Permanence and Temporal Interoperability of Metadata in the Linked Open Data Environment

Shigeo Sugimoto
University of Tsukuba, Japan
sugimoto@slis.tsukuba.ac.jp

Chunqiu Li
University of Tsukuba, Japan
licq.chunqiu@gmail.com

Mitsuharu Nagamori
University of Tsukuba, Japan
nagamori@slis.tsukuba.ac.jp

Jane Greenberg
Drexel University, USA
jg3243@drexel.edu

Abstract

This paper examines metadata longevity issues in the Linked Open Data (LOD) environment, where metadata, as a digital object, is transferred and shared on the open Web. Longevity is key for achieving metadata permanence, which allows metadata to remain interpretable by machines and humans over time. The discussion presented in this paper seeks to clarify risks in permanence of metadata, by focusing on metadata longevity challenges specific to metadata and metadata schemas in the LOD environment.

This examination addresses metadata longevity from several different viewpoints in order to clarify the requirements of metadata permanence in the LOD environment, and distinguish these needs from conventional document-like object environment or database-centric environment. A central theme in this work is that longevity of metadata is, in essence, the temporal interoperability of metadata. This paper uses the Metadata Application Profile methodology supported by the Dublin Core Metadata Initiative (DCMI) and DCMI's layered model of metadata interoperability to understand the nature of metadata in the LOD environment. Next, the paper discusses metadata longevity based on a set of facets of metadata entities such as metadata schemas; and the last part briefly discusses issues to use provenance description of metadata schemas and metadata schema registries from the viewpoint of long-term maintenance of metadata schemas.

Keywords: metadata longevity; digital preservation; provenance description; metadata schema registry; metadata schema maintenance

1. Introduction

The importance of metadata for digital object preservation is well recognized with metadata standards, such as PREMIS and METS, both supporting property sets for document object preservation. The longevity of metadata objects in the Web environment is, however, still largely unexplored. This is because metadata schemas are treated as instances conventional, operational database systems, or as document like objects, with longevity challenges addressed through common document or database preservation methods. In comparison, metadata in the open Web environment has different features and functionalities. Web resources may include metadata embedded in the headings and/or bodies; and there are many LOD conformant datasets available on the Web that may be used to link Web resources and other objects. A significant feature of metadata on the Web is that metadata instances may be transferred from site to site and saved for the future use. For example, metadata embedded in a HTML header and Cataloging In Publication (CIP) included in a digital book may be extracted and transferred as a digital object. Another example is RDF encoded metadata instances that can be downloaded from LOD datasets for different applications. A metadata schema, which defines structural, syntactic and semantic features of metadata, can also be transferred on the Web as well as the metadata instances because they are metadata about metadata. In fact, metadata transferrable as a digital object on the Web is First Class Object, as examined in this paper.

https://doi.org/10.23106/dcmi.952137567

DCMI Application Profiles (DCAP) provide an intellectual and structural framework for mixing-and-matching metadata vocabularies to define a metadata schema for an application. LOD recommends the use of standards such as OWL and RDF to define metadata schemas and vocabularies in order to make metadata interoperable. These conventions and best practices present a significant challenge when considering metadata longevity. On one hand, metadata application profiles use well-standardized, mature and fairly stable metadata vocabularies. On the other hand, the schema may rely on many components defined outside of the immediate schema. Changes to any of these external components can have a significant impact on the application schema's functionality.

Leading researchers have been hosting two metadata schema registries – DCMI Metadata Registry and MetaBridge[1]. These registries are developed based on RDF and LOD technologies. The DCMI Registry is dedicated to providing access to DCMI terms (Nagamori, Baker, Sakaguchi, Sugimoto, & Tabata, 2001). MetaBridge provides functions to store and provide application profiles in addition to metadata vocabularies and terms (Nagamori, Kanzaki, Torigoshi, & Sugimoto, 2011). These registries function differently than data and resource repositories. That is, they are generally not developed for long-term maintenance of metadata schemas but they have a large potential to serve as a long-term maintenance host, and could in fact be developed to address similar goals.

The Open Archival Information System (OAIS), which is a well-known standard for digital preservation, defines the Information Package model (CCSDS, 2012). Although the information package model does not address metadata longevity very well, it provides important insights. An information package is composed of an information object and Preservation Description Information (PDI). PDI is a metadata to describe attributes required to keep information object interpretable, i.e., renderable, playable, operable, and functional in various ways. This development focuses on digital objects. A serious, overlooked challenge here is that the PDI may contain or refer metadata embedded in the object. Thus, preservation of digital objects requires long-term maintenance of metadata. This paper recognizes this challenge, and highlights the need to address metadata longevity.

