DCPAPERS
Proc. Int'l Conf. on Dublin Core and Metadata Applications 2014

Provenance Description of Metadata using PROV with PREMIS for Long-term Use of Metadata

Chunqiu Li
Graduate School of Library, Information and Media Studies,
University of Tsukuba, Japan
licq.chunqiu@gmail.com

Shigeo Sugimoto
Faculty of Library, Information and Media Science,
University of Tsukuba, Japan
sugimoto@slis.tsukuba.ac.jp

Abstract

Provenance description is necessary for long-term preservation of digital resources. Open Archival Information System (OAIS) and Preservation Metadata: Implementation Strategies (PREMIS), which are well-known standards designed for digital preservation, define descriptive elements for digital preservation. Metadata has to be preserved as well as primary resource in order to keep the primary resources alive. However, due to the changing technology and information context, not only primary digital resources but also metadata are at risk of damage or even loss. Thus, metadata preservation is important as well as preservation of primary digital resources. Metadata preservation is a rather new research topic but critical for keeping metadata about preserved resources consistently over time. This paper focuses on provenance as an important issue in digital preservation. It discusses provenance description based on two major metadata standards—PROV and PREMIS. The goal of this study is to clarify a model for describing provenance for metadata preservation. This paper first describes some well-known standards—OAIS, PREMIS, PROV, and so forth, and then discusses a novel model of provenance description based on the PROV Ontology (PROV-O) and PREMIS OWL Ontology. The paper gives provenance description examples using PROV-O and PREMIS OWL Ontology respectively. Based on analysis and mapping among the basic classes of the PROV-O and PREMIS OWL Ontology, we propose an integrated, merged model. We discuss metadata schema provenance and some other open issues.

Keywords: digital provenance; metadata provenance; metadata longevity; PROV; PREMIS

1. Introduction

Metadata plays crucial roles in long-term use of digital resources and digital preservation. Damage or loss of metadata over time may cause serious problems in the long-term use of digital resources. Metadata schema changes may cause inconsistency in the use of metadata, which is also a risk for the long-term use of digital resources. Due to the high cost of re-creation of metadata, longevity of metadata is an important issue for long-term use of digital resources. Metadata schema, which defines a set of terms, structure of metadata instances and some related characteristics of metadata instances, has to be maintained as well as the metadata instances over time.

Provenance information is necessary for long-term use and preservation of digital resources. Provenance is a fundamental principle of archives (Pearce-Moses, 2005) and keeping provenance of every archived item is a fundamental archival function. Open Archival Information System (OAIS) and Preservation Metadata: Implementation Strategies (PREMIS) are widely accepted standards for digital preservation. They include provenance descriptions as primary information. Both OAIS and PREMIS state the importance of provenance description for preservation (Consultative Committee for Space Data System, 2012; PREMIS Editorial Committee, 2012).

As provenance is a general concept, provenance description is not limited to preservation of digital objects. There are several standards for provenance description such as PROV developed

https://doi.org/10.23106/dcmi.952136486

by the World Wide Web Consortium (W3C). PROV is defined as a general, high-level standard for provenance, whereas provenance descriptions in OAIS and PREMIS are defined for preservation of information resources. The primary goal of this paper is to study a model for describing provenance of metadata by combining PROV and PREMIS.

This study is primarily aimed at understanding the underlying model for the provenance of metadata for long-term use of metadata—in other words, the interoperability of metadata over time. Metadata preservation is purposed to assure the persistent availability, understandability, and usability of metadata. To make metadata interpretable correctly in the future context is a main goal of metadata preservation. Longevity of digital objects is well known as a crucial issue for the further progress of the networked information society. The technology standards for longevity of digital objects are applicable to the metadata instances because the metadata instances are mostly, but not necessarily, digital objects—e.g., an XML text file and an Excel file. Longevity of digital objects does mean that the objects can be correctly rendered over time. However, it does not necessarily mean that future users can properly understand the content of the object. For example, a table stored in an Excel file may be rendered over time but the attributes of the table cannot be properly understood without proper description of the meaning of the attributes and values. This table example shows a typical problem in metadata preservation—metadata as a digital object may be preserved; but metadata as a semantically meaningful entity may be lost. Even if a metadata instance is encoded in XML and stored in a plain-text file, semantics of XML elements may be lost if the meanings of the tags in the XML text are not properly preserved. Thus, preservation of metadata is not same as preservation of digital objects.

Metadata registries, which store the definitions of metadata terms and controlled vocabularies and provide them over the Internet, have crucial roles in making the metadata terms and controlled vocabularies usable across communities and over time. Moreover, maintaining application profiles is a crucial function for long-term use of metadata. However, management and use of provenance information of the metadata terms and vocabularies has not been discussed except for versioning and its control. Provenance of application profiles has been neither well discussed nor well recognized.

Based on this understanding about state-of-the-art of metadata provenance, this paper discusses a basic model for metadata provenance. The proposed model is defined based on PROV Ontology (PROV-O) and PREMIS OWL ontology. The rest of this paper is organized as follows. Section 2 describes provenance for the discussion in this paper followed by surveys of some major models and standards for preservation of digital resources and provenance description. Section 3 discusses the provenance description using PROV-O and PREMIS OWL ontology respectively. Section 4 shows mapping between PROV-O and PREMIS OWL ontology and proposes a novel model to combine them for provenance description oriented to digital preservation. Section 5 states metadata schema provenance issues for metadata longevity. Finally, Section 6 concludes the paper.

