An Ontology to Make the DELOS Reference Model and the 5S Model Interoperable

Maristella Agosti, Nicola Ferro(✉), and Gianmaria Silvello

Department of Information Engineering, University of Padua, Padua, Italy
{agosti,ferro,silvello}@dei.unipd.it

Abstract. This paper is an extended abstract of the paper published in the Future Generation Computer System journal [2] which takes into account two of the foundational models defining what a digital library is and how it should work: the *DELOS Reference Model* and the *Streams, Structures, Spaces, Scenarios, Societies (5S)* model.

The aim of this work is to enable a high-level interoperability at the model level. 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.

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 [3,12,17]. 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.

These compelling integration and collaboration needs have propelled the evolution of *Digital Library System (DLS)* [1] as systems that permit us to design and implement the overlapping set of functions of LAM. This evolution has been favored by the development of two foundational models of what DL are: the *Streams, Structures, Spaces, Scenarios, Societies (5S)* model [9] and the *DELOS Reference Model* [4]. They 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

M. Agosti et al. (Eds.): IRCDL 2016, CCIS 701, pp. 85–91, 2017.
DOI: 10.1007/978-3-319-56300-8_8

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 [1].

In [2], we addressed the open issue of making DL foundational models interoperable and in this work we outline the main results achieved by describing at an high-level the ontology making the DELOS Reference and the 5S Model interoperable.

The paper is organized as follows: Sect. 2 introduces the rationale behind this work; Sect. 3 outlines the main differences between the DELOS Reference Model and the 5S Model; Sect. 4 presents the general ontology allowing for the DELOS Reference Model and the 5S Model to interoperate; and Sect. 5 draws some final remarks.

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 [11]. 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.

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.

What is needed is a deeper and more abstract interoperability based on a commonly shared semantic view of what a DL is rather than a lower level one where data is just wrapped in a commonly understandable format. The quite ambitious goal of this paper is to propose a solution to this open problem. The proposed solution is based on the representation of each foundational DL model through ontologies, leveraging semantic Web and linked data technologies in order to ease their linking to other already existing ontologies and to achieve maximum interoperability.

Therefore, the proposed approach will pave the way for a deeper interoperability among operational DLS and lower the barriers between LAM. It is also opening up more advanced possibilities for the automatic processing of resources, since, for example, DLS could automatically exploit the link between the models they are built upon in order to exchange resources, interoperate and integrate functionalities. To the best of our knowledge, there is no previous work in the field which attempted to achieve interoperability among DLS at a high level of abstraction through a semantic description and mapping of their foundational models. We can only mention our very preliminary work [7], where we started to explore this idea in the context of quality in DL.

3 Background

The 5S [8,9] is a formal model which draws upon the broad digital library literature to produce a comprehensive base of support. It was developed largely bottom up, starting with key definitions and elucidation of digital library concepts from a minimalist approach.

The DELOS Reference Model [4] is a high-level conceptual framework that aims at capturing significant entities and their relationships within the digital library universe with the goal of developing a more robust model of it.

The DELOS Reference Model and the 5S model address a similar problem with different approaches. The 5S is a formal model providing mathematical definitions of the digital library entities that can be used to prove properties, theorems and propositions. The DELOS Reference Model does not provide formal definitions, but it does provide a way to model and manage the resources of the digital library realm by using concept maps [13] because of their simplicity and immediacy.

4 Semantic Mapping Between DL Models

In Fig. 1 we present the *Resource Description Framework (RDF)* graph of the unifying data model relating the DELOS Reference Model to the 5S model by means of a mapping between their most relevant high-level concepts. The presented RDF graph is composed by classes represented as circles and properties represented as directed edges between the classes.

The main constituents of the DELOS Reference Model are: the digital universe divided into DL, DLS and *Digital Library Management System (DLMS)*, the concept of `Resource` and six high-level main domains: User, Functionality, Content, Quality, Policy and Architecture.

A DL, represented by the `Digital Library` class in Fig. 1, is supported by a `Digital Library System` which is extended and deployed by a `Digital Library Management System`. These three classes manage `Resource`s, where a `Resource` is any identifiable entity in the DL universe and resembles the concept of resource used in the Web [16]. A `Resource` represents the class of everything that exists in the DL universe and it is related to the `rdfs:Resource` class. In addition to this general concept, the `Resource` in the DELOS Reference Model has some additional features: it can be arranged or set out according to a resource format which, for example, allows a `Resource` to be composed of or linked to other `Resource`s.

