A Semantic Model for Content Description in the Sapienza Digital Library

Angela Di Iorio(✉) and Marco Schaerf

DIAG - Department of Computer, Control, and Management Engineering
Antonio Ruberti, Sapienza University of Rome, Rome, Italy
{angela.diiorio,marco.schaerf}@uniroma1.it

Abstract. In this paper is presented the semantic model defined for descriptive metadata of resources, managed by the Sapienza Digital Library. The semantic model is derived from the Metadata Object Descriptive Schema, a digital library descriptive standard, for library applications. The semantic model can be used as top level conceptual reference model, in order to support the implementation of semantic web technologies for digital library's descriptive metadata. The semantic model is intended to be agnostic about the technology system to be adopted. The creation of resources' connections toward the linked data cloud, as well as the opportunity of exploiting the potential of services based on the ontology use, will rely on a well-defined semantic model, which has been widely tested by the implementation of a descriptive metadata profile.

Keywords: Digital libraries · MODS · Semantic modelling · Linked data · Ontologies

1 Introduction

The current scenario of Semantic Web technology challenges Organizations managing digital library systems. The increasing use of technologies oriented to the Semantic Web(SemWeb)[1] forces digital library managers to rethink the way of providing information about managed resources. The SemWeb provides a common framework which allows to share data and to re-use it across applications, enterprises, and community boundaries. Over the classic "Web of documents", the SemWeb is a technology stack[2] enabling to develop systems supporting trusted interactions over the network, and allowing computers to do more useful work. In respect to the "layer cake", the status of technologies, in the digital library management, is characterized by the wide use of the Extensible Markup Language (XML) technologies. By using semantics community-based, the digital library systems are able to exchange documents

[1] Semantic Web, http://www.w3.org/standards/semanticweb/.
[2] Semantic Web "layer cake", http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html.

D. Calvanese et al. (Eds.): IRCDL 2015, CCIS 612, pp. 36–47, 2016.
DOI: 10.1007/978-3-319-41938-1_4

in a coherent way. At present, the Resource Description Framework(RDF) data model has started to be experimented in library communities, under the motivation of the Linked Data Initiaves(LD) and it has been especially used over the web for open values (or controlled) vocabularies [12]. In addition, different ontologies related to library domain have been developed by designated communities[3,4,5]. The semantic model, presented in this paper, aims to support the SemWeb technologies implementation, over an existing digital library management system. Derived from a community-shared metadata profile, the semantic model encompasses conceptual descriptions expressing the knowledge base of the digital library descriptive metadata. The model is the conceptual guidance for considering the best solution to be adopted, among a variety of approaches for the implementation of SemWeb technologies for descriptive metadata.

2 Background and Motivation

The technologies, composing the first layers of the SemWeb stack, are here summarized in order to show their main characteristics and informational objectives. The XML technologies[6] have been conceived for managing data, structured as a hierarchical tree. XML is widely used for application contexts where is necessary to exchange data with a structure semantically pre-defined. One of the most important historical advantage is, having provided text information with a "well-formed" structure, that is named with human-readable semantics, that are community-based. The upper layers of the SemWeb stack use the XML as official syntax. The Resource Description Framework(RDF)[7], allows to build statements about web resources in the form of a subject-predicate-object expression. The statements can be interpreted by machines, that become capable to make connections between resources over the Web. The RDF schema(RDFS)[8] provides the framework to describe application-specific classes and properties and provides a data-modelling vocabulary for RDF data. The Web Ontology Language(OWL), currently at the second version[9], is a standard for defining ontologies that are used to capture knowledge about some domain of interest. It provides classes, properties, individuals, and data values, stored as SemWeb documents. The OWL is a semantic markup language for publishing and sharing ontologies on the Web. OWL is developed as a vocabulary extension of RDF, and every OWL ontology is a valid RDF document. The OWL may be categorised into three species or sub-languages in the 1st version (OWL-Lite, OWL-DL, OWLFull), and three profiles in the 2nd version. The OWL-EL, OWL-DL, OWL-RL, as profiles of the 2nd version, have been defined for shaping the ontology