2. Survey of Provenance Description Standards and Models

2.1. Digital Provenance and Metadata Provenance

We discuss provenance from the dual viewpoints of digital object provenance and that of metadata. Digital provenance and metadata provenance in this paper are defined as follows:

Digital provenance is chronology or chronological information related to management of a digital object. Digital provenance typically describes agents responsible for the custody and stewardship of digital objects, key events that occur over the course of the digital object's life cycle, and other information associated with the digital object's creation, management, and preservation (PREMIS Editorial Committee, 2012)—e.g., the organization responsible for eBook.

https://doi.org/10.23106/dcmi.952136486

Based on the definition above, we can define ***metadata provenance*** as chronology or chronological information about metadata, typically responsible agents, influencing actions, associated events and other related information about metadata over its lifecycle. Provenance about metadata schema is also metadata provenance, e.g., actions and events in the revision process of metadata schema, and so forth.

It is important for memory institutions to record and provide provenance information of their holdings. W3C Provenance Incubator Group listed provenance-related use cases, which include provenance in cultural heritage (W3C Provenance Incubator Group, 2010). Europeana provides access to resources held at cultural heritage institutions throughout Europe. Europeana is a use case of metadata provenance, in which metadata provenance is represented via Europeana Data Model using the OAI-ORE model (Eckert, 2012).

The paragraphs below summarize digital provenance and metadata provenance from the viewpoint of long-term use of digital objects:

(1) Metadata of preserved resources has to be consistently interpretable over time. It has to be recognized that preservation policy and environment of preserved resources may change over time and metadata interpretation may be affected by the changes. For example, in the case of recordkeeping, digital provenance could provide information about the origin, e.g., where, when, by whom, and how a resource was created and who are the successors of the preserved resource. This information will contribute to the interpretation of metadata by users in the future.

(2) Metadata provenance describes and keeps track of responsible agents, influencing actions, associated events that caused a change(s) in metadata. Change history of a metadata schema used in a service is crucial to keeping track of changes to metadata instances created based on that schema. Therefore, provenance of a metadata schema is crucial to keeping metadata correctly and consistently interpretable and may include change history of the schema as well as relationships to other entities such as base standards and system requirements.

2.2. Digital Preservation Standards—OASIS and PREMIS

The OAIS Reference Model is a widely used model for archiving and preserving digital resources. Provenance information in OAIS is defined as the history of the Content Information, which describes the origin of and changes on an archived resource, and agents who hold custody since its origination (Consultative Committee for Space Data System, 2012). The provenance description is a part of Preservation Description Information (PDI), and documents evolutionary processing history associated with the Content Information over its complete life cycle.

PREMIS is a widely used international metadata standard for the preservation of digital objects. The PREMIS Data Model defines five *Entities* for digital preservation, which are *Intellectual Entity*, (*Digital*) *Object*, *Event*, *Agent*, and *Right*. Documentation of actions on a digital object is critical for the maintenance of the object. The documentation, i.e., metadata about the actions, is aggregated as an *Event*. Thus, *Event* is crucial component for provenance description associated with *Object*. PREMIS Data Dictionary defines a set of descriptive elements of the five *Entities*. Those elements are called semantic units. Some of the semantic units associated with an *Event* record changes to a preserved digital object (PREMIS Editorial Committee, 2012). PREMIS OWL ontology defines classes and properties to describe preservation metadata in RDF.

2.3. Provenance Models—W3C PROV, Open Provenance Model and others

W3C PROV: The Provenance Working Group at W3C has published PROV family of documents, including the PROV Data Model (PROV-DM), PROV-O and so forth. The working group aims at the inter-operable interchange of provenance information in heterogeneous environments such as the Web. PROV-DM is a conceptual data model, which defines a set of concepts and relations to represent provenance (Moreau et al., 2013). PROV-O defines a set of

https://doi.org/10.23106/dcmi.952136486

classes and properties as an OWL2 ontology allowing mapping PROV-DM to RDF (Lebo et al., 2013).

Open Provenance Model (OPM): OPM is a research result of the International Provenance and Annotation Workshop (IPAW). Based on the OPM Core Specification (v1.1), the OPM is designed to meet six requirements, including: exchange of provenance information between systems, representation of provenance for any "thing" and so forth (Moreau et al., 2010). OPM Vocabulary (OPMV), OPM OWL Ontology (OPMO) and OPM for Workflows (OPMW) are defined pertaining to OPM. OPMV as an OWL-DL ontology designed to assist the interoperability between provenance information on the Semantic Web and to support provenance descriptions for datasets beyond those in the Web of Data (Zhao, 2010). OPMO as an OWL ontology allows full expressivity of OPM concepts and supports inferencing (Moreau et al., 2010). OPMW is also OWL-DL ontology developed to represent abstract workflows and workflow execution traces. OPMW extends and reuses OPM's core ontologies. In the latest release, OPMW also extends PROV to represent scientific processes (Garijo and Gil, 2014).

