Modelling DL Quality: a Comparison between Approaches: the DELOS Reference Model and the 5S Model

Maristella Agosti[1], Nicola Ferro[1], Edward A. Fox[2],
and Marcos A. Gonçalves[3]

[1]Department of Information Engineering, University of Padua
{agosti, ferro}@dei.unipd.it

[2]Department of Computer Science, Virginia Polytechnic Institute and State University
fox@vt.edu

[3]Department of Computer Science, Federal University of Minas Gerais
mgoncalv@dcc.ufmg.br

Abstract

Digital libraries have become a key part of the global information infrastructure, building upon over 15 years of R&D efforts. To support further advancement, and to aid education in the field, work on formal foundations has proceeded, leading to the DELOS Reference Model and the 5S Framework. This paper describes and compares these models, in particular with respect to the concept of quality, and shows how the models can interoperate. It presents mappings of similar concepts between the models. It also shows what (currently) is in one model but not the other. It is hoped that, based on this effort, each model will develop further – with better coverage of quality.

Categories and Subject Descriptors

H.3 [Information Storage and Retrieval]: H.3.7 Digital Libraries

General Terms

Design, Measurement, Theory

Keywords

DELOS Reference Model, 5S Framework, Quality

1 Introduction

This paper describes an ongoing cooperation and effort for defining a theory-based approach towards the assessment and evaluation of different quality aspects in Digital Libraries (DLs).

The conducted research stems from two formal models of DL, the DELOS Reference Model[1] (Agosti, et al., 2006) and the Streams, Structures, Spaces, Scenarios, Societies (5S) formal framework (Gonçalves, Fox, Watson, & Kipp, 2004), that both aim at laying a firm foundation for the DL field. The former was developed mostly top down, to cover existing and planned DL systems and operations, and to provide a reference resource that would guide future DL R&D, initially in the European Union. The latter, while drawing upon the broad DL literature in order to have a comprehensive base of support, was launched earlier and developed largely bottom up, starting with key definitions and with elucidation of the DL concept from a minimalist approach. This paper reflects an ongoing effort to develop a broad consensus regarding foundations for the DL field, to help with education (Pomerantz et al., 2006), research, development, and practice. This

[1] http://www.delos.info/ReferenceModel/

paper, in particular, focuses on a part of that work, namely the explication of the concept of quality with regard to DLs. It builds upon the initial findings reported in (Agosti, et al., 2007) and follows the proposed roadmap for developing a quality model for DLs.

More broadly, we aim not only to define a general quality model for DLs but also to compare the DELOS Reference Model and the 5S model in order to point out their strengths and weaknesses. That comparison allows us to perform an evaluation of the different approaches adopted in the two models and to assess the pros and cons of each strategy as well as its robustness and capability of properly modelling the domain of interest. Such a comparison also will help move the community toward consensus regarding DL foundations. While interoperability has long been a key aim of the DL community, and understanding it from a formal perspective has been a key goal (Shen, 2006), interoperability of models and frameworks is a particularly urgent concern. Thus, this paper aims to help not only the understanding of each of the two approaches, but also the mapping of terminology and the explication of scope of coverage of the two efforts. A key success will be when results from one approach can easily be adopted by those advancing the other approach.

The paper is organized as follows: Section 2 describes the main characteristics of the quality concept in the DELOS Reference Model. Section 3 discusses the approach to quality adopted in the 5S model. Section 4 compares the two approaches and presents commonalities and differences between them. Finally, Section 5 draws some conclusions and presents an outline for future work.

2 The Quality Main Concept in the DELOS Reference Model

Figure 1: The main concepts of the digital library universe.