[3] Bibliographic Ontology Specification, http://bibliontology.com/.
[4] FaBiO, the FRBR-aligned Bibliographic Ontology, http://purl.org/spar/fabio.
[5] Semantic Publishing and Referencing Ontologies, http://www.sparontologies.net/.
[6] XML technology, http://www.w3.org/standards/xml/.
[7] Resource Description Framework, http://www.w3.org/RDF/.
[8] RDF Schema 1.1, http://www.w3.org/TR/rdf-schema/.
[9] OWL 2 Web Ontology Language, http://www.w3.org/TR/owl2-overview/.

expressiveness and coherently supporting the computability degree. The implementation of the technologies, from the RDF up to OWL, over a legacy system, that uses XML as format for exchanging information, requires to state at least a point of reference for defining semantics, in order to model the implementation of SemWeb technologies: RDF as data model, RDFS as description of RDF data, OWL as more-formal language to be used by reasoning systems. The semantic model, presented in this paper, is indeed the reference model for implementing semantic web technologies in the domain of the management of resources, that belong to a digital library. In particular, it makes transparent main concepts associated with the resources' content description, used by the digital library system. The model can be used as semantic foundation for different approaches that can be undertaken in implementing SemWeb technologies, over the existing digital library system. The following list summarizes some of the feasible implementation cases that can be undertaken:

- creating data model based on RDF/RDFS [19]
- creating vocabularies/ontologies based on RDFS or OWL [14]
- creating mapping from XML to RDF [2,20]
- creating mapping from Database Management System to RDF[10]
- creating an ontology-based data access system [4,17]

3 Use Case Overview

The Sapienza University of Rome has established its digital library by means of a research project named Sapienza Digital Library(SDL) [10]. The vision of the SDL project was to provide Sapienza's community with a digital library, supporting the use of digital material managed, owned and produced by the Sapienza University [5]. The metadata framework, supporting the digital services of the produced SDL[11], was designed for managing heterogeneous resources, digitally representing the multidisciplinary community of the Sapienza University. The SDL metadata framework is the structured container for metadata managed by the SDL services. The SDL metadata framework is conceptually based on the Open Archival Information System(OAIS) [7]) Information Package (IP). The SDL-IP has to be structurally and semantically conforming with elements defined in the SDL metadata framework. At the present time, the SDL framework uses the metadata elements of the Metadata Objects Description Schema(MODS) for describing the intellectual contents of the resources [10], the PREservation Metadata Implementation Strategies[12](PREMIS), necessary to support the long-term digital preservation, and the Metadata Encoding and Transmission Standard[13](METS), for packaging and connecting the different metadata,

[10] R2RML: RDB to RDF Mapping Language, http://www.w3.org/TR/r2rml/.
[11] Sapienza Digital Library, sapienzadigitallibrary.uniroma1.it.
[12] PREMIS Preservation Metadata Maintenance Activity, http://www.loc.gov/standards/premis/.
[13] Metadata Encoding Transmission Standard, www.loc.gov/standards/mets/.

belonging to the digital resources. These digital library standards, maintained by the Library of Congress[14], are defined by XML schemas[15]. The conformance with standards is based on the production of XML files validated against the pertaining XML Schema. The semantic model, presented in the Sect. 4, is focused on the descriptive metadata section of the SDL framework(see next Sect. 3.1), and is the semantic representation of the MODS profile (see Sect. 3.3). The MODS profile has been defined for the SDL implementation needs, and has aligned descriptive metadata, coming from different Sapienza's communities (libraries, museums and archives).