Others: W7 model was developed o represent the semantics of data provenance in which provenance is conceptualized as a combination of seven interconnected elements including "what (occurring event)", "how (action leading to event)", "who (involved individuals or organizations)", "when (time of event)", "where (location of event)", "which (software or instrument that was used)" and "why (reason for why event happened)" (Liu, 2011). A Vocabulary for Data and Dataset Provenance (Voidp) defines terms to describe provenance relationships of data in linked datasets (Omitola et al., 2011). Provenance Vocabulary (PRV) as an OWL-DL ontology defines classes and properties for describing provenance of linked data on the Web. PRV is a domain specific specialization of PROV-O. It is notable that PRV defines terms for both data creation and data access (Hartig and Zhao, 2012). Provenance, Authoring and Versioning Ontology (PAV) is designed for the capture of essential descriptions for tracking the provenance, authoring and versioning of web resources (Ciccarese et al., 2013). BBC Provenance Ontology is designed to capture data about the provenance of data in an RDF Triple Store (BBC, 2012). Provenir Ontology (PO) defined in OWL-DL describes the classes and the properties to represent provenance metadata in eScience (Sahoo and Sheth, 2009).

2.4. Discussion on Provenance Description Standards and Models

Provenance may be about any resource, such as documents, rare books, web pages, datasets, transaction execution records, etc. This means that we need to use an appropriate vocabulary or vocabularies for provenance description in accordance with the type of resources and archiving purposes. Provenance description in OAIS and PREMIS is primarily for digital preservation whereas those standards shown in section 2.3 are defined for other purposes. Most of the ontologies are OWL-based; thus, the OWL-based definitions are useful for the reference to term definitions and reasoning of provenance.

PROV is designed generally and comprehensively for provenance description, referring to representation, interchange, query, access, and validation of provenance. PREMIS is widely used for digital preservation where provenance description is an important component. This study is primarily aimed at definition of a model of metadata provenance description for long-term use of metadata. We use PROV and PREMIS as a basis for general provenance description and provenance description for preservation in this study. Hereafter we will refer to PROV and PREMIS instead of PROV-O and PREMIS OWL Ontology unless we need to explicitly state the ontology.

3. Provenance Description Scenarios for Preservation

We use PROV-O and the PREMIS OWL Ontology to describe provenance information created during the lifecycle of digital objects and their metadata. Migration is a widely used method to assure digital objects accessible and usable over time. This section presents some instnaces of

https://doi.org/10.23106/dcmi.952136486

provenance description about the format migration shown below, referring to the *generationActivity*/*creationEvent* occurred to *Digital Object A*, responsible *Agent*, related date time, and also the derivation of *Digital Object A* in Format X to *Digital Object B* in Format Y via *migrationActivity* which caused the format change, and so forth.

3.1. Description of Activity and Event

Figure 1 shows a *generationActivity* leading to the generation of *Object A* by using PROV. The *generationActivity* (started at dateTime1, ended at dateTime2) resource is directed to *Object A*, which is linked to a generation Date-Time literal. PREMIS uses preservation-specific value vocabularies defined by Library of Congress. Those vocabularies provide terms expressed in SKOS vocabulary, e.g., *EventType*, *AgentType* and *RelationshipType.* Likewise, Figure 2 shows a *creationEvent* associated with *Object A* and the *creationEvent* happening during a period from dateTime1 to dateTime2. Meanwhile, the Figure also presents the *creationEvent* is linked to an *EventOutcomeInformation* resource, an *EventType* resource, and *EventDateTime* literal.

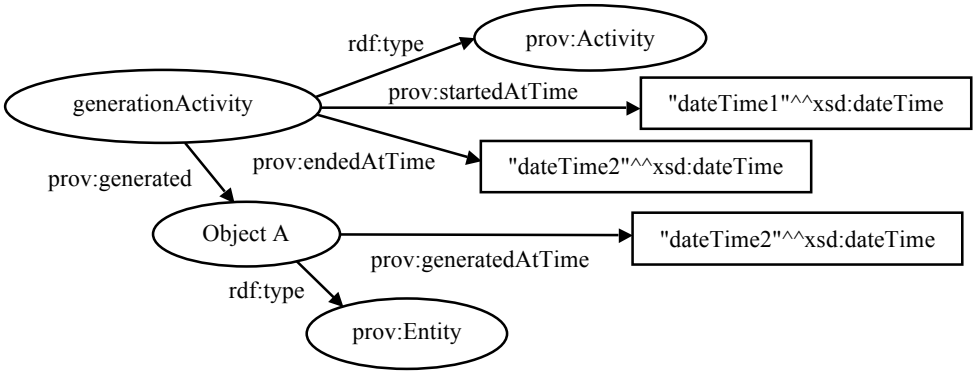

FIG.1. Provenance graph of generationActivity happened on Digital Object A using PROV