This paper focuses on metadata schema longevity chiefly in the LOD environment. The remainder of the paper is organized as follows: Section 2 discusses a framework for long-term maintenance of metadata schemas from the viewpoints of LOD and DCAP developments; Section 3, defines entities included in metadata preservation in the LOD environment; Section 4 discusses entities of metadata and risks in long-term use of the entities; and Sections 5 and 6 include a discussion and conclusion respectively.

2. Metadata and Metadata Schemas in the LOD Environment

2.1. Basic Concepts and Words

This section provides definitions of key words and concepts used in this paper. *Metadata* is defined as "(structured) data about data." Metadata about a metadata is called *meta-metadata*. A *metadata schema* is an expression of definitions, including structural, syntactic and semantic features of metadata for an application. Thus, a metadata schema is metadata about the application metadata--essentially the meta-metadata for the application. Controlled vocabularies used within metadata schema are referred to as *metadata vocabularies*, and loosely classified as *property vocabularies* or *value vocabularies*. Metadata vocabularies may be defined using a formal definitions scheme such as RDF Schema and OWL in the LOD environment. The Dublin Core Metadata Element Set (DCMES) of 15 elements, is a typical property vocabulary. Subject headings such as Library of Congress Subject Headings (LCSH) and Medical Subject Heading

[1] Metadata Registries. All sites retrieved May 27, 2016
The Dublin Core Metadata Registry, from http://dcmi.kc.tsukuba.ac.jp/dcregistry/
MetaBridge, from https://www.metabridge.jp/infolib/metabridge/menu/?lang=en

https://doi.org/10.23106/dcmi.952137567

(MeSH) are typical value vocabularies. A term included in a metadata vocabulary is called *metadata term*. DCMI Metadata Application Profiles (DCAP) define structural constraints for metadata and application specific requirements. The DCAPs and the DCMI's layered metadata interoperability model are key components for metadata longevity, and integral to issues in this paper.

Metadata is indispensable for searching, managing, and processing data object instances. Moreover, a metadata schema is indispensable for correctly creating and interpreting metadata. As defined above, a metadata schema is meta-metadata. In the LOD environment, metadata schemas may be shared on the Web along with metadata instances rendered with the existing schemas. As a result, the guiding metadata schemas must be interpretable by both humans and machines that comply with the LOD environment. There are schemes to define metadata schemas formally. These schemes that define metadata schemas are metadata about meta-metadata, i.e., meta-meta-metadata. Thus, this iteration of "meta" seems endless, although it is crucial to: 1) understand the levels of "meta-", 2) maintain and preserve metadata over time, and 3) keep descriptions of different levels of "meta" interpretable for machines and humans over time. Figure 1 shows relationships among "meta"-entities. In this Figure, it should be noted that all entities should be maintained for the long-term use of the object instance.

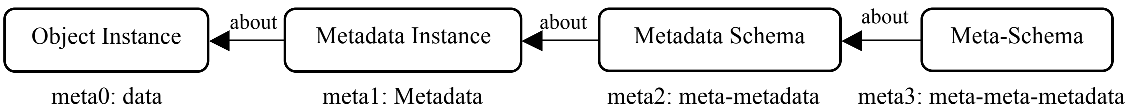

FIG.1. "meta-" relationships.

2.2. Metadata Preservation in the Linked Open Data Environment

There are various types of metadata, e.g., bibliographic descriptions, product descriptions, and rights notices. Researchers categorize these types, e.g., Lagoze (1996) and Greenberg (2005). In conventional library information systems, bibliographic records are stored in a database. Bibliographic databases are migrated from an old system to a new system many times. Metadata schemas are used for the migration, and system interoperability, and metadata schema documents need to be maintained along with the migrations. Thus, long-term maintenance of bibliographic records in a conventional system is carried out as long-term maintenance of the bibliographic database. Revision history of its bibliographic description scheme is generally recorded in its schema document when the database schema is revised.

This common maintenance practice does not, however, synchronize with the LOD metadata environment because of the fundamental difference of characteristics of metadata instances. In the LOD environment, a metadata instance is realized as an XML object which can be transferred and shared on the Web. Metadata schemas and vocabularies need to be maintained in order for metadata instances to be interpretable consistently over time. Figure 1 illustrates the requirement to maintain metadata in order to keep object instance interpretable by machine.