3.1 The SDL Descriptive Metadata

In the reference standard OAIS [7], the IP is a conceptual container of Content Information(CI) and Preservation Description Information(PDI), where the resulting package is viewed as being discoverable by virtue of the Descriptive Information. In the SDL project, during the digital resources life-cycle, the OAIS IP is considered the target package of the data management and the existence of the Descriptive Information is the pre-condition for building the SDL-IP. In the SDL implementation, the descriptive information is coded in MODS, for describing the intellectual content of the SDL-IP. The MODS uses libraries semantics, that are derived from MARC 21[16], the standard format created in 1999 (a revised version for the 21st century of the MAchine Readable Cataloguing (MARC)[17] created in 1960) and widely used by libraries information systems. The MODS semantics were used for describing the different kind of intellectual/creative works managed by the SDL. The SDL-IPs are characterized by different formats (still and moving images, texts, sounds, cartographics, etc.) and are differently structured (digital collections, books, images, videos, documents, hierarchies, maps). The intellectual/creative works, represented as SDL digital resources (SDL-IPs), can also be expressing multidisciplinary knowledge.

3.2 The MODS Standard and Linked Vocabularies

In [13] is reported how during the years, has raised the emergent need of having a standard less complex than MARC, but not as simple as the widely interoperable standard, Dublin Core Metadata Element Set (DC)[18]. In order to address these community's need, the Library of Congress developed the MODS, "a schema for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications"[19].

[14] Standards at the Library of Congress, http://www.loc.gov/standards/.
[15] W3C XML Schema - World Wide Web Consortium, http://www.w3.org/XML/Schema.
[16] MARC 21 Format for Bibliographic Data, www.loc.gov/marc/bibliographic/.
[17] MAchine Readable Cataloguing (MARC), http://en.wikipedia.org/wiki/MARC_standards.
[18] Dublin Core Metadata Element Set, http://dublincore.org.
[19] Metadata Object Description Schema, http://www.loc.gov/standards/mods/.

The MODS XML schema includes a subset of MARC fields, using language-based tags rather than numeric ones. The official website provides guidelines primarily intended to be used for assistance in creating original MODS records, as well as converted from MARC 21 or for use in developing detailed conversion specifications. The MODS standard claims that the element set is richer than Dublin Core, more oriented to the end-user than the full MARCXML schema, and simpler than the full MARC format. These main characteristics were, in general, deemed useful for describing at a sufficient granular level, the multi-faceted materials to be managed in the SDL.

The descriptive metadata for a MODS resource aggregates titles, names, subjects and other data elements that associated with the resource, further help to describe the resource. For systems, the MODS resource description and especially its descriptive metadata aids the indexing, searching, and displaying of information about the resource. For users, a MODS resource description assists with identifying, finding, selecting, and accessing a MODS Resource. Administrative metadata provides provenance information about the descriptive metadata. It includes information, such as the individual or organization responsible for the descriptive metadata and/or the date the descriptive metadata was last modified, as well as relationships expressed in the MODS resource description.

The XML schema of MODS is deployed on 20 top elements, that are variously structured with sub-elements. Except for the administrative description of the record (`recordInfo`), all MODS elements are repeatable. No element is mandatory, and the `relatedItem` element is recursive, because it can contain all the top elements of the schema, including itself.

The web availability of different mappings[20], initiaves [11], and studies extending the use of this standard over other information science disciplines, like archivist science [3] or toward museums' artifacts [8], supports the interoperability of the semantics adopted, as well as the probability of the information loss. The SDL metadata framework exploits the MODS's ability of providing a mechanism to dereference, in documented way, the controlled vocabularies' entries, used in the XML elements. In SDL, this ability has been exploited for using entries from the Library of Congress Linked Data Service[21], from other italian LD vocabularies[22], and from vocabularies locally defined[23]. The use of MODS schema attributes, like the controlled vocabulary's source authority name (`authority`), and the identifier (`authorityURI`) and the identifier of the vocabulary's entry value (`valueURI`), allows to create authoritative connections between the MODS element value and the authority maintaining the vocabulary's term as Linked Open Vocabulary(LOV)[24].

