

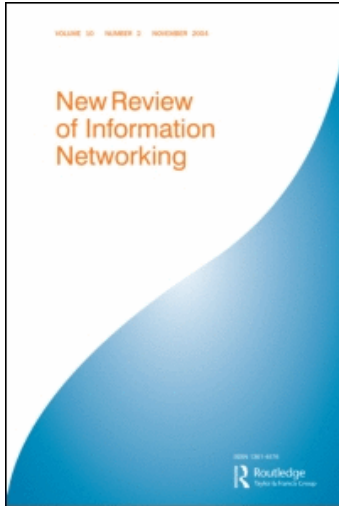
This article was downloaded by: [Rutgers University]

On: 3 January 2011

Access details: Access Details: [subscription number 921325574]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



New Review of Information Networking

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713703081>

Towards a Digital Library Policy and Quality Interoperability Framework: The DL.org Project

Perla Innocenti^a; Giuseppina Vullo^a; Seamus Ross^b

^a Human Advanced Technology & Information Institute (HATII), University of Glasgow, ^b Humanities Advanced Technology & Information Institute (HATII), University of Glasgow, University of Glasgow and Faculty of Information, University of Toronto,

Online publication date: 06 May 2010

To cite this Article Innocenti, Perla , Vullo, Giuseppina and Ross, Seamus(2010) 'Towards a Digital Library Policy and Quality Interoperability Framework: The DL.org Project', *New Review of Information Networking*, 15: 1, 29 – 53

To link to this Article: DOI: 10.1080/13614571003751071

URL: <http://dx.doi.org/10.1080/13614571003751071>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

TOWARDS A DIGITAL LIBRARY POLICY AND QUALITY INTEROPERABILITY FRAMEWORK: THE DL.ORG PROJECT

PERLA INNOCENTI and GIUSEPPINA VULLO

Human Advanced Technology & Information Institute (HATII),
University of Glasgow

SEAMUS ROSS

Humanities Advanced Technology & Information Institute (HATII),
University of Glasgow, University of Glasgow and Faculty of Information,
University of Toronto

Interoperability is a property referring to the ability of systems and organizations to work together. Today interoperability is recognized as a key step in the shift from isolated digital libraries toward a common information space that will allow users to browse through different digital libraries within a single integrated environment. In this paper, we discuss the premises underlying a novel Policy and Quality Interoperability Framework, taking into account the preliminary outcomes and the recommendations of the Policy and Quality Working Groups that are currently being run by the EU co-funded project Digital Library Interoperability, Best Practices, and Modeling Foundations (DL.org).

Keywords: *digital library, digital library organization, digital library systems, interoperability, modeling, policy, quality, framework*

Introduction

Digital libraries represent the confluence of many interdisciplinary fields, from data management, information retrieval, library sciences, document management to web services, information systems,

The work reported has been partially supported by the DL.org Coordination and Support Action, within FP7 of the European Commission, ICT-2007.4.3, Contract No. 231551.

We are grateful to the following participants of the DL.org Policy and Quality Working Groups (in alphabetical order) for their contribution to this research: Kevin Ashley, Genevieve Clavel-Merrin, Antonella De Robbio, John Faundeen, Nicola Ferro, Sarah Higgins, Wolfram Horstmann, Sarantos Kapidakis, Steve Knight, Hans Pfeiffenberger, Dirk Roorda, Tefko Saracevic, and Mackenzie Smith.

Address correspondence to Perla Innocenti, HATII at the University of Glasgow, 11 University Gardens, Glasgow, Scotland, G12 8QQ, UK. E-mail: p.innocenti@hatii.arts.gla.ac.uk

image processing, artificial intelligence, human-computer interaction, and digital curation. This multi-faceted nature has led researchers to offer a variety of definitions as to what a digital library is, often reflecting on different disciplinary perspectives (Borgman 1999; Fox et al. 1995; Ioannidis 2001, 2005; Ioannidis et al. 2005; Lagoze 2010). As Gonçalves et al. (2004) have explained the lack of unambiguous clarity on the boundaries of the term digital library arise because they are essentially complex multi-dimensional applications.

Ross (2003) pinpointed those aspects by characterizing a digital library as “the infrastructure, policies and procedures, and organisational, political and economic mechanisms necessary to enable access to and preservation of digital content” (p. 5). Among the current crop of digital libraries, there is a variety in character and type of content, with some being homogeneous collections on particular topics or media whereas others have a heterogeneous character (Ross 2003). All digital libraries are systems, and they instantiate particular systems and information architectures. The lack of agreement on the best design of digital library systems reflects, in part, a lack of agreement on the nature, functionality, and architecture of such applications.

The DELOS Digital Library Reference Model (Candela et al. 2008) aimed to address these lacunae. Starting with the DELOS Reference Model as its conceptual framework, the EU-funded Digital Library Interoperability, Best Practices and Modeling Foundations (DL.org, <http://www.dlorg.eu/>) project investigates interoperability issues in the context of digital libraries. Digital libraries are part of larger ecosystems and must be able to interrelate within the ecospace and with other infoecospaces. DL.org addresses digital library interoperability issues from the perspective of six core constituent parts (Architecture, Content, Functionality, Policy, Quality, and User) of the digital library. This paper focuses on two of these domains presenting the research hypotheses, theses, and the first outcomes on Policy and Quality Interoperability developed within the DL.org Policy and Quality Working Groups. As Ross (2008) noted interoperability in digital library infoecosystems depends upon reconciling heterogeneous policies in digital libraries, measuring quality of content and services, and addressing both these domains within the context of change (e.g., rising expectations of quality or policy drift).

Policy

Defining Policy in Digital Libraries

A policy can be understood as political, management, financial, and administrative mechanisms structured to ensure the delivery of certain consistent outcomes or behaviors.