Machine interpretability of metadata, a key contribution of this work, is defined as follows:

- A machine driven function supporting metadata search and display, or other processes, which can automatically identify a metadata term and select a function in accordance with the meaning of the term, the metadata is full-machine interpretable.
- In the case there are revisions which do not impact machine interpretability of metadata instances but affects human interfaces, e.g., human readable labels of metadata terms, the metadata is semi-machine interpretable.

Machine interpretability of metadata means full-interpretable and/or semi-interpretable metadata.

The goal of long-term metadata maintenance is to keep a metadata instance machine interpretable over time and as intended from when the instance was created. Therefore, keeping

https://doi.org/10.23106/dcmi.952137567

metadata machine interpretable consistently and over time is the primary issue in metadata longevity.

Thus, long-term maintenance of metadata instances in the LOD environment is fundamentally different from the long-term maintenance of digital objects modeled by the OAIS standard. The metadata longevity model introduced in this paper encompasses the OAIS standard for preservation of digital object instance. Instances of meta-0 level in Figure 1 are primarily included in this category. Some instances of meta-1, -2, and -3 level realized as a document for human readers should be preserved as a digital object instance. On the other hand, other instances in those levels realized as a first class object should be maintained without losing consistent machine interpretability.

2.3. Dublin Core Application Profiles and Metadata Interoperability Model

DCAP presents a generalized model of metadata schemas and their components (Heery & Patel, 2000). The Singapore Framework of DCAP shown in Figure 2 defines the components of a metadata schema for an application and related components such as metadata vocabularies. A definite separation of metadata terms and structural features is the key feature of DCAP. The Singapore Framework defines five components of an application profile – Functional Requirements, Domain Model, Description Set Profile, Usage Guidelines, and Encoding Syntax Guidelines. These components and metadata terms should be well maintained for interoperability of metadata across communities and over time.

DCMI defines a simple layered model to present levels of metadata interoperability shown in Figure 3. In the model, the lowest layer (Level 1) is interoperability given by shared informal term definitions and the highest layer (Level 4) is DSP interoperability given by shared formal vocabularies and constraints. Nagamori and Sugimoto (2004) defined a three-layered model for metadata interoperability based on the Application Profile concept shown in Figure 4.

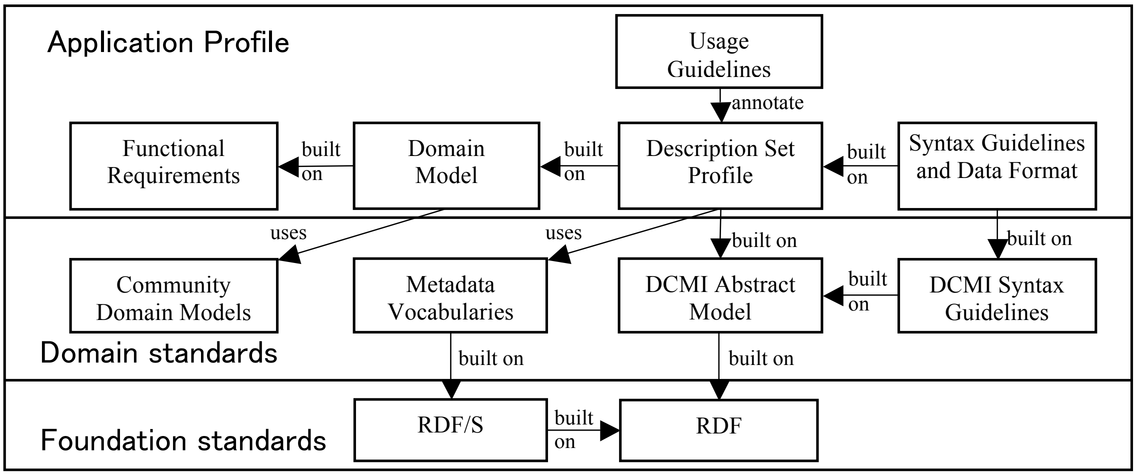

FIG.2. Singapore Framework of DCMI Application Profile (Nilsson, Baker, & Johnston, 2008).