The granularity and the flexibility of MODS, and its application in a multi-faceted environment, has required the definition of a profile for the SDL

[20] MODS conversions, http://www.loc.gov/standards/mods/mods-conversions.html.
[21] Library of Congress Linked Data Service, http://id.loc.gov/.
[22] Nuovo Soggettario, Biblioteca Nazionale di Firenze, http://purl.org/bncf/tid.
[23] Sapienza Digital Library vocabularies, http://sbs.uniroma1.it/sdl/vocabularies.
[24] Linked Open Vocabularies (LOV), http://lov.okfn.org/dataset/lov/.

implementation, in order to define the minimum obligation in using XML elements and attributes.

3.3 The SDL MODS Profile

The SDL MODS profile[25] encompasses metadata elements necessary to the MODS XML data coding, defines the structural and functional contexts where elements and attributes can be used, and defines the obligation constraints on the existence of specific MODS XML elements.

The principle, guiding the definition of the MODS profile, has been to maintain and to enrich as much as descriptive information about different content types, and to integrate XML elements by dereferencing URIs from LOV and the maintaining authority.

The definition of the MODS profile had followed two main reference guidelines. Firstly, the "Master List of Data Elements" is a framework of elements gathered from the profiles of the different digital library projects at the Library of Congress. The objective of the list is to work towards "compatibility of metadata usage throughout the institution, support the metadata use cases, and point to areas where metadata remediation for improved consistency or enhanced interoperability might be beneficial"[26]. Secondly, the implementation guidelines of the Digital Library Federation for the implementation of shareable MODS records[27].

The MODS profile is applied to the following content types: books, documents, images, videos, maps, hierarchical structures (like serial publications or archival collections), and digital collections.

During the project development the profile was incremented and enriched coherently to contents' descriptive needs. Each new integration of profile specifications has entailed the conformance check on the metadata relational structure, in order to verify the consistent coexistence of different materials and descriptions, inside of the same metadata framework.

Some specifications, indeed, have required the specific profiling of metadata elements with the inclusion of sub-elements and attributes, that have been set for the application context. The specifications have improved not only the granularity of the descriptive information, but also the data accuracy and the interoperability of the content re-use because the wide reference to local controlled vocabularies, and wherever possible to LOV.

Beyond the XML elements, whose semantics are defined in the MODS Schema, the profile had configured the following attributes for the local needs:

---

[25] The Sapienza Digital Library MODS profile (available by the end of January 2015), http://sbs.uniroma1.it/sdl/documentation.

[26] Metadata for Digital Content(MDC), Developing institution-wide policies and standards at the Library of Congress, http://www.loc.gov/standards/mdc/elements/MasterDataElementList-20120215.doc.

[27] Digital Library Federation/Aquifer Implementation Guidelines for Shareable MODS Records, https://wiki.dlib.indiana.edu/confluence/download/attachments/24288/DLFMODS_ImplementationGuidelines.pdf.

- @displayLabel is the label for displaying the XML element's content. In the SDL MODS profile specification, it was used for the portal implementation to show the italian label of the field.
- @xlink:href is the reference link associated with the XML element's content. In specific elements, it has been used for the portal implementation to activate the link over the XML elements content value.
- @authority, @authorityURI, @valueURI, these attributes are extensively used for referring, respectively to the authority controlled vocabulary's name, the URI/URL identifying the authority controlled vocabulary, the URI value exposed in the Linked Data Cloud.
- @xml:lang this element allows to label the XML elements' content with a language selector. It was extensively applied and coded in the existing descriptive record, and wherever it was possible, the Italian content was translated in English. The reference code used for language is the ISO-639-2[28].