In the world of digital libraries, a policy is typically described as a condition, term or regulation governing the operation of a digital library or some aspect thereof. People (such as digital library staff members, managers, and stakeholders) make policies for digital libraries. Sometimes, these policies can be expressed as rules. Rules provide mechanisms to express complex policies in ways that computer systems can interpret and apply them. At a user's level, digital library access policies must be enforced, and users often need to "be informed of the policies and educated as to what constitutes a reasonable behaviour" (Arms 2000), normally through usage policies. At a repository or at a collection level, formalized policies can be followed through trusted systems or through secure combiner (encryption, digital signatures, and public-key encryption). The subsequent user guidelines covers all the operations that can occur once materials are no longer under the direct management of the digital library.

Within the 5S Model (Gonçalves et al. 2004), policy is part of the socio-economic/legal aspects of the taxonomy of DL terms, and can be enforced by specific services, ranging from authentication and authorization, to particular application scenarios. Finally, a policy could be affected by quality parameters. This could require a quality assurance (QA) which would ensure "documented policies on the standards and best practices to be implemented and systematic procedures for measuring compliance with these policies" (Kelly 2004). This methodology was, for example, developed in 2004 to support JISC's digital library programs. A digital library policy can be considered as a specific case of *information policy*, which is defined as any law, regulation, rule, or practice that affects the creation, acquisition, disposition, organization, dissemination, use, or evaluation of information (Rubin 2000). In fact, policies and practices that are established and implemented by digital libraries regarding the creation, organization, use and dissemination of the knowledge are themselves *information policies* and they

have a tremendous impact on the accessibility of information contained in those digital libraries.

Policy Within the DELOS Digital Library Reference Model

The current definition of *policy* in the DELOS Digital Library Reference Model is that “the policy concept represents the set or sets of conditions, rules, terms and regulations governing interactions between the Digital Library and its users, whether virtual or real. Examples of policies include acceptable user behavior, digital rights management, privacy and confidentiality, charges to users, and collection delivery. Policies belong to different classes; for instance, not all policies are defined within the Digital Library or the organisation managing it. The policy [concept] supports the distinction between extrinsic and intrinsic policies. The definition of new policies and re-definition of older policies will be a feature of digital libraries” (Candela et al. 2008, p. 20).

In the same Reference Model, a *digital library* is defined as “an organisation, which might be virtual, that comprehensively collects, manages and preserves for the long term rich *Information Objects*, and offers to its *Actors* specialised *Functions* on those *Information Objects*, of measurable quality, expressed by *Quality Parameters*, and according to codified *Policies*” (Candela et al. 2008, p. 157). However, the current version of the Reference Model currently does not render nor explore the organizational context within which a digital library is instantiated and run; as many policies are extrinsic to the library itself and influence both content and architecture, this inter-relationship requires attention.

Policy Interoperability

The Policy Working Group agreed on the following definition of Policy Interoperability: Policy interoperability is seen as business level interoperability. This is a policy framework that makes it possible to compare and trust values and purposes of each organization. This type of interoperability is about peer-to-peer interoperability, but also about third-party service providers interoperable policies (i.e., data archives and the policies exchange with cloud providers). This is a high level policy interoperability that needs to be instantiated at process

level, whether those policies are being handled by human or machine.

Policy Interoperability Levels

In Library and Information Science, sadly, little work—and therefore few scientific publications—have been so far dedicated to investigated policies and policies harmonization in digital libraries. Only a firm understanding of the digital library's principles and aims, as expressed via policies, can provide the necessary rationale to connect users, contents, and functionalities. Digital libraries' policies implicate several economic, legal, and political issues, as the computer network services become more and more integrated into society, crossing their jurisdictional boundaries (Borgman 2000). Users' behaviors themselves are also influenced by those issues, as their role has become much more active and often the real boundaries are not perceived. In particular, legal boundaries have become a vital issue within the international computer networks and there is a need of governance, regulation, and Information Technology policy in order to link independent information systems. Among many aspects of DLs, new policy models are needed for protecting and managing intellectual property, privacy, and security, as well as new refined methods for authenticating users, documents, and transactions.

In 2004, a number of studies included investigations on policies for digital libraries. According to Dalton, Hartland-Fox and Thebridge (2004), evaluation data can be used in digital library policies to assess the extent to which the outcomes of the Electronic Information Services relate to the wider information service and institutional outcomes. It can also be used to devise policies (e.g., on electronic content and collection management). In a collective publication, a team of international experts from various countries including England, South Africa, The Netherlands, USA, Scotland, and Israel focused on digital libraries policies, planning, and practices, considering key policy issues that needed to be addressed (Andrews & Law 2004).

A policy could be affected by quality parameters which could require a QA. The QA would ensure that formally documented policies are systematically compliant with standards and best practices to be implemented. This methodology was, for example, developed in 2004 to support JISC's digital library programs

(Kelly 2004). Among the classification attempts, Lesk (2005) classified digital library policies from an internal point of view: funding policies (i.e., the decisions about the institutional support, the advertisers, the charges to apply to some services); legal policies (i.e., the protection which has to be given to the digital and the digitized contents, the decision about their long-term accessibility and preservation); privacy policies, which are involved in any interactive digital environment. More recently, in her 2008 overview of digital rights management in the library environment, Grace Agnew (2008) also discussed policies for digital asset management systems and digital library repositories, in order to identify and support the use of the authentic resource.

The 2009 DISK-UK Data Share Project Policy-making Guide is intended to be used as “a decision-making and planning tool for institutions with digital repositories in existence or in development that are considering adding research data to their digital collections” (Green, Macdonald, & Rice 2005). It is a collection of selected policy considerations compiled from multiple sources and organized in sections on research data quality, management, and preservation. With this guide “repository planners and developers can evaluate each set of requirements and choose the most applicable options for their policy and planning purposes” (Green, Macdonald, & Rice 2005, p. 4).