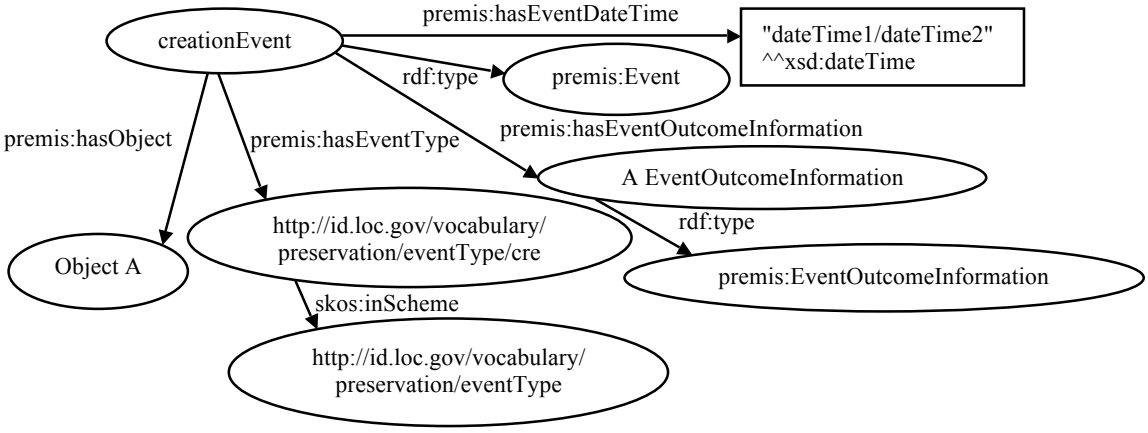

FIG.2. Provenance graph of creationEvent occurred to Digital Object A using PREMIS

3.2. Description of Responsible Agent

As shown in Figure 3, *Object A* is connected with a *Person* by property *wasAttributedTo* defined in PROV. The *generationAcitity* is linked to that *Person* via property *wasAssociatedWith*, from which we know the *Person* holds a responsibility for the generation of *Object A*. In PREMIS, *Agent* influences *Object* through *Event*. That is, *Agent* is not directly connected to *Object* as shown in Figure 4. However, PROV allows *Agent*, *Entity* and *Activity* to be related with each other directly.

https://doi.org/10.23106/dcmi.952136486

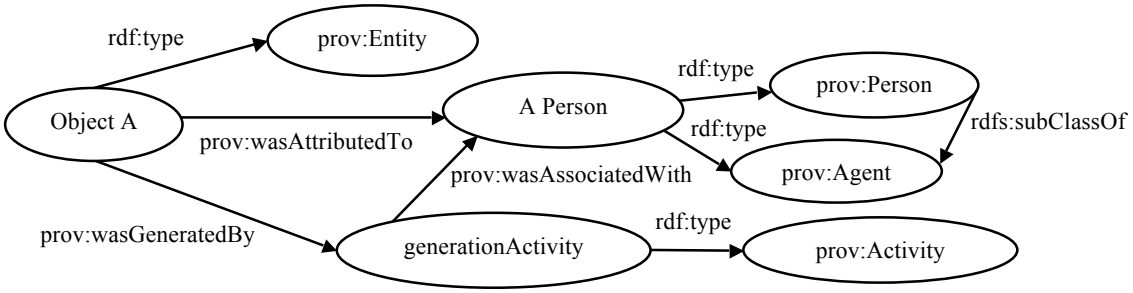

FIG.3. Provenance graph of Agent responsible for the generation of Digital Object A Using PROV

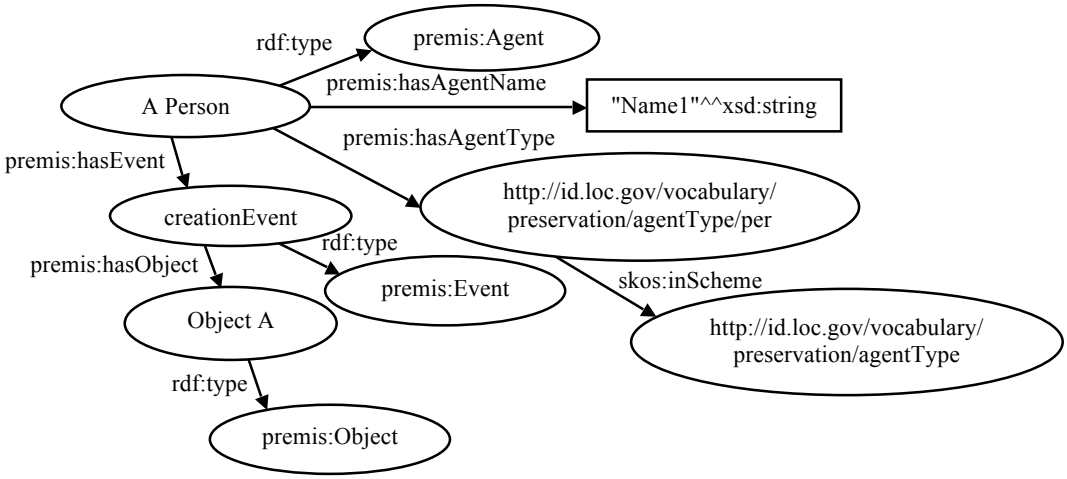

FIG.4. Provenance graph of Agent responsible for Event using PREMIS

3.3. Description of Relationships between Entities and Relationships between Objects

PROV describes the relationship between entities with the properties *wasDerivedFrom*, alternateOf, *specializationOf*, *wasQuotedFrom*, *wasRevisionOf*, *hadPrimarySource*, *hadMember*. Figure 5 shows that Object A is the primary source of Object B using PROV. PREMIS holds two types of relationship between Objects, including structural and derivation relationships defined in a SKOS vocabulary by Library of Congress. Using PREMIS, Figure 6 shows the derivation relationship between Object A and Object B due to the *migrationActivity*.