The SDL MODS profile provides specifications about 28 MODS metadata elements, that are necessary to the coherent description of the SDL resources. The profile provides references to 13 controlled vocabularies, locally defined, and 21 vocabularies of third party can be used. In addition the profile provides specifications for the customization needs of the SDL portal implementation.

The Table 1 shows the occurrences of specifications that were defined for each metadata elements. The higher number of specifications occurrences is significant because reveal the susceptibility of certain elements, in respect to others for the application context. For example, four over seven titleInfo specifications are destined to the portal visualization requirements. The eight specifications of the subject element have profiled most of the sub-elements, with the specification about the controlled vocabularies to be used and in some case, with anchors necessary to the SDL browsing services. The eight specifications of the relatedItem instead reveal the extensive use of this element, for building relationships between the SDL resources. The SDL MODS profile provides reference terms for qualifying the type of relationship, occurring between resources.

Table 1. MODS metadata elements and SDL profile specification occurrences

titleInfo	**7**	originInfo/date	1	recordInfo	1
name	**2**	originInfo/place	1	note	**2**
abstract	**4**	originInfo/publisher	1	genre	**2**
accessCondition	**2**	originInfo/edition	1	subject	**8**
language	1	originInfo/issuance	1	relatedItem	**8**
typeOfResource	1	originInfo/frequency	1	part	**2**
identifier	1	physicalDescr./digitalOrigin	1	extension	1
location	**2**	physicalDescr./internetMediaType	1	targetAudience	1
tableOfContents	1	physicalDescr./form	1	classification	1
		physicalDescr./extent	**2**		

[28] ISO 639-2 Language Codes, http://id.loc.gov/vocabulary/iso639-2.html.

4 The SDL Semantic Model for Descriptive Metadata

The SDL's semantic model(SDL-SM) summarizes the concepts specified in the SDL MODS profile. Specifically the SDL-SM represents the set of concepts