The DELOS Reference Model approaches the problem of modelling the DL universe by highlighting six main concepts (Candela, et al., 2006), as shown in Figure 1:

- *content*: the data and information that digital libraries handle and make available to their users;
- *user*: the actors (whether human or not) entitled to interact with digital libraries;
- *functionality*: the services that digital libraries offer to their users;
- *quality*: the parameters that can be used to characterize and evaluate the content and behaviour

of digital libraries;

- *policy*: a set of rules that govern the interaction between users and digital libraries;
- *architecture*: a mapping of the functionality and content offered by a digital library onto hardware and software components.

These six main concepts represent the high level containers that help organize the DELOS reference model. For each of these concepts, the fundamental entities and their relationships are clearly defined and discussed (Agosti, et al., 2006). Note that these six main concepts are not separate, but, on the contrary, are strongly inter-related; the entities within a main concept are often related to or influenced by the entities in the other main concepts.

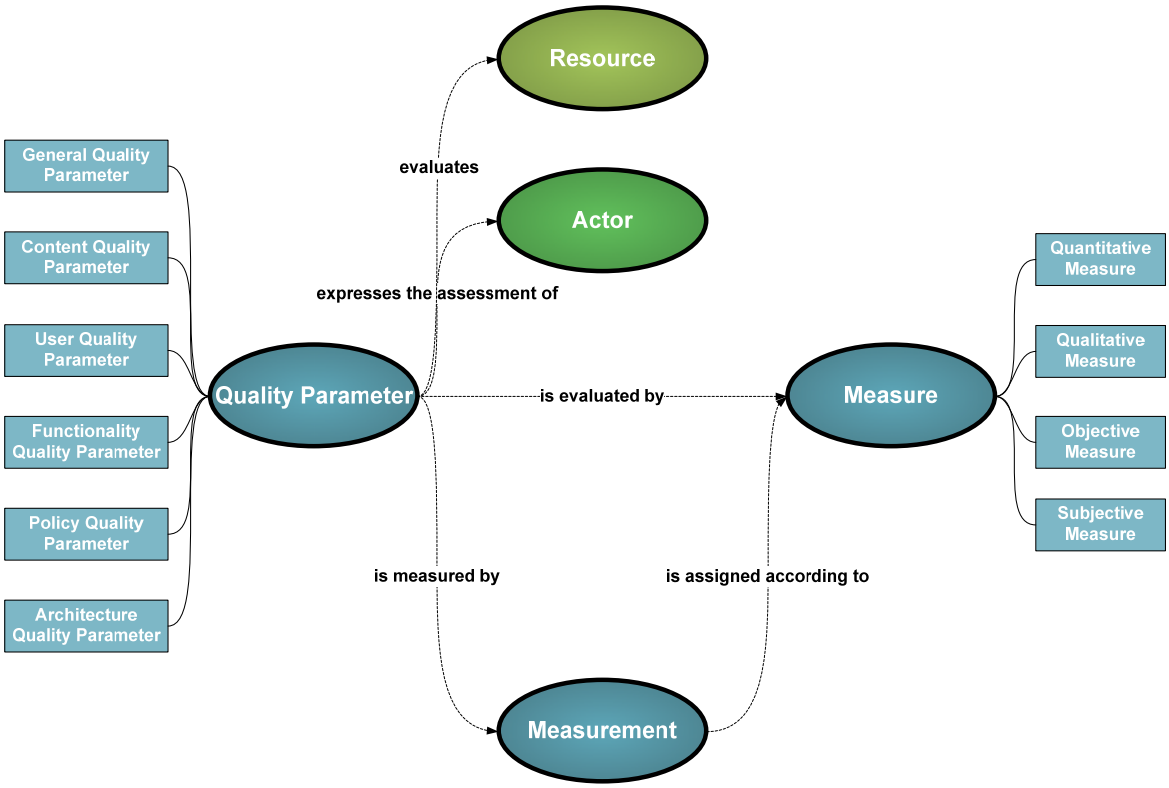

Figure 2: Main entities and their relationships involved in the quality main concept.