Description Set Profile Interoperability Shared formal vocabularies and constraints in records
Description Set syntactic interoperability Shared formal vocabularies in exchangeable records
Formal semantic interoperability Shared vocabularies based on formal semantics
Shared term definitions Shared vocabularies defined in natural languages

FIG.3. Interoperability Levels of DCMI (Nilsson, Baker, & Johnston, 2009).

Implementation Syntax Concrete syntax
Structural Constraints Abstract syntax
Metadata Vocabularies Term Semantics

FIG.4. A Layered Model of Metadata Schema (Nagamori & Sugimoto, 2004).

https://doi.org/10.23106/dcmi.952137567

A metadata schema registry, which is a repository of metadata schemas and terms, can be used as a basis for sharing metadata schemas and terms on the Web. Metadata registries are not a permanent service but have a crucial role to keep the meaning of metadata terms for metadata preservation. For instance, the Technical Registry PRONOM at the National Archives of UK collects and maintains file format information to help digital preservation.

3. Metadata Preservation as Temporal Interoperability of Metadata

3.1. Metadata Preservation Facets

The result of our work to date reveals a set of facets for the long-term maintenance of metadata – entities in different meta-levels, preservation description categories, requirements specific to metadata preservation in the LOD environment, and other aspects. Figure 5 summarizes the facets described in the paragraphs below. Versioning of metadata schemas is a crucial aspect for the long-term maintenance of metadata. We discuss versioning further, below, as a part of long-term preservation and provenance description of metadata and metadata schemas.

(1) Facet 1: Entity Format Types – Document Files, Databases, XML Encoded Texts

Longevity management of metadata entities depends on the implementation formats of entities to be preserved. For example, it may be often the case that a metadata instance is stored in a database and an XML encoded instance is created when downloading the instance from the database. In the LOD environment, any instance which should be identifiable as a resource has to be given a URI. Maintaining URIs consistent is one of the key issues for metadata permanence.

(2) Facet 2: Entity Types – Meta-Levels

As shown in Figure 1, there are instances of different meta-levels from level 0 to level 3. This paper assumes any instances of these four categories are realized in a digital form, although they may be realized as a non-digital instance, e.g., a printed document. The longevity of an Object Instance is a topic outside the scope of this immediate paper, given our focus metadata. Instances of meta-level 1, 2 and 3 may be implemented as a document-like instance, a database record, or an XML instance encoded in a metadata description standard, e.g., RDF. Metadata preservation may be done in three approaches – document preservation, database preservation and XML encoded instance preservation in accordance with requirements in each meta-level.

(3) Facet 3: Metadata Schema Components

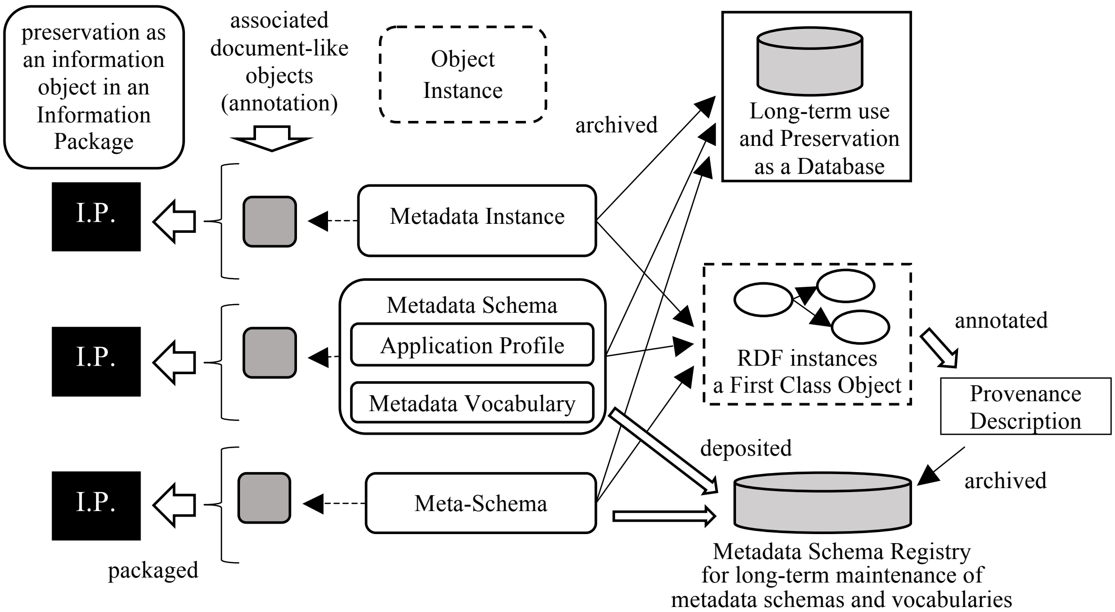

FIG.5. Metadata Entities and Preservation Options.