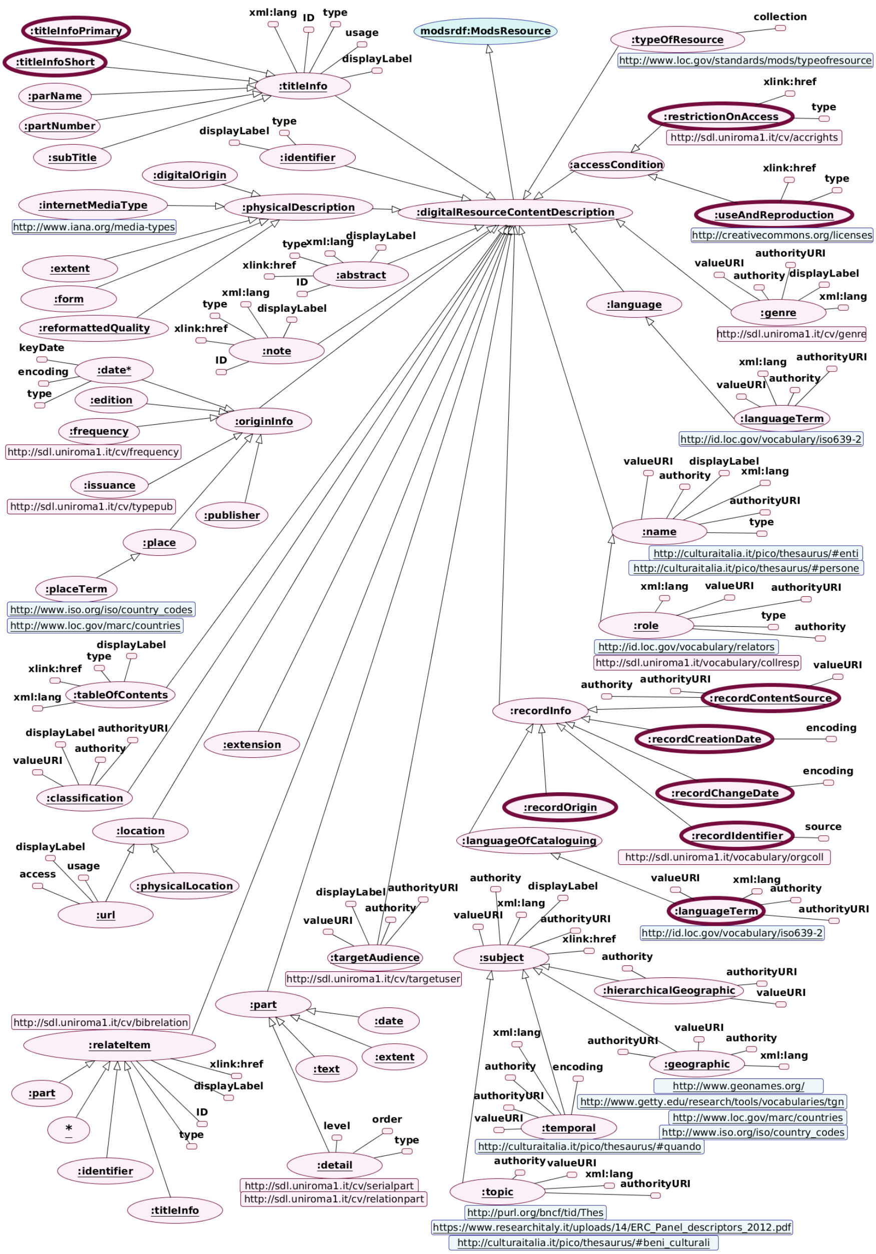

Fig. 1. The SDL's semantic model for contents' resource descriptions

underlying the knowledge domain related to the digital library descriptive metadata, about the intellectual content of digital library's resources.

The SDL-SM is meant to be agnostic about the technology system on which it will be implemented and it can be used as top level conceptual reference model, for implementing SemWeb technologies in the digital library system.

The SDL-SM can be considered as the reference model not only for developers, but also for domain experts. The pre-condition for re-using the concepts defined in this model is the acceptance of the core principles and best practices of the Semantic Web and Linked Data [6], in particular the use of specific policies for managing the SDL resource's identifiers building [9] is considered the point of reference for assigning the URI to the semantic concepts.

The Fig. 1 outlines the concepts laying down the semantics defined in the MODS XML schema hierarchy, that were used in the SDL MODS profile. The concepts are represented by the ovals and the bold line ovals represent the obligation constraints over the concepts. The repeatability constraint of concepts (only one concept is not repeatable: `recordInfo`) is distinguished by the ovals with underlined labels. Borrowing from the Description Logics the basic inference on concept expressions [1], each MODS top element is represented as a concept subsumption of the main concept `digitalResourceContentDescription` which in turn is a subsumption of the MODS `modsrdf:ModsResource` concept.

Analogously, the XML sub-elements (lower hierarchical levels), are considered as subsumptions of the concepts derived from the XML top elements of MODS. Differently from a data model design or from an ontology definition, the properties/relationships relating concepts, are placed in a subsumption hierarchy [18] leaving further definition open to the implementation choices.

The specifications, defined by the MODS profile, had leveraged on the customization of some attributes, that are relevant to the exhaustive description of the resources. The attributes used for MODS profile are represented in the figure, in Entity/Relationships diagrams notation style, for highlighting that the association of attributes to each related concept can be further modeled at implementation time.

The graphical representation of the SDL-SM respects the MODS XML schema hierarchical style, but makes evident the ambiguity points that have to be overcome in a SemWeb Technology implementation. The MODS schema has defined the use of same attributes for different XML elements[29], and indeed the figure highlights the wide re-use of the same attributes (i.e. `@displayLabel`, `@type`...) in different conceptual contexts. This graphical evidence warns that, in this case, the semantic disambiguation of the information is contained in the value of the attribute, and not in the XML schema data definition. Similarly, the homonymy of sub-elements (i.e. `@languageTerm`, `@extent`), need to be solved by inheriting the conceptual name of the higher level semantic elements. Mostly the same inheritance

[29] Attributes Used Throughout the MODS Schema, http://www.loc.gov/standards/mods/userguide/generalapp.html#list.

application coherently works for all sub-elements that cannot have sufficient semantic specification, like in the `role` case.

