will be picked up by the workgroup in the same development process.

The SETU standards are developed and maintained by the SETU organization; a not-for-profit organization. Its member contains temporary staffers and service providers. Part of the quality management is a document way how the standard is developed, reviewed and released. SETU standards are open, meaning that they are free to download and do not contain IPR. The board consisting of members and the ABU decides and assigns the workgroup with a specific task. Complaints will be handled by the board as well.

The financial model consists of a membership fee, in combination with funding of the sector organization (ABU). SETU does not have explicit long-term visions. It plans its activity on year-basis. Still it wants to deliver effective solutions for temporary staffing at minimal costs. Recently SETU has set up a promotion strategy to fasten the adoption process of the standard. This strategy contains the planning of events and publications. Part of this strategy was also the effort that was put on getting SETU standard on the comply or explain list of the Dutch Government, giving status to the SETU standards. SETU has a publication strategy which releases documents on four levels: To the public, to SETU participants, to SETU workgroup members, or to SETU Board members. The distinction between SETU participants and the public is made to give participants advantage in relation to the membership fee they pay. A mailing list is used for communication purposes.

The compliancy strategy is that SETU on purpose avoids this area, which means there is also no certification program. However SETU supports validation. With exception of the validation service SETU does not have components or tools available for implementation. In the past SETU supported the development of an open source component for time card communication based on the SETU standard. SETU has a highly knowledgeable helpdesk. Irregular, there is a SETU course available. Although SETU does not have preferred consultants, does not support pilots, or have reference implementations available, SETU does support incidental potential problematic implementation with high impact.

## Appendix N

# SETU Case Result of Applying QMSS (chapter 10)

The SETU quality measurement is based on three different sources:

1. Report of expert session of Dutch government (Van Hillegersberg & Minnecre, 2009).
2. All information available on [www.setu.nl](http://www.setu.nl), this includes
  - The specifications
    - SETU standard for ordering and selection 1.0
    - SETU standard for assignment 1.0
    - SETU standard for reporting time & expenses 1.0
    - SETU standard for invoicing 1.0
    - SETU standard for ordering and selection 1.1
    - SETU standard for assignment 1.1.
    - SETU standard for reporting time & expenses 1.1
    - SETU standard for invoicing 1.1
  - General documents
    - Statutes of SETU organization
    - By-laws of SETU organization
    - SETU operating procedures
    - Covering note on SETU standards
    - Readability guide SETU standards
    - Roadmap SETU 2007/2008
    - Recommended practices for Transmission and Security 1.0
  - Other documents on the website
    - The minutes of the workgroup meetings
    - Other
3. Information from SETU developers, not available in documents.

The sources that have influenced the opinion will be mentioned as evidence in Table 64 covering the complete scoring.

Measurable concept/ Attribute	Opinion	Evidence
1. Functionality	<p>1.1 Suitability</p> <p>It is a small focused set for which the business case is most evident. It could have had more added value with a broader scope. It is not a complete solution for all electronic transactions related to temporary staffing companies. It only presents a solution for the primary administrative process of hiring, but this is in line with the described scope. For instance it does not contain solutions for the exchange of planning information, or communication with job boards.</p> <p>There are hardly any irrelevant functions present in the standard. The standard is quite simple and only covers the main processes, without all the exceptions found in current practice. Because it was decided that all SETU standards should be implementable without implementation of all SETU standards (self-contained), each standard does contain some redundant information not needed when using the complete set of SETU standards.</p> <p>There are no doubts that the SETU standards solve a real life specific business problem. Attempts have been made to quantify the benefits in a business case.</p>	<p>Broader functionality: meeting minutes.</p> <p>Focus: Roadmap SETU (p2)</p> <p>Redundancy: Readability guide (p5).</p> <p>Business case: Excel sheet available.</p>
1.2 Accuracy	<p>Use of the SETU standard will lead to a certain degree of interoperability.</p> <p>The functional needs are satisfied in practice for the given scope. The only issue is that there are several requests (based on practice) for a broader scope of the standard.</p> <p>The specifications are consistent with the scope of the standards. They may even be useful in a broader sense than the current scope describes. For instance, although not described the standards may be useful for temporary employees in general, not just temporary staffing. The scope described in the expert advise is broader than the scope written in the standards documents.</p> <p>Most of the implementations are not 100% interoperable with each other, due to some freedom in the standards. There is some implementation freedom, to satisfy (not rational) customer demands.</p> <p>The standards also contain several process options (for instance how to deal with "multiplier"), which is standardized to several options. To be more interoperable the set has to be reduced to only one option.</p>	<p>Request for broader scope: Meeting minutes.</p> <p>Scope in expert advise, roadmap &amp; standards documents.</p> <p>Multiplier: Standard for Reporting time &amp; expenses version 1.1 (p.48).</p>
1.3 Compliance	<p>Implementations of SETU standard are also compliant with the international HR-XML standard (which implies that SETU is compliant to HR-XML). However HR-XML implementations do not necessarily comply to the SETU standards.</p> <p>There are several laws and regulations that have an impact on the SETU standard. For instance privacy laws related to the temporary worker and tax laws related to the invoice. The use of the Dutch BSN (unique social security number) number is limited by regulation. Although it is taken into consideration, the standard does not always explicitly state how it complies to laws and regulations.</p>	<p>Compliance HR-XML: SETU operating procedures (p.18).</p> <p>Document from tax services is available on SETU website.</p> <p>Compliance to privacy regulation: Standard for Reporting Time &amp; Expenses. (version 1.0 p.38) Also available in meeting minutes.</p>

<p>2. Reliability</p> <p>2.1 Maturity</p>	<p>The formulation of the text in the specification is solid. No unnecessary long sentences. By writing the specifications in English, more standardized sentences are used. For example "each timcard <i>must have a ...</i>"</p> <p>It may be possible to improve the standard by having a native English speaker review the text.</p> <p>The inclusion of "information rules", that are strictly formulated, avoids vagueness of possible weak sentences.</p> <p>Although there are some small differences between the four standards of SETU, in general the release schema generally results in one new release every 2 years, although one release per year is possible when there is a need for it. If there are no maintenance requests, then a period longer than 2 years is conceivable.</p> <p>The number of changes in general is somewhere between 5-20 per release. The implications of the changes differ to a large extent.</p> <p>All maintenance requests for a given standard will be taken into account when development of the new version of this given standard starts. None will be postponed. Although it might happen that the workgroup decides that the solution for a maintenance request will hurt backwards compatibility. Implying that it can not be solved in a point release, and have to wait for a major release.</p> <p>Apart from the specification itself, the following documents are available: A readability guide connecting all standards of SETU. A recommendation on how to handle transport of SETU standardized messages. The meeting minutes are also available providing background information on the reasoning behind certain solutions in the standards. And more. Even old, deprecated material is available in prominent places. Version 1.0 is listed in between the downloads of version 1.1 of the standard.</p> <p>For the user it is not visible what documents are outdated and which ones are valid. This leads to an overload of information, with a varying degree of quality.</p> <p>Measurement of the quality of the documentation is difficult. However we do notice that procedures have been created for quality purposes. Each document has gone through a double review process (both workgroup and SETU employees). Most documents are structured based on a template.</p> <p>The SETU is a relatively new organization specifically set up for maintaining the SETU standard. In the area of semantic standards this is not unusual. The organization still has to prove itself however. Standardization expertise is contributed by contracting consultants of TNO (a not for profit scientific organization with expertise in the field of standardization).</p> <p>The board of SETU currently consists of 6 experts from the temporary staffing industry. The chairman is COO of Randstad Holding, market leader in the Netherlands. USG (highest market share after Randstad) is also present in the board. The sector organization is also present in the board. Not present are small and middle size temporary staffing organizations, software vendors, or staffing customers.</p>	<p>Standard for reporting time and expenses.</p> <p>SETU operating procedures.</p> <p>Maintenance requests on SETU website.</p> <p>SETU statutes and bylaws.</p>
<p>2.2 Fault tolerance</p>	<p>The standards contain solutions for correction, to be used in processes defined by both trading partners. This could have been defined more strictly by one standardized correction process.</p> <p>The main philosophy of the SETU standard is that only 90% can be standardized. In many cases a project will be started between the trading partners for the remaining 10%. Incorrect implementations, ambiguity and other faults will be identified during this project. The validation service will also improve the chance of correct implementations.</p>	<p>Corrections: Standards.</p>
<p>2.3 Consistency</p>	<p>The use of the information elements is strict. SETU added a dictionary of elements on top of HR-XML, to avoid ambiguity. Because real process standardisation is not part of the SETU standards (SETU is much more focused on information), the standards can be used for multiple functions. Some are related to each other, for instance the use of the invoice standard for both a pro-forma invoice and self-billing. On the other hand the standard for ordering and selection serves multiple processes that are quite distinct. Within HR-XML the ambiguity is much more present. For instance between the Human Resource standard and the Employee standard.</p>	<p>Definitions: SETU data dictionary.</p> <p>Process standardization: SETU Standards.</p>

3. Usability	<p data-bbox="185 1470 229 1634">3.1 Understandability</p> <p data-bbox="185 409 619 1470">The documents are highly structured based on a template. There are no indications that the descriptions are not complete. To determine the readability of the standards, we decided to use the Flesch Kincaid Reading Ease Score and the Gunning Fog Score. The Flesch Kincaid Reading Ease Score is a formula based on number of words, sentences and syllables, while the Gunning Fog score is based on words, sentences and complex words. A low Gunning Fog score is preferable, just as a score close to 100 for the Flesch Kincaid Reading Ease. The scores are: Ordering &amp; Selection: 53.9 (Flesch) and 9.6 (Gunning) Assignment: 49.1 and 10.1 Reporting Time &amp; Expenses: 41.5 and 12.3 Invoicing: 59.1 and 8.8 This test was based on chapter 3 (the model of the standard), of each standard. Since we have no reference scores of any other standards, we can only compare the four SETU standards. The standard for invoicing has the best scores, while the standard for reporting time &amp; expenses has the lowest score by far. The difference is remarkable, and may suggest that the readability of the standard for reporting time and expenses can be improved. The time required to learn the overall standard is average. The amount of paper of the specifications is substantial. Furthermore, it is complex material, including the relation between SETU and HR-XML. The learning time can be reduced by offering more training and support. A first attempt has been made by the introduction of the readability guide in 2009, which is an easy and quick read. There is no training material available. Most of the functions are easily recognizable, even for non-experts of the domain of staffing industry. The functions are on the level of selection, ordering, assigning, time reporting and invoicing.</p>	<p data-bbox="185 136 229 409">Readability: Scores are determined by using: <a href="http://www.read-able.com/">http://www.read-able.com/</a> Learning time: Standards and Readability guide.</p>
3.2 Implementability	<p data-bbox="628 1470 672 1634">3.2 Implementability</p> <p data-bbox="628 409 908 1470">The effort needed to implement the standard in different kinds of software is limited. There are several running implementations and the number is slowly increasing. There are no projects halted that had the intention of implementing the standards. All projects succeeded. Implementations in COTS are fairly well presented. Many implementations are based on the examples provided. The implementations are often part of a larger project, restructuring organization processes both internally and externally. The standards are in line with current infrastructures within companies. Only problem is the support of ERP-kind of systems within the hiring organization. Currently for each implementation a custom made interface has to be implemented, which makes the implementation process costly. Not all functionality is implemented to a similar extent: implementations are mainly related to two (out of four) standards. This is mainly due to problem relevance. Although the standards can be used within regular (current) processes, to achieve maximal benefit the business processes have to be adapted as well.</p>	<p data-bbox="628 136 779 409">Implementation knowledge is gathered to looking at the meeting minutes, and secondly by SETU experts who also maintain the helpdesk for implementation questions/problems.</p>
4. Portability 4.1 Adaptability	<p data-bbox="917 1470 968 1634">4.1 Adaptability</p> <p data-bbox="917 409 989 1470">Within the scope of the standard, the standard is adaptable to deal with for instance specific customer demands. Adaptability to organizations outside the temporary staffing domain will be possible as well, but they make less sense.</p>	<p data-bbox="917 136 1036 409">The customer reporting requirements module is available. (Standard for reporting time and expenses version 1.1 p.53).</p>

4.2 Co-existence	<p>One of the strong points of SETU is the alignment of its standards to the international standard HR-XML. It co-exists with for instance HR-XML by adding a localization profile for the Dutch situation, making it applicable to use, while still remaining compliant to HR-XML. SETU has a policy in place that deals with new versions of HR-XML.</p> <p>One exception is the invoice standard; which is aligned with HR-XML and OAGIS. However in the Netherlands, the Dutch government has chosen the UN/CEFACT Cross Industry Invoice (CII) for future implementations. There are currently no conflicting issues with SETU, since there are no implementations of the CII.</p> <p>The SETU standards are unique, so there is never an implementation with co-existence of a different standard. It also has a policy to be unique; SETU will not develop anything that already exists.</p> <p>The modularity of the standard is high; each of the four standards is drafted to be used individually. Both HR-XML and SETU make use of building blocks of core components for reusable objects like "name".</p>	<p>SETU Operating Procedures. Standard for Invoicing version 1.1.</p> <p>Meeting minutes. SETU experts.</p>
4.3 Replaceability	<p>There is a strict policy on dealing with version numbering. In SETU defines that point releases (x) are always backwards compatible. This means that the old version can be translated to the new version, but not the other way around.</p> <p>The same functions are supported in version 1.0 and 1.1 of the standard. Even the same data elements are supported within version 1.0 and 1.1 although data definitions may have been changed (improved) based on indistinctness in practice.</p> <p>Version 1.1 is in this sense a replacement of version 1.0 and there is no functional reason for implementation of the old version.</p> <p>The SETU standard replaces the use of HR-XML in the Netherlands for the domain of temporary staffing.</p>	<p>SETU operating procedures: versioning procedures p.17.</p>
5. Maintainability		
5.1 Adaptability	<p>SETU will adapt according to the request it receives from the field. Not many adaptations have thus far been necessary in the history of SETU.</p> <p>Only once a bug has been reported that required immediate attention. Even in that case only one (outdated) file required replacement on the website, which means the action was not related to the content of the standard. Two other bugs mention faulty text descriptions in the specification, which are solved in the planned new version of the standards.</p> <p>While real bugs may require immediate updates of the standard, in the case of SETU this has not yet occurred. The impression exists that if there is a real bug, the SETU is capable of solving this bug rapidly.</p> <p>Normal updates have a very low frequency (once every two years), and is in line with the policy of SETU of only release a new version when there is a real business need.</p> <p>There is an important dependence on HR-XML. Although being an international standard, the future of HR-XML is uncertain since the departure of its managing director. HR-XML has a different release schema from SETU and releases many more versions of its standard than SETU does. In practice, the SETU standards are at least one version behind those of HR-XML. The current version of HR-XML is 3.0, while SETU is using HR-XML version 2.5</p> <p>Secondly, there is dependence on technical standards like XML, but consequences are low.</p> <p>Dynamic content like code lists are maintained separately outside the main specifications. Currently there are four code lists. It is remarkable to observe that these code lists are hardly dynamic and only one of them has been updated within the last 2 years.</p> <p>The SETU standard is not very strict; it standardized the meaning of the data, but not the process of exchanging the data. On data level it often includes several options to satisfy the needs of different processes. Together with the customer reporting requirement module, it facilitates the process of fine-tuning in implementations.</p>	<p>Maintenance requests tool on website: MR36, MR30, MR71.</p> <p>SETU operating procedures.</p> <p>Code lists on the SETU website.</p>
5.2 Stability	<p>Backwards compatibility is assured for point releases (x)</p> <p>Since 2008, there has only been one new (point) release for each of the four standards.</p> <p>In these 2 years, a total of 71 maintenance requests have been submitted. The standard for reporting time &amp; expenses and the standard for ordering and selection have the highest amount of maintenance request (both 20).</p> <p>The data dictionary contains strict and useful self-contained definitions. The definitions of SETU are more strict than definitions of the same elements in HR-XML, which may lead to interoperability problems.</p>	<p>SETU operating procedures: versioning procedures p.17.</p> <p>Maintenance requests on SETU website.</p> <p>Data dictionary on SETU website.</p>

5.3 Testability	<p>The SETU validation service supports implementation and test compliance. The validation service is more than XML schema validation; it also validates several business rules and has extensive error reporting. Finally, test messages may be created for testing purposes. There are reference implementations available represented in examples for each standard. Finally the helpdesk will also test implementations of the standard upon request.</p> <p>There is no certification program for the standards. There is no documentation on testing procedures either, nor are there requirements for a SETU compliant implementation.</p>	Validation Service on SETU website.
6. Adoptability		
6.1 Acceptance of users	<p>All major temporary staffing organizations support the SETU developments and have implemented the standards. Although there are no real figures available on the use of the SETU standard. However, it is clear that the standard for reporting time and expenses is the one mostly used, followed by the standard for invoicing. A rough estimate is that 10% of all timecards (including paper ones) are sent in the SETU format. Implementations of the other two standards are scarcely available. Still, the number of implementations is steadily, and slowly, increasing.</p> <p>Acceptance of end users, the staffing customers, is a major concern.</p>	Participation list on the SETU website. SETU experts.
6.2 Availability tools	<p>The standard is implemented in software, especially for staffing companies. Implementation in procurement software for staffing customers is lacking, and is currently often based on customization efforts. There are several procurement solutions available, acting as intermediary and based on SETU standards.</p> <p>The availability of tools or connectors that support implementation is lacking, with the exception of the SETU validation service. There are no lists available of services/products or tools that make use of the SETU standard.</p>	SETU experts.
6.3 Availability support	<p>The standards are maintained by SETU, which also has a helpdesk for support questions. If more support is required, the consultants of TNO are available for knowledge support, but unlike the free helpdesk, this support should be paid for. In the past a training course was organized by TNO, but this has not been the case recently.</p> <p>Currently there is only a limited set of consulting or implementation companies. Large implementation projects are often supported by the same consultants.</p> <p>Although the amount of consultants is fairly limited, when necessary, there is a sufficient amount of knowledge to support successful implementations.</p> <p>There is no community forum on Internet, for knowledge sharing. In the past there was a forum on the SETU website, but this was closed down because of lack of contributions.</p>	SETU experts.
7. Openness Process	<p>The SETU organization is a not for profit foundation solely aimed at the development and maintenance of the standard. SETU meetings are only open to participants that are appointed by the SETU board. Software vendors can not become participant, however by invitation of a participant they may join the meeting. SETU imposed these restrictions in order to avoid sessions that have massive attendance, but are not likely to provide valuable input the development of the standards.</p> <p>The SETU boards work based on consensus, although in practice solutions have always come up that are supported by the whole board. Different opinions are more often present in the workgroups, but this has thus far never led to a conflict. All solutions are prepared and designed by the independent consultants of TNO. Although sometimes workgroup members prefer a different solution they will ultimately agree on the solution proposed.</p> <p>The change procedure is described in the SETU operating procedures.</p>	Expert advice (p.7 &8). Meeting minutes. SETU operating procedures (p. 22).
7.2 Openness Specification	<p>The specifications are available on the SETU website (only free registration based on e-mail address is required). There are no costs related to the use of SETU, and no IPR such as patents are currently known.</p> <p>The specification documents all have a copyright statement, no creative commons. According to the expert advice there are no limitations on the re-use of the SETU standard.</p> <p>There is an open source implementation of SETU. There is however only one vendor supporting the open source tool (e-uur).</p>	Expert advice (p.7 &8). Standards documents.

Table 64 – Complete overview of SETU field test measurement results

## The Sources (chapter 11)

The literature review, Chapter 5, already identified several interesting sources containing measurable concepts often related to other domains, but might still be useful for the iQMSS. This section will present these sources in more detail, and focuses on practical models, measurable concepts and quality attributes. These practical studies related to quality have originated in different domains, of which some are more relevant than others. This part will give an overview based on the origin. The structure is as follows:

1. The software quality domain
2. The Information Systems (IS) quality & success domain
3. The data quality domain
4. The standards quality Domain
5. The XML standards quality domain
6. Evaluation frameworks
7. Other works

### 0.1 The software quality domain

Within this section we will describe which input from the software domain is useful for the selection of measurable concepts, as well as other input that is interesting for the quality model.

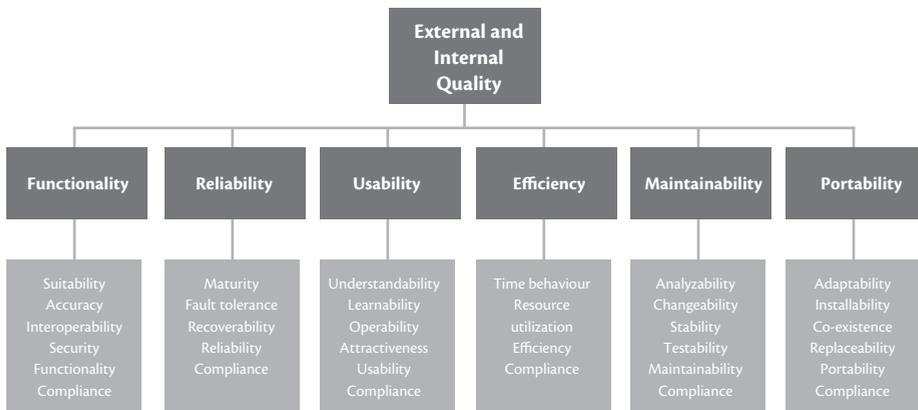


Figure 56 – External and internal quality model from ISO 9126 (ISO/IEC, 2001)

This appendix is based on:

(Folmer & Verhoosel, 2011): State of the Art on Semantic IS Standardization, Interoperability & Quality, Enschede: TNO, University of Twente, CTIT, NOiV.

The main quality factors can be defined as (Larrucea, 2008):

- **Functionality:** Are the required functions available in the software?
- **Portability:** How easy to transfer to another environment?
- **Reliability:** How reliable is the software?
- **Maintainability:** How easy is to modify the software?
- **Usability:** Is the software easy to use?
- **Efficiency:** How efficient is the software?

The external metric for this quality model are presented in ISO 9126 part 2 (ISO/IEC, 2003a), while the internal quality metrics are available within ISO 9126 part 3 (ISO/IEC, 2003b). These contain more detail and therefore more focused on software engineering and less applicable to the standards domain. ISO 9126 is the intended quality model, to be used in conjunction with ISO 14598 (ISO/IEC, 1999), which defines the process of software evaluation.



Figure 57 – ISO 14598: The process of software evaluation (simplified version)

ISO 9126 has spurred further developments by other scholars both focusing on improvements or adding specific elements for particular use cases. Improvements additions are suggested by the Quint model (Van Zeist et al., 1996; Van Zeist & Hendriks, 1996).

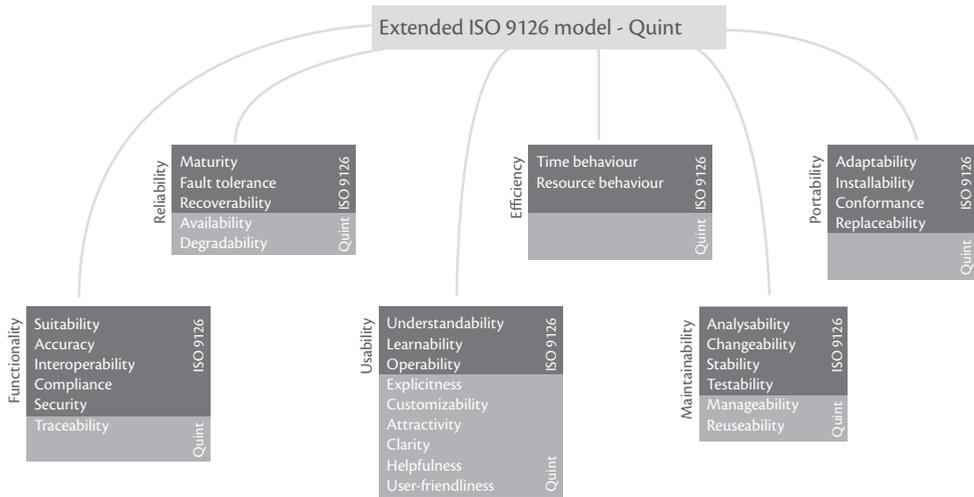


Figure 58 – Extended ISO 9126 model (Van Zeist et al., 1996; Van Zeist & Hendriks, 1996)

The follow up to ISO 9126 is the ISO 25000 (SQUARE) family, an extensive set of documents containing parts of the quality framework. The two main quality models are captured within ISO 25010 and cover both product quality (internal oriented), Figure 59, and quality in use (external oriented), Figure 60.

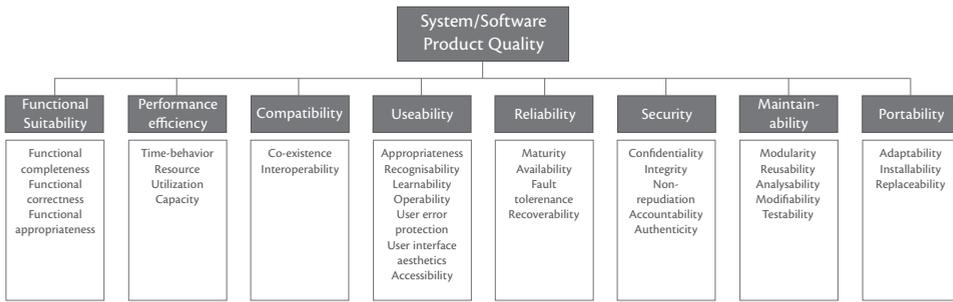


Figure 59 – System/Software product quality (ISO/IEC, 2011)



Figure 60 – Quality in use (ISO/IEC, 2011)

In addition to these models, ISO 25012 captures a data quality model, focusing on the quality of data a computer system uses. The quality characteristics from this model are (ISO/IEC, 2008): Accuracy, Completeness, Consistency, Credibility, Currentness, Accessibility, Compliance, Confidentiality, Efficiency, Precision, Traceability, Understandability, Availability, Portability and Recoverability.

Next to that all attributes have definitions, the ISO standards go one step beyond and define measures within the ISO 2502x standards, based on a template. Table 65 shows an example from ISO 25021 (ISO/IEC, 2007).

QME Category	Number of Faults
QME Name	Number of faults detected in review
QME ID	QME0403
Detail	Count the numbers of faults detected in review, during design/coding.
Input	Review report
Documentation	List of fault categories and their weights Lists of detected faults and their categories
Measurement scale	Ratio
Measurement focus	Internal
Measurement method	Objective
Used for	Fault removal

Table 65 – Example measure from ISO 25021 (ISO/IEC, 2007)

Again, just as with ISO 9216, several studies have led to extensions to the ISO quality model, for instance for specific software products. For instance the example of the extension in particular for web sites (Lew et al., 2010), as presented in Figure 61.

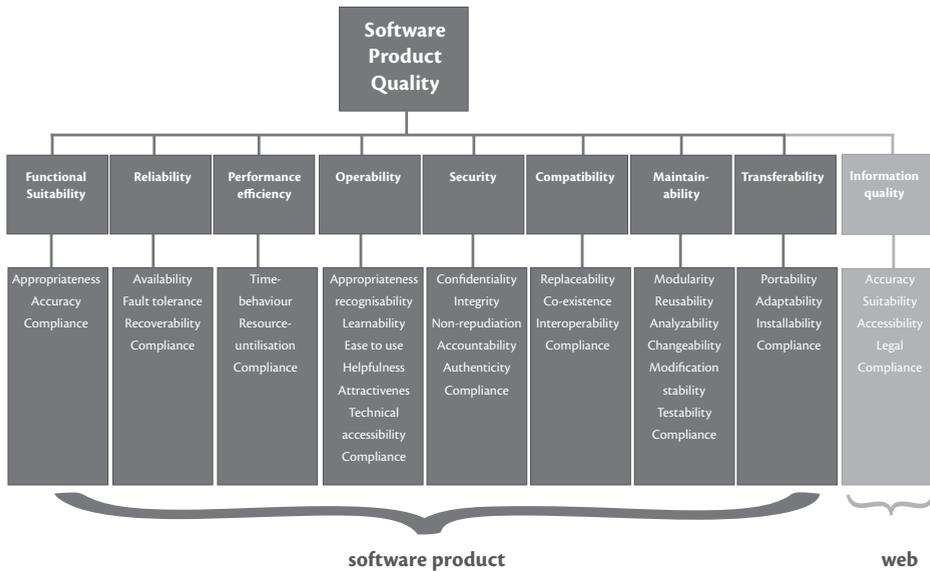


Figure 61 – Product quality model with an extension for web applications (Lew et al., 2010)

Related to the development processes CMM presents process maturity levels (Humphrey, 1989, 1997); initial, repeatable, defined, managed and optimizing. CMM is focused on making things measurable, and has been replaced by the CMMI since 2002 on which several other maturity models have been build like for software evaluation (Heck, Klabbers, & van Eekelen, 2010; Heck & Van Eekelen, 2008) or enterprise interoperability (Guedria et al., 2009).

