Future Generation Computer Systems 55 (2016) 129–146

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Digital library interoperability at high level of abstraction

Maristella Agosti, Nicola Ferro *, Gianmaria Silvello

Department of Information Engineering, University of Padua, Via Gradenigo 6/a, 35131 Padua, Italy

HIGHLIGHTS

- Foundational models for digital libraries.
- Interoperability among digital library systems and possible approaches.
- Unifying semantic model and ontology for high-level interoperability among digital library systems.
- In depth analysis of the user, content, functionality and quality domains in digital libraries.
- Concrete use case on annotation of illuminated manuscripts to illustrate how to apply the proposed methodology.

ARTICLE INFO

Article history:
Received 14 July 2015
Received in revised form
4 September 2015
Accepted 16 September 2015
Available online 3 October 2015

Keywords:
Digital library
Foundational models
5S model
DELOS Reference Model
Interoperability
Ontology

ABSTRACT

Digital Library (DL) are the main conduits for accessing our cultural heritage and they have to address the requirements and needs of very diverse memory institutions, namely *Libraries, Archives and Museums (LAM)*. Therefore, the interoperability among the *Digital Library System (DLS)* which manage the digital resources of these institutions is a key concern in the field.

DLS are rooted in two foundational models of what a digital library is and how it should work, namely the *DELOS Reference Model* and the *Streams, Structures, Spaces, Scenarios, Societies (5S)* model. Unfortunately these two models are not exploited enough to improve interoperability among systems.

To this end, we express these foundational models by means of ontologies which exploit the methods and technologies of Semantic Web and Linked Data. Moreover, we link the proposed ontologies for the foundational models to those currently used for publishing cultural heritage data in order to maximize interoperability.

We design an ontology which allows us to model and map the high level concepts of both the 5S model and the DELOS Reference Model. We provide detailed ontologies for all the domains of such models, namely the user, content, functionality, quality, policy and architectural component domains in order to make available a working tool for making DLS interoperate together at a high level of abstraction. Finally, we provide a concrete use case about digital annotation of illuminated manuscripts to show how to apply the proposed ontologies and illustrate the achieved interoperability between the 5S and DELOS Reference models.

1. Introduction

Digital Library (DL) have been steadily progressing since the early 1990s and they now determine how citizens and organizations study, learn, access and interact with their cultural heritage [1–8]. Despite their name, DL are not only the digital counter-part of traditional libraries but they are also concerned with other kinds of cultural heritage institutions, such as archives and museums, that is institutions typically referred to as *Libraries, Archives and Museums (LAM)*. In the context of LAM, unifying a variety of organizational settings and providing more integrated access to their contents are aspects of utmost importance. Although the type of materials may differ and professional practices vary, LAM share an overlapping set of functions and fulfilling them in "collaboration rather than isolation creates a win–win for users and institutions" [9].

These compelling integration and collaboration needs have propelled the evolution of *Digital Library System (DLS)* [10] as systems that permit us to design and implement the overlapping set of functions of LAM.

In the 1990s, DLS were monolithic systems, each one built for a specific kind of information resource – e.g. text, images, or

* Corresponding author.
E-mail addresses: maristella.agosti@unipd.it (M. Agosti), nicola.ferro@unipd.it (N. Ferro), gianmaria.silvello@unipd.it (G. Silvello).

http://dx.doi.org/10.1016/j.future.2015.09.020
0167-739X/

videos – and with very specialized functionalities developed ad hoc for those contents and their reference users. This approach caused a flourishing of systems where the very same functionalities, e.g. user management or repositories, were developed and redeveloped from scratch many times, causing them to be different and often incompatible one with each other. From the mid 2000s, DLS evolved towards service-oriented architectures, where components can be plugged into each other to provide the desired end-user functionalities, yet requiring careful and ad hoc configuration. This paradigm shift allowed DLS to become more and more user-centered systems, where the original content management functionality was partnered with new communication and cooperation functionalities such as user annotation [11], with the ultimate goal of acting as "a common vehicle by which everyone can access, discuss, evaluate and enhance information of all forms" [12]. As a consequence, DLS started to embody the above vision for LAM since they were no longer isolated systems but, on the contrary, they needed to cooperate together in order to improve the user experience in accessing information and to seamlessly integrate information resources of different cultural heritage institutions.

This evolution has been favored by the development of two foundational models of what DL are, namely the *Streams, Structures, Spaces, Scenarios, Societies (5S)* model [13] and the *DELOS Reference Model* [14], which made it clear what kind of entities should be involved in a DL, what their functionalities should be and how DLS components should behave, and fostered the design and development of operational DLS complying with them.

