Towards a Reference Quality Model for Digital Libraries

Maristella Agosti
Dept. of Information
Engineering
University of Padua
Via Gradenigo, 6/b – 35131
Padova, Italy
agosti@dei.unipd.it

Nicola Ferro
Dept. of Information
Engineering
University of Padua
Via Gradenigo, 6/b – 35131
Padova, Italy
ferro@dei.unipd.it

Edward A. Fox
Dept. of Computer Science
Virginia Tech
Blacksburg, VA 24061
fox@vt.edu

Marcos André Gonçalves
Dept. of Computer Science
Federal University of Minas
Gerais
Belo Horizonte, M.G., Brazil
mgoncalv@dcc.ufmg.br

Barbara Lagoeiro
Dept. of Computer Science
Federal University of Minas
Gerais
Belo Horizonte, M.G., Brazil
barbara@dcc.ufmg.br

ABSTRACT

This paper discusses the importance of defining a Reference Quality Model for Digital Libraries. Current approaches for Digital Library (DL) quality evaluation are presented. Our view of the steps necessary to achieve this goal is given and discussed.

Categories and Subject Descriptors

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

General Terms

Design, Measurement, Theory

Keywords

digital libraries, quality, evaluation, reference model

1. INTRODUCTION

In this paper, we discuss issues related to defining a quality model for digital libraries in the light of the recent efforts for building a Reference Model for Digital Library Management Systems [1] [2].

The idea of a Reference Model is to lay the foundations for the digital library field as a whole. The lack of agreement on these foundations has led to a number of uncoordinated efforts that are hard to combine and reuse to produce enhanced outcomes.

---

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

One very important aspect of building such a Reference Model is to capture the notion of quality in DL, i.e., how can one define what is a "good" or "successful" digital library? [9, 18]. Evaluation models for digital libraries have been proposed [9, 18, 6, 7], but some of them were not built having a common DL model as their foundation.

Defining quality and quality measures for digital libraries will allow:

- to detect problems in the system and obtain information to fix them;

- to follow the evolution of systems and their several components (e.g., collections, catalogs, services);

- to evaluate contents to be inserted in the system and check if their quality is compatible with contents already in the system; and

- obeying certain constraints, to compare two or more systems, with regard to some of their components.

This is a very complex task. Quality measures need to be formally defined so they can be captured and quantified. Even if successful in proposing a theoretical quality model, we need to take the appropriate steps to support its operativeness. The model needs to be evaluated and tested in many settings. Standards to capture the necessary information (e.g., log formats) to allow the computation of quality measures need to be established. Once all of this is done, tools can be developed to help with such evaluations. Standard collections can be built to test services, among many other possible tasks.

This paper is organized as follows. Section 2 presents current approaches to quality in digital libraries and DL quality evaluation. Section 3 presents what we think are the necessary steps to produce and give support to such a Reference Quality Model. Section 4 concludes the paper.

2. APPROACHES TO QUALITY IN DIGITAL LIBRARIES

2.1 Broad Studies and Conceptual Frameworks

In [17] and [15] DL evaluation challenges and requirements are enumerated and an evaluation conceptual framework is suggested. The evaluation requirements should answer questions such as: "Why to evaluate?", "What to evaluate?", and "How to evaluate?". The framework considers that the evaluation should deal with performance aspects of parts of the DL system. Thus, the performance can be evaluated according to the effectiveness (how well the system performs its tasks), efficiency (what are the costs for the system to perform the tasks), or a combination of these two factors.

In [16] Saracevic provides an overview of the work on DL evaluation. He analyzed about 80 evaluation studies along the lines of:

1. constructs that were evaluated: the evaluated construct can be a specific digital library or a DL related process;

2. context in which the evaluations were conducted: the evaluation can deal with human, system, usability, anthropological, ethnographic, sociological, or economic aspects;

3. criteria that were chosen as a basis for evaluation, i.e., the judgment standard that was defined for the evaluation. The criteria depend on the context, for instance, for an usability evaluation, the criteria "effort to understand" and "error rate" can be used as a basis for the evaluation;

4. methods that were used during the evaluation; some of those methods are: surveys, structured interviews, observations, case studies, focus groups, transaction log analyses, experimentation, and usage analysis.