### Earlier studies

There are many quality studies from the software domain. Some of which are overlapping with the ISO quality models, whilst others are presuccessors of ISO 9126, (e.g. (Boehm, 1973; Cavano & McCall, 1978a, 1978b)). Cavano & McCall (1978a) define software quality: Correctness, Reliability, Efficiency, Integrity, Usability, Maintainability, Testability, Flexibility, Portability, Reusability, and Interoperability.

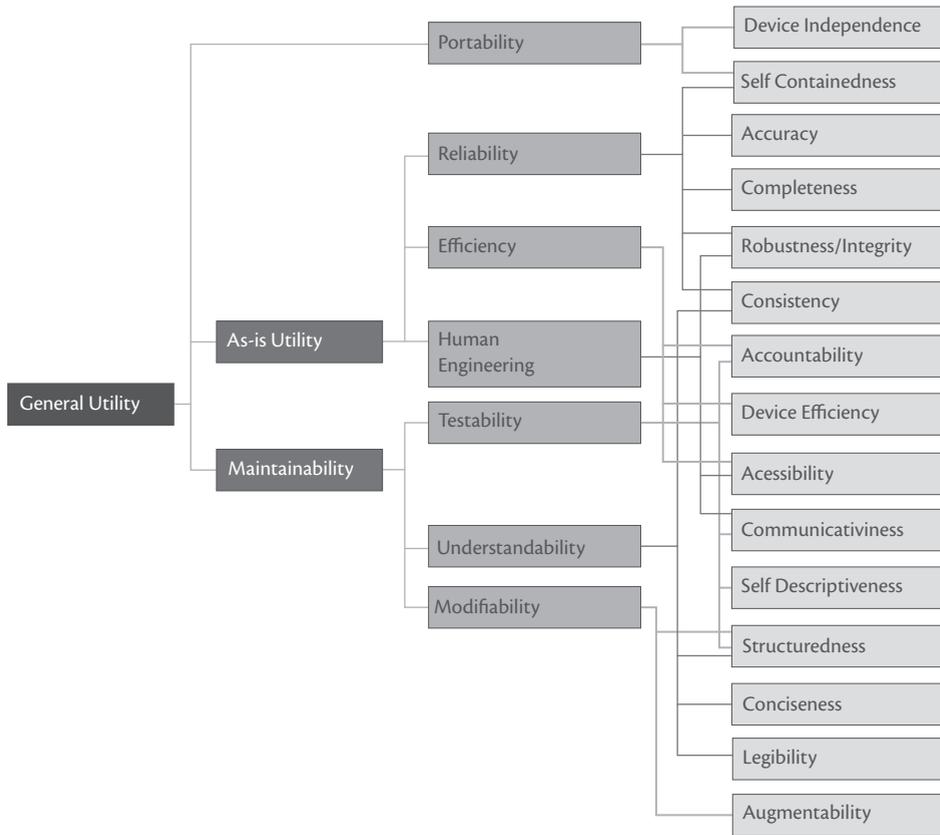


Figure 62 – Boehm's quality model (Boehm, 1973; Milicic, 2005)

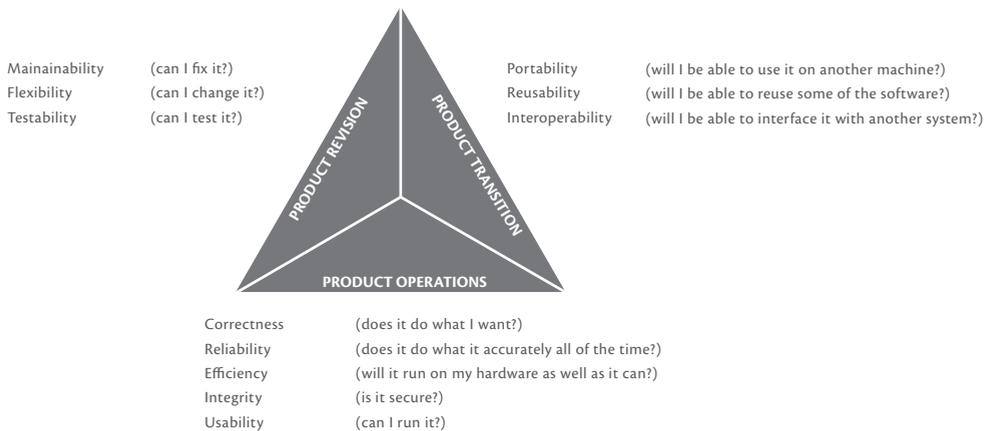


Figure 63 – McCall's quality model (Cavano & McCall, 1978b)

The Boehm quality model looks similar to the McCall quality model in the sense that there is a hierarchy of three layers called high-level, intermediate level, and primitive characteristics. The high-level characteristics address three main questions that a buyer of software has related to the general utility of software (Milicic, 2005):

- As-is utility: How well (easily, reliably, efficiently) can I use it as-is?
- Maintainability: How easy is it to understand, modify and retest?
- Portability: Can I still use it if I change my environment?

Other worth mentioning quality models are FURPS+, and Dromey's quality model. FURPS+ was developed by Robert Grady, and nowadays is part of the IBM Rational Software, and distinguishes functional (F) and non-functional (URPS) quality characteristics. FURPS stands for (Milicic, 2005):

- Functionality - which may include feature sets, capabilities and security
- Usability - which may include human factors, aesthetics, consistency in the user interface, online and context-sensitive help, wizards and agents, user documentation and training materials
- Reliability - which may include frequency and severity of failure, recoverability, predictability, accuracy and mean time between failure (MTBF)
- Performance - imposes conditions on functional requirements such as speed, efficiency, availability, accuracy, throughput, response time, recovery time and resource usage
- Supportability - which may include testability, extensibility, adaptability, maintainability, compatibility, configurability, serviceability, installability and localizability (internationalization)

Dromey's quality model focuses on the relation between the quality attributes and the sub-attributes, as well as attempting to connect software product properties with software quality attributes. For easy usage, it includes a 5-step process model as well.

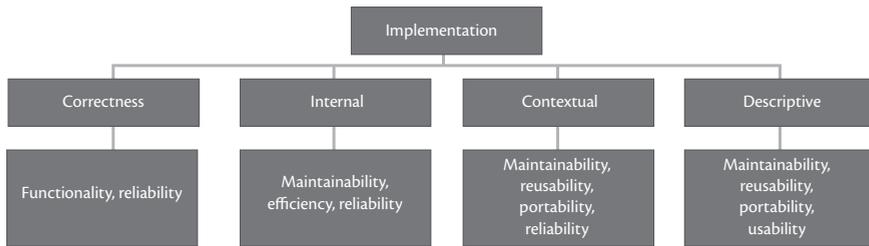


Figure 64 – Principles of Dromey's quality model (Milicic, 2005)

Because of its importance most of this work is already been included in the development of the earlier mentioned ISO standards on software quality. By including the ISO standards in our work, we have indirectly incorporated these sources. Although it seems that the software engineering domain has more than enough quality models, there are even more created by specific consultancy companies, like the TMap quality approach by Sogeti. And not to forget the quality models for specific software languages, for instance for JavaBeans (Washizaki, Hiraguchi, & Fukazawa, 2008).

### Different perspectives

There is however more than ISO 25000, some other scholars have taken different viewpoints. Fenton & Neill also introduced a set of attributes, related to the entities of software (Fenton & Neil, 2000).

ENTITIES	ATTRIBUTES	
	Internal	External
<b>Products</b>		
<b>Specifications</b>	size, reuse, modularity, redundancy, functionality, syntactic correctness, ...	comprehensibility, maintainability, ...
<b>Designs</b>	size, reuse, modularity, coupling, cohesiveness, inheritance, functionality, ...	quality, complexity, maintainability, ...
<b>Code</b>	size, reuse, modularity, coupling, functionality, algorithmic complexity, control-flow structuredness, ...	reliability, usability, maintainability, reusability
<b>Test data</b>	size, coverage level, ...	quality, reusability, ...
...	...	...
<b>Processes</b>		
<b>Constructing specification</b>	time, effort, number of requirements changes, ...	quality, cost, stability, ...
<b>Detailed design</b>	time, effort, number of specification faults found, ...	cost, cost-effectiveness, ...
<b>Testing</b>	time, effort, number of coding faults found, ...	cost, cost-effectiveness, stability, ...
...	...	...
<b>Resources</b>		
<b>Personnel</b>	age, price, ...	productivity, experience, intelligence, ...
<b>Teams</b>	size, communication level, structuredness, ...	productivity, quality, ...
<b>Organizations</b>	size, ISO Certification, CMM level	maturity, profitability, ...
<b>Software</b>	price, size, ...	usability, reliability, ...
<b>Hardware</b>	price, speed, memory size, ...	reliability, ...
<b>Offices</b>	Size, temperature, light, ...	comfort, quality, ...
...	...	...

Table 66 – Quality attributes (Fenton & Neil, 2000)

Another model from the software engineering domain: An instrument to measure the critical dimensions of the software quality, as perceived by customers (Issac, Rajendran, & Anantharaman, 2006). The respondents are requested to indicate their perception on a 7-point Likert scale, from extremely low to extremely high.

The product quality characteristics tested are (Issac et al., 2006): Functionality (Correctness, Security/Integrity, Instrumentation, Installability), Reliability (Error Tolerance), Usability (User-friendliness/Learnability), Efficiency (Availability, Resource Consumption), Maintainability (Expandability, Modifiability), Portability, Documentation, Service provisioning, Aesthetics.

Next to product quality other aspects are tested mainly on the organization that delivers the software product: Process quality (decomposed in organizational culture, management practice, benchmarking, quality standards, employee training), Customer focus, Competence of employees, Operational effectiveness and Infrastructure and facilities (Issac et al., 2006).

## 0.2 The IS quality & success domain

The scope of information systems (IS) domain is broader than software engineering, albeit that terminology within the software domain is changing to IS as well. See for example the latest ISO 25010 explicitly changed the title of the specification from software engineering to system and software engineering (ISO/IEC, 2011). Still, IS has a different history and focus. Just as in the area of software quality, in the world of IS the notion of quality is described by the introduction of 40 quality attributes (Delen & Rijsenbrij, 1992). It includes a

process model called the quality loop for incorporating quality in the IS development process. This quality loop consists of three steps (quality requirements, quality engineering, and characteristics). The quality notion is hierarchically decomposed into (4) dimensions, (21) aspects, and (40) quality attributes (Delen & Rijsenbrij, 1992).

DIMENSION	Aspects / attributes			
I PROCESS development and control of information systems	1. Quality conditions a. prof. skills b. account mgt. c. project mgt. d. system development	2. Quality control	3. Continuity	4. Completeness of services
II STATIC properties of the information system in maintenance & control	<b>P</b> 1. Flexibility	2. Maintainability	3. Testability	4. Portability
	<b>R</b> 5. Connectivity a. external b. internal	6. Reusability	7. Fitness of the infrastructure	
III DYNAMIC functioning of the system for the user	<b>O</b> 1. Reliability a. correctness b. completeness c. authorizedness d. timeliness	2. Continuity a. uninterrupted b. robustness c. restorability d. degradation possibility e. diversion possibility	3. Efficiency a. speed - internal - total b. user- friendliness c. economy d. match with manual proc. e. workability manual proc.	4. Effectiveness a. coverage of bus. processes b. availability - in time - on location c. usability d. decision support e. end user support
	<b>D</b>			
	<b>U</b>			
	<b>C</b>			
	<b>T</b>			
IV INFORMATION importance for company	1. Correctness	2. Completeness	3. Up-to-dateness	4. Accuracy
	5. Verifiability			

Table 67 – Quality attributes (Delen & Rijsenbrij, 1992)

Just like ISO within the software domain, within this framework a template is used to describe attributes.

Famous within the IS domain is the DeLone & McLean model for predicting success. Important to stress here is that the main conclusion is that IS success is multidimensional and an interdependent construct. The list of measures should not be used as is but by combining measures to study interdependencies and to create a comprehensive measurement instrument (DeLone & McLean, 1992; DeLone & McLean, 2003). When looking at the table the large number of quality attributes attracts attention.

System Quality	Information Quality	Information Use	User Satisfaction	Individual Impact	Organization Impact
Data accuracy	Importance	Amount of use/	Satisfaction with	Information	Application
Data currency	Relevance	duration of use:	specifics	understanding	portfolio:
Database contents	Usefulness	Number of	Overall satisfaction	Learning	Range and scope
Ease of use	Informativeness	inquiries	Single-item measure	Accurate	of application
Ease of learning	Usableness	Number of	Multi-item measure	interpretation	Number of
Convenience of access	Understandability	functions used	Information	Information	critical
Human factors	Readability	Number of	satisfaction:	awareness	applications
Realization of user requirements	Clarity	records	Difference	Information recall	Operating cost
Usefulness of system features and functions	Format	accessed	between	Problem	reductions
System accuracy	Appearance	Frequency of	information	identification	Staff reduction
System flexibility	Content	access	needed and	Decision	Overall
System reliability	Accuracy	Frequency of	received	effectiveness:	productivity
System sophistication	Precision	report requests	Enjoyment	Decision quality	gains
Integration of systems	Conciseness	Number of	Software satisfaction	Improved	Increased revenues
System efficiency	Sufficiency	reports	Decision-making	decision	Increased sales
Resource utilization	Completeness	generated	satisfaction	analysis	Increased market
Response time	Rehability	Charges for		Correctness of	share
Turnaround time	Currency	system use		decision	Increased profits
	Timeliness	Regularity of use		Time to make	Return on
	Uniqueness	Use by whom?		decision	investment
	Comparability	Direct vs.		Confidence in	Return on assets
	Quantitativeness	chauffeured		decision	Ratio of net
	Freedom from bias	use		Decision-	income to
		Binary use:		making	operating
		Use vs. nonuse		participation	expenses
		Actual vs. reported		Improved	Cost/benefit ratio
		use		individual	Stock price
		Nature of use		productivity	Increased work
		Use for intended		Change in decision	volume
		purpose		Causes	Product quality
		Appropriate use		management	Contribution to
		Type of		action	achieving
		information		Task performance	goals
		used		Quality of plans	Increased work
		Purpose of use		Individual power	volume
		Levels of use:		or influence	Service
		General vs.		Personal valuation	effectiveness
		specific		of IS	
		Recurring use		Willingness to pay	
		Institutionalization/		for	
		routinization		information	
		of use			
		Report acceptance			
		Percentage use vs.			
		opportunity for			
		use			
		Voluntariness of use			
		Motivation to use			

Table 68 – Success measures (DeLone & McLean, 1992)

Ten years later an update of the model was proposed by the same authors (DeLone & McLean, 2003). The most distinctive changes in the new model are the addition of both service quality and net benefits. Service quality is added to avoid too much focus on the product and on neglecting the service aspects. The 22-item SERVQUAL measurement instrument from marketing might be used as a fundament to measure this IS function. Net benefits have been included because the impact of the system goes beyond the individual user, but might have an impact on inter-organizational (industry), consumer or society aspects. Net benefits reflects the wide range of entities that might be affected by the IS function.

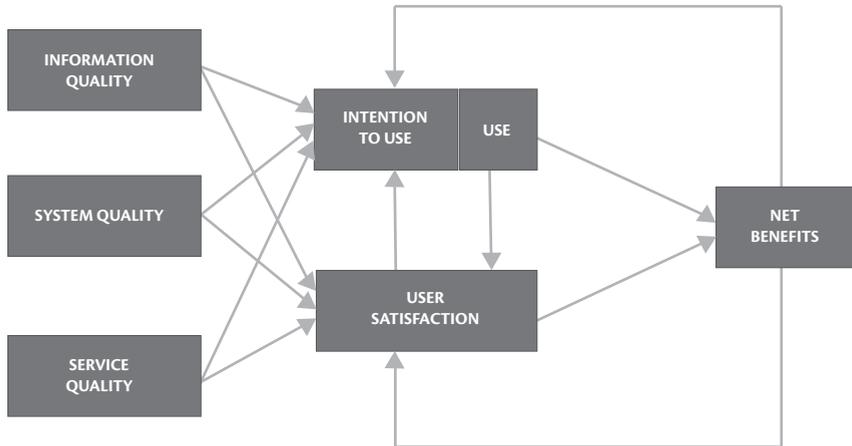


Figure 65 – Updated DeLone & McLean IS success model (DeLone & McLean, 2003)

Closer to the domain of semantic standards is e-business. The success metrics for e-business are therefore particular interesting (DeLone & McLean, 2003):

Systems quality	Information quality	Service quality
<ul style="list-style-type: none"> <li>• Adaptability</li> <li>• Availability</li> <li>• Reliability</li> <li>• Response time</li> <li>• Usability</li> </ul>	<ul style="list-style-type: none"> <li>• Completeness</li> <li>• Ease of understanding</li> <li>• Personalization</li> <li>• Relevance</li> <li>• Security</li> </ul>	<ul style="list-style-type: none"> <li>• Assurance</li> <li>• Empathy</li> <li>• Responsiveness</li> </ul>
Use	User satisfaction	Net benefits
<ul style="list-style-type: none"> <li>• Nature of use</li> <li>• Navigation patterns</li> <li>• Number of site visits</li> <li>• Number of transactions executed</li> </ul>	<ul style="list-style-type: none"> <li>• Repeat purchases</li> <li>• Repeat visits</li> <li>• User surveys</li> </ul>	<ul style="list-style-type: none"> <li>• Cost savings</li> <li>• Expanded markets</li> <li>• Incremental additional sales</li> <li>• Reduced search costs</li> <li>• Time savings</li> </ul>

Table 69 – Example of metrics (for E-Commerce success) within a model (DeLone & McLean, 2003)

Due to the success of the model, and just as with the ISO standards from the software domain, many scholars have developed extensions to the success model. Including a model for IS quality (Rodríguez & Casanovas, 2010), although the relations within the model have not been validated, the hypothesis and suggested measures are valuable input.

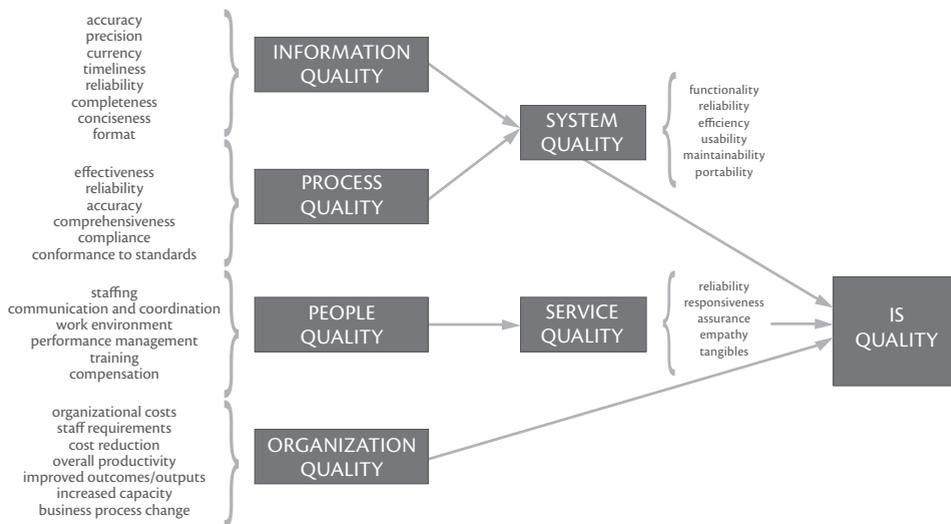


Figure 66 – Hypotheses for IS quality (Rodriguez & Casanovas, 2010)

The hypotheses suggest a positive relationship between the constructs. The suggested measures for each of the six qualities are as follows (Rodriguez & Casanovas, 2010):

- Information (or Data) quality should be measured in terms of accuracy, precision, currency, timeliness, reliability, completeness, conciseness, format of input/output and relevance.
- System quality: proposed are the ISO 9126 measures: functionality, reliability, efficiency, usability, maintainability and portability.
- Service quality: proposed is SERVQUAL (from the marketing domain), which consists of five dimensions: reliability, responsiveness, assurance, empathy and tangibles.
- Process quality: proposed is to use CMMI-DEV for selecting terms like: effectiveness, reliability, accuracy, comprehensiveness, compliance and conformance to standards.
- Organization quality: measures proposed by Sedera & Gable: organizational costs, staff requirements, cost reduction, overall productivity, improved outcomes/outputs, increased capacity and business process change.
- People quality: measurements are proposed from the People Capability Maturity Model, level 2 (P-CMM); these are related to staffing, communication and coordination, work environment, performance management, training and development and compensation.

Other research has focused on validating the measures to be used within the constructs, of the original DeLone & McLean model. This has led to the validated Enterprise System Success model (Sedera & Gable, 2004) presented within Figure 67.

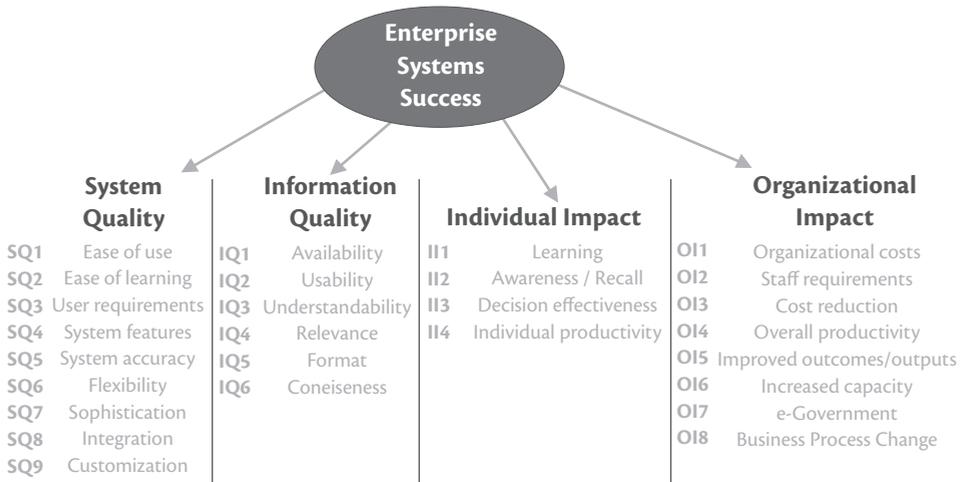


Figure 67 – Validated measures within the model for Enterprise System Success (Sedera & Gable, 2004)

Also quality attributes have been researched for specific purposes within the IS domain, like Knowledge Management (KM) (Owlia, 2010) and Service Oriented Architectures (SOA) (O'Brien et al., 2005), and even an instrument for measuring SOA maturity (Joachim et al., 2011). From the perspective of KM, a quality model has been set up with starting point the 8 quality dimensions from product engineering (Garvin, 1984). Based on ISO 9126, 11 factors have been selected for the model; the column on the left side of the table. Other quality attributes related to different knowledge products and services have been incorporated (Table 70). Based on this extensive research eight quality dimensions have been selected for Knowledge Management Systems presented in Table 71, with the characteristics presented in Table 72 (Owlia, 2010).

Software	Data	Information	Information systems	E-Services
Correctness	Correctness	Validity	Specification	
Reliability	Accuracy	Accuracy	Accuracy	Reliability
	Currency	Precision	Precision	Accuracy
	Volatility	Currency	Reliability	Currency
	Free of error		Upkeep	
Efficiency	Minimality	Value-Added	Effectiveness	Efficiency
	Value-Added		System usage	
Integrity	Security	Security		Security
				Privacy
Usability	Interpretability	Current position	Friendliness	Fulfillment
	Ease of understanding	Interpretability	Training	Ease of use
	Ease of manipulation	Ease of understanding	Learnability	Convenience
	Responsiveness		Memorability	Usability
			Reporting	
			Responsiveness	Responsiveness
			Understanding	Courtesy
			Attitude	Personalization
				Customization
				Empathy
Maintainability	Traceability	Disposition	Documentation	Recovery
Testability			Control	Support
			Format	Follow-up
			Technical assistance	Services
			Delivery and Installation	Serviceability
Expandability		Compatibility	Flexibility	Flexibility
Portability				
Reusability				
Interoperability				
	Usefulness	Relevance	Business alignment	Functionality
	Relevancy		Necessity	Performance
			Relevancy	
			Direction	
			Meaningfulness	
	Objectivity	Objectivity		
			Competence	Competence
	Completeness	Completeness	Integration	Completeness
	Amount of data	Amount of data	Complexity	Features
	Accessibility	Access	Accessibility	System availability
	Availability	Timeliness	Response time	Access
	Timeliness		Timeliness	Response time
				Timeliness
	Credibility	Reputation	Loyalty	Credibility
	Reputation	Believability		Assurance
	Believability	Authority		Trust
	Concise representation	Presentation	Front office	Aesthetics
	Consistent representation	Format	Marketing	Appearance
		Consistent representation		Web site design
		Coherence		Structure
			Participation	Collaboration
			Communication	Communication
				Transaction capability
				Storage capability

Table 70 – Quality attributes related to KM (Owlia, 2010)

Dimension	Definition
<b>Functionality</b>	The degree to which the system meets organizational objectives, operational standards and users' knowledge needs
<b>Completeness</b>	How sufficient and comprehensive is the system
<b>Reliability</b>	The degree to which knowledge is correct, accurate, consistent, and up to date
<b>Usability</b>	The effort required for using and involving in the system
<b>Access</b>	The extent to which knowledge is available for users
<b>Serviceability</b>	How well a KMS handles customers' enquiries
<b>Flexibility</b>	The degree to which acquiring knowledge in different situations/conditions is possible
<b>Security</b>	Confidentiality of information/knowledge shared when necessary

Table 71 – Quality dimensions for KM systems (Owlia, 2010)

Dimension	Characteristics
<b>Functionality</b>	Meeting organizational objectives Satisfying users' needs System usage Providing primary (core) knowledge Providing primary functions including knowledge creation, storage and retrieval, distribution and application
<b>Completeness</b>	Providing supplementary (advanced, innovative) knowledge Providing supplementary/more advanced functions and technologies e.g. artificial intelligence or expert systems Meeting established software, hardware, and communication standards
<b>Reliability</b>	Accuracy Fault free Consistency Currency Credibility, trustworthiness Legacy
<b>Usability</b>	Easy to use Friendliness Training, learnability Appearance Communication, knowledge conversation and sharing
<b>Serviceability</b>	Personalization Customization Handling users' enquiries Solving system problems Responsiveness, how well a KMS responds to demand by users.
<b>Access</b>	Accessibility Availability Response time Timeliness
<b>Flexibility</b>	Flexibility Compatibility Interoperability Scalability Future-proofed
<b>Security</b>	Security Privacy Control

Table 72 – Quality dimensions and their corresponding characteristics for KM systems (Owlia, 2010)

Research from the Carnegie Mellon University / Software Engineering Institute resulted in a set of quality attributes for Service Oriented Architectures (O'Brien et al., 2005):

1. Interoperability
2. Reliability
  - a. Message reliability
  - b. Service Reliability
3. Availability
4. Usability
  - a. Data Granularity
  - b. Normal Usability Operations
5. Security
6. Performance
  - a. XML in Web Services as a Performance Factor
7. Scalability
8. Extensibility
9. Adaptability
10. Testability
11. Auditability
12. Operability and Deployability
13. Modifiability

In the area of conceptual modeling, based on the theory of reasoned action, properties have been identified that address the perceived semantic quality of a conceptual model (Poels et al., 2005):

- Correct (accuracy): The conceptual schema represents the business process correctly.
- Relevant: All the elements in the conceptual schema are relevant for the representation of the business process.
- Complete: The conceptual schema gives a complete representation of the business process.
- Adequate (Consistent / Conflict-free): The conceptual schema contains contradicting elements.
- Realistic: The conceptual schema is a realistic representation of the business process.