Another critical aspect to be highlighted by the semantic point of view, is the recursive functionality held by the `relatedItem` element, that requires the semantic disambiguation of the contained elements. The disambiguation can be managed by a semantic inheritance mechanism where a combination of `relatedItem`'s attribute values and contained element's names can distinguish different concepts expressed by the recursive elements. The Fig. 1 also shows the reference to the controlled vocabularies, used by the SDL MODS profile, and exhibits the association to relevant concepts. The rounded rectangles containing an URI, and in close contact with some ovals (the concepts), specify that the concept's value is derived from the controlled vocabulary deferenced by URI. The URI in pink rounded rectangles specify SDL vocabularies, while the underlined URIs in the cyan rounded rectangles specify third party vocabularies.

The SDL-SM does not provide a further ontological modelling about relationships/properties because would imply implementation-dependent choices. The modelling of relationships/properties is a customization task and it depends on the SemWeb technology to be used. For example, the production of RDF triples can mostly depend on technology environment factors, the RDF triples could come from a Relational Database mapping or managed directly by a triple store system. Similarly, the conceptual modelling of an ontology, in order to be used by reasoning systems, cannot be expressed in OWL 2 because of the high complexity of computation, but it needs to be profiled into one of the three ontology language fragments, the OWL 2 profiles(EL, QL, RL)[30].

5 Related Work and Conclusion

The initiatives that can be coherently compared, because oriented toward the SemWeb, and related to the specific library application domain of the descriptive metadata, are briefly summarized.

The draft of the MODS RDF Ontology [16] is the semantic model expression of the MODS model and is a reference for shaping MODS resources as RDF triples. The basic conceptual assumption is similar to the SDL-SM because "a MODS resource description includes descriptive metadata about a MODS resource", but the ontology (coded in OWL 1 Lite[31]) defines more prescriptive constraints. The MODS XML elements are modelled 20 classes distinguished in: 1 main class (`modsrdf:ModsResource`), 9 MODS classes as ranges for properties whose domain is `modsrdf:ModsResource`, and 10 classes imported by an external ontology. Nevertheless, mostly MODS elements are modelled in 114 `owl:ObjectProperty`, with corresponding domain and range defined by RDFS statements. The second comparable initiative is the Bibliographic Framework Initiative (BIBFRAME) [15] where the focus is the translation of MARC 21 format to a Linked Open

[30] OWL 2 Web Ontology Language Profiles, http://www.w3.org/TR/owl2-profiles/.

[31] OWL Lite RDF Schema Features, http://www.w3.org/TR/2004/REC-owl-features-20040210/#s3.4.

Data model. The BIBFRAME has defined a model consisting of 4 core classes (Creative Work, Instance, Authority, and Annotation), and a vocabulary structured over different categories of information. The BIBFRAME metamodel is designed to be lightweight, flexible and able to accommodate the declarative needs of existing descriptive standards(RDA[32], DACS[33], VRA[34], etc.) and yet-to-be-developed community vocabularies. To best accomodate these communities the BIBFRAME RDF Schema is intentionally underspecified in terms of constraints such as domain and range. With regard to the SDL-SM, the choice of defining a conceptual reference model based on standardized and community-based semantics, and agnostic about technology, offers more flexibility and more consistency to further conceptualizations as implementation foundation. The use of digital library metadata, in a SemWeb scenario, will trespass on the hosting system boundaries, as well as the MODS conceptual assumption, and will match to other "world" like similar ontologies as BIBFRAME or others unpredictably similar.