Figure 2 shows the relationships between quality, the other main concepts, and the main entities involved in quality. Three main entities – *Quality Parameter*, *Measure*, and *Measurement* – belong to the quality main concept, while two main entities – *Actor* and *Resource* – belong, respectively, to the user and content main concepts. An Actor is someone or something which interacts with the DL universe, while a *Resource* is any identifiable entity in the DL universe.

Quality Parameters serve the purpose of expressing the different facets of the quality main concept. They provide information about how, and how well, a *Resource* performs with respect to some viewpoint. They express the assessment by an *Actor*, human or computational, about the *Resource* under examination. They can be evaluated according to different *Measures*, which provide alternative procedures for assessing different aspects of a *Quality Parameter* and assigning it a value. *Quality Parameter*s are actually measured by a *Measurement*, which represents the value assigned to a *Quality Parameter* with respect to a selected *Measure*.

Measures are further categorized according to the following specializations:

- *Objective Measure*, when it is obtained via a well defined process that does not depend on individual perception;
- *Subjective Measure*, when it is based on, or influenced by, personal feelings, tastes, or opinions;
- *Quantitave Measure*, when it is based on a unit of measurement which is expressed via numerical values; and
- *Qualitative Measure*, when it is based on a unit of measurement which is not expressed via

numerical values.

In addition, *Quality Parameters* are specialized and grouped according to the *Resource* under examination; we have:

- *General Quality Parameters* when the assessed *Resources* are a *Digital Library*, or a *Digital Library System*, or a *Digital Library Management System*;
- *Content Quality Parameters* when the assessed *Resources* belong to the content main concept;
- *User Quality Parameters* when the assessed *Resources* belong to the user main concept;
- *Functionality Quality Parameters* when the assessed *Resources* belong to the functionality main concept;
- *Policy Quality Parameters* when the assessed *Resources* belongs to the policy main concept;
- *Architecture Quality Parameters*, when the assessed *Resources* belong to the architecture main concept.

Finally, for each group a detailed list of *Quality Parameters* is given in order to provide actual indicators that have to be taken into consideration when dealing with and evaluating the DL universe.

It is important to note that the grouping described above is made from the perspective of the *Resource* under examination, i.e., the object under assessment. In any case, the *Actor*, meant as the active subject who expresses the assessment, is always taken into consideration and explicitly modelled, since he is an integral part of the definition of *Quality Parameter*. For example, the *User Satisfaction* parameter is grouped under the *Functionality Quality Parameter* because it expresses how much an *Actor* (the subject who makes the assessment) is satisfied when he uses a given *Function* (the object of the assessment). On the other hand, in the case of the *User Behaviour* parameter, the object of the assessment is an *Actor* together with his way of behaving with respect to some policy; for this reason, this parameter is put under the *User Quality Parameter* group.

3 The Quality Domain in the 5S Model

In (Goncalves et al., 2007), a quality model for digital libraries is proposed. The model is deeply grounded in the 5S formal framework for digital libraries (Goncalves et al., 2004). For each major DL concept in the 5S framework, a number of *Quality Dimensions* are formally defined and a set of *Numerical Indicators* for those quality dimensions are proposed, as shown in Figure 3.