### **0.3 The data quality domain**

Data or Information quality is arguably the area with most research in the area of quality. Let alone the resources at <http://mitiq.mit.edu/> and <http://iaidq.org> present a huge amount of studies in this area. This section will present some important work of that area that is focusing on attributes.

Data or information quality is part of the IS success models presented in the previous section. It is also studied within the domain of software engineering. But it is also a distinctive research area, with a strong research fundament, probably because of its importance: 60% of the surveyed firms (500 medium-size corporations with annual sales of more than US \$20 million) have problems with data quality (Wand & Wang, 1996; Wang & Strong, 1996). Within the domain of data quality, Juran's definition of fitness for use is commonly used (Wang & Strong, 1996; Zhu & Wu, 2010). To improve data quality the need was evident to understand what data quality means to data consumers, for which a conceptual frameworks of data quality has been constructed. One framework consist of data quality dimensions (Wand & Wang, 1996): completeness, meaningfulness, correctness, unambiguousness (representation and meaning). Which is applicable to standards as well. Another framework consists of 15 dimension within the following four categories (Wang & Strong, 1996):

- Intrinsic Data Quality (Believability, Accuracy, Objectivity, Reputation)
- Contextual Data Quality (Value-added, Relevancy, Timeliness, Completeness, Appropriate amount of data)
- Representational Data Quality (Interpretability, Ease of understanding, Representational consistency, Concise representation)
- Accessibility Data Quality (Accessibility, Access security)

This work was followed up with the development of a model (Table 73) (Kahn et al., 2002).

	<b>Conforms to Specifications</b>	<b>Meets or exceeds Consumer Expectations</b>
<b>Product Quality</b>	Sound Information <ul style="list-style-type: none"> <li>• Free-of-Error</li> <li>• Concise Representation</li> <li>• Completeness</li> <li>• Consistent Representation</li> </ul>	Useful Information <ul style="list-style-type: none"> <li>• Appropriate Amount</li> <li>• Relevancy</li> <li>• Understandability</li> <li>• <i>Interpretability</i></li> <li>• <i>Objectivity</i></li> </ul>
<b>Service Quality</b>	Dependable Information <ul style="list-style-type: none"> <li>• Timeliness</li> <li>• Security</li> </ul>	Useable Information <ul style="list-style-type: none"> <li>• Believability</li> <li>• Accessibility</li> <li>• Ease of Manipulation</li> <li>• Reputation</li> <li>• Value-Added</li> </ul>

Table 73 – Quality model (Kahn et al., 2002)

Including the following set of definitions (Table 74) (Kahn et al., 2002).

<b>Dimensions</b>	<b>Definitions</b>
<b>Accessibility</b>	The extent to which information is available, or easily and quickly retrievable
<b>Appropriate Amount of Information</b>	The extent to which the volume of information is appropriate for the task at hand
<b>Believability</b>	The extent to which information is regarded as true and credible
<b>Completeness</b>	The extent to which information is not missing and is of sufficient breadth and depth for the task at hand
<b>Concise Representation</b>	The extent to which information is compactly represented
<b>Consistent Representation</b>	The extent to which information is presented in the same format
<b>Ease of Manipulation</b>	The extent to which information is easy to manipulate and apply to different tasks
<b>Free-of-Error</b>	The extent to which information is correct and reliable
<b>Interpretability</b>	The extent to which information is in appropriate languages, symbols, and units, and the definitions are clear
<b>Objectivity</b>	The extent to which information is unbiased, unprejudiced, and impartial
<b>Relevancy</b>	The extent to which information is applicable and helpful for the task at hand
<b>Reputation</b>	The extent to which information is highly regarded in terms of its source or content
<b>Security</b>	The extent to which access to information is restricted appropriately to maintain its security
<b>Timeliness</b>	The extent to which the information is sufficiently up-to-date for the task at hand
<b>Understandability</b>	The extent to which information is easily comprehended
<b>Value-Added</b>	The extent to which information is beneficial and provides advantages from its use

Table 74 – Definitions of quality dimensions (Kahn et al., 2002)

Another set of data quality dimensions is presented in Table 75 (Wand & Wang, 1996) (the first mentioned dimensions are more cited than the latter mentioned dimensions).

	<b>Dimensions</b>
Internal View (design operation)	<p><b>Data-related</b> accuracy, reliability, timeliness, completeness, currency, consistency, precision</p> <p><b>System-related</b> reliability</p>
External View (use, value)	<p><b>Data-related</b> timeliness, relevance, content, importance, sufficiency, useableness, usefulness, clarity, conciseness, freedom from bias, informativeness, level of detail, quantitiveness, scope, interpretability, understandability</p> <p><b>System-related</b> timeliness, flexibility, format, efficiency</p>

Table 75 – Data quality dimension (Wand & Wang, 1996)

Many more information or data quality frameworks have been created. A comparison of 12 different quality models (including (Alexander & Tate, 1999; Dedeker, 2000; Katerattanakul & Siau, 1999; Naumann & Rolker, 2000; Shanks & Corbitt, 1999; Zhu & Gauch, 2000)) all containing different quality characteristics has been made by Knight and Burn (2005), while Byrd and Byrd (2012) compared 14 quality models. Based on the comparison, a summary of the most common dimensions including the occurrence frequency in the twelve frameworks is provided in (Table 76).

<b>Dimension</b>	<b># of times</b>	<b>Definitions</b>	<b>*conform Wang &amp; Strong (1996)</b>
1. Accuracy	8	extent to which data are correct, reliable and certified free of error*	
2. Consistency	7	extent to which information is presented in the same format and compatible with previous data*	
3. Security	7	extent to which access to information is restricted appropriately to maintain its security*	
4. Timeliness	7	extent to which the information is sufficiently up-to-date for the task at hand*	
5. Completeness	5	extent to which information is not missing and is of sufficient breadth and depth for the task at hand	
6. Concise	5	extent to which information is compactly represented without being overwhelming (i.e. brief in presentation, yet complete and to the point)*	
7. Reliability	5	extent to which information is correct and reliable*	
8. Accessibility	4	extent to which information is available, or easily and quickly retrievable*	
9. Availability	4	extent to which information is physically accessible	
10. Objectivity	4	extent to which information is unbiased, unprejudiced and impartial*	
11. Relevancy	4	extent to which information is applicable and helpful for the task at hand	
12. Useability	4	extent to which information is clear and easily used	
13. Understandability	5	extent to which data are clear without ambiguity and easily comprehended*	
14. Amount of data	3	extent to which the quantity or volume of available data is appropriate*	
15. Believeability	3	extent to which information is regarded as true and credible*	
16. Navigation	3	extent to which data are easily found and linked to	
17. Reputation	3	extent to which information is highly regarded in terms of source or content*	
18. Useful	3	extent to which information is applicable and helpful for the task at hand*	
19. Efficiency	3	extent to which data are able to quickly meet the information needs for the task at hand*	
20. Value-Added	3	extent to which information is beneficial, provides advantages from its use*	

Table 76 – Quality dimensions based on an analysis of 12 quality frameworks (Knight & Burn, 2005)

Next to the quality dimensions, a framework exists for information quality (IQ) assessment (Stvilia et al., 2007). The core of the framework is 22 quality dimensions, within three categories: intrinsic, relational and reputational. These quality dimensions are related to four sources of information quality problems:

1. Mapping-related IQ problems arise when there is incomplete, ambiguous, inaccurate, inconsistent, or redundant mapping between some state, event, or entity and an information entity.
2. Any changes in context (both culture and socio-technical structures) can change how IQ is understood and evaluated and can lead to an IQ problem.
3. Changes may occur in the information entity itself or in the real-world entity it represents.
4. The process of IQ change for an information entity can be passive or indirect, caused by changes to the underlying entity or context.

These four sources of problems are related to activity types (Stvilia et al., 2007):

1. Representation Dependent—activities that depend on how well one information entity represents another entity or some condition.
2. Decontextualizing—activities that use information outside its original context of creation (for instance, an activity may remove information entities from their original contexts and aggregate them into a new collection to support specific information needs or tasks).
3. Stability Dependent—activities that depend on how stable the information or its underlying entity is.
4. Provenance Dependent—activities that depend on the quality of metadata of the information's provenance, mediation, and upkeep.

In total 41 metrics have been set up to measure the 22 quality dimensions. Those 41 metrics should be situation dependently used. The general approach is to start with the IQ problems that might be expected based on the activity type at hand (for instance mapping or decontextualizing). The potential IQ problems are mapped on quality dimensions. Finally metrics are proposed for the selected quality dimensions (Figure 68).

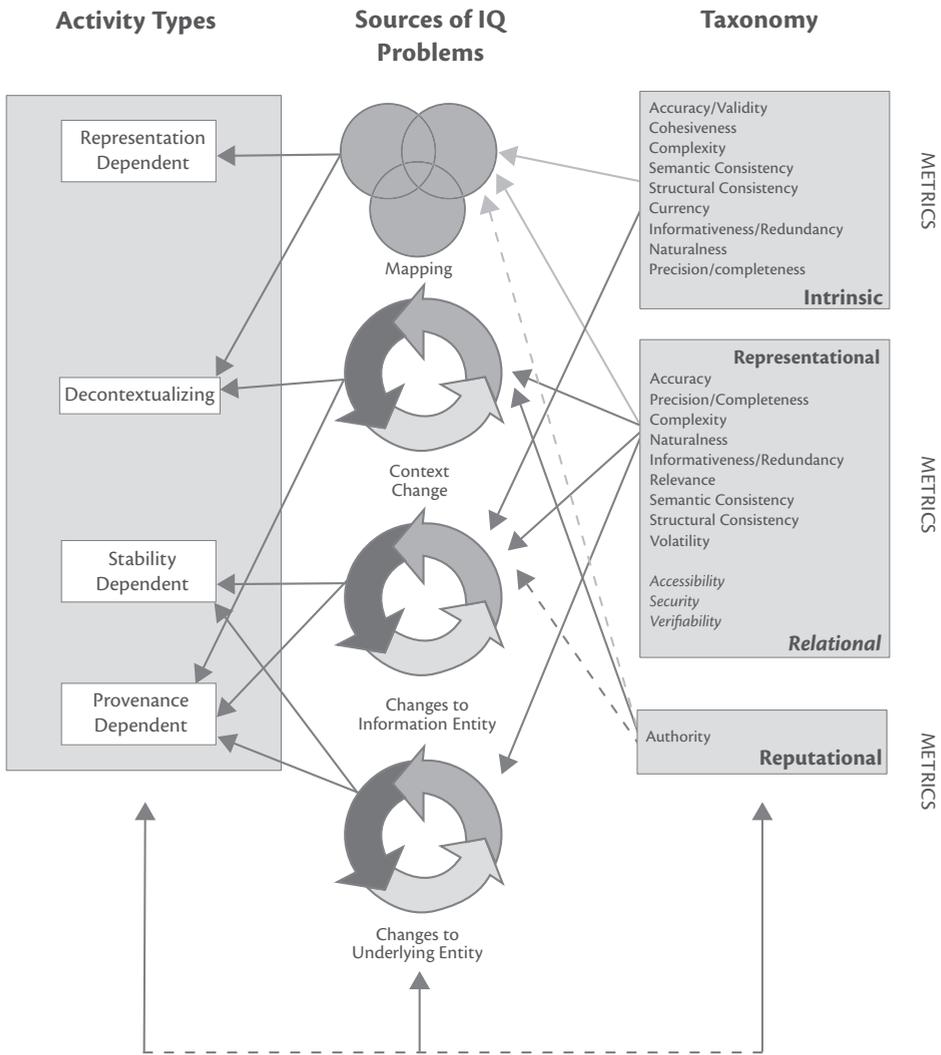


Figure 68 – Framework for information quality assessment (adapted from (Stvilia et al., 2007))

The DublinCore case study is particularly interesting since this is a semantic standard. Although it is important to notice that the quality of DublinCore standard itself is not at subject, but it is the data that is being transferred by making use of the DublinCore metadata standard. The proposed metrics for that data are presented in Table 77.

Dimension	Kinds of IQ problems counted	Possible metrics
<b>Intrinsic Precision/ Completeness</b>	Empty elements or element tags; less precision or completeness than expected for an element	Count of empty tags; count of incomplete values (circas); number of distinct elements
<b>Intrinsic Redundancy</b>	Repeated schema elements; repeated element values	Count of instances of repeated schema elements; Information Noise [content = 1 - (size of the term or token vector after stemming and stopping)/(object size before processing)]
<b>Intrinsic Semantic Consistency</b>	Contradicting values for the same elements	Count of instances of the same elements having different values
<b>Intrinsic Structural Consistency</b>	Inconsistent formatting or representation of the same elements	Count of instances of the same elements using different formatting
<b>Relational Accuracy</b>	Broken links to related objects	Counts of broken links
<b>Relational Completeness</b>	Missing elements from a recommended set of elements	Number of elements present from the WSDCMBP set of required elements (Title, Creator, Subject, Description, Date, Format, Identifier, Rights); FRBR Support Index for the DC schema defined according to a formula [skipped]
<b>Relational Semantic Consistency</b>	Elements containing inappropriate values according to a standard	Counts of instances of elements misuse
<b>Relational Structural Consistency</b>	Elements containing value codes that are not in a standard	Counts of instances of element formatting not matching recommended guidelines
<b>Relational Verifiability</b>	Original or related objects that are inaccessible or unrecoverable	(number of identifier + number of source + number of relation)/3

Table 77 – Metrics for information quality assessment for DublinCore metadata exchange (Stvilia et al., 2007)

#### **0.4 The standards quality domain**

This section will summarize what quality attributes have been described within the standards domain.

##### *0.4.1 General: The EU government perspective on quality*

Already in Chapter 1 & 5 the EU government policy was introduced that aims for inclusion of ICT standards maintained by others than the ESOs (EC, 2011b). The defined requirements for the EU recognition of ICT standards that might be used for selection purposes are in complete detail presented in Table 78.

## ANNEX II REQUIREMENTS FOR THE RECOGNITION OF TECHNICAL SPECIFICATIONS IN THE FIELD OF ICT

1. The technical specifications have market acceptance and their implementations do not hamper interoperability with the implementations of existing European or international standards. Market acceptance can be demonstrated by operational examples of compliant implementations from different vendors.

2. The technical specifications were developed by a non-profit making organization which is a professional society, industry or trade association or any other membership organization that within its area of expertise develops standards in the field of information and communication technologies and which is not a European, national or international standardization body,

through processes which fulfill the following criteria:

- (a) Openness: the technical specifications were developed on the basis of open decision-making accessible to all interested operators in the market or markets affected by the standard.
- (b) Consensus: the standardization process was collaborative and consensus based and did not favor any particular stakeholder. Consensus means a general agreement, characterized by the absence of sustained opposition to substantial issues by any important part of the concerned interests and by a process that involves seeking to take into account the views of all parties concerned and to reconcile any conflicting arguments. Consensus does not imply unanimity.
- (c) Transparency:
  - (i) All information concerning technical discussions and decision making was archived and identified.
  - (ii) Information on (new) standardization activities was widely announced through suitable and accessible means.
  - (iii) Participation of all interested categories of interested stakeholders was sought with a view to achieving balance.
  - (iv) Consideration and response were given to comments by interested parties.

3. The technical specifications reflect the following requirements:

- (a) Maintenance: Ongoing support and maintenance of published specifications are guaranteed over a long period.
- (b) Availability: Specifications are publicly available for implementation and use on reasonable terms (including for a reasonable fee or free of charge).
- (c) Intellectual property rights essential to the implementation of specifications are licensed to applicants on a (fair) reasonable and non-discriminatory basis ((F)RAND), which includes, at the discretion of the intellectual property right holder, licensing essential intellectual property without compensation.
- (d) Relevance:
  - (i) The specifications are effective and relevant.
  - (ii) Specifications need to respond to market needs and regulatory requirements.
- (e) Neutrality and stability:
  - (i) Specifications whenever possible are performance oriented rather than based on design or descriptive characteristics.
  - (ii) Specifications do not distort the market or limit the possibilities for implementers to develop competition and innovation based upon them.
  - (iii) Specifications are based on advanced scientific and technological developments.
- (f) Quality:
  - (i) The quality and level of detail are sufficient to permit the development of a variety of competing implementations of interoperable products and services.
  - (ii) Standardized interfaces are not hidden or controlled by anyone other than the organizations that adopted the technical specifications.

Table 78 – Requirements on standards in the field of ICT (EC, 2011b)

#### *O.4.2 General: What is a good standard?*

Several studies have set up some criteria for a good standard, in general. Good standards have the following characteristics (Simons & Vries, 2002), albeit that some are not valid for ICT standards:

1. Solution to communication problem (there has to be an interoperability problem for which the standard is a solution).
2. Customer demand: Every standard is developed by customer requirements. However often this is only a limited group, and additional focus is needed on the larger silent group of customers.
3. Manufacture, no handicraft.
4. Validity, minimal 3 years.
5. Compliant with other standards.
6. Backwards compatibility.
7. Future forwards compatibility.
8. Clarity (Readability).

Although with overlap, another list of properties that should apply to a good standard (Hesser, Czaya, & Riemer, 2007):

- The standard meets the needs of the users or the interested parties.
- The standard is available to the users in time.
- The standard is formulated in a way that is comprehensible and free of contradiction in terms of its scope (inner consistency).
- The standard can be implemented by the users.
- The standard does not contradict other existing standards (outer consistency). However this point is open to discussion because occasionally competition supports differing standards. In this way, the better standard is intended to assert itself, and innovation is not obstructed.
- The standard should tend to be more performance-based than prescriptive.
- The standard is sufficiently distributed amongst the users and is applied to an adequate extent by all interested parties. A standard that is not applied has no right to exist.
- If the development of a standard starts at an early stage, the lower the probability will be that economic interests might have already formed among the participants.

Teichman et al.(2008) cite de Vries: The standards shall be:

- Consistent, clear, and accurate.
- Comprehensible also to qualified persons who have not participated in their preparation.
- Fit for application and adoption without change as regional or national standards.
- Fit for use also by experts with mother tongues other than English or French (or, in the case of European standards, also German).

And Teichman et al. (2008) adds: Adequate language quality does not necessarily mean perfection. In fact, adequacy is determined by three mutually conflicting factors:

- Precision of the text.
- Intelligibility of the texts.
- Timely development of the standards.

### O.4.3 General: What is not a good standard?

Implementation might be tampered because of other reasons (Egyedi, 2008):

- The idea that underlies a standard may not be implementable (e.g. too comprehensive).
- The ideal of consensus decision-making may affect the standards process (e.g. lead to too many options) and, indirectly, the implementability of the standard.
- Different use of terminology in a standard specification may lead to problems of interpretation, implementation and interoperability.
- Modest user requirements and cost-constraints in the implementation process may lead to partial standard compliance and incompatible implementations.

These problems might be related to the standards specification (S), the conceptual idea (C), development process (SP) and its implementation process (IP), as different parts of the standardization ecosystem. Based on a discussion with a panel of experts, Egyedi (2008, 2009) studied causes of interoperability problems within the standardization ecosystems. The results are presented within Table 79.

<b>Causes of incompatibility</b>	<b>Locus</b>
Errors, ambiguities, inconsistencies	SP/S
Ambiguity of natural language	SP/S
Missing details, monopoly on tacit knowledge	S/IP
Ill-structured standards	S
Unclear how to handle options	S
Uncertain compatibility of non-binding recommendations	S
Complexity of comprehensive, ambitious standards	C
Too many options and parameters	SP/S/IP
Backward compatibility	C
Unclear official status of standards' companion book	S
Single company pushing for standard, weak specifications	SP
Overload of standards	C/IP
Deviation from and partial implementation of a standard	IP
Interference between standards	C/IP

Table 79 – Causes of incompatibility by a panel of experts (Egyedi, 2008, 2009)

Although the causes contain overlaps, and also the scoring seems a bit questionable (which might be explained by using a panel of experts as a research method), we still can say that 8 out of 14 causes relate to the specification document. Using the viewpoint of the standard as the combination of the idea, the process and the specification, one can argue that 13 out of 14 causes relate to the quality of the standard.

### O.4.4 Quality statements: Suggestions for better quality

The main critical success factors for standards are defined as (Simons & Vries, 2002):

1. Achievable compliance.
2. Acceptance.
3. Consistent implementation.

Farrell suggests that the relevant participants, the technical focus, and the internal processes of an SSO simultaneously influence its performance in terms of speed and quality (Zhao et al., 2005). This is also supported by the following three statements (Bernstein & Haas, 2008):

1. "To make effective use of [enterprise interoperability] standards, organizations need to commit substantial amounts of time and resources to create, implement and maintain the standards".
2. "The right stakeholders need to be involved to create a set of standards that meet the business needs of the organization".
3. "The scope of the architecture should also be carefully managed, as architectures too broad in scope can take years to build, only to become obsolete upon completion".

Specific for semantic standards is the critical importance of the availability of a data dictionary, in order to avoid standards drift (Steinfeld et al., 2007).

A study by Jakobs (2009b) shows that the quality of the standard is highly impacted by the position and the quality of workgroup members. This is supported by the work of Teichmann (2010) who agrees that the quality and quantity of the technical work produced by standards workgroups is affected by the management of the committees/workgroups, and on the individual effectiveness of the individual members. The selection of the participants within the workgroups will have an impact on the quality of the standard. Participants in working groups have a different background, but do need certain qualities (like familiarity with technical aspects, speak English, and have skills in technical writing) and motivation in order to be effective.

Institutional measures towards reducing standard-based interoperability problems (Egyedi, 2008):

<b>Institutional measures towards reducing standard-based interoperability problems</b>	
Drafting of standards	<ul style="list-style-type: none"> <li>• provide institutional support for editors and rapporteurs on standards engineering</li> <li>• involve technical editors</li> <li>• use pseudo-code or formal languages in a focused way</li> <li>• adopt a unified naming convention</li> <li>• clarify the type of options involved</li> <li>• specify how to deal with options (e.g. profiles)</li> <li>• specify the consequences of (not) implementing options</li> <li>• make the rationale that underlies choices in the specification explicit</li> <li>• issue a reference guide with the standard</li> <li>• organize wider scrutiny of the standard</li> <li>• translate the standard also to uncover ambiguities</li> <li>• co-ordinate interrelated standardization of different standards bodies</li> </ul>
Pre-implementation	<ul style="list-style-type: none"> <li>• validate standards before implementation in products ('walk-throughs')</li> <li>• develop a reference implementation/pre-implementation</li> <li>• develop a reference environment</li> <li>• include standard conformance and interoperability testing</li> <li>• organize interoperability events with different vendors (e.g. plug tests)</li> <li>• organize dialogue between standard developers and implementers</li> </ul>
Post-implementation	<ul style="list-style-type: none"> <li>• supply test suites</li> <li>• improve consistent use and integrity of standards with e.g. compliance and interoperability conformance statements compatibility logos, certification programmes</li> </ul>
Standards policy	<ul style="list-style-type: none"> <li>• prioritize implementability as a standard's requirement</li> <li>• reconsider desired level of consensus across all areas</li> </ul>

Table 80 – Improvement suggestions for the standardization system (Egyedi, 2007, 2008)

#### O.4.5 Quality attributes: Suggestions for better quality & measures

Related to standards the following (quality) measures are proposed by (Kasunic & Anderson, 2004):

- Standards explicitness
- Standards maturity
- Standards vendors supporting
- Standards feature coverage
- Standards sufficiency

Morell & Stewart (1995) studied process metrics and product metrics. Important process metrics are rather straight forward: is the agreed standardization process being carried out (e.g. Was there a formal need analysis? Was there feedback included, etc.)? Candidate process and product metrics are shown in Table 81.

Process Metrics	Product Metrics
Was a product specification put forward as a basis for the standard?	Were standards-conformant products on the market before the standard was finalized?
Is there multi-vendor interest in the standard?	Do the products meet user needs?
Was there an assessment of standard development time relative to the window of opportunity in the market?	Does the standard allow applications that are portable to different platforms, scalable in size, and interoperable with other applications?
How good is the management of end-user expectations relative to when vendors can actually deliver product?	Are products priced in terms of commodity costing?
Are there redundant or competing standards?	What impact has the standard had on the viability of the vendors' industry?
	Are diverse products interoperable?
	Are vendors building products?
	How large is the installed base?

Table 81 – Candidate process and product metrics (Morell & Stewart, 1995)

Related to telecom standards, Sherif et al. (2007) propose the following attributes to quality:

- Project management: use project management methodology in standards workgroups
- Scope Management, indications of poor scope management:
  - No, or superficial, terms of reference; which may indicate a lack of consensus among stakeholders.
  - Scope is not consistent with intended use.
  - Consensus decision making leading to overabundance of options.
  - Changes to term of reference.
  - No formal process to re-evaluate the need for the standard.
  - Existence of other overlapping standard groups.
  - Frustrated stakeholders that leave.
- Time Management:
  - (On time...)
- Quality Management, what helps:
  - Reviews.
  - Use of formal languages.
  - Use of technical editor.
  - Use of multiple languages to create unambiguous texts.
  - Reducing amount of tacit knowledge needed to implement.
  - Testing.

- New specification should be interoperate with legacy ones.
- Coordination with other standard bodies.
- Signs of troubles are:
  - No formal description.
  - Many options.
  - Missing information.
  - Inconsistency among different sections of the same document.
  - No tests for conformance or interoperability.
  - No pre-implementations.
- Cost Management:
  - Difficult but the costs of developing might not exceed the expected benefits.
- Resource Management:
  - Being dependent on volunteers, makes resource management difficult.
  - Hiring expertise as facilitators and independent mediators.
  - Signs of troubles:
    - One or more strong participants take over the agenda and drive it beyond the scope of the project.
    - Competing interests of the sponsors lead to a deadlock in the deliberations
    - Individual rivalries among participants prevent progress.
    - Group think and the rejection of new ideas.
- Documentation Management:
  - Readability.
  - Sign of trouble:
    - Not indexed not well organized.

In the end the standard has a specification document, often seen as core of the standard. This document should contain (Spivak & Brenner, 2001): Heading, Scope, References, Product qualification, Information to be supplied with bidders or tenders, Definitions, General, Construction, Tests and test procedures, Product identification, Packaging and package marking, Quality assurance provisions, Drawings. And some helpful hints (Spivak & Brenner, 2001): Revision status clearly marked, Identification of changes, Be specific, Keep wording simple and precise, Metric measurements according to measurement standards, The date in ISO standard.

Many specification documents are written by non-native English speakers. The linguistic quality will have an impact on the fitness for use of the standard in general (Teichmann, 2010). Broader than linguistic quality is the document quality that is addressed by the REVERE project (Rayson, Emmet, Garside, & Sawyer, 2001). This contains aspects like weak (ambiguous) sentences, including consistent use of MOSCOW. Although REVERE is intended to gather requirements by analyzing documents (including standards), the concepts and supporting toolset might also be useful to assess the quality of the standards' document (Sawyer, Rayson, & Garside, 2002).

However quality should not be limited to the specification document of the standard since users require more information to effectively make use of the standards. For instance there should be additional documentation (like education material, FAQ) available next to specification (Freericks, 2010).

De Vries (2008) has set up an extensive list of items for a best practice model for in company standardization. It includes best practices in different categories, ranging from policy to funding. The category "Company Standard Development Process" seems applicable and generalizable to other types of standardization, where company can be exchanged for SSO:

- Those who have asked for standards become involved in their development.
- Intended standards users become involved in their development.
- Suppliers and/or contractors can provide input in the company standard development process.

- Standards writers communicate with stakeholders during the development process.
- Each company standard is assessed on its expected fitness to contribute to business results.
- The company has a metastandard that provides criteria for its company standards.
- This metastandard is known by all involved in company standards development and they apply it.
- On a regular basis, the requirements in this metastandard are assessed on topicality and fitness for use.
- Company management has authorized this metastandard.
- The company standard is not just based on the personal opinion of one expert, but is broadly based.
- Participants in standard development consider their task as important and urgent.
- The status of writing standards equals the status of carrying out projects.
- There are enough competent employees for writing new standards and maintaining the quality and consistency of the existing standards collection.
- A “why” document is attached to each company standard. It provides the underpinning of the most important choice and decisions that have been made during standards development.
- A draft of each new company standard is sent out for comments to a relevant group of people within the company.
- There is a procedure for processing comments.
- Everybody is allowed to comment on draft standards.
- The CSD (Company Standards Department) coordinates comments processing.
- Comments, if any, are sent to the development team; they decide on adoption or rejection.
- In the case of rejection, they give reasons why.
- CSD checks the standard against the requirements in the metastandard.
- Company management authorizes the standard.
- ICT tools are used in developing and writing the standard.