In [6] is defined a DL conceptual model to develop test suits that would satisfy the needs of researchers in the DL evaluation area. This model is based on four main dimensions: data & collection, system & technology, users, and usage. The idea consists of using the relationships between these dimensions to create a set of evaluation criteria that, when answered, would generate DL detailed descriptions. These descriptions can be applied to define test-beds or to compare digital libraries.

Nicholson [14] presents a conceptual framework to guide holistic evaluations of library services, considering different points of view: from the user, library staff, and decision makers. Using a matrix of topics and perspectives for measurement, the evaluator can choose what to evaluate and how to evaluate it. This matrix presents the following views:

- Internal View of the System (what are the components of the system): compares components of the system against some type of standard. To evaluate it, staff interviews and surveys, and audits of collections, system, or staff can be used.

- External View of the System (how effective is the system): the user presents a query to the library and evaluates the usability of the system and the returned results. To evaluate it, interviews and focus groups can be used.

- External View of Use (how useful is the system): the user presents the overall usefulness of information obtained through the system. Surveys, interviews, focus groups, and user citation tracking can be used to evaluate it.

- Internal View of Use (how is the system manipulated): interactions between users and a system are analyzed to understand how a system is manipulated. This can be evaluated through the analysis of logs and user behavior.

Tsakonas et al. [19] developed a framework to evaluate the interaction between the user and the DL. An interaction is composed of three components: the user, the content and the system. The work considers three categories of evaluation criteria which define relationships among components: usability (the quality of the direct interaction between the user and the system), usefulness (whether the user needs are being fulfilled by the content), and performance (considering the system response). These categories can be applied to highlight requirements, parameters and metrics for the interaction evaluation.

In [7], the goal is to provide a set of flexible and adaptable guidelines for DL evaluation, outlining the main directions, methods, and techniques for assessing the components of a DL. Besides that, a study about existent DL evaluation approaches is performed, describing the main models to be applied during an evaluation. After this discussion, a framework based on [16] is described, trying to cover most of the aspects that can be found through the several levels of an evaluation process.

In [18], Shen proposes a model of DL success from the end user perspective, based on the integration of various research studies of different areas (e.g., digital libraries and information systems). This model helps to define when and how to measure the different quality aspects. In addition, numeric indicators for the quality of union catalogs and union services are specified.

Aimed at an evaluation from the user point of view, in [12] DigiQual is developed, a protocol based on a similar project for traditional libraries, which helps the DL administrators to understand the quality notion of the users of their system. The protocol defines that the users can answer about 12 quality themes throughout the time, systematically, to identify the best practices for a DL system.

Proposals that develop and present standards for log formats aimed at registering data for evaluation, such as [10] and [11], contribute towards a Reference Quality Model since they provide ways for storing information for assessment. In [11], a multi-level logging schema is proposed that accounts for a large amount of data about users, systems and user-system interactions. Some of this information is difficult to capture (e.g., information about the user behavior may require observing or interviewing the user, which may be very time consuming). Because of that, Klas et al. [11] focus their work on the concept level, which comprises general DL events such as *search*, *browse*, and *navigate*. They store information about services, like the timestamps for the start and end of an event, and the errors that may have occurred. Their proposal is built on top of the work of Gonçalves et al. [10], which describes an XML-based log format that captures detailed information about users and system behavior.

Gonçalves et al. [9] define an explicit formal/quantitative

quality model for digital libraries based on the 5S formal framework for digital libraries [8]. The model is validated through its application to several DLs in different scenarios. A tool implementing a portion of the model has been developed [13].

2.2 The DELOS Approach

2.2.1 The DELOS Reference Model

Digital Library is a complex concept which can be expressed using different perspectives and viewpoints. The DELOS[2] approach for the representation of this many-sided concept has been to start an effort for developing a Reference Model [2] where a framework of three tiers to represent three levels of abstraction is used to represent: the DL, the DLS and the DLMS. The DL is the level where the digital contents are kept, and the DLS is the level of all the organizational and software application components that are able to manage the contents, providing useful services to the interested users with the support of a DLMS.