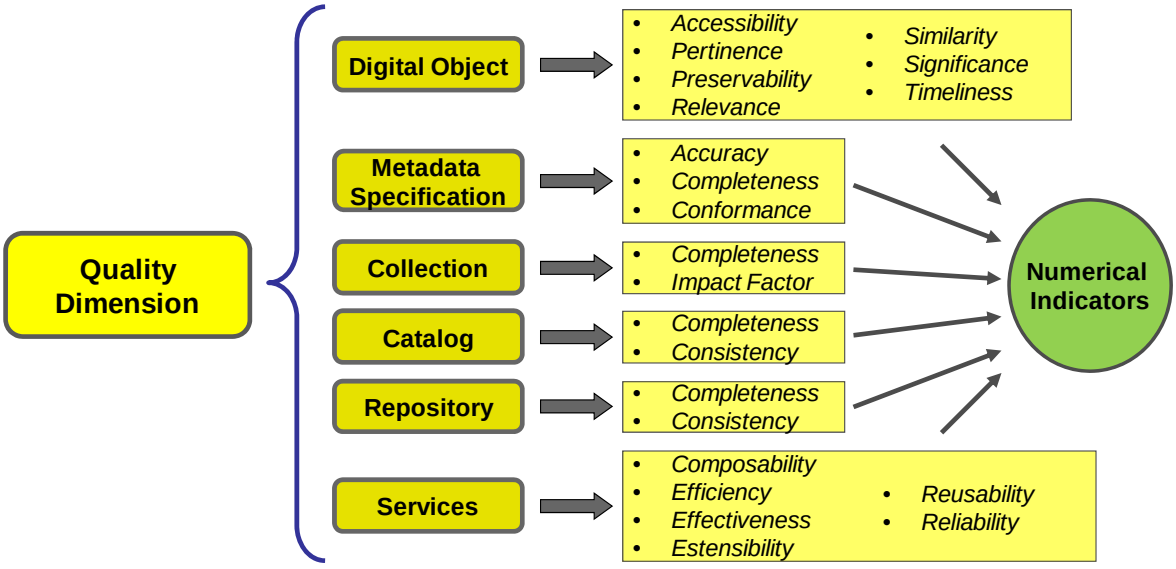

Figure 3: 5S quality model.

In particular, they consider key concepts of a minimal DL: *Digital Object*, *Metadata Specification*, *Collection*, *Catalog*, *Repository*, and *Services*. Regarding quality dimensions, they consider: accessibility, accuracy, completeness, composability, conformance, consistency, effectiveness, efficiency, extensibility, pertinence, preservability, relevance, reliability, reusability, significance, similarity, and timeliness. Regarding measurement, they consider characteristics like: response time (with regard to efficiency), cost of migration (with respect to preservability), and number of service failures (to assess reliability).

For some key DL concepts, pairs of form (quality dimension, numerical indicator) are illustrated through their application to a number of "real-world" digital libraries. The authors also discuss connections between the proposed dimensions of DL quality and an expanded version of a workshop's consensus view of the life cycle of information in digital libraries. Such connections can be used to determine when and where quality issues can be measured, assessed, and improved, as well as how possible quality problems can be prevented, detected, and eliminated.

To help operationalize this approach, a DL quality assessment toolkit has been developed and deployed (Moreira et al., 2006). The open source code that is available can be used by DL managers to assess the quality of their DL, based on the toolkit's processing of system logs and its access to DL content. To further extend the abovementioned 5S-based work, additional analysis has been undertaken to describe quality with regard to a union DL (Shen, 2006). Since many DLs are complex distributed systems (because of constraints related to politics, economics, intellectual property rights, etc.), and since users expect support to work with related content in a unified way, it is important to assess how easy it is to develop a union DL, as well as how useful is the integration of the disparate part.

Finally, it should be noted that the scope of the 5S-based work continues to expand, toward coverage comparable to the DELOS Reference Model (Murthy et al., 2007). In keeping with our adoption of a minimalist philosophy, we have defined a minimal DL in each of a number of key settings, leading to a set of different meta-models. One is for archaeological DLs (where there are real-world as well as digital objects). Another is for image DLs (where content-based image retrieval is supported). Others cover: union DLs, personal DLs, and DLs that facilitate knowledge management with superimposed information (e.g., annotations into parts of a document or image). Ultimately, as the 5S model continues to be used to support DL curricular development activities (Pomerantz et al., 2006), more analysis of the quality concept can proceed, with regard to the many areas in the DL domain.

4 Commonalities and Differences between the Two Approaches

To promote a deeper understanding of DL quality, to improve awareness of the differences between the two DL models, and to aid those seeking to interoperate between the two models, we have engaged in comparative analysis between the two models.