To conclude, for specific situations quality criteria have been selected that are generalizable and subsets of above mentioned criteria. For example we show the quality criteria for the Electronic Health Record (EHR) (Eichelberg et al., 2005):

1. The level of interoperability support: Does the EHR provide structured content suitable for automated processing? Does it specify content distribution rules?
2. Functionality: Does the standard allow for an explicit retrieval of records (or parts thereof) for a specific patient, based on an incoming request? Can it contain multimedia data? What kind of security mechanisms are supported for accessing healthcare records?
3. Complementarity: Since not all the standards provide all the necessary features, is it possible to combine them in a complementary way? Do the standard initiatives affect one another?
4. Market relevance: Is the standard accepted in the marketplace? Are there commercial implementations available or any signs of uptake by the industry?

### ***0.5 The XML standards quality domain***

Nowadays most semantic standards are ultimately expressed in the technical XML format. Although the technical format is a representation of the content of the semantic standard it still might be useful as an indicator of the quality of the semantic standard.

Based on ISO 9126, a set of XML Schema metrics were developed that measure the quality of the XML Schema and the exploitation of advanced features of XML Schema. These are (McDowell, Schmidt, & Yue, 2004):

- Number of: Complex Type Declarations, Simple Type Declarations, Annotations, Derived Complex Types, Global Type Declarations, Global Type References and Unbounded Elements.
- Average: Number of Attributes per Complex Type Declaration, Bounded Element Multiplicity Size, Number of Restrictions per Simple Type Declaration and Element Fanning.

Based on these 11 metrics two indices have been set up: Complexity and Quality. Element fanning is the average of the number of child elements and number of references each element has. Each of those measures are indicators of quality and complexity: for instance a large number of Complex Type Declarations will indicate a complex XML Schema, while a large number of annotations will indicate a well documented XML Schema.

Based on the analysis of quality of different XML specifications, the complexity of standards is assumed to have two parameters (Brutti et al., 2010):

1. **Uncertainty:** The number of distinct data containers that exist for a single specific type of information in a document (for example, the possible alternatives to specify the Order ID in an XML instance)
2. **Redundancy:** The total number of possible distinct data containers in a document to support a specific business example.

To improve the effectiveness, interoperability, of standards the availability of conformance test tools and the use of “use profiles” based on customization rules is recommended (Brutti et al., 2011). Customization rules deal with identifying subsets for specific context, coding of values, declaration of rules for context depended data structures and constraints arising from data dependencies. Customization rules depending on the dynamic execution of the data exchange are related to constraints based on the role of the actors involved, or based on the position of the current transaction in the running business process (Brutti et al., 2011).

Another practical measure is to check if all tags that are used in the XML Schema are listed in the data dictionary, based on a study of 26 semantic standards it was found that 15% of the tags are not listed (Bedini et al., 2011).

Instead of measuring within the specification itself, implementations might also be a valuable source of information to determine quality. This viewpoint was used to assess the concepts of completeness and relevancy (from the data quality domain) of the US GAAP XBRL-taxonomy based on defined quantifiable metrics (Zhu & Fu, 2009; Zhu & Wu, 2010, 2011). Completeness of a data standard is the extent to which the data standard specifies all the data elements needed by standard users. Relevancy of a data standard is the extent to which the data standard specifies only the specific data elements needed by standard users (Zhu & Wu, 2010). Both are measured by counting the number of custom added elements and used elements within implementations. Adding custom data elements might indicate that the standard does not specify all data elements needed by the standard users. The number of used elements, in relation to the number of available elements indicates the relevance.

For the Danish government some quality indicators have been gathered that are to a large extent related to XML specifics (Gottschick & Restel, 2010).

<b>Name</b>	<b>Document</b>	<b>Bad Smell Description</b>
<b>NDRs</b>	XML schema	Evaluates violations against OIOXML naming and design rules
<b>Spelling</b>	Free text	Evaluates spelling weakness in free text
<b>Readability</b>	Free text	Evaluates the readability using the “Flesch-Kincaid Reading Ease Score”
<b>Documentation coverage</b>	XML related	Evaluates comment coverage in XML and XML Schema documents
<b>Formatting rules</b>	XML, RFC	Evaluates violations against predefined format rules (e.g. after RFC 2223 or W3C Pubrules)
<b>Namespace rules</b>	XML Schema	Evaluates missing namespace declarations in XML schema documents
<b>Unfinished documents</b>	Free text, XML	Evaluates unfinished documents by checking for keywords like “todo”, “fixme”, etc.
<b>Modularization</b>	XML related	Evaluates poor modularisation (e.g. detects too long XML files)

Table 82 – Quality indicators (Gottschick & Restel)

In practice, semantic standards evolve in a fragmented and distributed fashion. To make integration and interoperability more efficient and scalable, the fragmented specifications need to fit into a coherent, semantic model (Kulvatunyou, Morris, Ivezic, & Frechette, 2008). They need to be logically consistent and contain minimal duplication. Additionally, semantically overlapping data structures should be related or annotated, because every term and data structure should have unique semantics.

Technically speaking, the following 'common sense' recommendations are made related to XML:

1. Reduction of the XML Schema elements in the library (delete unused components, and refine cardinalities) makes it much easier to manage and understand (Brutti et al., 2010).
2. Definition in the library, using the Schematron code, of constraints that are common for multiple standards (XML Schemas) (Brutti et al., 2010).
3. If the standard is encoded in XML Schema then its syntax and semantics must conform to W3C XML Schema specification (Kulvatunyou et al., 2008).
4. Best practices like the UN/CEFACT Naming and Design Rules (NDR) to be used (Kulvatunyou et al., 2008).

Tools for testing the XML design are available. For instance the XML Schema Quality of Design (QOD) developed by NIST. QOD may be used by an XML Schema developer to ensure compliance with a set of guidelines for XML Schema development, the Naming and Design Rules (NDR). The tool is configurable to use tests for NDR from any number of original sources It is distributed with tests of NDR from the following two groups:

- Open Application Group Naming and Design Rules 9.0
- United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) Applied Techniques Group (ATG) - Naming and Design Rules -- Version 2.0

Other tests for other NDR include OAG, UNCEFACT, UBL, DON, IRS, and UnitsML, and others are expected to be added.

### ***0.6 Evaluation framework for standards***

Appendix H contains description of frameworks that can be used for evaluation or comparison of standards. Although not specifically targeted at quality, there might be some links to quality, especially for those frameworks that contain much detail. Especially the extensive evaluation framework from Mykkanen and Tuomainen (2008) gives an in-depth understanding of a standard under subject, although it does not cover quality. However the evaluation framework (Figure 69) is more general applicable when steps 9 till 15 are aggregated as "performing evaluation".

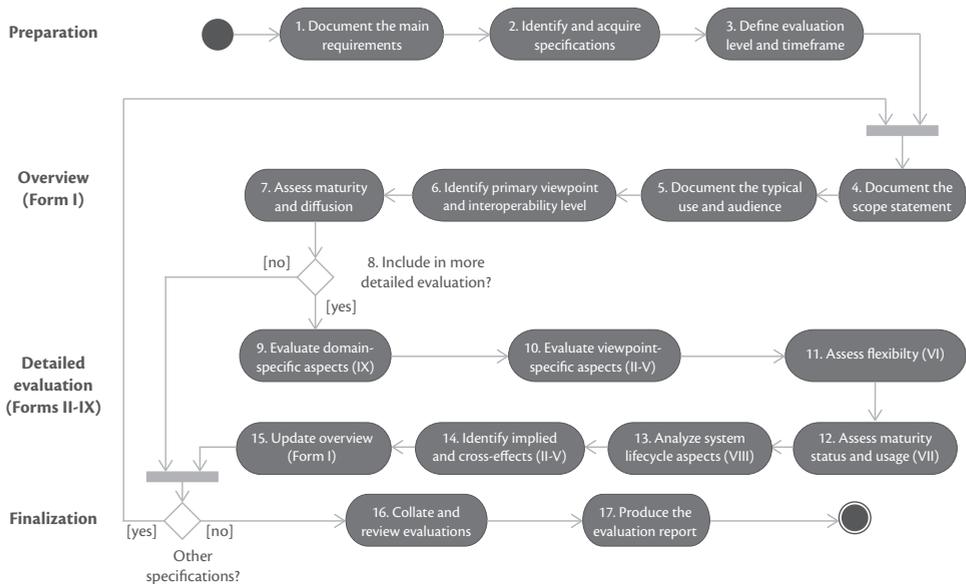


Figure 69 – The evaluation process (Mykkanen & Tuomainen, 2008)

Only the Reference Model Analysis Grid (RMAG) (Pawlowski & Kozlov, 2010) contains direct linkage with quality, as might be expected from its goal: Assessing, testing and validation of standards for achieving interoperability in the education domain. The RMAG is a methodology, including the framework, to assess standards with the aim to combine them in a harmonized framework. The main framework mentions the following quality attributes (Pawlowski & Kozlov, 2010): Seamlessness, Adaptation, Document Guidance, Compatibility, Extensibility. Next to the descriptive framework, a second framework for assessment is introduced. It consists of (Pawlowski & Kozlov, 2010):

1. Transformation and analysis: How can the standard be used for transformation of the organization (systems and processes)
2. Maintenance: How are the standards maintained?
3. Effectiveness and efficiency: How does the standard lead to effectiveness and efficiency?
4. Flexibility and integration: How flexible is the standard to be integrated within an organization?
5. Coordination and knowledge management: How are coordination mechanisms and knowledge exchange supported?
6. Interoperability: How does the standard lead to interoperability?
7. Understandability and usability: How understandable and usable is the standard?
8. Coherence: Is the standard coherent with other standards?

A framework with a dedicated focus on qualitative comparison is existing for the electronic health record (EHR). This Generic Component Model (GCM) (Blobel & Pharow, 2009) contains a generic structure, not only valid for EHR. Some of the characteristics in the model are descriptive (e.g. Service orientation), while other are qualitative by nature (e.g. Consistency of components). The full model is depicted in Table 83, including the scoring of EHR standards.

<b>GCM Characteristics</b>	<b>HL7</b>	<b>CCR</b>	<b>EN/ISO 13606</b>	<b>openEHR</b>	<b>IHE XDS</b>	<b>DICOM SR</b>
<b>Development Process Perspective</b>						
Unified process	Y	N	N	N	N	N
Business modeling	Y	N	N	N	P	N
Service orientation	P	N	N	N	P	N
View separation	P	N	N	N	N	N
Completion of ISO 10746-2	Y	N	N	N	N	N
Tooling available	Y	N	N	N	N	N
<b>System Architecture Perspective</b>						
Reference information model	Y	N	Y	Y	Y	Y
Meta model	P	N	N	N	N	N
Model transformation framework	Y	N	N	N	N	N
Model multiplicity	n	1	2	2	2	2
Concept representation	Y	P	Y	Y	N	Y
Consistency of components	N	Y	Y		N	N
Open concept repres. language	Y	Y	N	N	N	N
Composition/decomposition	Y	N	N	N	N	N
Signature/Certificate-enabled	N	Y	N	N	Y	Y
Formal logical reference	N	N	P	P		
Machine-processable	P	P	P	P	Y	Y
<b>Domain Perspective</b>						
Domain independence	N	N	N	N	N	N
Domain separation	N	N	N	N	N	P
Ontology driven	N	N	Y	Y	N	N
Vocabulary	Y	N	N	N	N	N
Reference to terminology	Y	Y	Y	Y	Y	Y
Communication security services	N	Y	N	N	Y	Y
Application security services	P	Y	P	N	N	N
Inclusion of medical devices	Y	N	F	N	F	Y
Specialty-related	N	Y	N	N	Y	N
Multimedia-enabled	P	P	N	N	Y	Y
<b>Feasibility of the Approach</b>						
Visualization support	Y	Y	N	N	N	N
Final specification available	N	Y	N	N	Y	Y
Implementation available	Y	Y	N	P	Y	Y
Commercial products available	Y	Y	N	N	Y	Y

Table 83 – GCM for EHR (Blobel & Pharow, 2009) (P-partial, Y-yes, N-no, F-future)

### 0.7 Other works

This section contains some other works, not fitting into the previous mentioned categories, but still valuable to take into account.

### *O.7.1 Requirements specification study*

Although not intended during the requirements workshop several quality characteristics were captured. The experts participating mentioned these aspects, but since the workshop was aimed at defining requirements, these quality attributes itself were noted but not thoroughly discussed. However these reflect their experience and needs for quality in practice. The mentioned quality criteria are:

1. The scope of the standard should be clear.
2. The standard should contain domain analysis.
3. The ease and efficiency of the technological implementation of the standard.
4. The amount of errors.
5. How many functions have explanations of when and why to use.
6. To test the readability and understandability of the standard.
7. To test the owner of the standard.
8. Patent test.
9. Should be independent of the language, but dependent on quality of the use of the language.
10. The amount of optional/mandatory items (too much optional is usually a sign of weakness).
11. The amount of examples (of implementations).
12. The amount of adaptors for the standard.
13. Should contain the structure of the standard (how many documents and their relations).
14. To test dependencies with other (dependent) standards (Quality is affected by depended standard).
15. To check interoperability and compatibility against other mandatory or commonly used standards.
16. To test the performance/operations of the SSO responsible for maintenance.

As example, one of the criteria should be if there is a patent check in place. Not meaning that the standard itself should contains patents. But there has to be check if concepts/materials are used in the design process of the standard that might be protected by patents or other barriers.

### *O.7.2 The product engineering quality domain*

The domain of software quality has extensively made reuse of existing knowledge from the product engineering domain. This implies that indirectly it is already accounted for, but due to its importance below two lists of quality dimensions are presented. First, Garvin identified eight dimensions of quality (Garvin, 1984):

1. Performance
2. Features
3. Reliability
4. Conformance
5. Durability
6. Serviceability
7. Aesthetics
8. Perceived Quality

Second, the dimensions of design quality are (Chase & Aquilano, 1995):

- Performance: Primary product or service characteristics.
- Features: Added touches, bells and whistles, secondary characteristics.
- Reliability: Consistency of performance over time.
- Durability: Useful life.
- Serviceability: Resolution of problems and complaints.

- Response: Characteristics of the human-to-human interface (timeliness, courtesy, professionalism, etc.).
- Aesthetics: Sensory characteristics (sound, feel, look, etc.).
- Reputation: Past performances and other intangibles.

In general terms, and in broader perspective, the following can be learned from quality guru's like Deming, Juran and Crosby (Ghobadian & Speller, 1994):

1. Controlling the process and not the product.
2. Not forgetting the human process.
3. The role of top management (responsible, commitment, etc.).
4. Education and training.
5. Prevention not inspection.
6. All aspects should be looked at; functional integration (Total Quality Management).

### 0.7.3 The Integrate sources

Integrate was the name of the project that sets the first steps into quality measurement of semantic standards. The research methodology involved several theories together with expert workshops. Their main source was ISO 9216, as discussed within the software engineering domain section. In addition several other sources were used:

#### EU Government sources: CAMSS

The project report CAMSS contains an overview of quality characteristics as implemented by governments (for the selection of standards) and others (Valayer, 2008). The results show a general harmony exists in the approach between European governments for assessing standards, resulting in four main criteria (for selection of standards):

- **Suitability criterion:** mainly related to meeting "public administration business" needs, identified criteria are: relevance, applicability, how the standard benefits the government, completeness functionality-wise. Also the relation with interoperability is seen as important, such as: "Well aligned with the Internet standards"; "relevant to an interoperability area", "proven effective for interoperability", "Is there a planned mechanism to assess the interoperability of different vendors and implementations of the standards?"
- **Potential criterion:** Assessment of standards with regard to scalability, maturity, stability and maintenance. And durability specifically in the area of archiving. Impact analysis is often done when assessing the use of a standard.
- **Openness criterion:** Openness is assessed in every case, but there is no universal way.
- **Market conditions criterion:** Market conditions also play a major role in assessing standards for eGovernment. Mainly, identification of best practices, reusability of standards, wide adoption and market support are to be assessed.

#### Other studies

Earlier work showed the following quality characteristics (Folmer & Bastiaans, 2008):

- Implementable:
  - Correct.
  - Univocal.
  - Complete within the scope of intended use:
    - No options.
    - Fitness for use.
    - Enough detail.

- Understandable:
  - Written for the right audience.
  - Terms of reference aligned with audience.
  - Not too complex.
  - Readable.
- Conformance assessment:
  - Present conformance guidelines.
  - Consistent: An exact implementation should yield full conformance.

The NASA has also developed a quality model for its Software Assurance Technology Center (SATC) (Hyatt & Rosenberg, 1996). From this model the requirements quality and product quality have quality attributes that relate to quality of standards. The requirements quality attributes are (Hyatt & Rosenberg, 1996): ambiguity, completeness, understandability, requirement volatility, traceability. Attributes of product quality are structure/architecture, reuse, maintainability, documentation. Ambiguity is measured by counting the number of weak phrases and the number of optional phrases. A weak phrase contains one of the wordings: adequate, as appropriate, as applicable, but not limited to, normal, if practical, timely, as a minimum. While an optional phrases is recognizable by using the words of can, may or optionally (Hyatt & Rosenberg, 1996).

#### 0.7.4 Quality in practice

Not described in literature, the problem survey (Chapter 3) already discovered that although quality is an issue, many standards already have some quality assurance organized. Although it is not researched in detail; by visiting websites of SSOs we learned that at times quality assurance (QA) is mentioned. For instance IHTSDO (maintaining the SNOMED CT standard) has set up a QA committee (<http://www.ihtsdo.org/about-ihtsdo/governance-and-advisory/standing-committees/qa-committee/>).

Based on a LinkedIn Group discussion we get some additional insight in current practice of these QA committees. OpenTravel ([www.opentravel.org](http://www.opentravel.org)) is reviewing the quality assurance part of their publication process (XML Schema), and they would like to understand how other SSOs handle this function. Questions raised by the Open Travel Alliance, that might be exemplary for many SSOs dealing with quality assurance for XML specifications are:

- Is QA formally (and documented) built into the development process, and does it include member involvement through formal structure or by informal review process?
- If members are not involved, how does your organization assure the standard is functionally fit-for-purpose?
- How is feedback collected? Is a staff person responsible for collecting feedback and curing whatever issues might be found? Do you use any sort of testing tools?
- Is your QA function the same for all standards published, regardless of originating work group?

The replies give an in-depth understanding of current options and solutions for quality assurance. Replies refer to policies of STAR (Standard for Technology in Automotive Retail) and DMTF (Distributed Management Task Force). Chuck Allen, based on experience from STAR, HR-XML and LETSI proposes a set of activities, which is summarized as follows:

1. Leverage user acceptance tests. Build in quality from the start. A simple user acceptance test usually means a brief narrative of what the new message or change is intended to accomplish. You pair this user story with an example of realistic data that would sufficient to execute the related business purpose. The output of that are the artifacts that reviewers can use in making qualitative assessments of whether the project achieved what it set out to achieve.
2. Facilitate "active review." Passive review generally is a waste of time. You can have very smart, dedicated people (staff and volunteers) reading documents and reviewing schema and still miss

defects. You need to create QA processes that ensure the standards are actually applied with some level of rigor and not merely looked at. This is the type of review finds problems that passive review simply does not uncover.

3. Test against your own design rules. For each rule, there is a specified rationale and an example. Testing against many design rules might be automated. A common way that design rules are tested is through the construction of Schematron assertions or simple scripts.
4. Build and maintain a test suite. Of course a lot of QA responsibility sits on the shoulders of the architect. It is wise (though not always done) to budget in some time and effort for developing a test suite and for automating testing of your library. What is in your test suite? (a) The collection of business-meaningful XML instances you've created in item. (b) In addition to the business-meaningful XML instances, the architect might also want to test against fully-expanded, auto-generated instances carried forward from the prior release. (c) The design rule test assertions referenced before. (d) Put the draft against a couple of the common code generation tools (assuming people in your community use them).
5. Leverage web 2.0 techniques, by embedded feedback forms in each page of component documentation.
6. Use a public review process for candidate releases.

#### 0.7.5 Other quality instruments

Many frameworks related to quality have been developed, such as a framework for evaluating the quality of web based sustainability reports (Freundlieb & Teuteberg, 2012), but some have become defacto standard instruments. We mention SERVQUAL and LORI.

#### SERVQUAL - RADER

Already mentioned as being proposed as part of the IS Quality Model (Rodriguez & Casanovas, 2010), is the SERVQUAL instrument from marketing. SERVQUAL (Parasuraman, Zeithaml, & Berry, 1988), later renamed to RATER, is an easy to use methodology to study the quality of services. Its key point, which is often criticized is that quality is defined as the gap between service expectations and experience. It measures the user satisfaction which is different than quality (Buttle, 1996). The basic dimensions are RATER (Buttle, 1996), with generic, somewhat customizable, questionnaires:

- Reliability: The ability to perform the promised service dependably and accurately
- Assurance: The knowledge and courtesy of employees and their ability to convey trust and confidence
- Tangibles: The appearance of physical facilities, equipment, personnel and communication materials
- Empathy: The provision of caring, individualized attention to customers
- Responsiveness: The willingness to help customers and to provide prompt service

The five gaps that organizations should measure, manage and minimize:

(<http://en.wikipedia.org/wiki/SERVQUAL>)

**Gap 1** is the distance between what customers expect and what managers think they expect - Clearly survey research is a key way to narrow this gap.

**Gap 2** is between management perception and the actual specification of the customer experience - Managers need to make sure the organization is defining the level of service they believe is needed.

**Gap 3** is from the experience specification to the delivery of the experience - Managers need to audit the customer experience that their organization currently delivers in order to make sure it lives up to the spec.

**Gap 4** is the gap between the delivery of the customer experience and what is communicated to customers - All too often organizations exaggerate what will be provided to customers, or discuss the best case rather than the likely case, raising customer expectations and harming customer perceptions.

**Gap 5** is the gap between a customer's perception of the experience and the customer's expectation of the service - Customers' expectations have been shaped by word of mouth, their personal needs and their own past experiences. Routine transactional surveys after delivering the customer experience are important for an organization to measure customer perceptions of service.

According to Nyeck et al. (2002) SERVQUAL definitely has value, but users use it as is, without looking at the applicability of the instrument for their specific case. For instance by starting with validating the instrument for their specific case would demise most criticism, and would be a good start point to use SERVQUAL. The lessons learned from SERVQUAL include use a simple structure and more attention on validation is justified, especially when developing an instrument for generic use.

### **LORI**

From the education domain stems the LORI (Learning Object Review Instrument) (Nesbit, Belfer, & Leacock, 2011) for assessing learning objects. The LORI measurement addresses nine quality aspects of learning objects:

1. Content Quality
2. Learning Goal Alignment
3. Feedback and Adaptation
4. Motivation
5. Presentation Design
6. Interaction Usability
7. Accessibility
8. Reusability
9. Standards Compliance

Each aspect is scored on a 5-points (stars) scale. For each of the nine quality aspect the extremities (1 and 5 stars) are defined, and an example is given for the state of 3 stars. Finally a one page score sheet, including the name of the learning object, the reviewer and general remarks, just as the score of each of the nine quality aspects.

# IMS Learning Resource Metadata (LRM) Description (chapter 12)

The description of IMS LRM v1.3 is created by applying the Semantic Standard Model (SSM), focusing on the contents and context part as these are related to the information need. By doing that we did not include a description of the development and maintenance organization, just as its application, because with the “internal” quality as focus, these parts seem irrelevant.

### Essential documents:

- IMS Meta-data Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata (Version 1.3 Final Specification; Date 31 august 2006)
- Guidelines for Using the IMS LRM to IEEE LOM 1.0 Transform (Version 1.0; 31 august 2006)
- 1484.12.1-2002 – IEEE Standard for Learning Object Metadata (13 June 2002, 13 may 2009)
- 1484.12.1-2002/Cor 1-2011 – Corrigendum to IEEE 1484.12.1-2002 (31 March 2011)
- 1484.12.3-2005 – IEEE Standard for Learning Technology – Extensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata (17 august 2005)

### Other relevant documents:

- IMS Application Profile Guidelines Overview - Part 1 – Management Overview Version 1.0 (10 October 2005)
- IMS Application Profile Guidelines Technical Manual - Part 2 – Technical Manual Version 1.0 (10 October 2005)
- IMS Vocabulary Definition Exchange (VDEX) - Version 1 Final Specification (February 2004), including the best practice guide, information model, binding, and conformance specification.
- Overview: <http://www.imsglobal.org/metadata/index.html>
- Wikipedia: [http://en.wikipedia.org/wiki/Learning\\_object\\_metadata](http://en.wikipedia.org/wiki/Learning_object_metadata)

#### 1. Targeted Audience

IMS LRM is intended to be localized for specific use in each country, by creating application profiles. The targeted audiences of IMS LRM to create these application profiles are country specific associations within the education domain. The suppliers of tools to whom the ‘learning resources’ play a role can be considered as the targeted audience of the application profiles. Examples of tools and applications are developing tools for learning resources, electronic learning environments, and repositories for learning materials. Teachers are the potential end users of this standard.

#### 2. Adopted Audience

Application profiles are created in many countries, including UK, Spain, Norway, etc. In the Netherlands at this moment two profile versions are defined. The CZP (Content ZoekProfiel) was defined in 2006; the latest version NL-LOM has been public since 2010. Currently this version is not yet widely adopted within the Netherlands. The CZP localization is based on the previous version of IMS LRM 1.2.4.

### 3. Active Community

The project group within IMS is the main actor within the community, although currently not active. Contributors in this project group were predominantly of national associations in the domain of education. Other groups include universities and standard development organizations. Commercial parties took part within the project group but in minority, for example Click2learn and Thomson Netg.

### 4. The Problem

Currently learning resources cannot uniformly be described which makes it difficult to find and assess appropriate learning resources. By standardizing a set of meta-data elements that can be used to describe learning resources it is aimed that the process of finding and using a learning resource will be more efficient.

### 5. Application Domain and Rules

The domain of the LRM standard is primarily the education sector and its learning resources for training, courses, etc. Other sectors are making use of the same IEEE LOM standard, like libraries, museums and image banks albeit that these domains are DublinCore dominated. There seems to be no restrictive rules, like legislation, to be relevant.

### 6. Business Case

The business case primarily related to the efficiency and effectiveness of searching, finding and using learning resources. It includes saving money on easy search, and avoidance costs of developing learning resources that are already available. It also relates to the quality of the education, because now high quality learning resources become better accessible, which might affect the quality of education. In other words the use of the best available learning resources will likely increase. Although the business case seems common sense, there are no calculated values or even a cost-benefits analysis available.

### 7. Paradigm

The main guiding principle is that it is seen as impossible to design a standard that fits all needs. Therefore the paradigm is to develop flexible structures, instead of strictly trying to define all elements.

### 8. Methods/Languages

Although other IMS standards make use of UML, this LRM standard does not contain traces of any method or language used. This might have to do with the relative simplicity of a meta-data standard, like IMS LRM. Other IMS project groups make use of a system that is able to produce documentation based on the information models. However for LRM only tables with definitions and other characteristics have been produced and that is for the IEEE LOM standard. The IMS LRM is a textual description of best practices based on IEEE LOM.

### 9. Architecture

The complexity lies in the fact of very complex situation of several standards that are related in different type of relations. First, several standards relate to meta-data: IMS LRM (v1.3), IEEE LOM (1484.12.1) and also DublinCore is relevant.

Second, the architecture distinguishes the information model and the technical bindings. The technical binding for IEEE LOM is textual standardized within 1484.12.3, without standardizing the technical representation within XSDs (although examples are given). IMS LRM defines XSDs (strict and loose). Both IEEE LOM and IMS LRM seem to support RDF (1484.12.4) as alternative technical binding, however this does not seem to be officially released.