All the DELOS domains are represented by key classes which are subclasses of `Resource`.

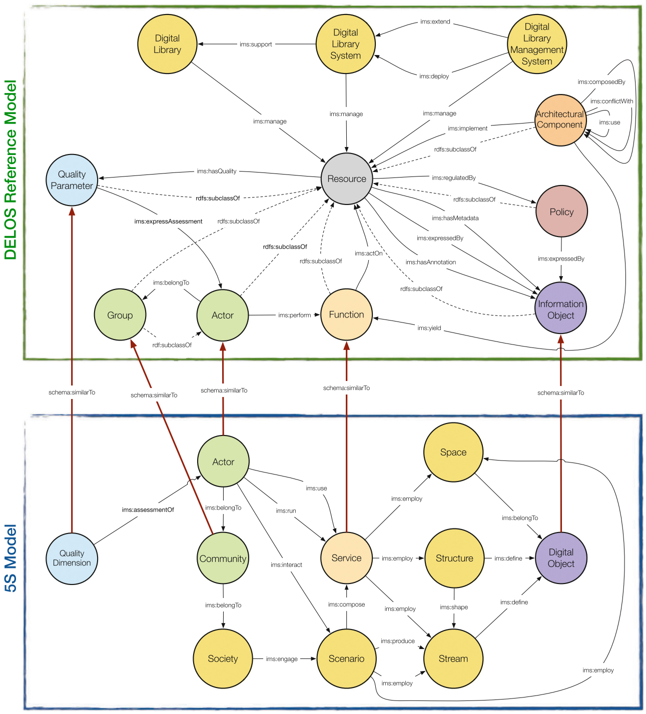

Fig. 1. Semantic mapping of the high level concepts in the 5S model and DELOS Reference Model and their relationships.

In the user domain these key classes are `Actor` and `Group`, where a `Group` may be a collection of actors and at the same time it is a subclass of `Actor` given that a group can be considered as a single actor itself. The user domain directly interacts with the functionality and the quality domains; indeed, an `Actor` may perform some `Function`s in the DL and may be assessed by some `Quality Parameter`.

The `Function` class always acts on a resource and the `Quality Parameter` is associated to a resource as happens in the case of `Actor`, which is related to `Quality Parameter`s by the `expressAssessment` property.

The main class of the content domain is `Information Object`, which represents any information managed by the DL and includes documents such as texts, images, videos, audio files, metadata and annotations.

The main class of the policy domain is `Policy`, which represents the single entity governing a resource with respect to a certain management point and it is connected to `Resource` by the `regulatedBy` property.

Lastly, the main class of the architecture domain is `Architectural Component`, which defines the organization or structure of the components of a given system or service; an architectural component may be `composedOf` other components or may `use` other components.

The RDF model of the 5S model is represented in the lower part of Fig. 1 where we can see five main classes representing the five S of the model: `Society`, `Scenario`, `Stream`, `Structure` and `Space`. A `Scenario` is engaged by a `Society` and `employs` or `produces` a `Stream` which is `shaped` by a `Structure`. The `Service` class is central to this model because it connects all of its fundamentals given that it employs a `Space`, a `Structure` and a `Stream`; furthermore, a `Service` is composed by one or more `Scenarios` which are related to a `Society`. A `Scenario` may employ a `Space`. The `Digital Object` class is defined by a set of streams and structures and employs some spaces.

The 5S RDF graph represents the user domain similarly to the DELOS Reference Model; indeed, the class `Actor` represents an agent (e.g. a human or a computer) which belongs to a `Community` belonging to a `Society`.

The mapping with DELOS is quite straightforward since `Actor` is mapped in the homonym class and the `Community` class is mapped into the `Group` class of DELOS. The central class of the quality domain in the 5S model is `Quality Dimension` which allows us to evaluate every major concept in a DL [10]; this class is mapped into the `Quality Parameter` of the DELOS Reference Model. The `Service` class is mapped into the `Function` class of the DELOS Reference Model and the `Digital Object` class is mapped into the `Information Object` class by the `schema:isSimilarTo` property.

As we can see, the 5S model has no explicit representation of the architecture and policy domain of the DELOS Reference Model.

5 Final Remarks

In [2] we proposed a common ontology which encompasses all the concepts considered by the two foundational models and creates explicit connections between their constituent domains.