https://doi.org/10.23106/dcmi.952137567

This facet is for metadata schema and meta-schema entities – application profiles, metadata vocabularies for certain domains and domain-neutral standards for metadata description such as XML and namespaces. For example, Description Set Profiles and Domain Models of Singapore Framework are encoded in a formal scheme and other components are expressed as natural language texts. Preservation strategy of these components depends on the entity format types.

(4) Facet 4: Dynamic Entities

Cases whereby metadata terms are removed from or added to a metadata schema, and when a new metadata schema is created by aggregating two existing schemas. In such cases, we often create a mapping table to map an old schema to a new schema. The mapping tables should be recorded as well as those schemas.

(5) Facet 5: Documentation

Any document entities and activities may be recorded for use in the future. The document entities have to be preserved as a part of metadata preservation. Contextual information, which may not be explicitly described in metadata schema entities, may be found in the documentation entities.

3.2. Related Research – Digital Preservation, Archiving and Provenance Description

Provenance is key for the maintenance of metadata. OAIS defines a package-based model for preserving digital objects. The OAIS model is applicable to any digital object of the meta-levels in the case we preserve it in a package. However, this model does not seem adequately support active entities directly accessed from other dynamic entities that change content.

Web archiving is a related area for this study. Internet Archive is a very large provider of archived Web resources. Memento defines a framework to keep old URIs consistently usable (Van de Sompel, Nelson, & Sanderson, 2013), and provides an exemplary way that may assist our work. Allowing the temporal tracking and overall path noting the history of a schema is significant for metadata longevity. This type of work can also inform URI management – one of the fundamental requirements. Keeping metadata as a Web page may be within the scope of Web archiving, but consistency management of metadata schemas is out of their scope.

In this paper, we focus on longevity of metadata entities as a first class object in the LOD environment, so that we do not focus on the longevity of metadata as a packaged object or Web page. Longevity of databases which store metadata entities is also out of the focus of this study.

Provenance descriptions track changes of metadata instances and metadata schemas. Provenance can include a series of descriptions of events for metadata instances, metadata schemas, vocabularies, and other related entities. An event description may be associated with metadata objects, e.g., agents, reasons, activities, etc. Provenance description is an important issue for metadata (Eckert, 2013). W3C has defined a provenance description model for the Web, i.e., W3C PROV (Groth & Moreau, 2013). The authors have developed a provenance description model for DSPs based on W3C PROV and RDF. This model defines Addition, Deletion and Revision activities based on W3C PROV and the DCMI Application Profile. We have experimentally applied the model to describe provenance among the versions of DPLA Metadata Application Profiles (Li, Nagamori, & Sugimoto, 2015; Li & Sugimoto, 2016).

Here, we point to metadata schema registries as services that keep metadata schemas interpretable by both machines and humans over time. There are registry services that provide metadata vocabularies and terms.[2] However, these services do not provide functions for storing a metadata schema defined for an application. DCAPs as a conceptual model for application

[2] Metadata schema and vocabulary services. All sites retrieved May 27, 2016
Linked Open Vocabularies (LOV), from http://lov.okfn.org/dataset/lov
Schema.org, from http://schema.org/docs/schemas.html
Open Metadata Registry, from http://metadataregistry.org/about.html

https://doi.org/10.23106/dcmi.952137567

metadata schema are being widely recognized. However, schema registry for DCMI AP is still a nascent area of research.

4. Risks in Metadata Longevity

4.1. Aspects for Longevity Risks of Metadata

Longevity of metadata and metadata schemas reveals a number of associated risks, which are reviewed in this section. What is risk for metadata longevity? Where do longevity risks exist? How can we find the risks? These are fundamental research questions for this study. The following sentences are typical risks:

Risk-A: Metadata schema of this metadata instance is unknown.
Risk-B: We cannot display a metadata creation guideline documents correctly.
Risk-C: An application uses a standardized metadata vocabulary but the name of the vocabulary is unknown.
Risk-D: Definition of a metadata term is not identifiable by URI given to the term.