Third, the flexible structure of LRM is implemented mainly by creating “application profiles” and using “vocabularies”. How to use application profiles and vocabularies is standardized within IMS Application Profile Guidelines and IMS Vocabulary Definition Exchange (VDEX).

Fourth, data elements are trying to make use of other existing standards whenever possible, like for instance date/time according to ISO standards.

Fifth, the relation with previous versions and transforms between different versions and bindings has been standardized as well: “Guidelines for Using the IMS LRM to IEEE LOM 1.0 Transform”, just as the availability of several XSLT files for transformation, and mapping between IMS LRM and DublinCore.

#### 10. Domain Model

No traces of a domain model, or an exemplary case, or other links to practice are found within the documents. Albeit that the IMS LRM 1.3, the best practices, contains somewhat more guidance on how to use the IEEE specifications within the domain, for instance by presenting the vocabularies that are often used within the domain.

#### 11. Constraints

There are no rules or constraints modeled explicitly, although some kind of constraints can be found within the information model. Like for instance the use of smallest permitted maximum, or element values that are constrained by another elements value.

#### 12. Process

Processes like ‘searching and finding’ or ‘tagging learning resources’ are not described and part of the standard.

#### 13. Data/Information

The IMS LRM completely uses the IEEE LOM information model that is the core of the standard. The nine main elements are: general, life cycle, meta-metadata, technical, educational, rights, relation, annotation and classification. Each, and their child-elements have been described within a table that captures the number, name, explanation, size, order, value space, data type and an example.

#### 14. Format

IEEE 1484.12.3 defines how XML-bindings can be created. The available IEEE XSDs are examples. Next to that IMS has released XSDs as well with unknown status. Although work has been done on a RDF binding (IEEE 1484.12.4), the results are not officially released. Different bindings can co-exist.

#### 15. Medium (Transport)

IMS LRM does not standardize the transportation of meta-data records; it is completely focused on the content of the meta-data. A variety of protocols are used in practice, including OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) and SRU (Search/Retrieval via URL) / SRW (Search/Retrieve Web services).



## Appendix Q

# Measurement Table IMS LRM (chapter 12)

Concept/Question	Definition	Further explanation	Questions (short name)
A3. Durability	The capability of the standard to have a long lasting life cycle; Connecting to current situations and future proof. (Garvin, 1984)	Includes concepts like changeability and flexibility.	
A3.1 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 considered standard. (ISO 9126)	Includes customizability.	
A3.1.1 Modularity	The logical structured design of the standard.	Re-use within the (content part of the) standard.	<ol style="list-style-type: none"> <li>1. Data dictionary</li> <li>2. Re-use components</li> <li>3. Independence of modules</li> </ol>
1. Assess if a data dictionary is available, and its use in different standard modules.	<p><i>Facts</i></p> <p>The information elements are being defined within IEEE LOM, but there is no data dictionary that is being used within multiple standards. IMS does not have a data dictionary shared by all its standards.</p> <p>Elements with the same name can have different explanations.</p>	<p><i>Feeling</i></p> <p>The definitions and further description of each element within IEEE LOM is of high quality.</p>	<p><i>Opinion</i></p> <p>Data dictionary might help to ease the understanding of the standard.</p> <p>Current presentation is fully focused on implementation.</p> <p>Unique naming of elements would be preferred.</p>
2. Assess if there is re-use of components, building blocks, within the standard.	<p><i>Facts</i></p> <p>There is re-use of certain parts, for instance the structure of vocabularies. Also a building block for identifiers is being reused.</p>	<p><i>Feeling</i></p> <p>The re-use of components might be somewhat over exposed, and not always that practical to use. It feels somewhat uncomfortable.</p>	<p><i>Opinion</i></p> <p>Seems good.</p>
3. Assess the possibilities of using certain modules of a standard independently.	<p><i>Facts</i></p> <p>There are cases available in which only element 5 (Educational) and 9 (Classification) are being used.</p>	<p><i>Feeling</i></p> <p>The coherence between the 9 main categories is limited. Element 9 and 5 together with an identifier, might be re-usable within other information exchange.</p>	<p><i>Opinion</i></p> <p>Somewhat average.</p>
A3.1.2 Dynamic Content	The capability of the standard to deal with dynamic, often changing content, without resulting in version explosion.	Keeping a stable version of the standard, while implementing flexible structures for dealing with dynamic content. (Product identification codes is an example of dynamic content)	<ol style="list-style-type: none"> <li>1. Dynamic content</li> </ol>

<p>1. Assess if code lists or other solutions are used for dynamic content resulting in a separation between dynamic content that is prone to changes and a standard core that proposes a stable representation of the business domain.</p>	<p><i>Facts</i> The standards make use of free to choose vocabularies. By using the flexible structures the standard is extremely stable.</p>	<p><i>Feeling</i> There might be some improvements by using RDF, by defining the tree structure.</p>	<p><i>Opinion</i> The use of vocabularies is good and makes it flexible.</p>
<p>A3.1.3 Extensibility</p>	<p>The extent to which a standard provides possibilities to extend the capabilities without affecting other parts of the implementation. (O'Brien et al., 2005)</p>	<p>Includes scalability, and forward compatibility, to support innovation.</p>	<p>1. Extension mechanism</p>
<p>1. Assess if there is an extension mechanism, such as additional data elements, with rules on how to use them.</p>	<p><i>Facts</i> Extension elements are available and there is guidance on how to use it.</p>	<p><i>Feeling</i> Organizations that want to add an additional data set can do so.</p>	<p><i>Opinion</i> In the application profiles the use of these elements is discouraged. However the current extension elements are a decent feature and the guidance avoids misuse to some extent.</p>
<p>A3.2 Maintainability</p>	<p>The capability of the standard to be modified in an efficient manner. (ISO 9126)</p>	<p>Modifications may include corrections, improvements or adaptation to changes in environment, requirements or functional specifications. This also covers re-usability, replaceability and co-existence.</p>	
<p>A3.2.1 Separation of Concerns</p>	<p>The level of separation of parts that are different in nature.</p>	<p>Separation of technology dependent and technology independent parts. But also separation between content addressing business people or technical people.</p>	<p>1. Separation of concerns</p>

Concept/Question	Definition	Further explanation	Questions (short name)
<p>1. Is separation of concerns applied within the standard (according to a defined policy)?</p>	<p><i>Facts</i> The standard consists of several types of artifacts:</p> <ul style="list-style-type: none"> <li>- Best practices (IMS LRM)</li> <li>- Information model (IEEE 1484.12.1)</li> <li>- Binding document to XML (IEEE 1484.12.3)</li> <li>- XSDs (as examples)</li> <li>- Transformations</li> </ul>	<p><i>Feeling</i> Although the technical and information aspects are separated, it seems not that strict. There is no real separation between different kind of readers of the documentation, like business people or implementers.</p> <p>And also there seems to be overlap, between the documents. There seems overlap as the best practice document summarizes the information model, just as the binding document includes parts of the information model as well. And which one is leading?</p>	<p><i>Opinion</i> Seems hard to improve as the origin lies in the separation of IMS and IEEE. Although the redundancy is troublesome.</p>
<p>A3.2.2 Localizations</p>	<p>The availability and use of localizations and extensions in the main standard</p>	<p>Several standards need localization for specific business/country use.</p>	<p>1. Option of localization 2. Localization rules</p>
<p>1. Are localization / profiles for specific purposes intended and are guidelines provided when to use localizations?</p> <p>2. Is there a strict policy of defining the contents of localizations?</p>	<p><i>Facts</i> There are two documents available for localization guidelines: A management overview, and a technical guideline. These documents are valid for localizations of all IMS specifications. Further on, it is specified within IMS LRM.</p> <p><i>Facts</i> There are general rules that localization should not contain elements that are already available in the main schema. There is also guidance on the technical content. Amongst others it is aimed at restriction of cardinalities. There is also detailed guidance on how to set up vocabularies.</p>	<p><i>Feeling</i> IMS LRM supports and promotes the use of localizations. It gives an impression on where localizations can be used for.</p> <p><i>Feeling</i> It might be possible to give more guidance on the functional contents of localizations. Also there are no checks known for localizations.</p>	<p><i>Opinion</i> Properly documented.</p> <p><i>Opinion</i> Seems appropriate. The level of strictness might be improved.</p>
<p>A3.2.3 Dependability</p>	<p>The dependences on other standards.</p>	<p>Co-existence with other standards, and optimal re-use of existing standards. (e.g. when the standard is a profile on top of other standard)</p>	<p>1. Dependencies</p>

<p>1. What kind of strong dependencies exist, or are they expected, regarding other standards. A strong dependence means that a version update of the other standard has an impact on this standard. Especially focus on generalization-specialization relations; is the dependency logical and managed?</p>	<p><i>Facts</i> IMS LRM is in fact IEEE 1482.12.1-2002 (also called IEEE LOM v1.0). IMS adds best practices and two bindings (XSDs) and transformations (XSLT) from previous IMS versions to IEEE LOM v1.0. The relations are actively managed.</p>	<p><i>Feeling</i> The relation with DublinCore is somewhat vague. Although there is a tendency that DC and IEEE LOM will be more aligned. Mappings from DC to IEEE LOM are included within the specification.</p>	<p><i>Opinion</i> IMS LRM is highly dependable on IEEE LOM and the work relation is not completely transparent and impacts openness. It also seems not efficient to have two maintenance organizations involved.</p>
<p>A3.2.4 Version Continuance</p> <p>1. Only assess the content of different versions, based on version history, and assess whether there is logic in the content changes, including backwards compatibility.</p>	<p><i>Facts</i> There are hardly any differences on functionality, if any. Next to that hardly any information about the functional differences between versions is available.</p>	<p>New versions support the same functions. Backwards compatible when possible.</p> <p><i>Feeling</i> From 1.2.4 to version 1.3 only the identifier element has been changed which introduces problems in practice. Further on mainly changes in naming that led to confusion. From 1.2.1.to version 1.2.4 many technical updates have been processed.</p>	<p>1. Version check</p> <p><i>Opinion</i> Backwards compatibility is taken into account. The changes seems logical. Although it seems mainly a technical view and improvements. Also the latest version IMS version is only needed to align the IEEE version and the IMS version.</p>
<p>A3.3 Advancedness</p> <p>A3.3.1 Installed Base</p>	<p>The state of the art notion of the standard, in relation to current practice and future innovation. The connection of the standard to the current ICT landscape in stakeholder organizations</p>	<p>Advancens in relation to the outside environment at stakeholder organizations. Currentness (ISO 25012): Fit with timeframe. In line with installed base.</p>	<p>1. Current ICT</p>
<p>1. How difficult is integration with existing systems (e.g. SAP) in stakeholders' organizations.</p>	<p><i>Facts</i></p>	<p><i>Feeling</i> Technical integration does not seem to be a problem. It is fairly basic XML technology in the end, and the education ICT domain is relatively innovative.</p>	<p><i>Opinion</i> Integration seems easy.</p>

Concept/Question	Definition	Further explanation	Questions (short name)
A3.3.2 Technical Advancedness	The position of the chosen technology in its life cycle	The maturity of the technology used.	1. Lifecycle position 2. Competing technology
1. Assess the position of the used technology on hype cycles. (e.g. The Gartner Hype Cycle on Technology.)	<i>Facts</i> LRM: Commonly used. IEEE LOM: Commonly used. XML: Extremely commonly used. RDF binding: Experimental.	<i>Feeling</i> LRM: "close to "Plateau of Productivity", although the decline in use of the IMS binding (in comparison with IEEE binding) is remarkable. IEEE LOM: close to "Plateau of Productivity" XML: "Plateau of Productivity" RDF binding: In between Peak of Inflated Expectations and Slope of Enlightenment: it has to prove itself.	<i>Opinion</i> With the exception of RDF, the technology used is very mature.
2. Assess the competing technology. Is the choice of technology justifiable?	<i>Facts</i>	<i>Feeling</i> Currently languages that are used to describe rules are quite popular, for instance SBVR or OCL. However this standard contains not that many conditional aspects, which lowers the advantage of these new approaches.	<i>Opinion</i> The choice of technology seems appropriate. Although the choices are made more than 5 years ago, also in current ICT landscape they seem to be appropriate.
A3.3.3 Business Processes	The connection of the standard with the current business processes of the stakeholders organizations	Commodity level of business processes. (e.g. the standard supports only self-billing, while within the domain this is highly uncommon.)	1. Commodity
1. Assess proposed business processes and compare them to the standardized business processes commonly used in practice.	<i>Facts</i> There are no use cases or any other description of business processes available.	<i>Feeling</i> There was the feeling that use cases were available, but these cannot be found within documentation.	<i>Opinion</i> Remarkable that the use cases cannot be found.
A3.3.4 Conceptual Advancedness	The expected life span of the conceptual solution	The complexity of the solution fits the complexity of the interoperability problem on a conceptual level. And does it fit the foreseen developments within the domain, or is it way ahead of its time (e.g. automated plug and play e-business in 2001)	1. Conceptual complexity 2. Future fit

<p>1. Assess the fit between the complexity of the solution and the complexity of the interoperability problem. Is it overkill?</p>	<p><i>Facts</i> Almost all parts are commonly used in practice.</p>	<p><i>Feeling</i> No overkill, probably more the opposite. More strictness (e.g. cardinality mandatory) would have been welcomed. The parts that are not always used are valid for certain specific applications.</p>	<p><i>Opinion</i> It's a good fit.</p>
<p>2. What is the fit with future views on the interoperability problem domain?</p>	<p><i>Facts</i></p>	<p><i>Feeling</i> Metadata will become even more important in next years with trends like open data. About 10 elements are used within all profiles in different countries. This might tamper future cross-border interoperability. These elements should be shifted to the IMS specification.</p>	<p><i>Opinion</i> Good.</p>

Table 84 – Measurement Table IMS LRM

A3.3.2: A hype cycle in Gartner's interpretation comprises five phases:

(<http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp>)

1. "Technology Trigger" — The first phase of a hype cycle is the "technology trigger" or breakthrough, product launch or other event that generates significant press and interest.
2. "Peak of Inflated Expectations" — In the next phase, a frenzy of publicity typically generates over-enthusiasm and unrealistic expectations. There may be some successful applications of a technology, but there are typically more failures.
3. "Trough of Disillusionment" — Technologies enter the "trough of disillusionment" because they fail to meet expectations and quickly become unfashionable. Consequently, the press usually abandons the topic and the technology.
4. "Slope of Enlightenment" — Although the press may have stopped covering the technology, some businesses continue through the "slope of enlightenment" and experiment to understand the benefits and practical application of the technology.
5. "Plateau of Productivity" — A technology reaches the "plateau of productivity" as the benefits of it become widely demonstrated and accepted. The technology becomes increasingly stable and evolves in second and third generations. The final height of the plateau varies according to whether the technology is broadly applicable or benefits only a niche market.



# Analysis Report of iQMSS Assessment on IMS LRM (chapter 12)

*Authors: Jasper Roes and Erwin Folmer (1-10-2011)*

This report presents the analysis of the iQMSS application results on IMS LRM

### Research Set up

According to the recipe on how to use the iQMSS, we took in principle 3 main steps; preparation, measuring, and analysis. Jos van der Arend (Kennisset) took the role of Client and Standards Developer, while Initiator and Principal roles were taken up by Jasper Roes and Erwin Folmer.

The information need is the durability, longevity expectations, for IMS LRM, in order to build Dutch localizations upon. In the context of localizations is important to know how localizations are dealt with; the adaptability of the standard, but also the current advancedness and the ease of maintenance. For that, the "A3. Durability" part of the QMSS has been selected.

### The results

Although IMS LRM is more complex than needed, and mistakes by taking wrong documents or bindings are easily made, it still is regarded as "save" choice regarding future-proofness. The chosen technology is mature and easy implementable. The standard is completely designed to build a localization profile upon, that can be developed by Kennisset. IMS LRM has quite extreme flexibility and extensibility in its structure that needs further development in specific domains.

Due to the large extent of extensibility and customizability Kennisset has to take into account that localization will be a severe effort, as many of the difficult decisions still have to be made. It is not limited to just change several cardinalities or add one or two local elements. It is far more fundamentally, for example on how to deal (and manage) the vocabularies.

### Improvement suggestions related to IMS

The improvement suggestions have been organized into four categories:

*A. The balance between the IMS LRM and the needed localizations might be shifted.*

1. The level of flexibility is that high that the interoperability in practice might be at stake, or too much dependable on the quality of the localization. It might be worthwhile to investigate the possibilities to include more standardization into the IMS LRM, by which less options and design choices have to be made within the localization profile.
2. Because of the current importance of the localization profiles, it might be an idea to let localizations be certified (by IMS) for quality assurance. Before doing this it is advisable to study some of the current localizations to see if there is an issue or not.

*B. Complex structure needs refinement and guiding.*

1. One of the main problems is the complex architecture of different standards, concepts, bindings, transformations, etc. It is hard to understand, and therefore hard to maintain. Studying simplification of the structure seems appropriate.
2. The added value of IMS LRM (the best practices, and the XML Schema's) seems limited and questionable why it exists and why it is not an informative appendix to IEEE LOM, or part of the IEEE family?
3. The separations of concerns are not that strict resulting in redundancy. For instance the IMS LRM best practice summarizes quite extensive IEEE 1284.12.1 and the IEEE binding document.
4. The relation between IMS and IEEE is not transparent and should be clarified and formalized in order not to interfere with openness principles.
5. There should be more guidance related to the documentation. One point of presence on the Internet is needed that contains all information regarding the standard. For instance the corrigendum of the IEEE specification is not mentioned at the IMS website. Even reduction of the number of documents should be considered.

*C. The content can be enhanced in clarity.*

1. The data dictionary might be improved. By creating a data dictionary the problem that the same names are used for information elements with a different meaning will be solved.
2. Adding context definitions, including the involved actors, systems, etc., would contribute to better understanding.
3. The coherence and logical structuring of the nine main categories within the standard can be improved albeit somewhat gold-plating; for instance group 5 (educational) and 9 (classification) would be more obvious to combine.

*D. Version management should be improved.*

1. Although there seems version continuance, information about version changes is inadequate. The IEEE corrigendum is a good example how changes can be documented. An overview of the changes and differences between the versions is much appreciated.
2. The number of versions, especially since there are hardly any functional improvements made, is questionable. The impression exist that less versions would have done the job just as well, and that there are probably better alignment options with IEEE (see also B2). A separation in functional and technical updates is also recommended.
3. Version naming is confusing: IEEE 1482.12.1-2002 contains LOM v1.0. What will a new version be called? IEEE 14.82.12.1-2013 and LOM v1.1 or LOM v1.0? It adds unnecessary complexity.

# References

- Abran, A., Khelifi, A., & Suryan, W. (2003). Usability Meanings and Interpretations in ISO Standards. *Software Quality Journal*, 11, 325-338.
- Akao, Y. (Ed.). (1990). *Quality Function Deployment: Integrating Customer Requirements into Product Design*. New York: Productivity Press.
- Al-Subaie, H. S. F., & Maibaum, T. S. E. (2006). *Evaluating the effectiveness of a goal-oriented requirements engineering method*. Paper presented at the Fourth International Workshops on Comparative Evaluation in Requirements Engineering, CERE'06, St. Paul, MN.
- Albrecht, C. C., Dean, D. L., & Hansen, J. V. (2005). Marketplace and technology standards for B2B e-commerce: Progress, challenges, and the state of the art. *Information and Management*, 42(6), 865-875.
- Alexander, J., & Tate, M. (1999). *Web Wisdom; How to Evaluate and Create Information Quality on the Web*. Mahwah, NJ: L. Erlbaum Associates Inc.
- Ambler, S., W. (2009). Enterprise Unified Process (EUP). Retrieved 11-11, 2009, from <http://www.enterpriseunifiedprocess.com/>
- Anh, V. T. (2007). An Introduction - The History of Standardization. In W. Hesser, A. Feilzer & H. J. De Vries (Eds.), *Standardisation in Companies and Markets* (2 ed., pp. 45-76). Hamburg: Helmut Schmidt University.
- Anicic, N., Ivezic, N., & Jones, A. (2006). An Architecture for Semantic Enterprise Application Integration Standards. In D. Konstantas, J.-P. Bourrières, M. Léonard & N. Boudjlida (Eds.), *Interoperability of Enterprise Software and Applications* (pp. 25-34). London: Springer-Verlag.
- ANSI. (2010). *United States Standards Strategy*. New York: ANSI.
- Asuncion, C. H., & Van Sinderen, M. J. (2010). Pragmatic Interoperability: A Systematic Review of Published Definitions. In P. Bernus, G. Doumeingts & M. Fox (Eds.), *Enterprise Architecture, Integration and Interoperability (EAI2N), part of WCC2010* (pp. 164-176). Brisbane: Springer.
- Backhouse, J., Hsu, C. W., & Silva, L. (2006). Circuits of Power in Creating De Jure Standards: Shaping an International Information Systems Security Standard. *MIS Quarterly*, 30(S1), 413-438.
- Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science*, 29(5), 530-545.
- Barresi, S., Rezgui, Y., Lima, C., & Meziane, F. (2008). Semantic Resources Integration and Interoperability in the Construction Domain. In Y. Manolopoulos, J. Filipe, P. Constantopoulos & J. Cordeiro (Eds.), *Enterprise Information Systems* (Vol. 3, pp. 336-347). Berlin Heidelberg: Springer.
- Basili, V. R., Caldiera, G., & Rombach, H. D. (1994). The Goal Question Metric Approach. In J. J. Marciniak (Ed.), *Encyclopedia of Software Engineering* (Vol. 2 Volume Set). Hoboken, NJ: John Wiley & Sons, Inc.
- Baskerville, R. L. (1997). Distinguishing action research from participative case studies. *Journal of Systems and Information Technology*, 1(1), 24-43.
- Beck, R., & Weitzel, T. (2005). Some Economics of Vertical Standards: Integrating SMEs in EDI Supply Chains. *Electronic Markets*, 15(4), 313-322.
- Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering*, 1(3), 213-222.
- Becker, J., Rosemann, M., & Schütte, R. (1995). Grundsätze ordnungsmäßiger Modellierung. *Wirtschaftsinformatik*, 37(5), 435-445.

- Bedini, I., Gardarin, G., & Nguyen, B. (2011). Semantic Technologies and E-Business. In E. Kajan (Ed.), *Electronic Business Interoperability: Concepts, Opportunities and Challenges*. (pp. 243-278). Hershey: IGI Global.
- Benbasat, I., & Zmud, R. (1999). Empirical Research in Information Systems: The Practice of Relevance. *MIS Quarterly*, 23(1), 3-16.
- Benbasat, I., & Zmud, R. W. (2003). The identity crisis within the IS discipline: Defining and communicating the discipline's core properties. *MIS Quarterly*, 27(2), 183-194.
- Benders, J., Batenburg, R., & Van Der Blonk, H. (2006). Sticking to standards; Technical and other isomorphic pressures in deploying ERP-systems. *Information and Management*, 43(2), 194-203.
- Bernstein, P. A., & Haas, L. M. (2008). Information integration in the enterprise. *Communications of the ACM*, 51(9), 72-79.
- Berre, A., Elvesæter, B., Figay, N., Guglielmina, C., Johnsen, S., Karlsen, D., et al. (2007). The ATHENA Interoperability Framework. In R. J. Gonçalves, J. P. Müller, K. Mertins & M. Zelm (Eds.), *Enterprise Interoperability II - New Challenges and Approaches* (pp. 569-580). London: Springer.
- Bhattacharyya, G. K., & Johnson, R. A. (1977). *Statistical Concepts and Methods*. New York: John Wiley & Sons.
- Blind, K. (2004). *The economics of standards; theory, evidence, policy*. Cheltenham: Edward Elgar.
- Blind, K. (2009). A Taxonomy of Service Standards and a Modification for E-Business. In K. Jakobs (Ed.), *Information Communication Technology Standardization for E-Business Sectors; Integrating Supply and Demand Factors* (pp. 24-30). Hershey: Information Science Reference.
- Blobel, B., & Pharow, P. (2009). Analysis and evaluation of EHR approaches. *Methods Inf Med*, 2009(48), 162-169.
- Blumberg, B., Cooper, D. R., & Schindler, P. S. (2011). *Business Research Methods* (3rd ed.). London: Mc-Graw Hill.
- Boehm, B. W. (1973). *Characteristics of software quality*. Redondo Beach, CA: Thompson Ramo Wooldridge Systems Group, Systems Engineering and Integration Division.
- Boh, W. F., Soh, C., & Yeo, S. (2007). Standards development and diffusion: A case study of RosettaNet. *Communications of the ACM*, 50(12), 57-62.
- Bovee, M., Ettredge, M. L., Srivastava, R. P., & Vasarhelyi, M. A. (2002). Does the Year 2000 XBRL Taxonomy Accommodate Current Business Financial-Reporting Practice? *Journal of Information Systems*, 16(2), 165-182.
- Branscomb, L. M., & Kahin, B. (1995). Standards Processes and Objectives for the National Information Infrastructure. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 3-34). Cambridge: The MIT Press.
- Bridgeland, D. M., & Zahavi, R. (2009). *Business modeling: a practical guide to realizing business value*. Burlington (MA): Morgan Kaufmann Publishers.
- Brinkkemper, S. (1996). Method engineering: engineering of information systems development methods and tools. *Information and Software Technology*, 38(4), 275-280.
- Brinkkemper, S. (1998). Assembly techniques for method engineering. *Advanced Information Systems Engineering*, 1413, 381-400.
- Brunnermeier, S. B., & Martin, S. A. (2002). Interoperability costs in the US automotive supply chain. *Supply Chain Management*, 7(2), 71-82.
- Brutti, A., Cerminara, V., D'Agosta, G., Sabbata, P., & Gessa, N. (2010). Use Profile Management for Standard Conformant Customisation. In K. Popplewell, J. Harding, R. Poler & R. Chalmeta (Eds.), *Enterprise Interoperability IV - Making the Internet of the Future for the Future of Enterprise* (pp. 449-459). London: Springer.
- Brutti, A., De Sabbata, P., Frascella, A., Novelli, C., & Gessa, N. (2011). Standard for eBusiness in SMEs networks: the increasing role of customization rules and conformance testing tools to achieve interoperability. In M. Zelm, M. Van Sinderen, G. Doumeingts & P. Johnson (Eds.), *Enterprise Interoperability, proceedings of the Workshops of the Third International IFIP Working Conference IWEI 2011*. Stockholm: Wiley.