But, to be fair to the relatively small amount of scientific publications in this field, it should be noted that the digital library domain is really a cross-disciplinary area in which many other disciplines converge. In more specific research domains—such as computer science theory, digital content management, data management, e-science, risk assessment, and digital repository certification, health care, and medical sector, Open Access Initiative—investigations have been conducted on a variety of aspects related to policies.

In the computer science domain, in the last decade policy-based systems have been used by diverse communities in a wide range of activities across the academic sector, the industrial sector and standardization bodies. Policy-based systems are widely employed in applications ranging from “quality of service management within networks, to security, access control policy, and enterprise modelling” (Lutfiyya, Garcia, & Moffett 2003). In these applications, policies are commonly defined as a set of rules governing choices in the behavior of a system, which can be changed modifying the policy rather than re-implement the system or amending

the requirements specification. A successful series of workshops focusing on policies for distributed systems (<http://www.policy-workshop.org/>) have been organized over the years by Institute of Electrical and Electronics Engineers (IEEE). The focus of the papers, published in IEEE proceedings, include policy-based networking, privacy and security management, storage area networking, enterprise systems, access control management, wireless networks, automation, and control pervasive environments, grid-computing and multi-agent systems, policy specifications, integration with management systems, trusted systems, and large scale systems.

Also, within the World Wide Web community there is considerable interest in Policy Based Networking. The World Wide Web Consortium (W3C) Web Services Policy Working Group (<http://www.w3.org/2002/ws/policy/>), for example, in 2007 produced a W3C (draft) documents on a Web Services Policy Framework (<http://www.w3.org/TR/2007/REC-ws-policy-20070904/>). This framework provides a general purpose model and corresponding syntax for expressing policies related to specific capabilities, requirements, and entities in a Web services-based system, encompassing a broad range of service requirements and capabilities.

In the e-science domain, Christine Borgman (Borgman et al., 2007) focused on digital library requirements for habitat ecology data. Habitat ecology is a discipline currently in a state of “transition from small science,” characterized by hand-crafted data collection to “big science,” with instrumented data collection, larger volumes of data, and distributed multi-disciplinary research teams. This is a new way of “doing science,” which requires new kinds of practices. The findings of this study have been used to identify design and policy considerations for digital libraries in e-science. Specifically for data policy, it highlighted how e-science scenarios contrasted with scientific and engineering views about data use policy.

Policy Interoperability Issues

If we were to sum up the areas in which policies for digital libraries are needed, upon the indications found in the scientific literature, these could be approximately indicated as follows: Access policies, Acquisition policies, Administration and Management, Cooperation policies, Digital Preservation policies, Disposal policies, Dissemination policies (Open Access), Distributed system and

network management, Internet policies for users, Personnel and staffing policies, Reference policies, Security and privacy policies, and Selection and collection development policies.

But this is evidently an incomplete list. And there is a MIT project, PLEDGE (http://pledge.mit.edu/index.php/Main_Page), which focused on the determination of a set of policies that affect operational digital preservation archives, with the goal of developing standardized means of recording and enforcing them using rules engines (Smith & Moore, 2006). The four main policy categories identified were: Organization, Environment and Legal Policies, Community and Usability Policies, Process and Procedure Policies, and Technology and Infrastructure Policies. An initial mapping between the PLEDGE policies list and the DELOS Reference Model was attempted during the first DL.org Working Group Meeting held in Tirrenia, Italy, July 1–2, 2009, (http://www.dlorg.eu/uploads/Working%20Groups/DL.org_Policy_WG_meeting_summary.pdf). The Policy Working Group has planned to move forward in its investigations and more activities also related to the PLEDGE Policies list.

A first study of policy interoperability areas resulted in the identification of the following exemplar area:

- access policy
- policy enforcement
- digital preservation and assessment policies.

For the access policies, we identified a first set of studies in the library and information, network, and medical literature (Arms 2000; Koulouris, Kapidakis, & Zhao 2003; Strassner 2003; Seto 2007; Marill & Luczak 2009; Nicholas et al. 2009). For the policy enforcement (Strassner 2003; Martin 1999) and digital preservation and assessment policies (Ross & Hedstrom 2005; Hitchcock et al. 2007; Wilson 2007; Beagrie et al. 2008; Dappert 2009; Jones 2009), we identified two projects (the MIT project PLEDGE, http://pledge.mit.edu/index.php/Main_Page and the EU-funded Sustaining Heritage Access through Multivalent Archiving-SHAMAN <http://shaman-ip.eu/shaman/>) which can provide useful guidance to the investigations of the DL.org Policy Working Group related to the interoperability requirements between Digital Libraries with respect to policies. These projects are

especially relevant for what concerns policy automation, rule-based systems and assessment frameworks (Innocenti et al., 2009). More details on the Policy Working Group research outputs and outcomes are being periodically released online at https://workinggroups.wiki.dlorg.eu/index.php/Main_Page.

Quality

Defining Quality Within Digital Libraries

Quality is “a broad term that encompasses notions of merit, worth, and significance. The word *quality* is used in ways that both describe and appraise” (Stake & Schwandt 2006, p. 405). According to the ISO standard 8402-1986, “quality is the totality of features and characteristics of a product or service that bears on its ability to satisfy stated or implied needs” (ISO 1986). This has since been refined by ISO 8402-1994 to “the totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs” (ISO 1994).

Quality “can be applied to products and processes, and is usually defined in relation to a set of guidelines or criteria” (Hofman et al. 1997, p. 33); it involves standards or best practices, for example, measures; thus, when quality is addressed then measures of quality have to be specified in addition to aspects, objectives and criteria. Only a small fraction of all the works on digital libraries are devoted to quality and a global approach is still lacking.

The relationships and the interdependencies between quality and interoperability can be extremely complex. Quality and interoperability affect each other and can be highly inter-related. Offering high quality services can require a high degree of interoperability among the different components of a system; similarly, poorly designed or low quality services can affect the degree of interoperability among different components that can be achieved, thus preventing the successful cooperation among different systems.