In particular, we highlight the important role of *service* in the 5S Model which is a broad concept comprising both the concepts of *function* and *architectural component* in the DELOS Reference Model. The presented ontology explicitly points out the connection between these concepts enabling connections between the two models which were not easy to recognize and establish otherwise. Furthermore, we investigated how policies are modeled and used by the two models; we pointed out that in the 5S model, the idea of *policy* is encompassed by the concept of *Society* even though it is not explicitly treated. Starting from this consideration, the common ontology enables the possibility of using policies defined in the DELOS Reference Model to regulate services in the 5S.

Much work is still ahead of us, since the proposed ontology needs to be operationalized into actual DLS and, probably, it will need to be extended both to accomplish specific details that arise when you make actual systems interoperate and to address peculiar needs of specialised domains, which may depart from the common general view. As a concrete example, we have started to work on an extension of the `Quality Domain` to model the scientific data generated by the experimental evaluation of *Information Retrieval (IR)* systems and to link them with expert profiles and expertise topics [14]. This specialisation also constitutes the starting point for dealing with the more complex problem of reproducibility and data citation in IR evaluation [5,6,15].

References

1. Agosti, M.: Digital libraries. In: Melucci, M., Baeza-Yaetes, R.A. (eds.) Advanced Topics in Information Retrieval. The Information Retrieval Series, vol. 33, pp. 1–26. Springer, Heidelberg (2011)
2. Agosti, M., Ferro, N., Silvello, G.: Digital library interoperability at high level of abstraction. Future Gener. Comput. Syst. (FGCS) **55**, 129–146 (2016)
3. Borgman, C.L.: What are digital libraries? Competing visions. Inf. Process. Manag. **35**(3), 227–243 (1999)
4. Candela, L., Castelli, D., Ferro, N., Ioannidis, Y.E., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., Agosti, M., Dobreva, M., Katifori, V., Schuldt, H.: The DELOS digital library reference model. Foundations for Digital Libraries. ISTI-CNR at Gruppo ALI, Pisa, Italy. http://delosw.isti.cnr.it/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf
5. Ferro, N.: Reproducibility challenges in information retrieval evaluation. ACM J. Data Inf. Qual. (JDIQ) **8**(2), 8:1–8:4 (2017)
6. Ferro, N., Fuhr, N., Järvelin, K., Kando, N., Lippold, M., Zobel, J.: Increasing reproducibility in IR: Findings from the Dagstuhl seminar on "reproducibility of data-oriented experiments in e-science". SIGIR Forum **50**(1), 68–82 (2016)
7. Ferro, N., Silvello, G.: Towards a semantic web enabled representation of DL foundational models: the quality domain example. In: Calvanese, D., De Nart, D., Tasso, C. (eds.) IRCDL 2015. CCIS, vol. 612, pp. 24–35. Springer, Cham (2016). doi:10.1007/978-3-319-41938-1_3
8. Fox, E.A., Gonçalves, M.A., Shen, R.: Theoretical Foundations for Digital Libraries: The 5S (Societies, Scenarios, Spaces, Structures, Streams) Approach. Morgan & Claypool Publishers, San Rafael (2012)
9. Gonçalves, M.A., Fox, E.A., Watson, L.T., Kipp, N.A.: Streams, structures, spaces, scenarios, societies (5S): a formal model for digital libraries. ACM Trans. Inf. Syst. (TOIS) **22**(2), 270–312 (2004)
10. Gonçalves, M.A., Lagoeiro, B., Fox, E.A., Watson, L.T.: What is a good digital library? A quality model for digital libraries. Inf. Process. Manag. **43**(5), 1416–1437 (2007)
11. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool Publishers, San Rafael (2011)
12. Lesk, M.: Practical Digital Libraries. Books Bytes & Bucks. Morgan Kaufmann Publishers, San Francisco (1997)
13. Novak, J.D.: Concept maps and Vee diagrams: two metacognitive tools to facilitate meaningful learning. Instr. Sci. **19**(1), 29–52 (1990)
14. Silvello, G., Bordea, G., Ferro, N., Buitelaar, P., Bogers, T.: Semantic representation and enrichment of information retrieval experimental data. Int. J. Dig. Libr. (IJDL) (2016)

15. Silvello, G., Ferro, N.: Data citation is coming. Introduction to the special issue on data citation. Bull. IEEE Tech. Comm. Dig. Libr. (IEEE-TCDL) **12**(1), 1–5 (2016)
16. W3C: Architecture of the world wide web, vol. 1 - W3C Recommendation, 15 December 2004. http://www.w3.org/TR/webarch/
17. Witten, I.H., Bainbridge, D.: How to Build a Digital Library. Morgan Kaufmann Publishers, San Francisco (2003)