These risks occur for many different reasons, e.g., insufficient documentation, insufficient information transfer when downloading metadata, inappropriate maintenance of documents, non-persistent URI, etc. We focus on the risks from three aspects in the following sections, (1) metadata instances, (2) application-based schemes, and (3) shared vocabularies and terms. The Singapore Framework, RDF and URI are the underlying framework for the discussion in this paper. As URI is the base identification scheme of any instances, persistency of URI is a fundamental requirement for the longevity of metadata. We discuss this issue later in this paper.

4.2. Metadata Instances

Temporal interpretability of metadata instances depends on the availability of the metadata schemas for those instances over time. Metadata instances can be classified into two classes – primary and secondary metadata. The primary metadata is a metadata instance stored in a metadata database, or embedded in a source resource such as books and Web pages. Non-primary metadata instances are secondary metadata, which may be created by copying and editing primary metadata. Provenance information of the secondary metadata is crucial in order to keep its consistency but is not always created.

When a metadata database is created by re-organizing downloaded metadata, the metadata instances become the primary data, with the underlying scheme guiding the new metadata database organization. In this case, provenance description should be included as contextual information of the newly created database, e.g., information about the source resource and schema.

Longevity risks of metadata instances depend on whether a metadata is primary or secondary. Longevity risks of primary metadata instances exist in the management of their metadata schemas. However, in the case of secondary metadata, their longevity risks are not only in the same factors but also in keeping provenance information of the metadata instances consistent. In the reality, provenance information may not be recorded in the most cases of copying metadata. This means that there is no way to keep those metadata consistent over time.

4.3. Application Profiles – Structural Constraints and Syntax

Long-term maintenance of metadata schemas is crucial for metadata longevity, a key point in this paper. The Singapore Framework (SF) offers support here, with five components – Functional Requirements, Domain Model, Description Set Profile, Usage Guidelines, and Encoding Syntax Guidelines. Although many existing metadata schemas do not specifically define each SF conformant application profile component, many schemas have these aspects integrated. This paper uses SF as the basis for discussion because it clearly states the aspects that should be included in the definition of a metadata schema for any application. SF explicitly

https://doi.org/10.23106/dcmi.952137567

defines dependency among the five components of SF and relationships between DSP and metadata vocabularies. This clear separation of the metadata schema components helps the maintenance of metadata schemas. From this viewpoint, we can reduce risks of metadata longevity by using SF for metadata schema definition.

Maintaining these five components depends on the formality of their descriptions. Natural language text documents should be maintained as a textual file, formal descriptions should be maintained in accordance with the formal schemes of the description, i.e., UML and RDF. Metadata schema instances presented in RDF may be stored as a set of triples, which means that we would need to maintain them as a set of metadata instances but not as a textual document.

4.4. Metadata Vocabularies and Terms

SF defines metadata vocabularies in a layer beneath the application profiles. Metadata vocabularies defined for a domain but neutral to particular applications are defined in that layer. This metadata vocabularies layer is defined above a layer in which domain neutral constructs for implementing metadata are defined, e.g., RDF, XML and other Internet standards. This separation is the fundamental issue from the viewpoint of metadata interoperability, meaning not only interoperability across communities, but also interoperability over time. URI, RDF and OWL are the standard schemes to define the terms in the LOD environment. Definitions of the terms and vocabularies in these schemes may be presented in a document, stored in a database, or realized as a first class object encoded in XML/RDF. From the viewpoint of machine interpretability of metadata, keeping the database or the first class object accessible is a key. From the viewpoint of human readability, any form of these forms is acceptable.

Versioning information for both terms and vocabularies is key for supporting long-term maintenance. There are different cases of versioning – versioning of DCMI terms is term-basis but versioning of decimal classifications is vocabulary-basis. As precise meaning of a metadata term depends on its versions, keeping version information is crucial for long-term use of metadata.

Maintenance authorities for metadata standards are supposed to be stable but may disappear over time. There are metadata schema registries and Web sites, which provide definitions of metadata vocabularies and terms. Multiple copies of descriptions are a double-edged sword – on one hand multiple copies are robust for keeping the content safe for the future, but on the other hand, multiple copies may cause troubles in consistent maintenance of versions.

4.5. Other Factors

Persistency of identifiers: URI, commonly used to identify Web and Internet resources, is used to identify metadata schema instances, metadata terms, and vocabularies in the LOD environment. Metadata term URIs are sometimes used, further, for term definitions, and making them accessible by the URI, i.e., resolvability of URI. Persistency of URL is the fundamental requirement for the longevity of metadata terms and vocabularies. Resolvability of URI is not mandated for the persistency of URI.