However, these two models are quite abstract and, still providing a unifying vision of what a DL is, they allow for very different choices when it comes to develop actual DLS. This has led to the growth of "ecosystems" where services and components may be able, at best, to interoperate together within the boundaries of DLS that have been inspired by just one of the two models for DL. However, there are no running examples of two DLS, one implementing the 5S model and the other the DELOS Reference Model, which are able to interoperate. Therefore, interoperability still represents one of the biggest challenges in the DL field [10,15].

In this work, we address a still open issue in the DL realm: to make DL foundational models interoperable in order to derive all the other kinds of interoperability, in particular, interoperability between operational DLS. The main contributions of the paper are:

- a detailed analysis of the 5S and DELOS Reference models pointing out common aspects and main differences;
- the definition of a common ontology which encompasses and links the concepts of the DELOS Reference Model and the 5S models, covering all the domains of such models: the user, content, functionality, quality, policy and architectural component;
- a concrete use case about digital annotation of illuminated manuscripts to show how to apply the proposed ontologies and illustrate the achieved interoperability between the 5S and DELOS Reference models.

The paper is organized as follows: Section 2 illustrates the rationale of the paper; Section 3 reports on some pertinent related works; Section 4 introduces the relevant aspects of the 5S Model and of the DELOS Reference Model; Section 5 presents the semantic mapping between the 5S Model and the DELOS Reference Model associating the high level concepts of one model to those of the other; Section 6 to 11 respectively show the correspondences between the notions and domains of user, content, functionality, quality, policy and architectural component in the two models; Section 12 presents a relevant case of application of the proposed approach where users who interact with two DLS, that manage illuminated manuscripts, are interested in annotating their contents to perform activities of their interest. Section 13 sums up the results presented in the paper.

2. Rationale

The current mainstream approach to bridge the interoperability gap between DLS and to provide comprehensive solutions able to embrace the full spectrum of LAM is to exploit semantic Web technologies and linked (open) data [16,17]. This allows for describing entities and information resources in a common way which enables their exchange, as for example happens in the case of library linked data [18].

This approach is both "external" and "bottom-up". It is "external" since it assumes that everything in a DL should be exposed on the Web rather than seeking direct interoperability among systems which may not necessarily be only Web-based. It is "bottom-up" because ontologies have been used only to describe the resources managed by a DLS and they are not used to represent the concepts themselves which constitute the DL model on which the DLS is based. Therefore, they allow for semantic interoperability and integration only at the data level, i.e. the lowest level possible in the architecture of a DLS. Indeed, at present different operational DLS, even if they are based on the same DL model, may not fully interoperate because the mapping between the abstract foundational model, either the DELOS or the 5S one, and the actual one implemented by the DLS is not explicitly provided. As a consequence, each operational DLS may not actually "know" the way in which another operational DLS calls and refers to the same operations.

Consider Fig. 1(a) which depicts the current approach. If we need two operations in two different DLS to interoperate, we need to create some kind of link between them at the business logic level. This link is typically manual and hard-coded, e.g. by direct invocation of the respective functionalities, since each DLS has no knowledge or understanding of each others internals. Each operation in each DLS makes use and processes some kind of data at the data logic level. Usually, these data need to be mapped to some common format, through some external ontology, to be exchanged between the two systems, due to lack of reciprocal knowledge. Because of that, we consider this as a "bottom-up" approach since the data level, the one where interoperability is achieved, is the lowest level in the architecture of a DLS.

As a concrete example of what is depicted in Fig. 1(a), suppose that two different operational DLS, one built using the 5S model and the other built using the DELOS Reference Model, need to enrich a user profile by exploiting their own specific service. To this end, they both need to exchange user data and to access their specific enrichment functionalities. Since a common view of user is lacking among the two DLS, we could use the class `Agent` in *Friend of a Friend (FOAF)*[1] to represent the notion of user of a DL in order to allow the two systems to exchange the profiles. Nevertheless, the `Agent` class is neither related to the concept of *Society* nor to the concept of *Actor* which represent users in the 5S model and in the DELOS Reference Model, respectively. Therefore, to exchange the user profiles, the two different operational DLS would need a set of (hard-coded) rules instructing them how to translate the `Society` ψ and the `Actor` `W` into the `Agent` class. Moreover, the `Agent` class is neither related to the notion of *Scenario* in the 5S model nor to the notion of *Function* in the DELOS Reference Model, whose specializations define the services for profile enrichment. Moreover, there is no ontology which tells the operational DLS based on the 5S model that its own `Scenario` δ for profile enrichment corresponds to a specific `Function` `D` for profile enrichment in the DELOS Reference Model. Therefore, to enrich user profiles with reciprocal information, the two different DLS would need to somehow manually (hard) code

[1] http://www.foaf-project.org/.