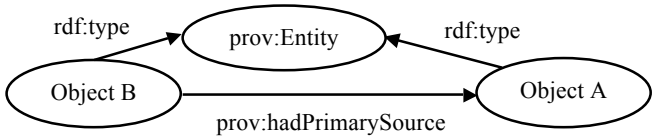

FIG.5. Derivation Relationship between Digital Object A and Digital Object B using PROV

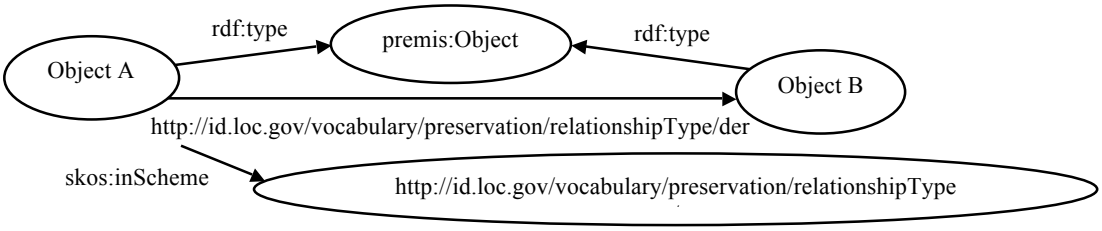

FIG.6. Derivation relationship between Digital Object A and Digital Object B using PREMIS

https://doi.org/10.23106/dcmi.952136486

Furthermore, PROV also defines relationships between *Activities* and relationships between *Agents*, whereas PREMIS does not include those relationships. Figure 7 shows the relationship expressed by property *wasInformedBy* between the *migrationActivity* and *generationActivity*, which means the *migrationActivity* used *Object A* created by the *generationActivity*.

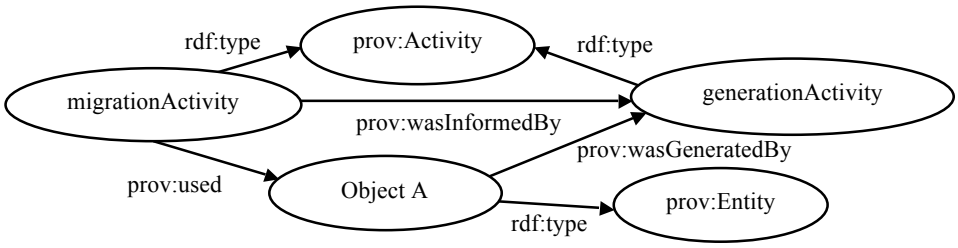

FIG.7. Relationship between Activities in PROV

4. A Merged Model for Provenance Representation by Integrating PROV-O with PREMIS OWL Ontology

4.1. Mapping of Basic Classes between PROV-O and PREMIS OWL Ontology

PROV has the three base classes, i.e., *prov:Entity*, *prov:Agent* and *prov:Activity*. PREMIS defines classes, including *premis:IntellectualEntity*, *premis:Object*, *premis:Agent*, *premis:Event*, and so forth. Based on the interpretation in PROV (Lebo et al., 2013) and PREMIS (PREMIS Editorial Committee, 2012), the paragraphs below discuss mappings between them.

premis:IntellectualEntity is a set of content items as a single intellectual unit, e.g., book, map, photograph, or database. *premis:Object* is a discrete unit of information in digital form. *prov:Entity* can be in physical or digital or conceptual or imaginary thing. We can conclude that *prov:Entity* has a broader meaning than *premis:IntellectualEntity* and *premis:Object*. Hence, we map *premis:IntellectualEntity* and *premis:Object* as subclass of *prov:Entity*.

premis:Event indicates a description about an action (or activity) impacting an *Object*. *prov:Activity* means actions or processes performed by *Agent(s)* or acted on *Entity (-ies)*. *premis:Event* is oriented to preservation actions, and only important *Events* are recorded. On the other hand, *prov:Activity* does not have limitation of action domain or types. That is, the meaning of *premis:Event* is narrower than *prov:Activity*. Therefore, we map *premis:Event* as subclass of *prov:Activity*.

premis:Agent can be a person, or an organization, or a software program/system associated with *Events* in the life of an *Object*. *prov:Agent* bears responsibility for occurred *Activity*, or the existence of *Entity*. However, their *Agent* types are almost the same. In a sense, *premis:Agent* can be seen to be equal to *prov:Agent*. And the relation can be described using *owl:equivalentClass*.

4.2. A Proposed Model Integrating PROV-O with PREMIS OWL Ontology

Both PROV and PREMIS have properties to describe provenance, and they are defined based on RDF and OWL. PROV is designed for generalized provenance description and interchange among different systems, whereas PREMIS is primarily for preservation metadata description used for digital preservation. The specialized PREMIS terms used to describe preservation could enrich expressive power of PROV. By introducing the controlled vocabulary for event types suggested in PREMIS, interoperability of *Activity* descriptions in PROV could be enhanced.