Metadata mapping tables: Metadata mapping tables are often created for many purposes – merging two or more metadata datasets, metadata harvesting, federated search, and so forth. Metadata mapping table is a crucial resource in the long time line of metadata maintenance. As a metadata mapping table is a kind of metadata, i.e., description of metadata mapping, we can apply the model proposed in this paper for the metadata mapping table.

Contexts: Contextual information of metadata and metadata schemas is crucial for their longevity. However, it is hard to describe the contextual information perfectly. Every metadata schema designed for an application has its contexts that may or may not be described as a part of the schema. For instance, descriptions about selection process of metadata terms from standard vocabularies are crucial to know the context of the metadata schema and to correctly interpret

https://doi.org/10.23106/dcmi.952137567

metadata. In theory, it is feasible to keep such description because SF includes components that may include contextual information. However, those descriptions tend to be not provided

5. Discussions – Provenance Description for Risk Management for Metadata Longevity

Over the last few years there have been national and international calls targeting the archiving and preservation of research data. National and global agencies require data deposition and they require researchers with funding to make their research data accessible, and reusable. In connection with this significant development, it seems equally if not more important to call for the publication, preservation, and archiving of metadata standards, and the levels and facets reviewed above, to support long-term interpretability.

Metadata schema documentation is required for proper maintenance of metadata instances. We learned that not many LOD datasets provide information about their metadata schema (Honma, Tanaka, Nagamori, & Sugimoto, 2014). Proper versioning information of metadata schemas is necessary but application schemas tend to loose consistent maintenance of the information. These sorts of information should be maintained in a machine interpretable form rather than a human readable form from the viewpoint of keeping metadata machine interpretable over time.

Formal description of provenance of metadata schemas is essential to cope with this maintenance problem. The authors have developed a provenance description model based on DCMI Description Set Profiles and W3C PROV in order to formally express provenance of description set profiles in RDF and use the description for automated consistency checking.

Another crucial issue is to use metadata schema registries for long-term maintenance of metadata schemas. Current metadata schema registries and related services provide current information about metadata vocabularies. MetaBridge provides a function to store/provide description set profiles in RDF but it has only a simple versioning function to replace an old version by a new version. Our study on provenance description of DSP shows that RDF-based provenance description helps maintenance of metadata schemas. It is necessary to be able to identify a version of a metadata term used in a DSP. A DSP is linked to a metadata term by URI of the term. However, as URI does not convey any version information of the term, we need to use provenance information of the DSP to maintain the linkage between the DSP and its corresponding version of the metadata term, which may be implemented in a metadata schema registry for metadata schema preservation.

URI is not persistent but metadata terms have to be consistently identifiable over time in the LOD environment. Thus, persistency of URI is essential for long-term maintenance of metadata schemas. Persistent URIs rely on persistent URI resolvers. Metadata schema may be able to keep the definition of a metadata term associated with its URI. LOCKSS metaphor may work for keeping definitions of metadata terms and application profiles consistent. We need collaborating metadata registries for keeping metadata schemas safe over time.

6. Concluding Remarks

In this paper, we have discussed risks in metadata longevity by analyzing various entities of metadata from different aspects. We have proposed to use provenance description and metadata schema registry for the risk management in this paper. Preservation of digital objects is a well-known research topic for digital curation and archiving. Conventional digital preservation is oriented to preservation of a primary entity such as documents, games, pictures, etc. Metadata preservation has been discussed within the scope of conventional digital preservation. However, in the LOD environment, there are many new issues for long-term and consistent use of metadata. Conventional OAIS-based preservation is a frozen preservation because we need to retrieve and open information packages. On the other hand, in the LOD environment, we will have many metadata instances stored in our files as a first class object. Keeping these instances consistently interpretable is crucial in such an environment, which may be called unfrozen archive. The

https://doi.org/10.23106/dcmi.952137567

development of the concept of DCAP has contributed to clarify requirements for metadata interoperability. However, temporal interoperability of metadata is still not well studied yet.

It is widely known that term definitions and term usage changes over time, and can further change due to domain use. In this paper, we mentioned this issue as contexts. We understand that it is important to include the context explicitly in the metadata schema management process but it is challenging to explicitly and consistently describe the contexts based on the underlying data model of LOD. Management of contextual information for the longevity of metadata and metadata schemas is a fundamental issue but is left for our future study.