- Brzezinski, K. M. (2010a). On Aligning the Properties of Standards with the Needs of Their Direct Users, Network Operators. In K. Jakobs (Ed.), *New Applications in IT Standards - Development and Progress* (pp. 70-94). Hershey - New York: Information Science Reference.
- Brzezinski, K. M. (2010b). Standards are Signs. In J.-C. Graz & K. Jakobs (Eds.), *15th EURAS Annual Standardisation Conference "Service Standardization"*. Lausanne: The EURAS Board Series.
- Bucciarelli, P. (1995). The Current Debate on IT Standardization Policy in the European Union. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 421-429). Cambridge: The MIT Press.
- Bundschuh, M., & Dekkers, C. (2008). *The IT Measurement Compendium - Estimating and Benchmarking Success with Functional Size Measurement*. Berlin-Heidelberg: Springer-Verlag.
- Burton-Jones, A., & Gallivan, M. J. (2007). Toward a deeper understanding of system usage in organizations: a multilevel perspective *MIS Quarterly*, 31(4), 657-679.
- Buttle, F. (1996). SERVQUAL; review, critique, research agenda. *European Journal of Marketing*, 30(1), 8-32.
- Byrd, L. W., & Byrd, T. A. (2012). *Developing an Instrument for Information Quality for Clinical Decision Making*. Paper presented at the 45th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- CabinetOffice. (2011). *Procurement Policy Note (PPN) Use of Open Standards when specifying ICT requirements*. London: UK Cabinet Office.
- Cargill, C. F. (1989). *Information Technology Standardization: theory, process, and organizations*. Bedford: Digital Press.
- Cargill, C. F. (1995). A Five-Segment Model for Standardization. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 79-99). Cambridge: The MIT Press.
- Cargill, C. F., & Bolin, S. (2007). Standardization: a failing paradigm. In S. Greenstein & V. Stango (Eds.), *Standards and Public Policy* (pp. 296-328). Cambridge: Cambridge University Press.
- Cash Jr, J. I., & Konsynski, B. R. (1985). IS redraws competitive boundaries. *Harvard Business Review*, 63(2), 134-142.
- Cathomen, I., & Klein, S. (1997). The development of FEDI in Switzerland: A life-cycle approach. *International Journal of Electronic Commerce*, 1(4), 129-145.
- Cavano, J., P., & McCall, J., A. . (1978a). A framework for the measurement of software quality. In S. Jackson & J. A. Lockett (Eds.), *Proceedings of the software quality assurance workshop on Functional and performance issues* (pp. 133-139).
- Cavano, J., P., & McCall, J., A. (1978b). A framework for the measurement of software quality. *ACM Sigsoft Software Engineering Notes*, 3(5), 133-139.
- Chang, C., & Jarvenpaa, S. (2005). Pace of Information Systems Standards Development and Implementation: The Case of XBRL. *Electronic Markets*, 15(4), 365-377.
- Chari, K., & Seshadri, S. (2004). Demystifying integration. *Communications of the ACM*, 47(7), 58-63.
- Chase, R. B., & Aquilano, N., J. (1995). *Production and Operations Management: Manufacturing and Services* (7th ed.). Chicago: Irwin.
- Chen, D., Doumeingts, G., & Vernadat, F. (2008). Architectures for enterprise integration and interoperability: Past, present and future. *Computers in Industry*, 59(7), 647-659.
- Chen, M. (2003). Factors affecting the adoption and diffusion of XML and web services standards for e-business systems. *International Journal of Human Computer Studies*, 58(3), 259-279.
- Chikofsky, E. J., & Cross, J. H. I. (1990). Reverse Engineering and Design Recovery: A Taxonomy. *IEEE Software*, 7(1), 13-17.
- Choi, B., Raghu, T. S., & Vinze, A. (2004). Addressing a standards creation process: A focus on ebXML. *International Journal of Human Computer Studies*, 61(5), 627-648.
- Chong, A. Y.-L., & Ooi, K.-B. (2008). Adoption of interorganizational system standards in supply chains - An empirical analysis fo RosettaNet standards. *Industrial Management & Data Systems*, 108(4), 529-547.
- Chwelos, P., Benbasat, I., & Dexter, A. S. (2001). Research Report: Empirical Test of an EDI Adoption Model. *Information Systems Research*, 12(3), 304-321.

- Cole, J. (2004). Raising the bar. *Computer*, 37(9), 112-113.
- Cooney, M. (2011). US revamps science, technology standard-setting efforts. Retrieved 06-10, 2011, from <http://www.networkworld.com/community/blog/us-revamps-science-technology-standard-setting>
- Coulin, C. (2006). A situational method engineering approach to requirements elicitation workshops in the software development process. *Software process improvement and practice*, 11(5), 451.
- Creswell, J. W. (2009). *Research Design - Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Los Angeles: Sage.
- Crosby, P. M. (1979). *Quality is free, The art of making quality certain*. New York: McGraw-Hill.
- Damsgaard, J., & Truex, D. (2000). Binary trading relations and the limits of EDI standards: The Procrustean bed of standards. *European Journal of Information Systems*, 9(3), 173-188.
- Davenport, T. H. (2005). The Coming Commodization of Processes. *Harvard Business Review*, 83(6), 100-108.
- David, P. A., & Greenstein, S. (1990). The economics of compatibility standards: an introduction to recent research. *Economics of Innovation and New Technology*, 1, 3-41.
- Davies, J., Harris, S., Crichton, C., Shukla, A., & Gibbons, J. (2008). Metadata standards for semantic interoperability in electronic government. In *Proceedings of the 2nd International Conference on Theory and Practice of Electronic Governance* (pp. 67-75). New York, NY: ACM.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982-1003.
- De Vries, H. J. (2006). IT Standards Typology. In K. Jakobs (Ed.), *Advanced Topics in Information Technology Standards and Standardization Research* (Vol. 1, pp. 1-26). Hershey: Idea Group Publishing.
- De Vries, H. J. (2007). Fundamentals of Standards and Standardization. In W. Hesser, A. Feilzer & H. J. De Vries (Eds.), *Standardisation in Companies and Markets* (2 ed., pp. 1-43). Hamburg: Helmut Schmidt University.
- De Vries, H. J. (2008). Best Practice in Company Standardization. In K. Jakobs (Ed.), *Standardization Research in Information Technology: New Perspectives* (pp. 27-47). Hershey: Information Science Reference.
- Dedeke, A. (2000). A conceptual framework for developing quality measures for information systems. Paper presented at the Proceedings of 5th International Conference on Information Quality.
- Delen, G. P. A. J., & Rijsenbrij, D. B. B. (1992). The specification, engineering, and measurement of information systems quality. *The Journal of Systems and Software*, 17(3), 205-217.
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60-95.
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of management information systems*, 19(4), 9-30.
- Deming, W. E. (1986). *Out of the crisis: quality, productivity and competitive position*. Cambridge: Cambridge University Press.
- Dogac, A., Kabak, Y., Namli, T., & Okcan, A. (2008). Collaborative business process support in eHealth: Integrating IHE profiles through ebXML business process specification language. *IEEE Transactions on Information Technology in Biomedicine*, 12(6), 754-762.
- Dogac, A., Pattenden, S., & Zelm, M. (2010). *Future Internet Enterprise Systems (FIeS) Standardisation Task Force Report*.
- DuBois, F. L., & Reeb, D. (2000). Ranking the international business journals. *Journal of International Business Studies*, 31(4), 689-704.
- EC. (2004). *European Interoperability Framework for Pan-European eGovernment Services*. Brussels: European Commission.
- EC. (2008). *The European e-Business Report 2008; The Impact of ICT and e-business on firms, sectors and the economy*. Brussels: European Commission.
- EC. (2009). *Modernising ICT Standardisation in the EU - The Way Forward* (No. COM(2009) 324 final). Brussels: European Commission.

- EC. (2010a). *2010-2013 ICT Standardisation Work Programme for industrial innovation*. Brussels: European Commission.
- EC. (2010b). *A Digital Agenda for Europe* (No. COM(2010) 245). Brussels: European Commission.
- EC. (2010c). *Europe 2020 Flagship Initiative Innovation Union* (No. SEC(2010) 1161). Brussels: European Commission.
- EC. (2010d). *The European eGovernment Action Plan 2011-2015* (No. COM(2010) 743). Brussels: European Commission.
- EC. (2010e). *European Interoperability Framework (EIF) for European public services* (No. COM(2010) 744 final). Brussels: European Commission.
- EC. (2011a). *European Council 4 february 2011 Conclusions* (No. EUCO 2/1/11). Brussels: European Commission.
- EC. (2011b). *Proposal for a Regulation of the European Parliament and of the Council on European Standardisation and amending Council Directives 89/686/EEC and 93/15/EEC and Directives 94/9/EC, 94/25/EC, 95/16/EC, 97/23/EC, 98/34/EC, 2004/22/EC, 2007/23/EC, 2009/105/EC and 2009/23/EC of the European Parliament and of the Council* (No. COM(2011) 315 final). Brussels: European Commission.
- EC. (2011c). *A strategic vision for European standards: Moving forward to enhance and accelerate the sustainable growth of the European economy by 2020* (No. COM(2011) 311 final). Brussels: European Commission.
- Economides, N. (1996). The Economics of networks. *International Journal of Industrial Organization*, 14(6), 673-699.
- Egyedi, T. M. (2000). Institutional Dilemma in ICT standardization: Coordinating the diffusion of technology? In K. Jakobs (Ed.), *Information Technology Standards and Standardization: A Global Perspective* (pp. 48-62). Hershey: Idea Group Publishing.
- Egyedi, T. M. (2003). Consortium problem redefined: Negotiating 'democracy' in the actor network on standardization. *Journal of IT Standards and Standardization Research*, 1(2), 22-38.
- Egyedi, T. M. (2007). Experts on Causes of Incompatibility between Standard-Compliant Products. In G. Doumeings, J. Müller, G. Morel & B. Vallespir (Eds.), *Enterprise Interoperability* (pp. 553-563): Springer London.
- Egyedi, T. M. (2008). An Implementation Perspective on Sources of Incompatibility and Standards' Dynamics In T. M. Egyedi & K. Blind (Eds.), *The Dynamics of Standards* (pp. 28-43). Cheltenham: Edward Elgar.
- Egyedi, T. M. (2009). Between Supply and Demand: Coping with the Impact of Standards Change. In K. Jakobs (Ed.), *Information Communication Technology Standardization for E-Business Sectors; Integrating Supply and Demand Factors* (pp. 171-185). Hershey: Information Science Reference.
- Egyedi, T. M., & Blind, K. (2008). General Introduction. In T. M. Egyedi & K. Blind (Eds.), *The Dynamics of Standards* (pp. 1-12). Cheltenham: Edward Elgar.
- Egyedi, T. M., & Heijnen, P. (2008). How stable are IT standards? In T. M. Egyedi & K. Blind (Eds.), *The Dynamics of Standards* (pp. 137-154). Cheltenham: Edward Elgar.
- Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., & Laleci, G. B. (2005). A survey and analysis of electronic healthcare record standards. *ACM Computing Surveys*, 37(4), 277-315.
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, 14(4), 532-550.
- Elgarah, W., Falaleeva, N., Saunders, C., C. , Ilie, V., Shim, J. T., & Courtney, J., F. . (2005). Data exchange in interorganizational relationships: review through multiple conceptual lenses. *ACM SIGMIS Database*, 36, 8-29.
- EP. (2010). *Report on the future of European standardisation* (No. 2010/2051 (INI)). Brussels: European Parliament.
- Ernst, D. (2010). *Indigenous Innovation and Globalization: the Challenge for China's Standardization Strategy*. Honolulu: East-West Center.

- Falge, C., Otto, B., & Österle, H. (2012). *Data Quality Requirements of Collaborative Business Processes*. Paper presented at the 45th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- Fenton, N., E. , & Neil, M. (2000). *Software metrics: roadmap*. Paper presented at the Proceedings of the Conference on The Future of Software Engineering.
- Fodor, O., & Werthner, H. (2004). Harmonise: A step toward an interoperable e-tourism marketplace. *International Journal of Electronic Commerce*, 9(2), 11-39.
- Folmer, E. (2011). *The Quality Model of Semantic IS Standards*. Paper presented at the 8th International Conference "Standardization, Prototypes and Quality: A means of Balkan Countries' Collaboration". Thessaloniki.
- Folmer, E., & Bastiaans, J. (2008). Quality of Electronic Messaging Standards' Specifications. In K.-D. Thoben, K. S. Pawar & R. Gonçalves (Eds.), *Proceeding of the 14th International Conference on Concurrent Enterprising: ICE2008*. Lisbon: Centre for Concurrent Enterprise.
- Folmer, E., Berends, W., Oude Luttighuis, P., & Van Hillegersberg, J. (2009). Top IS research on quality of transaction standards, a structured literature review to identify a research gap. In K. Jakobs (Ed.), *Proceedings of 6th International Conference on Standardization and Innovation in Information Technology (SIIT 2009)*. Tokyo: Verlag Mainz.
- Folmer, E., Berends, W., Oude Luttighuis, P., Van Hillegersberg, J., & Lammers, R. (2010). *Research on Quality of Transaction Standards: The Maturity of a Research Topic*. Paper presented at the Enterprise Architecture, Integration and Interoperability (EAI2N), part of WCC2010, Brisbane.
- Folmer, E., Hinderer, H., & Otto, B. (2003). The openXchange framework for ebXML based business process integration. *Proceedings of the 10th ISPE International Conference on Concurrent Engineering Research and Application, Enhanced Interoperable Systems*, 447-453.
- Folmer, E., Krukkert, D., Oude Luttighuis, P., & Van Hillegersberg, J. (2010). *Requirements for a quality measurement instrument for semantic standards*. Paper presented at the 15th EURAS Annual Standardisation Conference, Lausanne.
- Folmer, E., Oude Luttighuis, P., & van Hillegersberg, J. (2011a). Do semantic standards lack quality? A survey among 34 semantic standards. *Electronic Markets*, 21(2), 99-111.
- Folmer, E., Oude Luttighuis, P., & Van Hillegersberg, J. (2011b). *A Model for Semantic IS Standards*. Paper presented at the 16th EURAS Annual Standardisation Conference.
- Folmer, E., Oude Luttighuis, P., & Van Hillegersberg, J. (2011c). A Reference Model of an Instrument for Quality Measurement of Semantic Standards. In K. Blind & K. Jakobs (Eds.), *Proceedings of the 7th International Conference on Standardisation and Innovation in Information Technology (SIIT 2011)* (pp. 69-78). Berlin: IEEE.
- Folmer, E., & Punter, M. (2011). *Management and Development Model for Open Standards (BOMOS) version 2*. The Hague: Netherlands Open in Connection.
- Folmer, E., Van Bekkum, M., Oude Luttighuis, P., & Van Hillegersberg, J. (2011). *The Measurement of Quality of Semantic Standards: The Application of a Quality Model on the SETU standard for eGovernment*. Paper presented at the 16th EURAS Annual Standardisation Conference.
- Folmer, E., & Van Soest, J. (2011). *Towards a quality model for Semantic IS Standards*. Paper presented at the International Conference on Information Quality, Adelaide.
- Folmer, E., & Verhoosel, J. (2011). *State of the Art on Semantic IS Standardization, Interoperability & Quality*. Enschede: TNO, University of Twente, CTIT, NOiV.
- Fomin, V. V., Su, J., & Gao, P. (2011). Indigenous standard development in the presence of dominant international standards: the case of the AVS standard in China. *Technology Analysis & Strategic Management*, 23(7), 745-758.
- Fowler, M. (2004). *UML Distilled: A Brief Guide to the Standard Object Modelling Language* (3rd ed.). Boston, MA: Pearson Education.
- Freericks, C. (2010). *Workable Service Standard Guides Through Meta-Standards*. Paper presented at the 15th EURAS Annual Standardisation Conference "Service Standardization", Lausanne.

- Frenkel, K. A. (1990). Politics of standards and the EC. *Communications of the ACM*, 33(7), 40-51.
- Freundlieb, M., & Teuteberg, F. (2012). *Evaluating the Quality of Web Based Sustainability Reports - A Multi-Method*. Paper presented at the 45th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- Frick, N., & Schubert, P. (2011, June 12-15, 2011). *A Maturity Model for B2B Integration (BIMM)*. Paper presented at the 24th International Bled eConference eFuture, Bled, Slovenia.
- Friesen, N. (2005). Interoperability and Learning Objects: An Overview of E-Learning Standardization. *Interdisciplinary Journal of Knowledge and Learning Objects*, Volume 1, 23-31.
- Gallaher, M. P., O'Conner, A. C., & Phelps, T. (2002). *Economic Impact Assessment of the International Standard for the Exchange of Product Model Data (STEP) in Transportation Equipment Industries* (No. RTI Project Number 07007.016).
- Gallaher, M. P., O'Connor, A. C., Dettbarn Jr, J. L., & Gilday, L. T. (2004). *Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry*. Gaithersburg: NIST.
- García, F., Bertoa, M. F., Calero, C., Vallecillo, A., Ruiz, F., Piattini, M., et al. (2006). Towards a consistent terminology for software measurement. *Information and Software Technology*, 48(8), 631-644.
- García, F., Ruiz, F., Calero, C., Bertoa, M. F., Vallecillo, A., Mora, B., et al. (2009). Effective use of ontologies in software measurement. *The Knowledge Engineering Review*, 24(SI 01), 23-40.
- García, F., Ruiz, F., & Piattini, M. (2004). Definition and empirical validation of metrics for software process models. *Product Focused Software Process Improvement*, 3009, 146-158.
- García, F., Serrano, M., Cruz-Lemus, J., Ruiz, F., & Piattini, M. (2007). Managing software process measurement: A metamodel-based approach. *Information Sciences*, 177(12), 2570-2586.
- Garvin, D. A. (1984). What Does "Product Quality" Really Mean? *Sloan management review*, 26(1), 24-43.
- Gauch, S. (2008). + vs -: Dynamics and Effects of Competing Standards of Recordable DVD-Media. In T. M. Egyedi & K. Blind (Eds.), *The Dynamics of Standards* (pp. 47-67). Cheltenham: Edward Elgar.
- Gerst, M., Bunduchi, R., & Williams, R. (2005). *Social shaping & standardization: A case study from auto industry*. Paper presented at the 38th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- Gerst, M., Iversen, E., & Jakobs, K. (2009). An Integrated View of E-Business and the Underlying ICT Infrastructure. In K. Jakobs (Ed.), *Information Communication Technology Standardization for E-Business Sectors; Integrating Supply and Demand Factors* (pp. 1-7). Hershey: Information Science Reference.
- Ghobadian, A., & Speller, S. (1994). Gurus of quality: a framework for comparison. *Total Quality Management*, 5(3), 53-69.
- Glass, R. L. (2008). The notion of an information system Apgar score. *Information Systems Management*, 25(3), 290-291.
- Gol. (2010). *Policy on Open Standards*. New Delhi: Government of India.
- Gonzalez-Barbone, V., & Llamas-Nistal, M. (2007). eAssessment: Trends in content reuse and standardization. In *Frontiers In Education Conference - Global Engineering: Knowledge Without Borders, Opportunities Without Passports, 2007. FIE '07. 37th Annual* (pp. 11-16). Milwaukee, WI: IEEE.
- Gottschick, J., & Restel, H. (2010). An Empirical Evaluation of the Quality of Interoperability Specifications for the Web. In *Software Engineering and Advanced Applications (SEAA)* (pp. 398-405).
- Grant, S., & Young, R. (2010). Concepts and Standardization in Areas Relating to Competence. *International Journal of IT Standards and Standardization Research*, 8(2), 29-44.
- Gregor, S. (2006). The Nature of the Theory in Information Systems. *MIS Quarterly*, 30(3), 611-642.
- Grindley, P. (1995). *Standards, Strategy, and Policy: Cases and Stories*. Oxford: Oxford University Press.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). London: Sage.
- Guedria, W., Chen, D., & Naudet, Y. (2009). A Maturity Model for Enterprise Interoperability. In R. Meersman, P. Herrero & T. Dillon (Eds.), *Lecture Notes in Computer Science (LNCS): OTM 2009 Workshops* (Vol. 5872, pp. 216-225). Berlin-Heidelberg: Springer-Verlag.

- Hanseth, O., Jacucci, E., Grisot, M., & Aanestad, M. (2006). Reflexive standardization: Side effects and complexity in standard making. *MIS Quarterly: Management Information Systems*, 30(SPEC. ISS.), 563-581.
- Hardwick, M., Spooner, D. L., Rando, T., & Morris, K. C. (1996). Sharing Manufacturing Information in Virtual Enterprises. *Communications of the ACM*, 39(2), 46-54.
- Harrison, G. W., & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, XLII, 1009-1055.
- Hart, P., & Saunders, C. (1997). Power and Trust: Critical Factors in the Adoption and Use of Electronic Data Interchange. *Organization Science*, 8(1), 23-42.
- Hasselbring, W. (2000). The Role of Standards for Interoperating Information Systems. In K. Jakobs (Ed.), *Information Technology Standards and Standardization: A Global Perspective* (pp. 116-130). Hershey: Idea Group Publishing.
- Hawkins, R. (2009). Business Models and the Dynamics of Supply and Demand for Standards. In K. Jakobs (Ed.), *Information Communication Technology Standardization for E-Business Sectors; Integrating Supply and Demand Factors* (pp. 31-51). Hershey: Information Science Reference.
- Heck, P., Klabbers, M., & van Eekelen, M. (2010). A software product certification model. *Software Quality Journal*, 18(1), 37-55.
- Heck, P., & Van Eekelen, M. (2008). *LaQuSo Software Product Certification Model*. Eindhoven-Nijmegen: LaQuSo.
- Henderson-Sellers, B., Simons, A., & Younessi, H. (1998). *The OPEN Toolbox of Techniques*. Harlow: Addison-Wesley.
- Hesser, W., & Czaya, A. (2007). The Standardization Policy of the European Union. In W. Hesser, A. Feilzer & H. J. De Vries (Eds.), *Standardisation in Companies and Markets* (2 ed., pp. 455-510). Hamburg: Helmut Schmidt University.
- Hesser, W., Czaya, A., & Riemer, N. (2007). Development of Standards. In W. Hesser, A. Feilzer & H. J. De Vries (Eds.), *Standardisation in Companies and Markets* (2 ed., pp. 123-169). Hamburg: Helmut Schmidt University.
- Hevner, A., & Chatterjee, S. (2010). *Design Research in Information Systems - Theory and Practice* (Vol. 22). New York-London: Springer.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly: Management Information Systems*, 28(1), 75-105.
- Hoel, T., & Hollins, P. A. (2008). *Learning technology standards adoption - how to improve process and product legitimacy*. Paper presented at the ICALT - The 8th IEEE International Conference on Advanced Learning Technologies.
- Hoel, T., Hollins, P. A., & Pawlowski, J. M. (2010). On the Status of Learning Technology Specifications and Standards. *International Journal of IT Standards and Standardization Research*, 8(2), i-vi.
- Hoel, T., & Mason, J. (2011). Expanding the Scope of Metadata and the Issue of Quality. In T. Hirashima et al. (Ed.), *Proceedings of the 19th International Conference on Computers in Education*. Chang Mai, Thailand: Asia-Pacific Society for Computers in Education.
- Horrocks, I., Patel-Schneider, P., Boley, H., Tabet, S., Grosof, B., & Dean, M. (2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. Retrieved 10-10, 2010, from <http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/>
- Hovav, A., Patnayakuni, R., & Schuff, D. (2004). A model of Internet standards adoption: The case of IPv6. *Information Systems Journal*, 14(3), 265-294.
- Humphrey, W. S. (1989). *Managing the Software Process*. Reading: Addison-Wesley Publishing Company.
- Humphrey, W. S. (1997). *A Discipline for Software Engineering* (6th ed.). Reading: Addison-Wesley Publishing Company.
- Hyatt, L., & Rosenberg, L. (1996). *A Software Quality Model and Metrics for Identifying Project Risks and Assessing Software Quality*. Paper presented at the Product Assurance Symposium and Software Product Assurance Workshop.

- IEEE. (1990). *IEEE Standard Glossary of Software Engineering Terminology* (No. 610.12-1990 ). New York: IEEE Computer Society.
- ISO/IEC. (1984). *International vocabulary of basic and general terms in metrology*.
- ISO/IEC. (1999). *ISO/IEC 14598-1 Information technology - Software product evaluation - Part 1: General overview*.
- ISO/IEC. (2001). *ISO/IEC 9126-1 Software engineering - Product quality - Part 1: Quality model*.
- ISO/IEC. (2003a). *ISO/IEC 9126-2 Software engineering - Product quality - Part 2: External metrics*.
- ISO/IEC. (2003b). *ISO/IEC 9126-3 Software engineering - Product quality - Part 3: Internal metrics*.
- ISO/IEC. (2007). *ISO/IEC 25021 Software engineering - Software product Quality Requirements and Evaluation (SQuARE) - quality measure elements*.
- ISO/IEC. (2008). *ISO/IEC 25012 Software engineering - Software product Quality Requirements and Evaluation (SQuARE) - Data quality model*.
- ISO/IEC. (2011). *ISO/IEC 25010 Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuARE) - System and software quality models*.
- Issac, G., Rajendran, C., & Anantharaman, R. N. (2006). An instrument for the measurement of customer perceptions of quality management in the software industry: An empirical study in India. *Software Quality Journal*, 14(4), 291-308.
- Jain, H., & Zhao, H. (2003). *A Conceptual Model for Comparative Analysis of Standardization of Vertical Industry Languages*. Paper presented at the Workshop on Standard Making: A Critical Research Frontier for Information Systems, Seattle, WA.
- Jakobs, K. (2009a). ICT Standardisation in China, the EU, and the US. In *Innovations for Digital Inclusions. ITU-T Kaleidoscope*. Mar del Plata: IEEE.
- Jakobs, K. (2009b). Perceived Relation between ICT Standards' Sources and their Success in the Market. In K. Jakobs (Ed.), *Information Communication Technology Standardization for E-Business Sectors; Integrating Supply and Demand Factors* (pp. 65-80). Hershey: Information Science Reference.
- Jakobs, K., & Kritznier, J. (2009). How to Select the Best Platform for ICT Standards Development. In K. Jakobs (Ed.), *Information Communication Technology Standardization for E-Business Sectors; Integrating Supply and Demand Factors* (pp. 81-96). Hershey: Information Science Reference.
- Jansen, S., & Brinkkemper, S. (2009). Applied Multi-Case Research in a Mixed-Method Research Project: Customer Configuration Updating Improvement. In A. Cater-Steel & L. Al-Hakim (Eds.), *Information Systems Research Methods, Epistemology and Applications* (pp. 120-139). Hershey: IGI Global.
- JCGM/WG2. (2008). *International vocabulary of metrology - Basic and general concepts and associated terms (VIM)*.
- Joachim, N., Beimborn, D., & Weitzel, T. (2011). *An Instrument for Measuring SOA Maturity*. Paper presented at the International Conference on Information Systems (ICIS), Shanghai.
- Johnston, H. R., & Vitale, M. A. (1988). Creating Competitive Advantage With Interorganizational Information Systems. *MIS Quarterly*, 12(2), 153-165.
- Juran, J. M., & Gryna, F. M. (Eds.). (1988). *Juran's quality control handbook* (4th ed.). New York: McGraw-Hill.
- Kabak, Y., & Dogac, A. (2010). A survey and analysis of electronic business document standards. *ACM Computing Surveys (CSUR)*, 42(3), 1-31.
- Kaefer, F., & Bendoly, E. (2000). Adoption of Electronic Data Interchange: A model and practical tool for managers. *Decision Support Systems*, 30(1), 23-32.
- 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.
- Kasunic, M., & Anderson, W. (2004). *Measuring Systems Interoperability: Challenges and Opportunities*: Carnegie Mellon University.
- Katerattanakul, P., & Siau, K. (1999). Measuring information quality of web sites: development of an instrument. In *Proceedings of the 20th international conference on Information Systems*. Atlanta, GA: Association for Information Systems.