References

1. Baader, F.: The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, Cambridge (2003)
2. Bohring, H., Auer, S.: Mapping XML to OWL ontologies. Leipzig. Informatik-Tage **72**, 147–156 (2005)
3. Bountouri, L., Gergatsoulis, M.: Interoperability between archival and bibliographic metadata: an EAD to MODS crosswalk. J. Libr. Metadata **9**(1–2), 98–133 (2009)
4. Calvanese, D., Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: the DL-Lite family. J. Autom. Reason. **39**(3), 385–429 (2007). http://link.springer.com/10.1007/s10817-007-9078-x
5. Catarci, T., Di Iorio, A., Schaerf, M.: The Sapienza Digital Library from the holistic vision to the actual implementation. Procedia Comput. Sci. **38**, 4–11 (2014)
6. Hyland, B., Atemezing, G., Villazón-Terrazas, B.: Best practices for publishing linked data. W3C Working Group Note (2014)
7. Consultative Committee for Space Data: Reference Model for an Open Archival Information System (OAIS), Recommended Practice CCSDS 650.0-M-2 Magenta Book (2012). http://public.ccsds.org/publications/archive/652x0m1.pdf
8. Di Iorio, A., Schaerf, M.: Applicability of digital library descriptive metadata to the contemporary artworks. In: Nesi, P., Santucci, R. (eds.) ECLAP 2013. LNCS, vol. 7990, pp. 78–89. Springer, Heidelberg (2013)
9. Di Iorio, A., Schaerf, M.: The organization information integration in the management of a digital library system. In: 2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL), pp. 461–462. IEEE, poster and Extended Abstract (2014)

[32] Resource Description & Access(RDA), http://www.rdatoolkit.org/.
[33] Describing Archives: A Content Standard, Second Edition (DACS), http://files.archivists.org/pubs/DACS2E-2013.pdf.
[34] Visual Resources Association(VRA), http://www.loc.gov/standards/vracore/.

10. Di Iorio, A., Schaerf, M., Bertazzo, M.: Establishing a Digital Library in Wide-Ranging University's context. In: Agosti, M., Esposito, F., Ferilli, S., Ferro, N. (eds.) IRCDL 2012. CCIS, vol. 354, pp. 172–183. Springer, Heidelberg (2013)
11. Europeana Libraries consortium: Europeana libraries: Aggregating digital content from Europe's libraries (2013). http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=270933
12. Isaac, A., Waites, W., Young, J., Zeng, M.: Library linked data incubator group: datasets, value vocabularies, and metadata element sets. W3C Incubator Group Report **25**, (2011)
13. Guenther, R., McCallum, S.: New metadata standards for digital resources: MODS and METS. Bullet. Am. Soc. Inf. Sci. Technol. **29**(2), 12–15 (2003)
14. Hogan, A., Delbru, R., Umbrich, J.: RDFS & OWL reasoning for linked data (2013)
15. Library of Congress: Bibliographic framework as a web of data: Linked data model and supporting services (2012). http://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf
16. Library of Congress: The MODS RDF ontology primer (2012). http://www.loc.gov/standards/mods/modsrdf/primer.html#vocabularies
17. Noy, N.F.: Semantic integration: a survey of ontology-based approaches. ACM Sigmod Rec. **33**(4), 65–70 (2004)
18. Rector, A.L.: Modularisation of domain ontologies implemented in description logics and related formalisms including OWL. In: Proceedings of the 2nd International Conference on Knowledge Capture, pp. 121–128. ACM (2003)
19. Shadbolt, N., Hall, W., Berners-Lee, T.: The semantic web revisited. IEEE Intell. Syst. **21**(3), 96–101 (2006)
20. Van Deursen, D., Poppe, C., Martens, G., Mannens, E., Walle, R.: XML to RDF conversion: a generic approach. In: 2008 International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution. AXMEDIS 2008, pp. 138–144. IEEE (2008)