Based on the mapping shown in section 4.1, we propose a provenance description model for preservation of digital resources and metadata, by integrating the PROV with PREMIS. The merged model shown in Figure 8 introduces the *premis:Object* and *premis:IntellectualEntity* as the subclass of *prov:Entity*, *Collection*, *Bundle*, and *Plan* are also subclasses of *Entity*. Meanwhile, *premis:Event* is mapped to the subclass of *prov:Activity, premis:Agent* is equivalent to *prov:Agent*. In the Figure, the classes in PROV are written in italic, and the classes in PREMIS

https://doi.org/10.23106/dcmi.952136486

are shown with underline. Moreover, as shown in Figure 8, the relationships between classes, the generation or invalidation time of *Entity*, and the start or end time of *Activity*/*Event* can also be described via properties (written with namespace prefix, i.e., prov) from PROV.

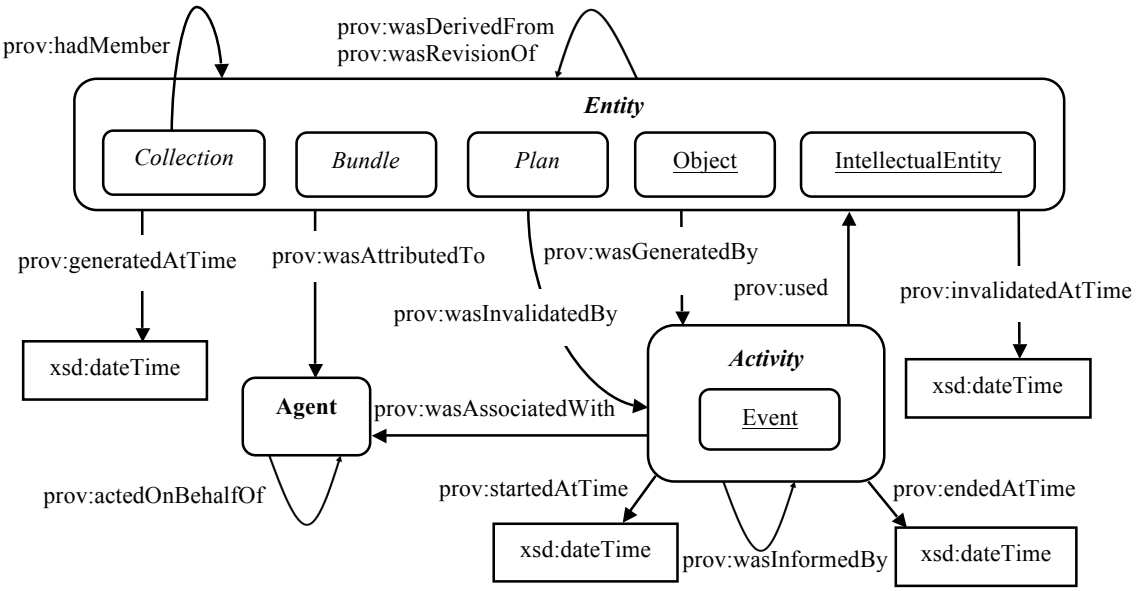

FIG.8. The merged model for provenance description oriented to digital preservation

4.3. Provenance Description Using the Proposed Model

Eckert presented the concept of Provenance Context. A Provenance Context can be seen as a Named Graph about identified resource (Eckert, 2013). Named Graph may be used for tracking provenance of RDF data, replication of RDF graphs, and versioning (Dodds and Davis, 2012). PROV allows grouping of provenance description and defines *Bundle* as a named set of descriptions (Lebo et al., 2013).