The DELOS Reference Model [2] aims at providing a representation which characterizes existing and future DLMS from at least the four perspectives: DL end-users, designers, system administrators, and application developers. It introduces the main concepts, the relationships between these concepts, and the constraints that hold among them. It also prescribes aspects that are mandatory for this type of information system. Figure 1 (extracted from [4]) represents the highest level concepts of the DELOS reference model:

content is the entry point for all the concepts related to the content that is managed and disseminated by the DL, e.g., collections, information space model, metadata, ontologies;

user is the root for concepts like roles, communities, and profiles, that represent aspects of the DL users;

functionality is the entrance to that part of the model which concerns DL functions;

architecture regards software components, hosting nodes and how these are linked and constrained;

quality groups qualitative parameters characterizing the digital library behavior within a given operational domain;

policy covers all the concepts that are related to established procedures or plans of actions governing the DL, such as collection management, preservation, and access rights.

From a final user point of view, a DLS is the collection of tools he can use to access and browse the collection of digital information objects – the Digital Library – that is of interest, where the management and the keeping over time of the objects is done by a DLMS and the maintenance of the collection of objects is secured by an organization in charge of it. As outlined in the DELOS Digital Library Manifesto [4], at least three types of conceptually different "systems" can characterize the digital library universe: the DL, the DLS and the DLMS, which are hierarchically related; so are their models, i.e., the DL model is included in the DLS one, and the latter is included in the DLMS model.

<hr/>

[2] http://www.delos.info/

2.2.2 Notion of Quality in the DELOS Reference Model

The notion of "quality", which is one of the highest level concepts of the DELOS Reference Model [2], as seen in Section 2.2.1, can be considered at each of the three levels of abstraction: DL, DLS, and DLMS. This means that we can define quality parameters for the information objects, for the services given to the users, and for the system that supports the management of the services. Once the quality parameters are defined, the control of them can be pursued making use of specific control tools and mechanisms. Quality encompasses the characteristics of a DL and the resources that it contains that can benefit from being measured and monitored. The quality is expressed by a set of quality parameters; each parameter can be measured; those measurements are mostly related to the contents and the functionalities.

As far as the content is concerned, the quality of each information object needs to be verified over acquisition and lifetime, because of that it becomes necessary to define a set of quality parameters that can be expressed via a value assigned as result of measurement, where the act of measuring includes a quality parameter in accordance with a selected process and a unit of measurement. The value of a quality parameter is obtained via the selected process, that does not depend on individual perception.

One parameter related to the content assesses the information object quality of being complete. This parameter encompasses the extent to which an information object is of sufficient breadth, depth, and scope for the task at hand, as pointed out in [3].

Authenticity is a content quality parameter which measures whether an information object retains the property of being what it purports to be; this definition takes into account the results and experience of the InterPARES project [5]. The provenance content quality parameter concerns the origin or earliest known history of an information object. This parameter is particularly important when dealing with scientific data. The provenance of data must be tracked since a scientist needs to know where the data came from and what cleaning, rescaling, or modelling was done to arrive at the data to be interpreted [1].

The DLS is the system in charge of implementing the DL. It is composed of components and hosting nodes. As a consequence, the DLS inherits quality related concepts from the DL while it needs new relationships that make it possible to assign such parameters to the entities it deals with, e.g., a hosting node. In particular, the quality assigned to a component supports the DL system administrator during the component selection and configuration phases.

The DLMS is a software system with diverse components. Like other well-constructed software systems, the DLMS has been conceived and developed applying principles and methods of software engineering. Taking into account that the fundamental principles of software engineering are applicable throughout the software life cycle, the DL designer, the DL system administrator, and the DL application developer need to make reference to those general principles. In particular they need to refer to software engineering best practices regarding software quality measurement.