From a broad modelling point of view, the notion of Quality Parameter in the DELOS reference model corresponds to the Quality Dimension in the 5S quality model and both models further specialize these notions according to relevant DL facets: the six main concepts in the former case and the major DL concepts in the latter case. As it is shown in Figure 4, many of the major DL concepts of the 5S quality model refer to what is called Content Quality Parameter in the DELOS reference model, while the Services major DL concept corresponds to the Functionality Quality Parameter. This shows the broader coverage of the DELOS reference model with respect to the 5S quality model, since areas as General Quality Parameters, User Quality Parameters, Policy Quality Parameters, and Architecture Quality Parameters are not covered in the present version of the 5S quality model. It should be noted that this is also a consequence of the two different modelling approaches adopted: top-down for the DELOS reference model and bottom-up for the 5S quality model. It is thus natural that the DELOS reference model is broader in its scope but, as we will see later on, we expect that the 5S quality model is more detailed and precise with respect to specific areas. For example, the Metadata Specification major DL concept falls under the Content Quality Parameter area but, while the 5S quality model provides specific quality dimensions for this concept, the DELOS reference model provides only the means for covering this area but not actual quality parameters.

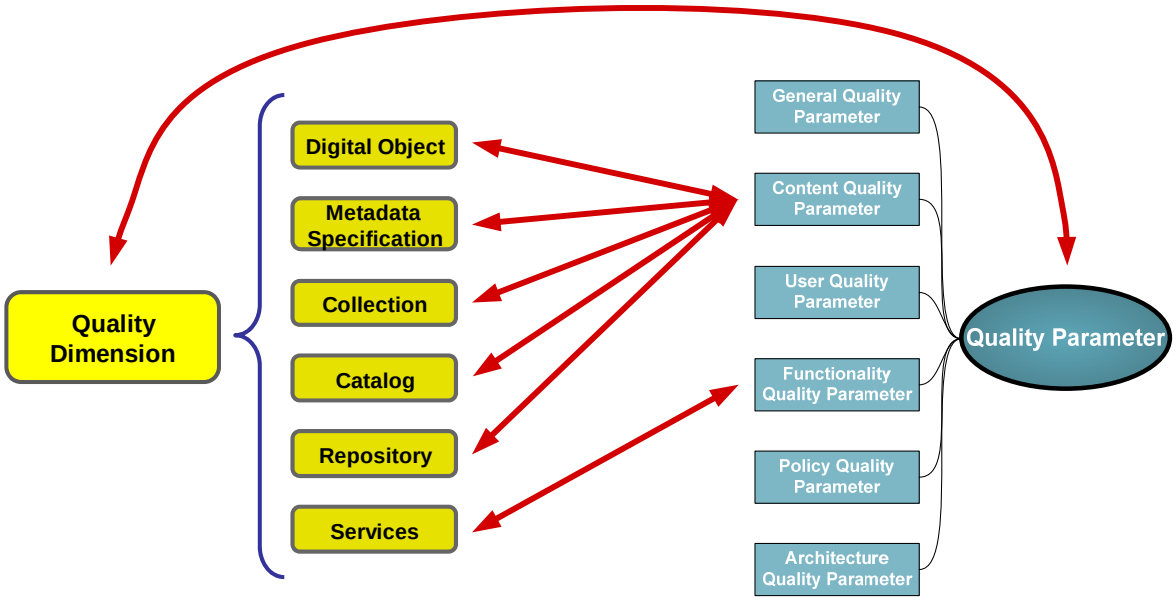

Figure 4: Comparison between the "major DL concepts" and the "six main concepts".