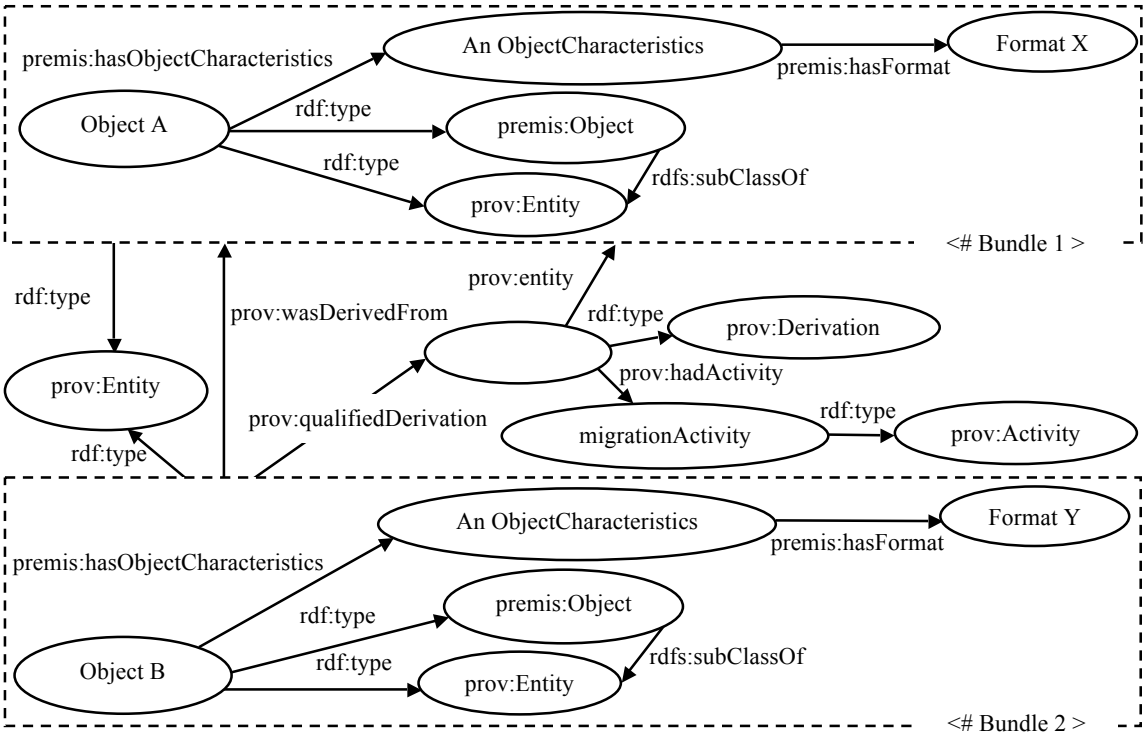

FIG.9. Provenance graph of the format change from Digital Object A to Digital Object B using Bundle

https://doi.org/10.23106/dcmi.952136486

Through the definition of *Bundle*, we can describe the provenance of *Bundle*. For the assumed example, *Digital Object A* in Format X is migrated to *Digital Object B* in Format Y. Here, we define two *Bundles*, i.e., *Bundle 1* and *Bundle 2*. *Bundle 1* and *Bundle 2* respectively describes the format feature of *Digital Object A* and *Digital Object B* as shown in Figure 9, which shows the format change caused by *migrationActivity*. As *Bundle* is an *Entity* in PROV, we can also express the derivation between *Bundle 1* and *Bundle 2*. In PROV, by using property *qualifiedDerivation*, we can qualify how *Bundle 2* was derived from *Bundle 1*. In Figure 9, *Bundle 2* is linked to a blank node through property *qualifiedDerivation*. And from the blank node, the *migrationActivity* caused the format change is expressed.

5. Provenance Description for Long-term Use of Metadata

Metadata schema longevity is a vital aspect of metadata longevity. Given to the necessity of provenance in preservation, metadata schema provenance should be documented and managed with a purpose for metadata preservation. On one hand, a metadata is a digital object, and on the other hand, a metadata is a logical data entity neutral to any particular physical representation as a digital object. There are widely accepted standards for the longevity of digital objects, e.g., OAIS and PREMIS. However, there is no well-established model or standards for the longevity of metadata as a logical data entity. In this paper, the authors propose a model for provenance description of metadata from the viewpoint of metadata longevity.

By the nature of metadata, there is meta-metadata and meta-meta-metadata which mean "data about metadata" and "data about meta-metadata". Metadata schema is a typical meta-metadata because it is a description of metadata from the viewpoint of structural and/or semantic definition. Because of the nature of metadata, meta-metadata and meta-meta-metadata are metadata.

Metadata instances are created as (1) a digital instance of metadata, e.g., a text file describing a book, a CSV file of bibliographic records, or (2) a logical data instance expressed as a self-contained digital object or embedded in a digital object, e.g., a metadata expressed as an RDF/XML instance and an RDFa expression embedded in an HTML document. In both cases, provenance is an important issue for the longevity of metadata - they require both digital object provenance and metadata provenance, i.e., metadata instance as a file and a written instance in the file.

Provenance of the metadata schema is one of the key issues for the long-term use of metadata instances. Metadata schema provenance can be categorized using DCMI application profile – (1) Vocabulary Provenance, (2) Structural Provenance (i.e., provenance of description set profiles), (3) Provenance of other components: Encoding Syntax Guidelines, User Guidelines, and Functional Requirements. Vocabulary provenance is for recording semantic change of terms. Structural provenance includes revision history of terms used in the schema as well as the revision history of structural constraints. Other provenance descriptions are crucial for readers in the future to understand contextual information to process metadata. From another viewpoint, a vocabulary mapping table created for a metadata schema mapping is a metadata instance about the metadata schema mapping, e.g., conversion from an old schema to a new schema, and merger of two schemas. Provenance description for the table should be given to record a change history of metadata terms used in the schema(s).

6. Discussion and Future Work

Although many projects have made great efforts for digital preservation, there is no efficient method proposed for metadata preservation. Metadata provenance for metadata longevity in the Semantic Web is an important issue. It is easier to collect and merge open metadata from various sources. Given to the dynamic factors, e.g., URI, linkage relation, and RDF vocabulary, the representation of provenance of metadata and metadata schema is necessary.

There is a challenge in how to make metadata provenance interoperable and semantic even preservation environment changes during a long time period. Interoperability in provenance

https://doi.org/10.23106/dcmi.952136486

description is useful for the interchange among various domains or systems. Semantic provenance is required to make the meaning of provenance easily and correctly understandable by both humans and machines. In any event, preservation context and provenance context for metadata need further research.