Acknowledgements

This work was supported in part by JSPS Kaken Grant-in-Aid for Scientific Research (A) #16H01754 and #25240012.

References

CCSDS. (2012). Reference Model for an Open Archival Information System (OAIS). Retrieved May 27, 2016, from http://public.ccsds.org/publications/archive/650x0m2.pdf.

Eckert, Kai. (2013). Provenance and Annotations for Linked Data. Proceedings of the International Conference on Dublin Core and Metadata Applications, 2013, 9-18. Retrieved May 27, 2016, from http://dcpapers.dublincore.org/pubs/article/viewFile/3669/1892.

Greenberg, Jane. (2005). Understanding metadata and metadata schemes. Cataloging & classification quarterly, 40(3-4), 17-36.

Groth, Paul, and Luc Moreau (Eds.). (2013). PROV-Overview. Retrieved May 27, 2016, from https://www.w3.org/TR/prov-overview/.

Heery, Rachel, and Manjula Patel. (2000). Application Profiles: Mixing and Matching Metadata Schemas. Ariadne, 25. Retrieved May 27, 2016, from http://www.ariadne.ac.uk/issue25/app-profiles.

Honma, Tsunagu, Kei Tanaka, Mitsuharu Nagamori, and Shigeo Sugimoto. (2014). Extracting Description Set Profiles from RDF Datasets using Metadata Instances and SPARQL Queries. Proceedings of the International Conference on Dublin Core and Metadata Applications, 2014, 109-118. Retrieved May 27, 2016, from http://dcpapers.dublincore.org/pubs/article/view/3706.

Lagoze, Carl. (1996). The Warwick Framework: A container architecture for aggregating sets of metadata. D-Lib Magazine, July/August,1996. Retrieved May 27, 2016, from http://www.dlib.org/dlib/july96/lagoze/07lagoze.html.

Li, Chunqiu, Mitsuharu Nagamori, and Shigeo Sugimoto. (2015). Temporal Interoperability of Metadata: an Interoperability-based View for Longevity of Metadata. Proceedings of the 6th International Conference on Asia-Pacific Library and Information Education and Practice, 2015, 199-209.

Li, Chunqiu, and Shigeo Sugimoto. (2016). Provenance description of metadata for long-term maintenance of metadata application profiles, unpublished draft.

Nagamori, Mitsuharu, and Shigeo Sugimoto. (2004). A Metadata schema framework for functional extension of metadata schema registry. Proceedings of the International Conference on Dublin Core and Metadata Applications, 2004. Retrieved May 27, 2016, from http://dcpapers.dublincore.org/pubs/article/viewFile/764/760.

Nagamori, Mitsuharu, Thomas Baker, Tetsuo Sakaguchi, Shigeo Sugimoto, and Koichi Tabata. (2001). A Multilingual Metadata Schema Registry Based on RDF Schema.Proceedings of the International Conference on Dublin Core and Metadata Applications, 2001, 209-212. Retrieved May 27, 2016, from http://dcpapers.dublincore.org/pubs/article/view/660.

Nagamori, Mitsuharu, Masahide Kanzaki, Naohisa Torigoshi, and Shigeo Sugimoto. (2011). Meta-Bridge: A Development of Metadata Information Infrastructure in Japan. Proceedings of the International Conference on Dublin Core and Metadata Applications, 2011, 63-68. Retrieved May 27, 2016, from http://dcpapers.dublincore.org/pubs/article/view/3632.

Nilson, Mikael, Thomas Baker, and Pete Johnston. (2008). The Singapore Framework for Dublin Core Application Profiles.Retrieved May 27, 2016, from http://dublincore.org/documents/2008/01/14/singapore-framework/.

Nilson, Mikael, Thomas Baker, and Pete Johnston. (2009). Interoperability Levels for Dublin Core Metadata. Retrieved May 27, 2016, from http://dublincore.org/documents/2009/05/01/interoperability-levels/.

Van de Sompel, Herbert, Michael L. Nelson, and Robert Sanderson. (2013). HTTP Framework for Time-Based Access to Resource States-Memento. No. RFC 7089 (Status: Informational). Retrieved August 31, 2016, from https://tools.ietf.org/html/rfc7089.

https://doi.org/10.23106/dcmi.952137567