In the digital library field, we can have four types of approaches to quality interoperability:

- Content-based approach: quality interoperability of data, meta-data and collections
- Services-based approach: quality interoperability of digital libraries considered as organizations

- Technical approach: multilingual access quality interoperability, IR quality interoperability
- User-based approach: qualitative interoperability, user studies evaluation

The different approaches previously mentioned depend on the interdisciplinary nature of digital library field which involves heterogeneous research areas: LIS studies (mainly focused on services, organization, metadata), IR studies (search engines, metadata management), computer science studies (digital libraries as systems and architectures), and HCI studies (interfaces, user studies) (Su 1992; Shneiderman 1998).

A quality model for digital libraries was elaborated in 2007 within the 5S (Streams, Structures, Spaces, Scenarios, and Societies) theoretical framework (Gonçalves et al. 2007; 2004): the model, which was addressed to digital library managers, designers and system developers, defined a number of dimensions of quality proposing a set of numerical dimensions, which were illustrated with real case studies.

Within the DELOS Digital Library Reference Model, quality is described as one of the six core domains of the Digital Library Universe as follows:

The Quality concept represents the parameters that can be used to characterize and evaluate the content and behavior of a Digital Library. Quality can be associated not only with each class of content or functionality but also with specific information objects or services” (Candela et al. 2008, p. 20).

The Quality domain is very broad and dynamic by nature. The representation provided by this model is therefore extensible with respect to the myriad of specific quality facets each institution would like to model. Quality Parameter is actually a class of various types of quality facets, for example, those that currently represent common practice.

Quality Interoperability

Digital library theoretical frameworks should help researchers and professional to have a common reference to classify and

compare heterogeneous digital libraries, identifying differences and commonalities more easily. However, a digital library can be classified as being interoperable with another one at different levels (e.g., semantic level, service level, object level).

The progressive development of models for defining the different facets of digital libraries and for investigating the various factors which affect interoperability represents a key step to be able to deal with the complexity of interactions between different digital library entities—users, information resources, added-value services, policies, and so on—and the quality of a digital library, in a systematic and exhaustive way.

For the DL.org Quality Working Group, quality interoperability first means the possibility for digital libraries to share a common quality framework. As quality is still a low-prioritized aspect of digital libraries, the Quality Working Group is investigating both the research areas and the real-world cases in which quality issues are explored and examined.

The research investigation is taking into account the definitions of quality; for example, what and how to measure, the digital library theoretical models and multiple interoperability approaches. The Quality Working Group has been involved in the scientific literature analysis, which has helped to identify three research areas in which the quality issues have been most developed:

- Data quality
- Digital libraries evaluation
- Quality parameters.

Data Quality

As digital libraries' functions and activities revolve around “collections of digital works” (Waters 1998), or “rich digital content” (Candela et al. 2008), research on data quality has a pivotal role.

Data represent “real world objects, in a format that can be stored, retrieved, and elaborated by a software procedure, and communicate through a network” (Batini & Scannapieco 2006, p. 6). Data can be:

- structured, if each data element has fixed structure
- semi-structured, if the data has a flexible structure
- unstructured, when data are expressed in natural language and are not structured (Batini & Scannapieco 2006)

Dimensions and techniques for data quality “are progressively more complex to conceive and use from structured to unstructured data” (Batini & Scannapieco 2006).

Any data quality-related activity starts with the selection of the dimensions to measure the data quality level. Quality dimensions can be referred “either to the extension of data, i.e., to data values, or to their intension, i.e., to their schema” (Barone, Cabitza, & Grega 2007, p. 421).

The data quality literature provides a thorough classification of data quality dimensions; however, due to the contextual nature of quality, no general agreement exists on which set of dimensions defines the quality of data or on the exact meaning of each dimension (Batini et al. 2009).

An analysis of the six most important classifications of quality dimensions (Batini & Scannapieco 2006) indicated that the focus of the majority of authors is concentrated on the following basic set of data quality dimensions:

- accuracy;
- completeness;
- consistency; and
- timeliness

Within the digital library field, data quality dimensions correspond to those digital library concepts describing the digital library as a collection of digital objects, for example, data.

In the 5(S) Model-grounded quality model for digital libraries (Gonçalves et al. 2007), the digital library concepts are listed hierarchically (Digital objects, Metadata specification, Collection, Catalog, Repository, Services. Considering “Collection” (whose quality dimension is completeness) as a set of digital objects and metadata specifications, the core data elements within this model are:

- “Digital object”, which has the following quality dimensions: accessibility, pertinence, preservability, relevance, similarity, significance, timeliness; and
- “Metadata specification”, which has the following quality dimensions: accuracy, completeness, and conformance.

The DELOS Reference Model, instead, proposes a more generic classification of data quality dimensions within the Content Quality Parameter, without distinguishing the data/metadata/collection levels. Content is one of the six domains of the Digital Library Universe and “is composed of a set of Information Objects organised in Collections” (Candela et al. 2008 p. 19). “Metadata evaluation” is the only quality dimension that refers explicitly to metadata, whereas the other dimensions could be theoretically applied to any Digital Library Content forms of information.

The DELOS Reference Model Content Quality Parameter includes the following dimensions: integrity, authenticity, preservation performance, trustworthiness, perceivability, size, freshness, scope, viability, provenance, fidelity, and metadata evaluation.

Quality Within Digital Library Evaluation Studies

As fundamental to discerning quality is comparison, the concept of quality is strictly related to evaluation and measurement. A subtle distinction exists between quality and evaluation is the judgment-making or appraisal of quality of both procedures and outcomes (Stake & Schwandt 2006).

Judging quality is described within the logic of evaluation: “(1) establish criteria of merit, (2) construct standards, (3) measure performance and compare to standards, and (4) synthesize and integrate results into a judgment of merit, worth, or significance” (Scriven 1991).

Digital library evaluation constitutes the main research area in which digital library quality dimensions have been developed. In this context, digital libraries are considered as complex but measurable entities.