References

BBC. (2012).Provenance Ontology. Retrieved March 18, 2014, from http://www.bbc.co.uk/ontologies/provenance.

Consultative Committee for Space Data System. (2012,June). CCSDS 650.0-M-2.Reference model for open archival information system (OAIS), Recommended Practice, Issue 2.Retrieved March 18, 2014, from http://public.ccsds.org/publications/archive/650x0m2.pdf.

Ciccarese, Paolo, Stian Soiland-Reyes, Khalid Belhajjame, Alasdair JG Gray, Carole Goble, and Tim Clark. (2013).PAV 2.0 - Provenance Authoring and Versioning ontology. Journal of Biomedical Semantics 2013, 4:37. Retrieved March 18, 2014, from http://www.jbiomedsem.com/content/4/1/37.

Dodds, Leigh and Ian Davis. (2012, May 31). Chapter 5. Data Management Patterns. Linked Data Patterns: A pattern catalogue for modeling, publishing, and consuming Linked Data. Retrieved March 18, 2014, from http://patterns.dataincubator.org/book/named-graphs.html.

Eckert, Kai. (2012). Metadata Provenance in Europeana and the Semantic Web. Retrieved July 25, 2014, from http://edoc.hu-berlin.de/series/berliner-handreichungen/2012-332/PDF/332.pdf.

Eckert, Kai. (2013). Provenance and Annotations for Linked Data. Proceedings of the International Conference on Dublin Core and Metadata Applications.2013, 9-18.

Garijo, Daniel and Yolanda Gil. (2014, July 11). The OPMW-PROV Ontology. Retrieved July 29, 2014, from http://www.opmw.org/model/OPMW/.

Hartig, Olaf and Jun Zhao. (2012, March 14).Provenance Vocabulary Core Ontology Specification. Retrieved March 18, 2014, from http://trdf.sourceforge.net/provenance/ns.html.

Liu, Jun. (2011). W7 Model of Provenance and its Use in the Context of Wikipedia.Ph.D. Dissertation. The University of Arizona.

Lebo, Timothy, Satya Sahoo, Deborah McGuinness, Khalid Belhajjame, James Cheney, David Corsar, Daniel Garijo, Stian Soiland-Reyes, Stephan Zednik, and Jun Zhao.(2013, April 30). PROV-O: The PROV Ontology. Retrieved March 18, 2014, from http://www.w3.org/TR/prov-o/.

Moreau, Luc, Paolo Missier, Khalid Belhajjame,Reza B'Far,James Cheney, Sam Coppens,Stephen Cresswell, Yolanda Gil,Paul Groth,Graham Klyne, Timothy Lebo,Jim McCusker,Simon Miles,James Myers, Satya Sahoo,and Curt Tilmes.(2013, April 30).PROV-DM: The PROV Data Model. Retrieved March 18, 2014, from http://www.w3.org/TR/prov-dm/.

Moreau, Luc, Ben Clifford, Juliana Freire, Joe Futrelle, Yolanda Gil, Paul Groth, Natalia Kwasnikowska, Simon Miles, Paolo Missier, Jim Myers, Beth Plale, Yogesh Simmhan, Eric Stephan and Jan Van den Bussche. (2011). The Open Provenance Model Core Specification (v1.1). Future Generation Computer Systems, 27, (6), 743-756.

Moreau, Luc, Li Ding,Joe Futrelle,Daniel Garijo Verdejo,Paul Groth,Mike Jewell,Simon Miles,Paolo Missier,Jeff Pan,and Jun Zhao. (2010, October 12). Open Provenance Model (OPM) OWL Specification. Retrieved March 18, 2014, from http://openprovenance.org/model/opmo.

Omitola, Tope, Christopher Gutteridge, and Nicholas Gibbins.(2011).voidp: A Vocabulary for Data and Dataset Provenance. Retrieved March 18, 2014, from http://www.enakting.org/provenance/voidp/.

Pearce-Moses, Richard. (2005). A Glossary of Archival and Records Terminology (pp. 317). Chicago: The Society of American Archivists. Retrieved March 18, 2014, from http://files.archivists.org/pubs/free/SAA-Glossary-2005.pdf.

PREMIS Editorial Committee. (2012). PREMIS Data Dictionary for Preservation Metadata, version 2.2. July 2012. Retrieved March 18, 2014, from http://www.loc.gov/standards/premis/v2/premis-2-2.pdf.

Sahoo, S. Satya, and Amit P. Sheth. (2009). Provenir ontology: Towards a Framework for eScience Provenance Management. Retrieved March 18, 2014, from http://corescholar.libraries.wright.edu/knoesis/80.

W3C Provenance Incubator Group.(2010).Use Case Report. Retrieved July 25, 2014, from http://www.w3.org/2005/Incubator/prov/wiki/Use_Case_Report.

Zhao, Jun.(2010, October 6).Open Provenance Model Vocabulary Specification. Retrieved March 18, 2014, from http://open-biomed.sourceforge.net/opmv/ns.html.

https://doi.org/10.23106/dcmi.952136486