From a final user point of view, taking note that the final user mostly uses an access function to search and browse the DL with the final aim of having delivered a copy of information objects of interest and of certified quality, the final user is interested in quality and quality control over
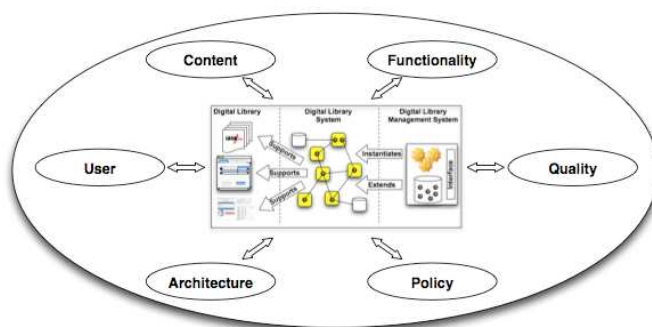
Figure 1: Highest Level Concepts of the DELOS Reference Model.

the contents of the DLS, and the notion of quality of the search for contents, where the searching is implemented by a searching service that has to be evaluated before, to be given for use to the final user.

3. STEPS TOWARDS A REFERENCE QUALITY MODEL FOR DLS

We envision the following steps in order to achieve the goal of building a Reference Quality Model for DL:

1. Contribute to the definition of a Reference Model for DL

   A set of concepts (at least a minimal set), defining what aspects, that have to be taken into account in a digital library, are going to be defined.

2. Formalization of the Model

   To support precision and accuracy in the definition of the concepts in the Reference Model, there is the need to formalize pertinent aspects.

3. Definition and Formalization of Quality Indicators

   Quality dimensions for several of the concepts defined in the Reference Model need to be defined. Numeric indicators for each quality dimension will then be proposed based on the formalization of the concepts in the Reference Model provided in Step 2.

4. Defining the context for each quality dimension in light of the Information Life Cycle

   Each quality dimension needs to be associated with one phase of the Information Life Cycle (i.e., Creation, Distribution, Seeking, and Utilization). This will set the context for specifying when we can apply and compute the respective numeric indicators for each quality dimension and how to use the results of the quality analysis.

5. Discussion with the community and reformulation

   The model needs to be discussed with the community and to be validated by it. Several reformulations to accomodate several different perspectives may be necessary.

6. Providing Support for the Model

   Once we have a solid version of the Reference Quality Model, tools implementing the numeric indicators for

each dimension and supporting the envisioned evaluation process need to be build. We will also need standards such as a standard log format to help to capture the necessary information for evaluation.

4. CONCLUSIONS

A co-operative work to make some steps towards the definition of a complete Reference Quality Model for DL has been initiated having in mind the objective of defining, and developing, a model where all previous relevant experiences come together in a synergistic way.

Acknowledgments

5. REFERENCES

[1] S. Abiteboul, R. Agrawal, P. Bernstein, M. Carey, S. Ceri, B. Croft, D. DeWitt, M. Franklin, H. Garcia-Molina, D. Gawlick, J. Gray, L. Haas, A. Halevy, J. Hellerstein, Y. Ioannidis, M. Kersten, M. Pazzani, M. Lesk, D. Maier, J. Naughton, H.-J. Schek, T. Sellis, A. Silberschatz, M. Stonebraker, R. Snodgrass, J. D. Ullman, G. Weikum, J. Widom, and S. Zdonik. The Lowell Database Research Self-Assessment. *Communications of the ACM (CACM)*, 48(5):111–118, 2005.

[2] M. Agosti, L. Candela, D. Castelli, N. Ferro, Y. Ioannidis, G. Koutrika, C. Meghini, P. Pagano, S. Ross, H.-J. Schek, and H. Schuldt. A Reference Model for DLMSs – Interim Report. In L. Candela and D. Castelli, editors, *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 – `http://146.48.87.122:8003/OLP/Repository/1.0/Disseminate/delos/2006_WP1_D142/content/pdf?version=1` [last visited 2007, March 23], September 2006.

[3] C. Batini and M. Scannapieco. *Data Quality.* Springer-Verlag, Berlin, Germany, 2006.

[4] L. Candela, D. Castelli, Y. Ioannidis, G. Koutrika, P. Pagano, S. Ross, H.-J. Schek, H. Schuldt,and C. Thanos. Setting the Foundations of Digital Libraries: The DELOS Manifesto. *D-Lib Magazine,* Vol. 13 No. 3/4, March/April 2007.

[5] L. Duranti. The long-term preservation of accurate and authentic digital data: the INTERPARES project. *Data Science Journal,* 4:106–118, 2005.