Every digital library evaluation model depends on a digital library global framework that can take very different approaches.

Several digital library aspects cannot be measured and assessed just by librarians or system managers: the evaluation of digital libraries requires interdisciplinary competences and need heterogeneous skills.

In 2000, Marchionini proposed the application of the same techniques and indicators used for traditional libraries such as circulation, creation, and growth of collections, users data, users satisfaction, and financial stability indicators (Marchionini 2000).

Reviewing the evaluation criteria identified by Lancaster (1993) and by Saracevic & Kantor (1997), Saracevic systematized the issue within a continuative approach, highlighting the need to focus on the digital library mission and objectives (Saracevic 2000).

As Saracevic (2000) has cogently argued, evaluation must involve selections and decisions related to the:

- construct (what to evaluate);
- context (which level, which objectives);
- criteria reflecting performance as related to selected objectives (e.g., performance aspects on which to focus, features to assess); and
- methodology (which instruments to use, what samples and procedures).

He indicated also two evaluation levels:

- user-centered level, which can be social, institutional, individual or focused on the interface; and
- system-centered level, which can be focused on the engineering, on the processing or on content (Saracevic 2000).

Reeves, Apedoe, & Woo (2005) proposed some guidelines to evaluate DLs, focusing on the decision process that is behind any evaluation. Chowdhury and Chowdhury (2003) highlighted the need to focus on the global impact the DL has on its users and on the society in general, integrating LIS, IR, and HCI criteria.

Through the analysis of eighty digital library case studies Saracevic (2004) outlined what approaches and methodologies

are concretely used, observing the small quantity of “real data” compared to the explosion of meta-literature. Saracevic concluded that there’s no “best” methodology: different aims can lead up to different methods.

The first important step for a collaborative reference model has been done by the D-Lib Working Group on Digital Library Metrics (1998–2002) <http://www.dlib.org/metrics/public/>, whose work has been continued by DELOS within WP7 “Evaluation” http://www.delos.info/index.php?Itemid=52&id=26&option=com_content&task=view. The aim of the MWG (Metrics Working Group) was to allow comparisons between large-scale DLs, through the specification of the key functions of a DL, which were considered as an IR environment.

The development of an evaluation model was carried forward by DELOS. Its evaluation schema initially had these dimensions:

- data/collection,
- system/technology, and
- users/uses (Fuhr et al. 2001)

Subsequently (Fuhr et al. 2007), the schema was integrated to Saracevic’s evaluation questions (Saracevic 2004).

The DL.org Quality Working Group has started its research on digital library quality agreeing on the need to avoid the fragmentation of quality models and considering quality as the theoretical basis for any DL evaluation framework.

Quality Parameters: Towards a Quality Core Model

The Quality Working Group is currently working on a quality parameter pattern, whose structure and description will be tested and enhanced along the project duration, which is thought to be the most characteristic for DLs and shall help to identify best practices. This simplified pattern should help DLs to interoperate in the quality domain.

This pattern includes Policy Quality Parameters, Content Quality Parameters, and Generic Quality Parameters.

A first evaluation of existing approaches and best practices in relation to quality interoperability within digital libraries raises the identification of key-issues and challenges, which include:

- the investigation on the Quality Parameters definitions and relationships within the DELOS Reference Model;
- the production of examples and user scenarios to move the Quality domain from theory to practice, in order to involve the professional community to the enhancement of the model;
- the feasibility test of the Quality Core Model as a valid interoperability framework; and
- the elaboration of best practices and recommendations.

The Quality Working Group is currently working on the Quality Core Model refinement and on its feasibility test, which will imply the collection of feedbacks from the digital libraries community.

Towards a Digital Library Policy and Quality Interoperability Framework

In order to achieve interoperability between digital libraries, common policy and quality frameworks are needed. An organizational approach to policy and quality interoperability requires the identification of the core policy and quality properties affecting digital libraries, in order to set up a shared integrated framework. Selection, organization, and dissemination policies determine, in large part, a digital libraries' scope and effectiveness, which need to be measured according to selected quality criteria. Indeed, the spread of quality indicators to evaluate information systems needs an upper framework to foster cooperation and exchange of quality data. Considering the digital libraries' context, policy and quality interoperability can be faced with an organizational approach.

Findings

This first evaluation of existing approaches and best practices, in relation to policy and quality interoperability within digital libraries and fruitful discussions within the DL.org Policy and

TABLE 1 Findings on Policy and Quality Interoperability Within the DL.Org Policy and Quality Working Groups

	Common findings	Policy WG findings	Quality WG findings
Broadness of the DELOS Digital Library Reference Model and need of an organizational approach	X		
Definition of policy interoperability as “business level interoperability”		X	
Repositioning within the DELOS Reference Model		X	
Development of a Quality Core Model			X
Lack of formalization within digital libraries	X		
Interdisciplinary approach	X		
Production of user scenarios and case studies	X		
Feasibility tests needed	X		

Quality Working Groups (WGs) members, allowed the identification of the following key issues, challenges, and preliminary findings as summarized in Table 1 and detailed in the following paragraphs:

BROADNESS OF THE DELOS DIGITAL LIBRARY REFERENCE MODEL AND NEED OF AN ORGANIZATIONAL APPROACH

The Policy Working Group agreed that the policy domain is broader than how it is currently represented in the Reference Model. The current Model focused heavily on system architecture and does not clearly address the issue of the context of the digital library systems. Therefore, the Policy WG will place more focus on the policy organizational context. This issue will also be addressed by the overall project participants, in order to clarify the position of the Reference Model. Considering that there is an organization beyond a Digital Library that defines the policy of the overall system in which a Digital Library is operating, one of the main recommendations of the Quality Working Group is to take into account the “organization” that is wrapping the existing levels of the Digital Library, the Digital Library System, and the Digital Library Management System. The underlying rationale of

this extension is that the concept of the Digital Library might not be sufficient to address all interoperability issues that are under investigation in DL.org.