- Katz, M. L., & Shapiro, C. (1985). Network Externalities, Competition, and Compatibility. *American Economic Review*, 75(3), 424.
- Kauffman, R. J., & Mohtadi, H. (2004). Proprietary and open systems adoption in E-procurement: A risk-augmented transaction cost perspective. *Journal of Management Information Systems*, 21(1), 137-166.
- Kellehear, A. (1993). *The Unobtrusive Researcher - A guide to methods*. Sydney: Allen & Unwin.
- Kelly, D., Feller, J., & Finnegan, P. (2006). Complex Network-Based Information Systems (CNIS) Standards: Toward an Adoption Model. In *The Transfer and Diffusion of Information Technology for Organizational Resilience* (pp. 3-20).
- King, W. R., & Epstein, B. J. (1982). Assessing Information System Value: An Experimental Study. *Decision Sciences*, 13(4), 34-45.
- Klein, H. K., & Myers, M. D. (1999). A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly: Management Information Systems* 23(1), 67.
- Kleppe, A., G., Warmer, J., & Bast, W. (2003). *MDA Explained: The Model Driven Architecture: Practice and Promise*. Boston: Addison-Wesley Longman.
- Knight, S. A., & Burn, J. (2005). Developing a framework for assessing information quality on the World Wide Web. *Informing Science*, 8, 159-172.
- Knothe, T., & Jochem, R. (2007). Quality Criteria for Enterprise Modelling in the Context of Networked Enterprises. In G. Doumeingts, J. Müller, G. Morel & B. Vallespir (Eds.), *Enterprise Interoperability* (pp. 149-158). London: Springer.
- Kosanke, K. (2006). ISO Standards for Interoperability: a Comparison. In D. Konstantas, J.-P. Bourrières, M. Léonard & N. Boudjlida (Eds.), *Interoperability of Enterprise Software and Applications* (pp. 55-64). London: Springer-Verlag.
- Kotonya, G., & Sommerville, I. (1998). *Requirements Engineering: Processes and Techniques*. Chichester: Wiley.
- Krechmer, K. (2006). Open Standards Requirements. In K. Jakobs (Ed.), *Advanced Topics in Information Technology Standards and Standardization Research* (Vol. 1, pp. 27-48). Hershey: Idea Group Publishing.
- Krechmer, K. (2008). Open Standards Requirements. In K. Jakobs (Ed.), *Standardization Research in Information Technology: New Perspectives* (pp. 49-65). Hershey: Information Science Reference.
- Krechmer, K. (2009). Open Standards: A Call for Change. *IEEE Communications Magazine*, 47(May), 88-94.
- Kreger, H. (2003). Fulfilling the web services promise. *Communications of the ACM*, 46(6), 29-34.
- Kroes, N. (2010). How to get more Interoperability in Europe: Openness at the heart of the EU Digital Agenda. Retrieved 06-17, 2010, from <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/10/300&format=HTML>
- Krukkert, D., & Punter, M. (2008). *Kwaliteitsraamwerk voor Standaarden*. Enschede: TNO.
- Kubicek, H., & Cimander, R. (2009). Three dimensions of organizational interoperability. *European Journal of ePractice*(6 January 2009).
- Kulvatunyou, B., Morris, K. C., Ivezic, N., & Frechette, S. (2008). Development Life Cycle for Semantically Coherent Data Exchange Specification. 16(4), 279-290.
- Lammers, R., Folmer, E., & Ehrenhard, M. (2010). Narrowing the Gap between Open Standards Policy and Practice: The Dutch E-Government Experience. In M. Janssen, W. Lamersdorf, J. Pries-Heje & M. Rosemann (Eds.), *EGES 2010 (part of WCC2010)* (pp. 7-18). Brisbane: Springer.
- Lampathaki, F., Mouzakitis, S., Gionis, G., Charalabidis, Y., & Askounis, D. (2009). Business to business interoperability: A current review of XML data integration standards. *Computer Standards and Interfaces*.
- Larrucea, X. (2008). MDSOA quality evaluation. In K.-D. Thoben, K. S. Pawar & R. Gonçalves (Eds.), *Proceeding of the 14th International Conference on Concurrent Enterprising: ICE2008* (pp. 977-982). Lisbon: Centre for Concurrent Enterprise.
- Lau, F. (1999). Toward a framework for action research in information systems studies. *Information Technology & People*, 12(2), 148-175.

- Lebreton, B., & Legner, C. (2007). Interoperability Impact Assessment Model: An Overview. In R. J. Gonçalves, J. P. Müller, K. Mertins & M. Zelm (Eds.), *Enterprise Interoperability II - New Challenges and Approaches* (pp. 725-728). London: Springer.
- Lee, A. S. (1989). A Scientific Methodology for MIS Case Studies. *MIS Quarterly*, 13(1), 33-50.
- Legner, C., & Lebreton, B. (2007). Preface to the Focus Theme Section: 'Business Interoperability' Business Interoperability Research: Present Achievements and Upcoming Challenges. *Electronic Markets*, 17(3), 176-186.
- Levina, N., & Vaast, E. (2005). The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems. *MIS Quarterly* 29(2), 335-363.
- Levitt, S. D., & List, J. A. (2006). What Do Laboratory Experiments Tell Us About the Real World? *Journal of Economic Perspectives*, 21(June), 153–174.
- Lew, P., Olsina, L., & Zhang, L. (2010). *Quality, Quality in Use, Actual Usability and User Experience as Key Drivers for Web Application Evaluation*. Paper presented at the Web Engineering (ICWE 2010).
- Libicki, M. C. (1995). Standards: The Rough Road to the Common Byte. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 35-78). Cambridge: The MIT Press.
- Lier, B., van. (2009). *Luhmann ontmoet "The Matrix"*. Erasmus University, Rotterdam.
- Löwer, U. M. (2005). *Interorganisational Standards - Managing Web Services Specifications for Flexible Supply Chains*. Heidelberg: Physica-Verlag - A Springer Company.
- Lu, X. H., Huang, L. H., & Heng, M. S. H. (2006). Critical success factors of inter-organizational information systems - A case study of Cisco and Xiao Tong in China. *Information and Management*, 43(3), 395-408.
- Lucking-Reiley, D., & Spulber, D. F. (2001). Business-to-Business Electronic Commerce. *The Journal of Economic Perspectives*, 15(1), 55-68.
- Lyytinen, K., & King, J. L. (2006). Standard Making: A Critical Research Frontier for Information Systems Research. *MIS Quarterly*, 30, 405-411.
- Maciaszek, L. A. (2001). *Requirements Analysis and System Design; Developing Information Systems with UML*. Harlow: Addison-Wesley.
- Madnick, S., Wang, R., Lee, Y., & Zhu, H. (2009). Overview and Framework for Data and Information Quality Research. *Journal of Data and Information Quality*, 1(1), 1-22.
- Mähönen, P. (2000). The Standardization Process in IT - Too Slow or Too Fast. In K. Jakobs (Ed.), *Information Technology Standards and Standardization: A Global Perspective* (pp. 35-47). Hershey: Idea Group Publishing.
- Maiden, N. A. M., & Rugg, G. (1996). ACRE: selecting methods for requirements acquisition. *Software Engineering Journal*, 11(3), 183-192.
- Markus, M. L. (2011). Historical Reflections on the Practice of Information Management and Implications for the Field of MIS. In R. D. Galliers & W. Currie (Eds.), *The Oxford Handbook of Management Information Systems. Critical Perspectives and New Directions* (pp. 3-15). Oxford: Oxford University Press.
- Markus, M. L., & Gelinas Jr, U. J. (2008). Comparing the Standards Lens with Other Perspectives on IS Innovations: The Case of CPFR. In K. Jakobs (Ed.), *Standardization Research in Information Technology: New Perspectives* (pp. 185-201). Hershey: Information Science Reference.
- Markus, M. L., Majchrzak, A., & Gasser, L. (2002). A Design Theory for Systems that Support Emergent Knowledge Processes. *MIS Quarterly*, 26(3), 179-212.
- Markus, M. L., Steinfield, C. W., Wigand, R. T., & Minton, G. (2006). Industry-wide Information Systems standardization as collective action: The case of U.S. residential mortgage industry. *MIS Quarterly*, 30(SI), 439-465.
- Maxwell, J. A. (2005). *Qualitative Research Design - An Interactive Approach* (2nd ed. Vol. 42). Thousand Oaks: Sage.
- McCall, J. A., Richards, P. K., & Walters, G. F. (1977). *Factors in Software Quality* (Vol. 1-3).

- McDonald, M. P. (2010). Leading in Times of Transition: the 2010 CIO Agenda. Retrieved 01-12-2010, from [http://blogs.gartner.com/mark\\_mcdonald/2010/01/19/leading-in-times-of-transition-the-2010-cio-agenda/](http://blogs.gartner.com/mark_mcdonald/2010/01/19/leading-in-times-of-transition-the-2010-cio-agenda/)
- McDowell, A., Schmidt, C., & Yue, K. B. (2004). Analysis and metrics of XML schema. In H. R. Arabnia & H. Reza (Eds.), *Proceedings of the International Conference on Software Engineering Research and Practice, SERP'04* (Vol. 2, pp. 538-544). Las Vegas, NV.
- McNichol, T. (2006). *AC/DC: The Savage Tale of the First Standards War*. San Francisco: Jossey-Bass.
- Mendoza, R., A. , & Ravichandran, T. (2007). *Organizational Assimilation of Vertical Standards: An Integrative Model*. Paper presented at the 40th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- Merriam-Webster. (2011). Instrument. Retrieved 04-07, 2011, from <http://www.merriam-webster.com/dictionary/instrument>
- Meuser, M., & Nagel, U. (1994). Expertenwissen und Experteninterview. In R. Hitzler, A. Honer & C. Maeder (Eds.), *Expertenwissen. Die institutionelle Kompetenz zur Konstruktion von Wirklichkeit* (pp. 180-192). Opladen: Westdeutscher Verlag.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis* (2nd ed.). Thousand Oaks: Sage.
- Milicic, D. (2005). Software Quality Models and Philosophies. In L. Lundberg, M. Matsson & C. Wohlin (Eds.), *Software quality attributes and trade-offs*. Karlskrona: Blekinge Institute of Technology.
- Mingers, J. (2001). Combining IS Research Methods: Towards a Pluralist Methodology. *Information systems research*, 12(3), 240.
- Mora, B., García, F., Ruiz, F., & Piattini, M. (2008). SMML: Software Measurement Modeling Language.
- Mora Rivas, B., Garcia Rubio, F., Ruiz González, F., & Piattini, M. (2008). SMML: Software Measurement Modeling Language: University of Castilla La Mancha.
- Moraga, C., Moraga, M. A., Calero, C., & Caro, A. (2009, 24-25 Aug. 2009). *SQuaRE-Aligned Data Quality Model for Web Portals*. Paper presented at the Quality Software, 2009. QSIC '09. 9th International Conference on.
- Morell, J. A., & Stewart, S. L. (1995). A Five-Segment Model for Standardization. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 198-219). Cambridge: The MIT Press.
- Mori, A. R., & Consorti, F. (1998). Integration of clinical information across patient records: a comparison of mechanisms used to enforce semantic coherence. *IEEE Transactions on Information Technology in Biomedicine*, 2(4), 243-253.
- Mutkoski, S. (2011). Defining Open Standards: A Comparison of Policy and Practice. In K. Blind & K. Jakobs (Eds.), *Proceedings of the 7th International Conference on Standardisation and Innovation in Information Technology (SIIT 2011)* (pp. 219-230). Berlin: IEEE.
- Mykkanen, J. A., & Tuomainen, M. P. (2008). An evaluation and selection framework for interoperability standards. *Information and Software Technology*, 50(3), 176-197.
- Mylonopoulos, N. A., & Theoharakis, V. (2001). Global perceptions of IS journals - Where is the best IS research published? *Communications of the ACM*, 44(9), 29-33.
- Naudet, Y., Latour, T., Guedria, W., & Chen, D. (2010). Towards a systemic formalisation of interoperability. *Computers in Industry*, 61(2), 176-185.
- Naumann, F., & Rolker, C. (2000). *Assessment Methods for Information Quality Criteria*. Paper presented at the International Conference on Information Quality, Cambridge, MA.
- Nelson, M., Shoonmaker, M., Shaw, M., Shen, S., & Wang, R. (2002). Modularized Interoperability in Supply-Chains: A Coadoption study of RosettaNet's XML-based Interorganizational Systems. In M. J. Shaw (Ed.), *E-Business Management - Integration of Web Technologies with Business Models* (pp. 417-438). Boston: Kluwer.
- Nelson, M. L., Shaw, M. J., & Qualls, W. (2005). Interorganizational System Standards Development in Vertical Industries. *Electronic Markets*, 15(4), 378-392.
- Nentwig, L., Adametz, H., Bittins, S., Gottschick, J., Reichling, K., & Meyer, S. (2008). *Quality Framework for Interoperability Assets*. Brussels: Semantic Interoperability Centre Europe.

- Nesbit, J., Belfer, K., & Leacock, T. (2011). Learning Object Review Instrument (LORI) User Manual version 1.5. Retrieved 05-24, 2011, from <http://www.elera.net/eLera/Home/Articles/LORI%201.5.pdf>
- Nickerson, J. V., & Zur Muehlen, M. (2006). The ecology of standards processes: Insights from internet standard making. *MIS Quarterly*, 30(SI), 467-488.
- NOiV. (2007). *The Netherlands Open in Connection - An action plan for the use of Open Standards and Open Source Software in the public and semi-public sector*. The Hague: Ministry of Economic Affairs.
- Nunamaker, J. F., Jr., Chen, M., & Purdin, T. D. M. (1990). Systems development in information systems research. *Journal of Management Information Systems*, 7(3), 89-106.
- Nurmilaakso, J. M., & Kotinurmi, P. (2004). A review of XML-based supply-chain integration. *Production Planning and Control*, 15(6), 608-621.
- Nurmilaakso, J. M., Kotinurmi, P., & Laesvuori, H. (2006). XML-based e-business frameworks and standardization. *Computer Standards and Interfaces*, 28(5), 585-599.
- Nyeck, S., Morales, M., Ladhari, R., & Pons, F. (2002). 10 years of service quality measurement: reviewing the use of the SERVQUAL instrument. *Cuadernos de Diffusion*, 7(13), 101-107.
- O'Brien, L., Bass, L., & Merson, P. (2005). *Quality Attributes and Service-Oriented Architectures*. Pittsburgh: Carnegie Mellon University, Software Engineering Institute.
- OGC. (2009). *The Specification Model — A Standard for Modular specifications*. Wayland: Open Geospatial Consortium Inc.
- OMG. (2010). *Model Driven Message Interoperability (MDMI)* (No. formal/2010-03-01). Needham: Object Management Group.
- Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, 2(1), 1-28.
- Otto, B., Folmer, E., & Ebner, V. (2011). A Characteristics Framework for Semantic Information Systems Standards. *Information Systems and E-Business Management* (Online First - 24 november 2011).
- Oude Luttighuis, P., & Folmer, E. (2010). Equipping the Enterprise Interoperability Problem Solver. In Y. Charalabidis (Ed.), *Interoperability in Digital Public Services and Administration; Bridging E-Government and E-Business* (pp. 339-354). Hershey - New York: Information Science Reference.
- Owlia, M. S. (2010). A framework for quality dimensions of knowledge management systems. *Total Quality Management*, 21(11), 1215-1228.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 64(1), 12-40.
- Park, J., & Ram, S. (2004). Information systems interoperability: What lies beneath? *ACM Transactions on Information Systems*, 22(4), 595-632.
- Park, S. (2006). Standardization and Network Externalities. In K. Jakobs (Ed.), *Advanced Topics in Information Technology Standards and Standardization Research* (Vol. 1, pp. 251-281). Hershey: Idea Group Publishing.
- Pawlowski, J. M., & Kozlov, D. (2010). Analysis and Validation of Learning Technology Models, Standards and Specifications: The Reference Model Analysis Grid (RMAG). *International Journal of IT Standards and Standardization Research*, 8(2), 1-19.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45-77.
- Perera, C. (2007). Standardization and Innovation. In W. Hesser (Ed.), *Standardisation in Companies and Markets* (pp. 307-380). Hamburg: Helmut Schmidt University.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences a practical guide*. Malden: Blackwell.
- Pindar, A. (2010). *Standardization for a competitive and innovative Europe: a vision for 2020 - Report of the expert panel for the review of the European Standardization System* (No. EXP 384 final).
- Poels, G., Maes, A., Gailly, F., & Paemeleire, R. (2005). Measuring the Perceived Semantic Quality of Information Models. In J. Akoka, S. Liddle, I.-Y. Song, M. Bertolotto, I. Comyn-Wattiau, S. Si-Said Cherfi, W.-J. Heuvel, B. Thalheim, M. Kolp, P. Bresciani, J. Trujillo, C. Kop & H. Mayr (Eds.), *Perspectives in Conceptual Modeling* (Vol. 3770, pp. 376-385). Berlin-Heidelberg: Springer.

- Quint, V. (2010). Amaya Home Page. Retrieved 11-02, 2011, from <http://www.w3.org/Amaya/>
- Rada, R. (1993). Standards: the language for success. *Communications of the ACM*, 36(12), 17-23.
- Rada, R. (1995). Consensus versus speed. *Communications of the ACM*, 38(10), 21-23.
- Rada, R. (1998). Corporate shortcut to standardization. *Communications of the ACM*, 41(1), 11-15.
- Rada, R. (2000). Consensus versus Speed. In K. Jakobs (Ed.), *Information Technology Standards and Standardization: A Global Perspective* (pp. 19-34). Hershey: Idea Group Publishing.
- Rada, R., & Berg, J. (1995). Standards: free or sold? *Communications of the ACM*, 38(2), 23-27.
- Rada, R., Cargill, C., & Klensin, J. (1998). Consensus and the Web. *Communications of the ACM*, 41(7), 17-22.
- Rada, R., & Ketchell, J. (2000). Sharing standards: standardizing the European information society. *Communications of the ACM*, 43(3), 21-25.
- Rayson, P., Emmet, L., Garside, R., & Sawyer, P. (2001). The REVERE Project: Experiments with the Application of Probabilistic NLP to Systems Engineering. In M. Bouzeghoub, Z. Kedad & E. Métails (Eds.), *Natural Language Processing and Information Systems* (Vol. 1959, pp. 288-300). Berlin-Heidelberg: Springer.
- Rebstock, M. (2009). Semantic ambiguity: Babylon, Rosetta or beyond? *Communications of the ACM*, 52(5), 145-146.
- Reimers, K., Johnston, R., & Klein, S. (2010). The difficulty of studying inter-organisational IS phenomena on large scales: critical reflections on a research journey. *Electronic Markets*, 20(3), 229-240.
- Reimers, K., Johnston, R. B., Guo, X., Klein, S., Xie, B., & Li, M. (2011). Novice-based Data Collection Methods for the Study of IOIS: Practice Probes and Learning Communities. In *Proceedings of the 24th Bled eCommerce Conference "eFuture"*.
- Reinstaller, A. (2008). Economic Analysis of Technological Standards and Standard Setting. In A. Hommels, J. Schueler & A. Fickers (Eds.), *Bargaining Norms Arguing Standards* (pp. 70-85). The Hague: STT.
- Robertson, S., & Robertson, J. (1999). *Mastering the Requirements Process*. Harlow: Addison-Wesley.
- Robey, D., Im, G., & Wareham, J. D. (2008). Theoretical Foundations of Empirical Research on Interorganizational Systems: Assessing Past Contributions and Guiding Future Directions. *Journal of the Association for Information Systems*, 9(9), 497-518.
- Rodriguez, N., & Casanovas, J. (2010). *A structural model of information system quality; an empirical research*. Paper presented at the Americas Conference on Information Systems (AMCIS), Lima, Peru.
- Roe, B. E., & Just, D. R. (2009). Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data. *American Journal of Agricultural Economics*, 91(5), 1266-1271.
- Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: Free Press.
- Rukanova, B. (2005). *Business transactions and standards: towards a system of concepts and a method for early problem identification in standard implementation projects*. University of Twente, Enschede.
- Rukanova, B., Wigand, R. T., & Tan, Y. H. (2009). From national to supranational government inter-organizational systems: An extended typology. *Lecture Notes in Computer Science (LNCS)*, 5693, 317-327.
- Rukanova, B. D., Slooten, K., van, & Stegwee, R. A. (2006). Business Process Requirements, Modeling Technique and Standard: how to Identify Interoperability Gaps on a Process Level. In D. Konstantas, J.-P. Bourrières, M. Léonard & N. Boudjlida (Eds.), *Interoperability of Enterprise Software and Applications* (pp. 13-23). London: Springer-Verlag.
- Rumsey, S. (2008). *How to find informationa guide for researchers*. Maidenhead: Open University Press.
- Sahay, R., Akhtar, W., & Fox, R. (2008). PPEPR: Plug and play Electronic Patient Records. In *Proceedings of the ACM Symposium on Applied Computing* (pp. 2298-2304). Fortaleza, Ceara: ACM.
- Samuelson, P. (2006). Copyrighting standards. *Communications of the ACM*, 49(6), 27-31.
- SAP. (2009). Industry-Specific Business Maps. Retrieved 11-09, 2009, from <http://www.sap.com/solutions/businessmaps>
- Sawyer, P., Rayson, P., & Garside, R. (2002). REVERE: Support for Requirements Synthesis from Documents. *Information Systems Frontiers*, 4(3), 343-353.

- Schuldt, R. (2011). *UDEF - Six Steps to Cost Effective Data Integration*. Charleston, USA.
- Schwartz, R. B., & Russo, M. C. (2004). How to quickly find articles in the top IS journals. *Communications of the ACM*, 47(2), 98-101.
- Schwind, M., Stockheim, T., & Weiss, K. (2008). A Diffusion Model for Communication Standards in Supply Networks. In K. Jakobs (Ed.), *Standardization Research in Information Technology: New Perspectives* (pp. 105-121). Hershey: Information Science Reference.
- Sedera, D., & Gable, G. (2004). *A Factor and Structural Equation Analysis of the Enterprise Systems Success Measurement Model*. Paper presented at the International Conference on Information Systems (ICIS), Charlottesville.
- Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. *MIS Quarterly*, 35(1), 37-56.
- Shanks, G., & Corbitt, B. (1999). Understanding data quality: Social and cultural aspects. In *Proceedings of the 10th Australasian Conference on Information Systems* (pp. 785-797).
- Sherif, M. H. (2006). Standards for Telecommunication Services. In K. Jakobs (Ed.), *Advanced Topics in Information Technology Standards and Standardization Research* (Vol. 1, pp. 183-205). Hershey: Idea Group Publishing.
- Sherif, M. H., Jakobs, K., & Egyedi, T. M. (2007). Standards of quality and quality of standards for Telecommunications and Information Technologies. In M. Hörlesberger, M. El-nawawi & T. Khalil (Eds.), *Challenges in the Management of New Technologies* (pp. 427-447). Singapore: World Scientific Publishing Company.
- Silverman, D. (2006). *Interpreting Qualitative Data* (3rd ed.). London: Sage.
- Silverman, D. (2010). *Doing Qualitative Research* (3rd ed.). London: Sage.
- Simcoe, T. (2007). Delay and de jure standardization: exploring the slowdown in Internet standards development. In S. Greenstein & V. Stango (Eds.), *Standards and Public Policy* (pp. 260-295). Cambridge: Cambridge University Press.
- Simons, C. A. J., & Vries, H. J., de (2002). *Standaard of Maatwerk, Bedrijfskeuzes tussen uniformiteit en verscheidenheid*. Schoonhoven: Academic Services.
- Söderström, E. (2004). Formulating a general standards life cycle. *Lecture Notes in Computer Science (LNCS)*, 3084, 263-275.
- Software Process Engineering Meta-Model, version 2.0. (2008). Object Management Group.
- Sommerville, I., & Sawyer, P. (1997). *Requirements Engineering A Good Practice Guide*. Chichester: John Wiley & Sons
- Song, M., Jiang, H., & Wu, H. (2007). International Standardization. In W. Hesser, A. Feilzer & H. J. De Vries (Eds.), *Standardisation in Companies and Markets* (pp. 347-380). Hamburg: Helmut Schmidt University.
- Spivak, S. M., & Brenner, F. C. (2001). *Standardization essentials: principles and practice*. New York: Dekker.
- Stango, V. (2004). The Economics of Standards Wars. *Review of Network Economics*, 3(1), 1-19.
- Staron, M., Meding, W., & Nilsson, C. (2009). A framework for developing measurement systems and its industrial evaluation. *Information and Software Technology*, 51(4), 721-737.
- Steen, H. U. (2011). Limits to the regulatory state in the rule-making of digital convergence: a case study of mobile TV standards governance in the European Union and China. *Technology Analysis & Strategic Management*, 23(7), 759-772.
- Stegwee, R. A., & Rukanova, B. D. (2003). Identification of Different Types of Standards for Domain-Specific Interoperability. In *Proceedings of MIS Quarterly Special Issue Workshop on Standard Making: A Critical Research Frontier for Information Systems* (pp. 161-170).
- Steinfeld, C. W., Markus, M. L., & Wigand, R. T. (2011a). *Cooperative Advantage and Vertical Information System Standards: An Automotive Supply Chain Case Study*. Paper presented at the 44th Hawaii International Conference on System Sciences (HICSS), Hawaii.
- Steinfeld, C. W., Markus, M. L., & Wigand, R. T. (2011b). Through a Glass Clearly: Standards, Architecture, and Process Transparency in Global Supply Chains. *Journal of Management Information Systems*, 28(2), 75-108.

- Steinfeld, C. W., Wigand, R. T., Markus, M. L., & Minton, G. (2007). Promoting e-business through vertical IS standards: lessons from the US home mortgage industry. In S. Greenstein & V. Stango (Eds.), *Standards and Public Policy* (pp. 160-207). Cambridge: Cambridge University Press.
- Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103(2684), 677-680.
- Stewart, J., Shen, X., Wang, C., & Graham, I. (2011). From 3G to 4G: standards and the development of mobile broadband in China. *Technology Analysis & Strategic Management*, 23(7), 773-788.
- Stvilia, B., Gasser, L., Twidale, M. B., & Smith, L. C. (2007). A framework for information quality assessment. *Journal of the American Society for Information Science and Technology*, 58(12), 1720-1733.
- Swann, P. (2000). *The Economics of Standardization*. Manchester: Manchester Business School.
- Swann, P. (2010). *The Economics of Standardization: An update*. Nottingham: Innovative Economics Limited
- Swatman, P. M., Swatman, P. A., & Fowler, D. C. (1994). A model of EDI integration and strategic business reengineering. *Journal of Strategic Information Systems*, 3(1), 41-60.
- Tanenbaum, A. S. (1989). *Computer networks*. Englewood Cliffs, NJ: Prentice-Hall.
- Teichmann, H. (2010). *Effectiveness of Technical Experts in International Standardization*. Paper presented at the 15th EURAS Annual Standardisation Conference "Service Standardization", Lausanne.
- Teichmann, H., Vries, H. J., de, & Feilzer, A. (2008). Linguistic Qualities of International Standards. In K. Jakobs (Ed.), *Standardization Research in Information Technology: New Perspectives* (pp. 86-103). Hershey: Information Science Reference.
- Thissen, W. A. H., & Stam, W. J. (1992). Electronic data interchange in an industrial sector: The case of The Netherlands' building industry. *Information and Management*, 23(1), 15-30.
- Thomas, J. W., Proberts, S., Dawson, R., & King, T. (2008). A Case Study of the Adoption and Implementation of STEP. In T. M. Egyedi & K. Blind (Eds.), *The Dynamics of Standards* (pp. 117-134). Cheltenham: Edward Elgar.
- Thomas, J. W., Proberts, S., Dawson, R., & King, T. (2010). A Step Towards the Adoption of Standards Within the UK Ministry of Defence. In K. Jakobs (Ed.), *New Applications in IT Standards - Development and Progress* (pp. 54-69). Hershey - New York: Information Science Reference.
- TOG. (2011). About the UDEF. Retrieved 12-10, 2011, from <http://www.opengroup.org/udedef/>
- Tolk, A. (2003). *Beyond Technical Interoperability – Introducing a Reference Model for Measures of Merit for Coalition Interoperability*. Paper presented at the International Command and Control Research and Technology Symposium, Washington.
- Ullberg, J., Chen, D., & Johnson, P. (2009). Barriers to Enterprise Interoperability. In R. Poler, M. Sinderen & R. Sanchis (Eds.), *Enterprise Interoperability* (Vol. 38, pp. 13-24). Berlin Heidelberg: Springer.
- Updegrove, A. (1995). Consortia and the Role of the Government in Standard Setting. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 321-350). Cambridge: MIT Press.
- Updegrove, A. (2007). The Essential Guide to Standards. Retrieved 01-05, 2011, from <http://www.consortiuminfo.org/essentialguide/forming1.php>
- US. (2006). *Executive Order: Promoting Quality and Efficient Health Care in Federal Government Administered or Sponsored Health Care Programs* (No. 13410). Washington: The White House.
- US. (2010). *America Competes Reauthorization Act* (No. H.R.5116). Washington: US Government.
- Valayer, C. (2008). *Project Report - "CAMSS" Common Assessment Method for Standards and Specifications*. Brussels: Strateqo Consortium.
- van Aken, J. E. (2004). Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. *Journal of Management Studies*, 41(2), 219-246.
- Van Blommestein, F. (2007). Decentralized Metadata Development for Open B2B Electronic Business. In G. Doumeingts, J. Müller, G. Morel & B. Vallespir (Eds.), *Enterprise Interoperability* (pp. 355-364). London: Springer.
- Van der Linde, H., Hofstad, C. J., Geurts, A. C. H., Postema, K., Geertzen, J. H. B., & van Limbeek, J. (2004). A systematic literature review of the effect of different prosthetic components on human functioning with a lower-limb prosthesis. *Journal of Rehabilitation Research and Development*, 41(4), 555-570.