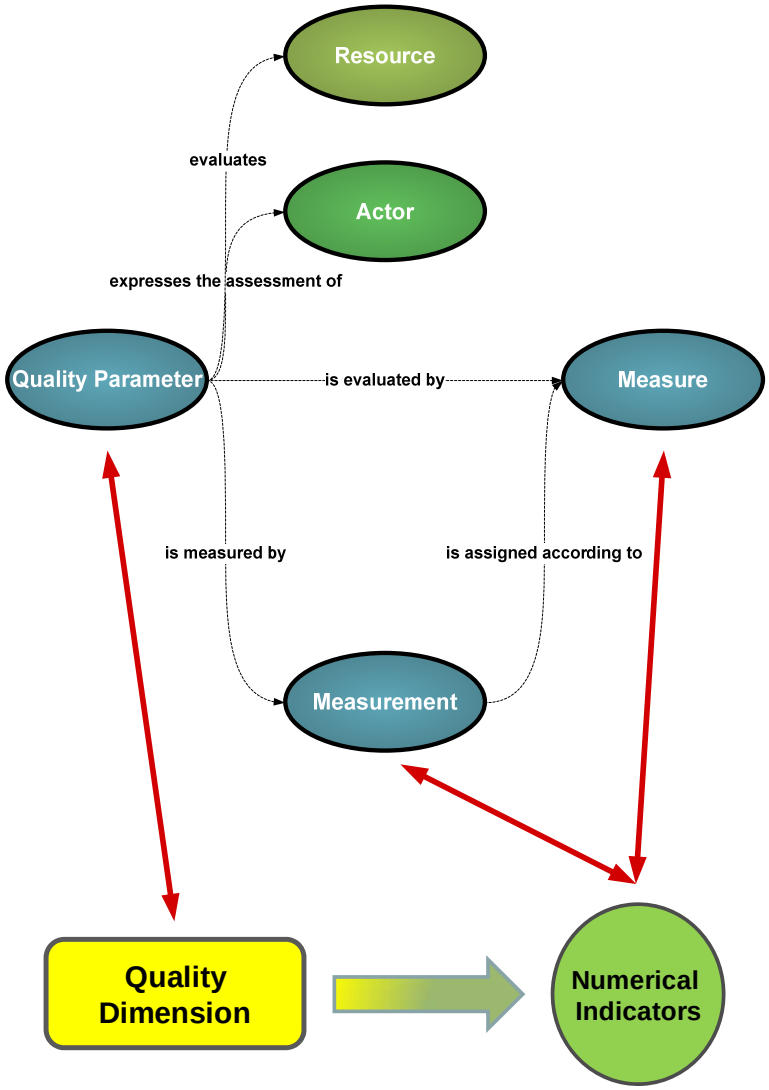

Figure 5: Comparison between the "numerical indicators" and the "measure/measurement" pair.

Figure 5 shows a more detailed comparison concerning the other entities involved in the quality realm. The notion of numerical indicator in the 5S quality model corresponds to the (measure, measurement) pair in the DELOS reference model, which in addition gives us finer control in modelling the distinction between the process adopted for measuring a quality parameter and the actual value assigned to a given measurement. Furthermore, the DELOS reference model explicitly takes into account both the Resource under assessment and the Actor who is carrying out the assessment, aspects which are not dealt with in the 5S model. Nevertheless, it should be noted that the notion of resource under assessment is somewhat implicit in the 5S quality model; indeed, each major DL concepts specifies which is the object of the assessment, e.g. a digital object, a service, and so on. These differences can still be brought back to the top-down and bottom-up approaches which each model has respectively adopted.

Therefore, as discussed above, the wider breadth and the more systematic modelling of the DELOS reference model with respect to the current version of the 5S quality model impact different areas of the quality realm. On the other hand, the 5S quality model gains much more depth in certain areas where the DELOS reference model only provides support for further investigation and deepening. For example, the 5S quality model provides actual numerical indicators for all the analysed quality dimensions while the DELOS reference models provides the means for introducing numerical indicators via the (measure, measurement) pair but it does not contain any actual numerical indicator. Moreover, the 5S quality model has been implemented in a software toolkit which can be used for assessing the quality of existing DLs, while for the DELOS reference model, being a more recent effort, there is no software implementation of the quality realm yet.