REPOSITIONING OF POLICY WITHIN THE DELOS REFERENCE MODEL

During these investigations, we produced a proposal on the repositioning of Policy in the overall Reference Model, as a metalayer above and beyond the digital library, and also an operational layer in the digital library. The rationale behind our proposal is that digital libraries represent the confluence of vision, mandate, and the imagined possibility of content and services constructed around the opportunity of use. Underpinning every digital library is a policy framework. It is the policy framework that makes the digital library viable—without a policy framework a digital library is little more than a container for content—for even the mechanisms of structuring the content within a traditional library building as container (e.g., deciding what will be on what shelves where) are based upon policy. So, policy governs how a digital library is instantiated and run; a library without policy, therefore, is similar to a Ferrari in a world without roads and populated only by blind drivers. The policy domain is, therefore, a metadomain that is situated both outside the digital library and any technologies used to deliver it and is within the digital library. That is, policy exists as an intellectual construct that is deployed to frame the construction of the digital library and its external relationships, and then these, and other more operational policies, are represented in the functional elements of the digital library. Therefore, policy permeates the digital library from conceptualization through operation and needs to be represented as such in the model at these various levels.

DEVELOPMENT OF A QUALITY CORE MODEL

In order to broaden the applicability of the Quality framework within the DELOS Digital Library Reference Model, the Quality Working Group identified some selected aspects to define and develop objectives and criteria for their evaluation. The final simplified pattern will help the constitution of a shared vocabulary and the identification of best practices towards Digital Libraries quality interoperability. The Quality Working Group is currently working on the refinement of the Quality Core Model and on

testing its effectiveness, establishing connections with real-world digital libraries and planning a survey/interview on quality interoperability issues in order to collect feedback and best practices from the professional community.

LACK OF FORMALIZATION WITHIN DIGITAL LIBRARIES

The PLEDGE project provided evidence that very few current digital libraries have formal policies in place. They do when there are business concerns (e.g., the commercial digital libraries) and they do, for example, for access control, but for many types of policies there is very little written down and none of it is machine-readable. In general, it seems that it is too early to expect formally-encoded DL policies in actual DLs and there are no formal standard policy languages for the Web, as yet; although, there are ongoing efforts to map research languages like AIR (Kagal 2009) to the new W3C recommendation standards for the Rules Interchange Format http://www.w3.org/2005/rules/wiki/RIF_Working_Group. While some policies are published within digital libraries, it is even more difficult to access to formalized digital library quality frameworks, which normally focus only on specific facets of the library (such as data or standards); quality changes over time and is still a low priority issue within digital libraries.

INTERDISCIPLINARY APPROACH

For the policy domain this is an almost uncharted territory, with interdisciplinary studies on policies taking place outside the traditional digital library's domains and disciplines, such as computer science theory, digital content management, data management, e-science, risk assessment, digital repository certification, health care and medical sector, and Open Access Initiative. The following areas were identified as particularly relevant for the goals of DL.org: policy classification (focusing on the Policy by scope section of the DELOS Reference Model); manual vs. automated policies (and in particular how to encode those policies for machine discovery, and which languages can be used to represent policies and make them functional, with particular attention to semantic web technologies); policy management (in particular how policies are appraised and enforced); policies evolution over time; and interconnectedness of policy and quality.

Relevant studies on Digital Libraries' quality are taking place within LIS disciplines, computer science, human-computer interaction, data management, digital repositories assessment, and research on open access; in order to investigate and indentify interoperability patterns, an interdisciplinary approach is needed.

PRODUCTION OF USER SCENARIOS AND CASE STUDIES

Policy user scenarios are being produced to support investigations in this field and the collection and definition of best practices for developing a policy interoperability framework in the digital library domain.

Through the elaboration of the Quality Core Model, the Quality Working Group experts have been involved in the analysis of the selected quality core parameters: their definitions have been examined and revised; some key-questions have been added; and more examples and user scenarios have been provided in view of the most urgent quality interoperability challenges.

FEASIBILITY TESTS NEEDED

The Policy Working Group team is currently engaged in surveying a selected representative sample of cross-domain international large and medium scale digital libraries whose automated and manual policies will be analyzed.

In parallel, the Quality Working Group will be involved in the identification of formalized quality frameworks that are currently used by digital libraries as a result of collaborative efforts; these examples will be analyzed according to the DELOS RM quality taxonomy and the Quality Core Model.

Next Steps

With the goal of defining a Policy Interoperability Framework, the Policy Working Group is focused on identifying a selected set of policies, collecting more user scenarios, providing example encodings, and reviewing the standards development process in the area of interoperable policies. These tasks include scenarios on Open Access policies, data harmonization, funding bodies, and policy comparability. This examination also examines the expectations of consumers, whether human or machine(s). A crucial task

is the selection of international digital library/data centers to analyze policy statements in terms of interoperability and formalization. The development of these scenarios and the investigation within real-life digital libraries policy frameworks will support the grouping of policy statements and scenarios; provide the basis for further enhancement of the Policy domain within the DELOS DL Reference Model; allow the mapping between the PLEDGE policies and the enhanced attributes of Policy in the DELOS DL Reference Model and the connection with the SHAMAN Assessment Framework; and provide the basis for policy interoperability guidelines for the DL.org Cookbook.

The next steps of the Quality Working Group activities will include the identification and integration digital library quality frameworks to underpin the creation of the Quality Core Model. In parallel, the selection of a small number of ongoing large digital libraries that are using formalized quality models is already in progress in order to analyze and compare those “real-world” frameworks according to the Quality Core Model and the DELOS RM Quality concept map. The whole process should allow mobilization of the digital library community, raising the interoperability issue and testing the Quality Core Model feasibility.