- Van der Veer, H., & Wiles, A. (2006). *Achieving Technical Interoperability - the ETSI approach*. Sophia Antipolis Cedex: ETSI.
- Van Hillegersberg, J., & Minnece, P. H. (2009). *Expert advies SETU versie 1*. Retrieved 11-11, 2009. from [http://www.gbo.overheid.nl/fileadmin/os/documenten/OS\\_Expertadvies\\_SETU\\_v1.0.pdf](http://www.gbo.overheid.nl/fileadmin/os/documenten/OS_Expertadvies_SETU_v1.0.pdf).
- Van Lamsweerde, A. (2004). *Goal-Oriented Requirements: A Roundtrip from Research to Practice*. Paper presented at the 12th International Requirements Engineering Conference, Kyoto.
- Van Wegberg, M. (1999). *The Design of Standardisation Processes in ICT: An evolutionary transaction cost approach*. Maastricht: University of Maastricht.
- Van Wegberg, M. (2006). Standardization and Competing Consortia: The Trade-Off between Speed and Compatibility. In K. Jakobs (Ed.), *Advanced Topics in Information Technology Standards and Standardization Research* (Vol. 1, pp. 111-127). Hershey: Idea Group Publishing.
- Van Wessel, R. M. (2008). *Realizing business benefits from company IT standardization; case study research into the organizational value of IT standards, towards a company IT standardization management framework*. CentER, Tilburg University, Tilburg.
- Van Zeist, B., Hendriks, P., Paulussen, R., & Trienekens, J. (1996). *Kwaliteit van softwareproducten; praktijkervaringen met een kwaliteitsmodel*. Amsterdam: Kluwer Bedrijfsinformatie.
- Van Zeist, R. H. J., & Hendriks, P. R. H. (1996). Specifying software quality with the extended ISO model. *Software Quality Journal*, 5(4), 273-284.
- Venkatraman, S., Bala, H., Venkatesh, V., & Bates, J. (2008). Six strategies for electronic medical records systems. *Communications of the ACM*, 51(11), 140-144.
- Von Westarp, F., Weitzel, T., Buxmann, P., & Köning, W. (2000). The Standardization Problem in Networks - A General Framework. In K. Jakobs (Ed.), *Information Technology Standards and Standardization: A Global Perspective* (pp. 168-185). Hershey: Idea Group Publishing.
- Wagner, C. S., Cargill, C. F., & Slomovic, A. (1995). Open Systems Standards in Manufacturing: Implications for the National Information Infrastructure. In B. Kahin & J. Abbate (Eds.), *Standards Policy for Information Infrastructure* (pp. 178-197). Cambridge: The MIT Press.
- Walsham, G. (1993). *Interpreting Information Systems in Organizations*. Chichester, UK: Wiley.
- Wand, Y., & Wang, R. Y. (1996). Anchoring Data Quality Dimensions in Ontological Foundations. *Communications of the ACM*, 39(11), 86-95.
- Wang, E. T. G., & Seidmann, A. (1995). Electronic data interchange: competitive externalities and strategic implementation policies. *Management Science*, 41(3), 401-418.
- Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5-33.
- Wapakabulo Thomas, J. (2010). *Data-Exchange Standards and International Organizations - Adoption and Diffusion*. Hershey: Information Science Reference.
- Wareham, J. (2005). Critical themes in electronic commerce research a meta-analysis. *Journal of Information Technology (JIT)*, 20(1), 1.
- Warmer, J., & Kleppe, A. (1998). *The Object Constraint Language: Precise Modeling With UML*. Boston: Addison-Wesley Professional.
- Washizaki, H., Hiraguchi, H., & Fukazawa, Y. (2008). A metrics suite for measuring quality characteristics of javabeans components. *Lecture Notes in Computer Science (LNCS)*, 5089, 45-60.
- Weber, S. (2010). *Design Science Research: Paradigm or Approach?* Paper presented at the Americas Conference on Information Systems (AMCIS), Lima, Peru.
- Weitzel, T., Beiborn, D., & König, W. (2006). A unified economic model of standard diffusion: The impact of standardization cost, network effects, and network topology. *MIS Quarterly*, 30(S1), 489-514.
- Weitzel, T., Wendt, O., Beiborn, D., & König, W. (2006). Network Effects and Diffusion Theory: Extending Economic Network Analysis. In K. Jakobs (Ed.), *Advanced Topics in Information Technology Standards and Standardization Research* (Vol. 1, pp. 282-305). Hershey: Idea Group Publishing.
- Werle, R., & Iversen, E. J. (2006). Promoting Legitimacy in Technical Standardization. *Science, Technology & Innovation Studies*, 2(March), 19-39.

- West, J. (2007). The economic realities of open standards: black, white, and many shades of gray. In S. Greenstein & V. Stango (Eds.), *Standards and Public Policy* (pp. 87-122). Cambridge: Cambridge University Press.
- Wieringa, R. J. (1996). *Requirements Engineering: Frameworks for Understanding*. Amsterdam: Wiley.
- Wigand, R. T., Markus, M. L., & Steinfield, C. W. (2005). Preface to the Focus Theme Section: Vertical Industry Information Technology Standards and Standardization. *Electronic Markets*, 15(4), 285 - 288.
- Wigand, R. T., Steinfield, C. W., & Markus, M. L. (2005). Information technology standards choices and industry structure outcomes: The case of the U.S. home mortgage industry. *Journal of Management Information Systems*, 22(2), 165-191.
- Wikipedia. (2011a). Instrumentation Engineering. Retrieved 04-07, 2011, from [http://simple.wikipedia.org/wiki/Instrumentation\\_engineering](http://simple.wikipedia.org/wiki/Instrumentation_engineering)
- Wikipedia. (2011b). Triage. Retrieved 12-10, 2011, from <http://en.wikipedia.org/wiki/Triage>
- Wybo, M. D., & Goodhue, D. L. (1995). Using interdependence as a predictor of data standards. Theoretical and measurement issues. *Information and Management*, 29(6), 317-329.
- Yin, R. K. (2009). *Case study research design and methods* (4th ed.). London: Sage.
- Zachman, J. A. (1997). Enterprise Architecture: The Issue of the Century. *Database Programming & Design*, 10(3), 44-53.
- Zhao, K., Xia, M., & Shaw, M. J. (2005). Vertical e-business standards and standards developing organizations: A conceptual framework. *Electronic Markets*, 15(4), 289-300.
- Zhao, K., Xia, M., & Shaw, M. J. (2007). An integrated model of consortium-based e-business standardization: Collaborative development and adoption with network externalities. *Journal of Management Information Systems*, 23(4), 247-271.
- Zhao, K., Xia, M., & Shaw, M. J. (2011). What motivates firms to contribute to consortium-based E-business standardization? *Journal of Management Information Systems*, 28(2), 305-334.
- Zhu, H., & Fu, L. (2009). *Towards Quality of Data Standards: Empirical Findings from XBRL*. Paper presented at the International Conference on Information Systems (ICIS), Phoenix.
- Zhu, H., & Wu, H. (2010). *Quality of XBRL US GAAP Taxonomy: Empirical Evaluation using SEC Filings*. Paper presented at the Americas Conference on Information Systems (AMCIS), Lima, Peru.
- Zhu, H., & Wu, H. (2011). Quality of data standards: framework and illustration using XBRL taxonomy and instances. *Electronic Markets*, 21(2), 129-139.
- Zhu, K., Kraemer, K. L., Gurbaxani, V., & Xu, S. X. (2006). Migration to open-standard interorganizational systems: Network effects, switching costs, and path dependency. *MIS Quarterly*, 30(SI), 515-539.
- Zhu, X., & Gauch, S. (2000). *Incorporating quality metrics in centralized/distributed information retrieval on the World Wide Web*. Paper presented at the Research and Development in Information Retrieval.
- Zur Muehlen, M., Nickerson, J. V., & Swenson, K. D. (2005). Developing web services choreography standards - The case of REST vs. SOAP. *Decision Support Systems*, 40(SI 1), 9-29.

# Glossary of Terms

**Semantic Information System (IS) standards** (in short **semantic standards**): Are designed to promote communication and coordination among the organizations; these standards may address product identification, data definitions, business document layout, and/or business process sequences (Steinfeld et al., 2007).

**Inter-organizational interoperability** (in short: **interoperability**): The ability of two or more organizational systems to exchange information, to interpret the information that has been exchanged and to act upon it in an appropriate and agreed upon manner" (adapted from (Rukanova, 2005)).

**Inter-Organizational information System (IOS)**: Is defined as an automated information system (IS) shared by two or more companies (Cash Jr & Konsynski, 1985) to facilitate the creation, storage, transformation and transmission of information (Johnston & Vitale, 1988). IOS is a broad term including concepts like data integration, but it differs from normal internal distributed systems by its ability to exchange information with the outside world (Johnston & Vitale, 1988).

**Standard Setting Organization (SSO)**: Includes every organization that is involved in developing and maintaining standards, including formal bodies, industry consortia and anything else that can be present in practice.

**Quality of semantic standard**: The ability to achieve its intended purpose – semantic interoperability – effectively and efficiently.

**iQMSS (instrument for Quality Model of Semantic Standards – in short: Quality Instrument)**: A measuring device for determining the quality value of a semantic standard based on applying a quality model.

**QMSS (Quality Model of Semantic Standards)**: The set of measurable concepts and the relationships between them, which provide the basis for specifying quality requirements and evaluating the quality of the entities of semantic standards (Adapted from (García et al., 2009)).

**SSM (Semantic Standard Model)**: A conceptual model of semantic standards that can be used for identification of the attributes of a semantic standard.

**Information need**: Insight necessary to manage objectives, goals, risks, and problems (García et al., 2009).

**Measurable concept**: Abstract relationship between attributes of semantic standards and information needs (García et al., 2009).

**(Quality) attribute**: A measurable physical or abstract property of a semantic standard (Adapted from (García et al., 2009)).

**Measure**: The measurement approach defined and the measurement scale (a measurement approach is either a measurement method, a measurement function or an analysis model) (García et al., 2009).

**Measurement**: A set of operations whose objective is to determine the value of a measurement result, for a given attribute of a semantic standard, using a measurement approach (García et al., 2009).



# Abstract

Little scientific literature addresses the issue of quality of semantic standards, albeit a problem with high economic and social impact. Our problem survey, including 34 semantic Standard Setting Organizations (SSOs), gives evidence that quality of standards can be improved, but for improvement a quality measurement instrument is needed. 81% of the survey respondents is interested in using such instrument. It can be expected that improved quality of semantic standards will lead to improved interoperability and improved economic welfare and social life.

Our main research question is: *What are the characteristics of an instrument to measure quality of semantic standards that will aid standard developers in improving their standards?*

Based on design science methodology this research developed such an instrument, called iQMSS (instrument for Quality Model of Semantic Standards). The iQMSS involves three types of artifacts: artifacts related to the Quality Model of Semantic Standards (QMSS), the Semantic Standard Model (SSM), and the instrumentalization. On the highest abstraction level SMO (Software Measurement Ontology) from the software engineering domain was selected and used for the creation of both QMSS and SSM with consistent concepts and definitions.

The extensive research process included coverage of both scientific studies and practical experiences, and led to the highly customizable final version of the iQMSS, in line with the gathered requirements.

The main artifact QMSS consists of in total 100 quality aspects structured within three hierarchical trees; product quality (intrinsic), process quality (the organization of the standard), and quality in practice (application of the standard).

The SSM is applied upfront to gather knowledge about the standard, create overview and understanding about what we are able to measure in practice for a specific standard. The main concepts within the model are the standards' context, its content, its development and maintenance processes, and its application. On the lowest level it contains 33 aspects of a semantic standard.

The glue between the artifacts is the usage model, which describes how the iQMSS should be applied including four roles. The end result after each application is the analysis report which provides the basis for a standards' improvement project.

But does the iQMSS aid standard developers in improving their standards? Our validation cases, in the temporary staffing and education domains, suggest that the iQMSS identifies improvement suggestions. Based on the validation workshops we held in both the United States and the Netherlands, we believe that the iQMSS will be used in practice as aid for standards developers.



# Nederlandse Samenvatting

Een leven zonder standaarden is nauwelijks voor te stellen. De “meter” en “kilogram” zijn standaarden die we dagelijks gebruiken, maar ook de euromunt, de euro95 benzine, A4 papier, etc. zijn onmisbaar. In de ICT wereld is het dat niet veel anders; sterker nog het Internet was er niet geweest zonder ICT standaarden zoals IP (Internet Protocol) en HTML (HyperText Markup Language). Een specifiek soort standaarden zijn semantische standaarden, deze houden zich bezig met de betekenis van informatie. Een type standaard die van groot belang is om interoperabiliteit, succesvolle samenwerking tussen organisaties te bereiken. Interoperabiliteit is noodzakelijk voor effectief en efficiënt samen te werken voor zowel economische als maatschappelijke doelstellingen. Onderzoek heeft al aangetoond dat de voordelen in verschillende industrie sectoren in de miljarden lopen en daarnaast door interoperabiliteit levens in de zorg bespaard kunnen worden.

Deze semantische standaarden zijn een bijzondere groep van standaarden, veelal ontwikkeld in het eigen domein en een relatief korte ontstaansgeschiedenis. Bekende problemen van deze standaarden zijn met name gerelateerd aan de langzame adoptie in de praktijk waaraan het nodige onderzoek is verricht. Een onderzoekshiaat is de kwaliteit van deze standaarden en hoe deze verbeterd kan worden zodat een standaard een effectievere en efficiëntere bijdrage levert aan interoperabiliteit. Dit onderzoek richt zich op de ontwikkeling van een kwaliteitsinstrument voor semantische standaarden. Doelstelling daarbij is dat de resultaten van de kwaliteitsmeting gebruikt kunnen worden om de kwaliteit van de standaard mee te verbeteren.

De praktijk heeft behoefte aan een dergelijk kwaliteitinstrument, zo bleek uit een enquête onder 34 semantische standaarden (hoofdstuk 3). De deelnemers geven aan dat de kwaliteit van hun standaarden verbeterbaar is en dat een kwaliteitsverbetering zal leiden tot verbetering in interoperabiliteit. Daarnaast geven ze in ruime mate aan geïnteresseerd in het gebruik van het kwaliteitinstrument.

Vooraf is er uitvoerig onderzoek verricht naar de aanwezige literatuur om zowel het onderzoekshiaat te kunnen bevestigen als wel een fundament van studies te zoeken die gebruikt kunnen worden bij de ontwikkeling van het kwaliteitinstrument (hoofdstuk 4 en 5). Daarnaast zijn er workshops georganiseerd die een set van wensen en eisen hebben opgeleverd voor het kwaliteitinstrument (hoofdstuk 7). In meerdere iteraties, gebruik makend van een rijke set van onderzoeksmethodieken, is het kwaliteitinstrument (iQMSS) ontwikkeld. De belangrijkste onderdelen zijn een model van een semantische standaard (SSM) en het kwaliteitmodel (QMSS). Daarnaast zijn praktische hulpmiddelen gewenst om het instrumentele karakter te benadrukken en het gebruik in de praktijk eenvoudig te maken. Dit onderzoek laat zien hoe deze hulpmiddelen er uiteindelijk uit kunnen zien (hoofdstuk 8). Tot slot is er ook een gebruiksmodel ontwikkeld als leidraad voor het uitvoeren van een kwaliteitsmeting.

Het model van een semantische standaard (hoofdstuk 9) is een praktisch hulpmiddel om inzicht te krijgen in de complexe omgeving van een semantische standaard. Op hoofdniveau bestaat een semantische standaard uit het organisatorische probleemdomein, een oplossing op verschillende niveaus, een ontwikkel en beheer proces, en aspecten gerelateerd aan het gebruik in de praktijk. Deze vier onderwerpen zijn verder uitgewerkt in het SSM. Het SSM kan ook gebruikt worden voor andere doeleinden dan een kwaliteitsmeting.

De kern van de iQMSS is het kwaliteitsmodel (hoofdstuk 11) waarin een onderscheid wordt gemaakt tussen productkwaliteit (intrinsiek), proceskwaliteit (organisatie) en de kwaliteit-in-gebruik (praktijk), met een focus op productkwaliteit. Voor elk is een hiërarchische kwaliteitsboom ontwikkeld waarin vele kwaliteitsaspecten zoals bruikbaarheid, duurzaamheid, aanpasbaarheid, onderhoudbaarheid, etc. zijn opgenomen. Voor productkwaliteit is daarnaast een ruime set van “metingen” ontwikkeld die relatief abstracte kwaliteitsattributen meetbaar maken in de praktijk. Zo wordt de Gunning-Fog Index voorgesteld om de leesbaarheid te bepalen als onderdeel van de begrijpbaarheid en bruikbaarheid van een standaard. Of wordt de technische complexiteit meetbaar gemaakt door ondermeer te toetsen of bekende naamgevingregels (NDR) gebruikt zijn.

Het gebruiksmodel (hoofdstuk 8.1 en 11.4) beschrijft de rollen en benodigde kennis en ervaring die nodig is voor het uitvoeren van een kwaliteitsmeting met behulp van het iQMSS. Het iQMSS is generiek opgezet zodat het voor een brede set van semantische standaarden bruikbaar is, maar waardoor het ook eerst “specifiek” gemaakt moet worden voor een standaard, voordat de meting en analyse kan starten. Daarbij is betrokkenheid van een ontwikkelaar van de standaard essentieel evenals kennis van het iQMSS. Het gebruiksproces is erop gericht dat slechts een beperkt deel van het QMSS gebruikt gaat worden op basis van de informatie behoefte van klant. Zo blijft de inzet van de iQMSS laagdrempelig.

Tijdens en na de ontwikkeling van het iQMSS is ervaring in de praktijk opgedaan (hoofdstuk 10 en 12). Deze praktijkervaringen laten zien dat het iQMSS al snel toegevoegde waarde biedt. Niet alleen bieden de resultaten een helder en bruikbaar overzicht van mogelijke verbeterpunten, de meerwaarde zit hem vooral ook in de “kwaliteitsbril” die het iQMSS biedt. Dat wil zeggen dat de gestructureerde manier om naar kwaliteit te kijken op zichzelf al van toegevoegde waarde blijkt te zijn.

De iQMSS met in het bijzonder het SSM en het QMSS, is de bijdrage van dit onderzoek aan de wetenschap. Om de praktische relevantie van het eindresultaat te toetsen zijn twee sessies met potentiële gebruikers georganiseerd waarin het onderzoeksresultaat, de iQMSS, is gepresenteerd. Op basis van deze sessies kunnen we concluderen dat er vanuit verschillende semantische standaarden interesse is om het iQMSS te gaan gebruiken. Er zijn inmiddels stappen gezet om het iQMSS in te zetten voor meer kwaliteitsmetingen, maar ook om het vast onderdeel te laten zijn van kwaliteitsbewakingprocessen bij standaardisatie-organisaties.

In de toekomst zullen deze gebruikservaringen mogelijk tot verdere verbeteringen van de iQMSS kunnen leiden. Het iQMSS kan dan verder evolueren en mogelijk op termijn zelf een standaard worden.

# Dankwoord

Ik heb veel geleerd de afgelopen drie jaren. Niet alleen over het boeiende onderwerp semantische standaarden, maar ook over het uitvoeren van wetenschappelijk onderzoek, en de voor- en nadelen van een parttime promotietraject. Het was een boeiend traject, maar ook een zwaar traject. Dat had ik niet alleen kunnen uitvoeren, maar gelukkig hebben velen een steentje bijgedragen.

Allereerst ben ik veel dank verschuldigd aan Paul Oude Luttighuis. Paul heeft het promotievuurtje opnieuw aangewakkerd. Maar daarnaast heeft Paul door de jaren heen een grote bijdrage geleverd met de vele discussies en het vele reviewen. De gang van zaken tussen TNO en Novay hebben het niet altijd eenvoudig gemaakt, maar toch zijn we erin geslaagd om goed samen te werken.

De eerste kennismaking met Jos van Hillegersberg was op Madeira lang geleden. De hernieuwde kennismaking in Twente heeft geleid tot een succesvolle en prettige samenwerking, die we hopelijk ook in de toekomst blijven voortzetten. Jos, bedankt voor het bieden van de kansen!

De commissieleden Ramses Wessel, Robert Stegwee, Peter Apers, Jan van den Ende en Sjaak Brinkkemper: Dank voor het plaatsnemen in de commissie. Sjaak, bijzonder en fijn dat je nadat je 13 jaar geleden mijn afstudeerbegeleider bij Baan was, nu weer een belangrijke rol op je wilde nemen.

Many thanks to Boris Otto for being there as part of the committee and also for the long standing friendship. It is always a pleasure to meet up with Boris either at a bar in any southern European location or in St. Gallen to exchange ideas for a paper, discuss European Commissions framework programs, or simply to talk about football.

Binnen TNO heeft Gerlof Bosveld ervoor gezorgd dat mijn promotieonderzoek mogelijk werd en een vliegende start kreeg. Helaas heeft Gerlof het eindresultaat niet meer mogen meemaken. Met name in de laatste periode heb ik binnen TNO veel ondersteuning gehad van Gert Kruithof, Jean-Louis Roso en Erik Fledderus. Erik ook bedankt voor het plaatsnemen in de commissie.

Veel TNO collega's hebben een steentje bijgedragen. Vanuit mijn werkveld: Dennis, Jack, Jasper, Matthijs, Diederik en Paul Brandt: Dank voor de deelname aan de pizza-workshop. Laura, Michael en Jack voor het reviewen van stukken tekst. Menno voor het knutselwerk met de demonstrator en last-minute figuren maken vanuit Brazilië (het krat bier komt). Dennis en Roel voor het werken aan de requirements en aan het metamodel, en niet te vergeten Catelijne voor te veel om op te noemen. Ook niet te vergeten zijn de TNO-collega's die meegedaan hebben aan de requirements workshops, waarvoor dank! Anderen wist ik te chanteren met pizza's en hebben een avond geofferd om mee te denken: Michiel Stornebrink, Lex Heerink, Maria Iacob en Lucas Meertens: Dank hiervoor!

Op de UT dank ik mijn collega's voor de talloze (gang-/lunch-)discussies over van alles en nog wat. Daarnaast de studenten Joris van Soest, Wouter Berends, Roel Veneberg en Joris Bastiaans die ook aan onderdelen hebben meegewerkt.

Ook niet te vergeten zijn alle collega's van het programma Nederland Open in Verbinding, waarvoor ik jaren met veel plezier heb gewerkt. Bijzonder woord van dank aan Ineke Schop voor het financieel ondersteunen, en daarnaast het accepteren van mijn bijzondere rol binnen het NOiV-team.

Het leuke van parttime promotieonderzoek is dat je de link kan leggen met de normale werkzaamheden. Daardoor kan je veel praktijkervaring en toegang tot praktijksituaties meenemen in het onderzoek en je netwerk gebruiken, en dat heb ik zeer uitgebreid gedaan. Ik ben daarbij dank verschuldigd aan alle deelnemers van masterclasses, BOMOS werkgroep sessies, World standards day bijeenkomst, workshops, en niet te vergeten de deelnemers aan de surveys. Al jullie input en ervaringen heb ik meegenomen in mijn onderzoek.

Specifiek wil ik daarbij de volgende personen nog noemen: Erik Wijnen en Matthijs Punter voor de EL&I case. Verena Ebner en Boris Otto voor de Siemens case. Scott Wilson, Frans de Liagre Böhl, Dennis Krukkert en Paul Oude Luttighuis voor het XCRI experiment. Voor de SETU studie: Michael van Bekkum, Dennis Krukkert en Jasper Roes. Voor de IMS LRM studie: Jos van der Arend en Jasper Roes. And many thanks to Richard Soley for giving me the opportunity to present my results and to organize a workshop at the OMG Technical Meeting in Orlando.

Daarnaast ben ik dank verschuldigd aan de organisaties die mij SETU-implementatie materiaal hebben aangereikt: Randstad, Adecco, Manpower, USG, Flexservice, CTB en Netive. Dit heeft geleid tot zeer interessante resultaten die echter niet zijn opgenomen in dit proefschrift, maar ik beloof dat ik hier nog een publicatie aan ga wijden.

Ook heb ik de kans gehad om aan een viertal doctoral consortia deel te nemen (AMCIS, I-ESA, KIN (Amsterdam Business Research Institute) en UT Executive PhD course), en zo mijn ideeën te kunnen bediscussieren met een gelouterd wetenschappelijk panel. Hierdoor kwam ik al vroeg in aanraking met kritische vragen waardoor de kwaliteit van mijn werk ongetwijfeld is verbeterd. Panelleden dank hiervoor! Tot slot, discussies met het EURAS/SIIT werkveld heb ik zeer weten te waarderen en hoop ik in de toekomst verder te continueren.

Een parttime promotietraject is zwaar, vooral het gebrek aan vrije weekenden en avonden. Het heeft meer opgeleverd dan alleen wetenschappelijke verrijking. Letterlijk niet te missen zijn de 10 kilo extra gewicht (kennis is blijkbaar zwaar). Het laatste jaar was extra zwaar, helemaal als je het eerste concept hebt ingediend en van anderen afhankelijk wordt. Voor een control freak is dat zwaar; Jos, sorry voor het pushen.

De (schoon-)familie is de stabiele basis. Mijn ouders ben ik dankbaar dat ze me de vrijheid en mogelijkheid hebben gegeven om te studeren. Daarnaast hebben ze me geleerd wat hard werken is. Alleen ben ik daar misschien iets in doorgesloten. Mijn broer en zus hebben het pad geëffend. De logeerpartijtjes en spelletjes met Brigitte, Emiel, Tirza en Iris zorgden voor de nodige ontspanmomenten, evenals andere uitjes met vrienden. Maar natuurlijk ben ik de meeste dank verschuldigd aan Elvira. Ze heeft me vanaf dag 1 in dit traject gesteund en is dat blijven volhouden tot het einde. Daarbij nooit geklaagd over mijn vele avonden van huis, of dat ik vele avonden en weekenden thuis aan het werk was, en daardoor menig feestje/verjaardag alleen bezocht zodat ik door kon werken. Daarnaast heeft ze me ook de beginselen van statistiek en onderzoeksterminologie bijgebracht, en geholpen bij het reviewen van mijn werk, het layouten van figuren/tabellen, en zelfs ondersteuning aan studenten. Niks was teveel gevraagd! Het laatste jaar was zeker niet makkelijk voor ons, maar ook daar komen we weer sterker uit. Nu dit project is afgerond wordt het tijd voor ons andere project, en niet te vergeten meer tijd om te genieten!

Met een gepromoveerde zus en een gepromoveerde vrouw, was voor mij een promotie een must om nog iets in de melk te brokkelen te hebben. Erg leuk dat ik bij beide jullie paranimf mocht zijn, en dat ik nu jullie als paranimfen mag hebben.

Tot slot wil ik mijn grote vriend Pierre (en zijn collega's) bedanken, die me de betrekkelijkheid van dingen laat zien, en me de afgelopen jaren ook rust heeft gegeven.

Erwin Folmer

# About the Author

Erwin Folmer was born on July 10, 1975 in Zevenaar, the Netherlands. He received his Masters degree in Technical Business Administration from the University of Twente. In 1999 Erwin started working for KPN Research as an innovator, and was especially involved in designing order entry and billing systems for new KPN services such as ADSL. From 2001 Erwin joined TNO to lead the European OpenXchange project focusing on business to business integration. Since then he has been involved in standardization and interoperability research with a special interest in semantic standards.

He is involved in a number of standard domains such as temporary staffing industry, waste, invoices, transportation and education, amongst others. Erwin is the founder of the SETU Timecard standard, which is used nowadays by the whole Dutch staffing industry. He chaired many standardization workgroups, including HR-XML SIDES, SETU and STOSAG. Currently he chairs the Standards steering board for the Dutch Semantic Invoice Model. Erwin was involved in setting up the Integrate consortium that led to the development of several instruments for semantic standards, such as the adoption instrument.

As a standardization expert he was a member of the Netherlands Open in Connection (2009-2011) project team to improve the adoption of open standards (and open source software) in the Netherlands. As part of this role, Erwin authored the BOMOS (management and development of open standards) book that many standardization organizations use as a guidance when organizing the management of standards.

In 2009 he joined the University of Twente part-time to embark on a PhD research program on standardization, while continuing his work for TNO. As well as being involved in teaching activities, Erwin has also presented his work at many international scientific conferences and in 2010 participated in the doctoral consortium of the American Conference on Information Systems. Erwin's research has led to many publications in books, magazines, conference proceedings, and journals; Electronic Markets and Information Systems and E-Business Management, amongst others.