In the following, we discuss a more detailed comparison of the various quality parameters/quality dimensions proposed by the two models.

In particular, we perform a mapping between quality dimensions in the 5S-based quality model and the quality parameters of the DELOS Quality Domain. The mapping of equivalent quality parameters/dimensions in both models is shown in Table 1. Given as parentheticals are the types of concepts to which the quality dimensions/parameters should be applied.

Table 1: Mapping between quality dimensions in 5S and the DELOS Reference Model.

DELOS Reference Model	5S
Preservability (Content)	Preservability (Digital Object)
Authoritativeness (Content)	Significance (Digital Object)
Scope (Content)	Completeness (Collection/Catalog)
Freshness (Content)	Timeliness (Digital Object)
Availability (Functionality)	Reliability (Service)
User Satisfaction (Functionality)	Relevance/Pertinence (Digital Object)
Response Time (Functionality)	Efficiency (Service)

While in the overall comparison shown in Figure 4 we can find an almost direct and unique mapping between the “major DL concepts” of the 5S quality model and the “six main concepts” of the DELOS reference model, Table 1 shows that in a comparison of the detailed quality parameters and quality dimensions we can find mappings that cross the boundaries of the single major/main concept. For example the User Satisfaction quality parameter, which concern the Functionality main concept in the DELOS reference model, can be, for some respects, matched to the Relevance and Pertinence quality dimensions, which belong to the Digital Object major concept in the 5S quality model. These differences in the lower levels of the classification hierarchy needs to be further investigated, since they can be the clue of implicit assumptions in both models which can go beyond the simple adoption of the two different top-down and bottom-up approaches.

Table 2 provides a different comparison. We show the quality dimensions that are present in the 5S quality model but not in the DELOS Quality Domain. Table 3 provides the opposite comparison, giving what is covered in the DELOS Reference Model, but not (yet) in the 5S quality model.

Table 2: Quality Dimensions present in 5S and not in the DELOS Reference Model.

5S – DELOS Reference Model
Digital Object {Accessibility, Similarity}
Metadata {Accuracy, Completeness, Conformance}
Catalog, Repository {Consistency}
Service {Extensibility, Reusability}

Table 3: Quality Dimensions present in the DELOS Reference Model and not in 5S.

DELOS Reference Model – 5S
General Quality Parameter {All}
Content Quality Parameter {Integrity, Authenticity, Provenance, and Size}
Functionality Quality Parameter = {Usability, Documentation, Availability, Orthogonality, Fault Management, Robustness, Awareness of Service, Expectation of Service, Impact of Service}
User Quality Parameter {All}
Policy Quality Parameter {All}
Architecture Quality Parameter {All}

It can be seen from these two tables that the DELOS Reference Model in its Quality Domain is much broader than 5S (so far) in its coverage of quality aspects. This is due to the bottom up and minimalist approach used in 5S. Table 2 also shows that there is an opportunity for including some new quality dimensions in the Reference Model coming from 5S, which currently are not incorporated in the Reference Model.

Ultimately, quality models are meant to be used. We note, based on the work with deploying the 5S Quality Model through a toolkit (Moreira et al., 2006), that it would be helpful to have a similar aid to help with measurement of DL quality with respect to the DELOS Reference Model. One approach would be to extend the existing toolkit to cover additional concepts as listed in Table 3. On the other hand, for concepts covered in both models, we could use Table 1 to support mappings of terminologies.

5 Conclusions

This paper has discussed DL quality, from a model-based approach, but aimed toward practical measurement in support of assessment and evaluation. At the same time, it has provided a partial comparative analysis of the DELOS and 5S models, aiming to support interoperability of the models, and to help each advance toward a broader coverage of the DL domain. We hope this will be of benefit to those working toward establishing a formal foundation for the DL field, as well as be beneficial to those developing and deploying DL systems and services.