References

- Agnew, G. *Digital rights management: A librarian's guide to technology and practise*. Oxford: Chandos, 2008.
- Andrews, J., and D. Law. Eds. *Digital libraries. Policies planning and practices*. Ashgate: Aldershot, 2004.
- Arms, W. Y. *Digital libraries*. Cambridge: MIT Press, 2000.
- Barone, D., F. Cabitza, and S. Gega. “HDQ: A meta-model for the quality improvement of heterogeneous data.” *ICDIM 2007*: 418–423.
- Batini, C., and M. Scannapieco. *Data quality: Concepts, Methodologies and Techniques*. Berlin: Springer, 2006.
- Batini, C., et al. “Methodologies for Data Quality Assessment and Improvement.” *ACM Computing Surveys* 41.3 Article 16 (2009).
- Beagrie, N., et al. *Digital preservation policy study*. Part 1: Final Report October 2008, JISC. Web. Feb 2010. <http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf>
- Borgman, C. L. “What are digital libraries? Competing visions.” *Information Processing and Management* 35.3 (1999): 227–243. Print.

- Borgman, C. L. *From Gutenberg to the global information infrastructure: access to information in the networked world*, 2000. Cambridge, MA: MIT Press. Print.
- Borgman, C. L., J. C. Wallis, and N. Enyedy. "Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries." *International Journal on Digital Libraries* 7.1-2 (2007): 17–30.
- Candela, L., et al. *The DELOS Digital Library Reference Model - Foundations for Digital Libraries, Version 0.98, DELOS Network of Excellence - Project no. 507618*, 2008. Web. Feb 2010. <http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf>
- Chowdhury, G. G., and S. Chowdhury. *Introduction to digital libraries*. London: Facet, 2003.
- D-Lib Working Group on Digital Library Metrics (1998–2002)*. <<http://www.dlib.org/metrics/public/>>
- Dalton, P., R. Hartland-Fox, and S. Thebridge. "Evaluating electronic information services." *Digital libraries. Policies Planning and practices*. Eds. J. Andrews, and D. Law. Ashgate: Aldershot, 2004.
- Dappert, A. Report on policy and strategy models for libraries, archives and data centres, PP2-D3, Planets External Deliverable, 2007. Web. Feb 2010. <http://www.planets-project.eu/docs/reports/Planets_PP2_D3_ReportOnPolicyAndStrategyModelsM36_Ext.pdf>
- DELOS within WP7 "Evaluation."* <http://www.delos.info/index.php?Itemid=52&id=26&option=com_content&task=view>
- DL.org Policy Working Group*. <https://workinggroups.wiki.dlorg.eu/index.php/Main_Page>
- Fox, E. A., et al. "Digital Libraries." *Commun ACM* 38.4 (1995): 23–28.
- Fox, E. A., L. T. Watson, and N. A. Kipp. "Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries." *ACM Trans Inf Syst* 22 (2004): 270–312.
- Fuhr, N. et al. *Digital libraries: A generic classification and evaluation scheme*. Proceedings of the 5th European conference on Digital Libraries 2001. Berlin: Springer, 187–199.
- Fuhr, N., et al. "Evaluation of digital libraries." *International Journal on Digital Libraries* 8.1 (2007): 21–38.
- Gonçalves, M.A. et al. "Streams, Structures, Spaces, Scenarios, Societies (5s): A formal model for digital libraries." *ACM Trans. Inf. Syst.* 22.2 (2004): 270–312. Web. Feb 2010. <<http://portal.acm.org/citation.cfm?id=984325>>
- Gonçalves, M. A. et al. "A quality model for digital libraries." *Information Processing and Management* 43(2007):1416–1437.
- Green, A., S. Macdonald, and R. Rice. *policy-making for research data in repositories: a guide. DISK-UK Data Share Project, 2009*. Web. Feb 2010. <<http://www.disc-uk.org/docs/guide.pdf>>
- Hofman, P., et al. *Selection Criteria for Quality Controlled Information Gateways. Version 1.1, DESIRE – Project no. RE 1004*. 1997. Web. Feb 2010. <<http://www.ukoln.ac.uk/metadata/desire/quality/>>
- Hitchcock, S., et al. *Survey of repository preservation policy and activity*. Preserv project, DRAFT paper, 9 January 2007, updated 21 February 2007. Web. Feb 2010. <<http://preserv.eprints.org/papers/survey/survey-results.html>>