[6] N. Fuhr, P. Hansen, M. Mabe, A. Micsik, and I. Solvberg. Digital libraries: A generic classification and evaluation scheme. In *Proc. of the European Conf. on Digital Libraries,* pages 187–199, Heidelberg, 2001. Springer.

[7] N. Fuhr, G. Tsakonas, T. Aalberg, M. Agosti, P. Hansen, S. Kapidakis, C.-P. Klas, L. Kovas, M. Landoni, A. Micsik, C. Papatheodorou, C. Peters, and I. Solvberg. Evaluation of digital libraries. *Int. Jour. of Digital Libraries,* 2007.

[8] M. A. Gonçalves. *Streams, Structures, Spaces, Scenarios, and Societies: A Formal Framework for Digital Libraries and Its Applications: Defining a Quality Model for Digital Libraries.* PhD thesis, Virginia Tech CS Department, Blacksburg, Virginia, 2004. URL - http://scholar.lib.vt.edu/theses/available/etd-12052004-135923/.

[9] M. A. Gonçalves, B. L. Moreira, E. A. Fox, and L. T. Watson. What is a good digital library? - defining a quality model for digital libraries. To appear in Information Processing and Management, 2007.

[10] M. A. Gonçalves, G. Panchanathan, U. Ravindranathan, A. Krowne, E. A. Fox, F. Jagodzinski, and L. Cassel. The XML log standard for digital libraries: analysis, evolution, and deployment. In *Proc. of the 3rd ACM/IEEE-CS Joint Conf. on Digital Libraries,* pages 312–314, Washington, DC, USA, 2003. IEEE Computer Society.

[11] C.-P. Klas, N. Fuhr, S. Kriewel, H. Albrechtsen, G. Tsakonas, S. Kapidakis, C. Papatheodorou, P. Hansen, L. Kovacs, A. Micsik, and E. Jacob. An experimental framework for comparative digital library evaluation: the logging scheme. In *Proc. of the 6th ACM/IEEE-CS Joint Conf. on Digital Libraries,* pages 308–309, New York, NY, USA, 2006. ACM Press.

[12] M. Kyrillidou and S. Giersch. Developing the DigiQual protocol for digital library evaluation. In *JCDL '05: Proc. of the 5th ACM/IEEE-CS Joint Conf. on Digital Libraries,* pages 172–173, New York, NY, USA, 2005. ACM Press.

[13] B. Lagoeiro, M. A. Gonçalves, and E. A. Fox. 5squal: A quality tool for digital libraries. In *Proc. of the 7th ACM/IEEE Joint Conf. on Digital Libraries,* page (demonstration accepted), New York, NY, USA, 2007. ACM Press.

[14] S. Nicholson. A conceptual framework for the holistic measurement and cumulative evaluation of library services. *Jour. of Documentation,* 60(2):164–182, 2004.

[15] T. Saracevic. Digital library evaluation: Toward evolution of concepts. *Library Trends - Special issue on Evaluation of Digital Libraries,* 49(3):350–369, 2000.

[16] T. Saracevic. Evaluation of digital libraries: an overview. Presentation at the DELOS WP7 workshop on the evaluation of digital libraries, Department of Information Engineering, University of Padua, Italy, October 2004. URL - http://www.scils.rutgers.edu/~tefko/DL_evaluation_Delos.pdf.

[17] T. Saracevic and L. Covi. Challenges for digital library evaluation. In *Proc. of the 63rd Annual Meeting of the American Society for Information Science,* volume 37, pages 341–350, 2000.

[18] R. Shen. *Applying the 5S Framework to Integrating Digital Libraries.* PhD thesis, Virginia Tech CS Department, Blacksburg, Virginia, 2006. URL - http://scholar.lib.vt.edu/theses/available/etd-04212006-135018/.

[19] G. Tsakonas, S. Kapidakis, and C. Papatheodorou. Evaluation of user interaction in digital libraries. In M. Agosti, N.Fuhr, ed.: Notes of the DELOS WP7 Workshop on the Evaluation of Digital Libraries, 2004. URL - http://dlib.ionio.gr/wp7/workshop2004/_program.html.