Acknowledgements

The authors thanks Milena Dobreva for the time spent in discussing and elaborating the different categories of measures. The work reported in this paper has been partially supported by the DELOS Network of Excellence on Digital Libraries, as part of the Information Society Technologies (IST) Program of the European Commission (Contract G038-507618). Support also was provided in a grant on Quality Metrics to Emory and Virginia Tech, and by NSF (IIS-0736055, 0535057, 0325579; DUE-0532825, 0435059).

References

Agosti, M., Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Schek, H.-J., & Schuldt, H. (2006). A Reference Model for DLMSs Interim Report. In L. Candela, & D. Castelli (Eds.), *Deliverable D1.4.2 - Reference Model for Digital Library Management Systems [Draft 1]. DELOS, A Network of Excellence on Digital Libraries -- IST-2002-2.3.1.12, Technology-enhanced Learning and Access to Cultural Heritage*. Online at: http://146.48.87.122:8003/OLP/Repository/1.0/Disseminate/delos/2006_WP1_D142/content/pdf?version=1

Agosti, M., Ferro, N., Fox, E. A., Gonçalves, M. A., & Lagoeiro, B. (2007). Towards a Reference Quality Model for Digital Libraries. In D. Castelli, & E. A. Fox (Eds.), *Proc. 1st International Workshop on Digital Libraries Foundations*. Online presentation at: http://fox.cs.vt.edu/talks/2007/20070623JCDLdlfWkshpAgosti-Quality.ppt

Candela, L., Castelli, D., Ioannidis, Y., Koutrika, G., Pagano, P., Ross, S., Schek, H.-J., & Schuldt, H.. (2006). The Digital Library Manifesto. *DELOS, A Network of Excellence on Digital Libraries - IST-2002-2.3.1.12, Technology-enhanced Learning and Access to Cultural Heritage.* Online at: http://146.48.87.122:8003/OLP/Repository/1.0/Disseminate/delos/2006_other_0081/content/pdf?version=1

Gonçalves, M. A., Fox, E. A., Watson, L. T., & Kipp, N. A. (2004). Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries. *ACM Transactions on Information Systems (TOIS)*, 22 (2), 270-312.

Gonçalves, M. A., Lagoeiro, B., Fox, E. A., & Watson, L. T., (2007). "What is a Good Digital Library? A Quality Model for Digital Libraries. *Information Processing & Management*, 43 (5), 1416-1437.

Moreira, Barbara Lagoeiro, Goncalves, Marcos Andre, Laender, Alberto H.F., and Fox, Edward A. 5SQual - A Quality Assessment Tool for Digital Libraries, in Proc. WDL 2006, 20 Oct. 2006, Florianopolis, SC, Brasil

Murthy, Uma, Gorton, Douglas, Torres, Ricardo, Goncalves, Marcos, Fox, Edward, and Delcambre, Lois. Extending the 5S Digital Library (DL) Framework: From a Minimal DL Towards a DL Reference Model. 1st Workshop on Digital Library Foundations, ACM/IEEE-CS Joint Conference on Digital Libraries, June 18-23, 2007, Vancouver, British Columbia, Canada, http://fox.cs.vt.edu/talks/2007/20070623JCDLdlfWkshpMurthy-5S.ppt

Pomerantz, Jeffrey, Oh, Sanghee, Yang, Seungwon, Fox, Edward A., and Wildemuth, Barbara M. The Core: Digital Library Education in Library and Information Science Programs. D-Lib Magazine, Nov. 2006 (Vol. 12 No. 11), http://www.dlib.org/dlib/november06/pomerantz/11pomerantz.html

Shen, Rao. Applying the 5S Framework to Integrating Digital Libraries, PhD dissertation, Virginia Tech Dept. of Computer Science, April 2006. Online at: http://scholar.lib.vt.edu/theses/available/etd-04212006-135018/