- Innocenti, P. et al. *Assessing Digital Preservation Frameworks: the approach of the SHAMAN project*. Proceedings of the International Conference on Management of Emergent Digital EcoSystems, France 2009, ISBN 978-1-60558-829-2, 2009. Web. Feb 2010. <<http://portal.acm.org/citation.cfm?id=1643823.1643899&coll=GUIDE&dl=GUIDE&CFID=63081623&CFTOKEN=61810568>>
- Institute of Electrical and Electronics Engineers (IEEE). <http://www.policy-workshop.org/>
- Ioannidis, Y., Ed. *Digital libraries: Future directions for a european research programme*. DELOS Brainstorming Report. San Cassiano, Italy, 2001. Web. Feb 2010. <delos-noe.iei.pi.cnr.it/activities/researchforum/. . ./brainstorming-report.pdf>
- Ioannidis, Y. "Digital libraries at a crossroads." *International Journal on Digital Libraries* 5.4 (2005): 255–265.
- Ioannidis, Y., et al. "Digital library information-technology infrastructures." *International Journal on Digital Libraries* 5.4 (2005): 266–274.
- ISO 8402:1986 Quality – Vocabulary. 1986.
- ISO 8402:1994. Quality management and quality assurance – Vocabulary. 1994.
- Jones, S. *A report on the range of policies required for and related to digital curation*. Digital Curation Centre, University of Glasgow. 2009. Web. Feb 2010. <http://www.dcc.ac.uk/docs/reports/DCC_Curation_Policies_Report.pdf>
- Kagal, L. *AIR Policy Language*, 2009. Web. Feb 2010. <<http://dig.csail.mit.edu/2009/AIR/>>
- Kelly, B. "Interoperable Digital Library Programmes? We must have a QA!" Eds. R. Heery, and L.Lyon, *ECDL 2004, LCNS 3232* (2004): 80–85.
- Koulouris, A., S. Kapidakis, and Y. Zhao. "Collections and access policies of the digital material of ten national libraries." *Proceedings of the 3rd ACM/IEEE Joint Conference on Digital Libraries (JCDL 2003)*. New York: ACM Press, 2003. 409.
- Lagoze, C. "Lost Identity: The Assimilation of Digital Libraries into the Web." PhD thesis, Information Science, Cornell University, 2010. Web. Feb 2010. <<http://www.cs.cornell.edu/lagoze/dissertation/dissertation.html>>
- Lancaster, F. W. *If you want to evaluate your library*, 2nd edn. San Francisco: Morgan Kaufman. 1993.
- Lesk, M. *Understanding Digital Libraries*, 2nd ed. San Francisco, CA: Morgan Kaufmann, 2005, 321–323.
- Lutfiyya, H., F. Garcia, and J. Moffett. "Policy 2003: Workshop on policies for distributed systems and networks." *Journal of Network and Systems Management* 11.3 (2003): 373–376.
- Marchionini, G. "Evaluating digital libraries: a longitudinal and multifaceted view." *Library Trends* 49.2 (2000): 304–333.
- Marill, J. L., and E. C. Luczak. "Evaluation of Digital Repository Software at the National Library of Medicine." *D-Lib Magazine* 15 (2009): 5–6.
- Martin, J. C. *Policy-based networks: Sun BluePrints™ OnLine*. Palo Alto: Sun Microsystems, 1999.
- Nicholas, N. N. Ward, and K. Blinco. "A Policy Checklist for Enabling Persistence of Identifiers." *D-Lib Magazine* 15.1/2. 2009. Web Feb 2010. <<http://www.dlib.org/dlib/january09/nicholas/01nicholas.html>>

- PLEDGE. <http://pledge.mit.edu/index.php/Main_Page>
- Reeves, T. C., X. Apedoe, and Y. H. Woo. *Evaluating digital libraries: a user-friendly guide*. The National Science Digital Library. 2005. Web. Feb 2010. <<http://www.dpc.ucar.edu/projects/evalbook/EvaluatingDigitalLibraries.pdf>>
- Ross, S. *Digital library development review*. Final report. National Library of New Zealand, 2003. Web. Feb 2010. <http://eprints.erpanet.org/50/01/ross_report.pdf>
- Ross, S. *Preservation of interoperability and interoperability of preservation*. Third Workshop on Foundations of Digital Libraries, 2008. ECDL2008, Aarhus. Web. Feb 2010. <http://www.delos.info/files/pdf/DLFoundations2008/10_RossDLFoundations08.pdf>
- Ross, S, and M. Hedstrom . "Preservation research and sustainable digital libraries." *International Journal on Digital Libraries* 5.4 (2005): 317–324.
- Rubin, R. E. *Foundations of library and information science*. New York: Neuman-Schuman Publishers, 2000.
- Saracevic, T. "Digital library evaluation: toward an evolution of concepts." *Library Trends* 49.3 (2000): 350–369.
- Saracevic, T. Evaluation of digital libraries: an overview. In: DELOS Workshop on the evaluation of digital libraries. 2004. Web Feb 2010. <http://www.dlib.ionio.gr/wp7/ws2004_Saracevic.pdf>
- Saracevic, T., and P. B. Kantor. "Studying the value of library and information services." *Journal of American Society for Information Science* 48 (1997): 543–563.
- Scriven, M. *Evaluation thesaurus*, 4th edn. Newbury Park: Sage. 1991.
- Seto, B. *Implementing the NIH Data Sharing Policy: Expectations and Challenges*, National Institute of Health. Presentation in Berlin 5 Open Access: From Practice to Impact: Consequences of Knowledge Dissemination. 2007. Web. Feb 2010. <<http://www.aepic.it/conf/viewpaper.php?id=268&cf=10>>
- Shneiderman, B. *Designing the user interface strategies for effective human-computer interaction*, 3rd edn. Reading: Addison Wesley. 1998.
- Smith, M., and R. Moore. "Digital Archive Policies and Trusted Digital Repositories." *International Journal of Digital Curation* 2.1: 92–101.
- Stake, R. E., and T. A. Schwandt. "On discerning quality in evaluation." Eds. I. F. Shaw, J. C. Greene, and M. M. Mark, *The Sage handbook of evaluation*. London: Sage. 2006. 404–418.
- Strassner, J. S. *Policy-based network management: solutions for the next generation*. San Francisco: Morgan Kaufmann, 2003.
- Su, L. T. "Evaluation measures for interactive information retrieval." *Information Processing and Management* 28 (1992): 503–516.
- Sustaining Heritage Access through Multivalent Archiving-SHAMAN*. <<http://shaman-ip.eu/shaman/>>
- Waters, D. J. "What are digital libraries?" *CLIR issues*, 4. Web. Feb 2010. <<http://www.clir.org/pubs/issues/issues04.html#dlf>>
- Wilson, A. InSPECT: Significant Properties Report, Second version. 2007. London: ADHS. Web. Feb 2010. <http://www.significantproperties.org.uk/documents/wp22_significant_properties.pdf>

World Wide Web Consortium (W3C). *Web Services Policy Working Group*. <<http://www.w3.org/2002/ws/policy/>> and <http://www.w3.org/TR/2007/REC-ws-policy-20070904/>

World Wide Web Consortium (W3C). *Recommendation standards for the rules interchange format*. <http://www.w3.org/2005/rules/wiki/RIF_Working_Group>

World Wide Web Consortium (W3C). *The Web Services Policy 1.5 – Framework*, 2007. Web. Feb 2010. <<http://www.w3.org/TR/2007/REC-ws-policy-20070904/>>