

Meta-Bridge: A Development of Metadata Information Infrastructure in Japan

Mitsuharu Nagamori
Graduate School of
Library, Information
and Media Studies,
University of Tsukuba,
Japan
nagamori@slis.tsukuba.a
c.jp

Masahide Kanzaki
Xenon Limited
Partners, Japan
mkanzaki@gmail.com

Naohisa Torigoshi
Infocom
Corporation
Digital Archive
System Dept.,
Japan
tori@infocom.co.jp

Shigeo Sugimoto
Graduate School of
Library, Information
and Media Studies,
University of
Tsukuba, Japan
sugimoto@slis.tsukub
a.ac.jp

Abstract

This report presents the metadata information infrastructure project funded by the Ministry of Internal Affairs and Communications, Japan. The goal of this project is to build a metadata schema registry to support sharing of metadata schemas on the Internet and promote reuse of metadata schemas and metadata interoperability. The registry is designed based on the Singapore Framework. Partners in this project include several major memory organizations, research institutions, and private sector entities in Japan. This project has collected several metadata schemas from the partners and converted them into the formal project schema based on RDF. This report first presents the background and goal of the project, and then the requirements and design of the registry system named Meta-Bridge.

Keywords: Metadata Schema Registry, Application Profile, Metadata Interoperability

1. Introduction

Many metadata schemas are used in the current Internet and Web environment. A conventional schema is designed in accordance with the requirements and standards for interoperability within a domain. However, cross-domain interoperability is still a problem. Even though there are some metadata schemas which are used commonly across domains, e.g. Dublin Core and FOAF, the barrier between domain-dependent schemas is a crucial problem from the viewpoint of metadata interoperability over the Internet.

Dublin Core Application Profiles are intended to promote metadata interoperability across domains. It is well understood that metadata schema registries play an important role in promoting interoperability. Current metadata schema registries, however, are primarily designed to share metadata vocabularies, e.g. the DCMI registry at Tsukuba (2007) and the Open Metadata Registry at JES & Co. (2006). Not many of the existing and well-known schemas are available as application profiles in Web-friendly formats such as RDF. There is no well-known metadata schema registry that can handle both metadata vocabularies and application profiles. We suggest that the lack of application profiles and registries capable of managing application profiles is a chicken-and-egg problem. Based on these thoughts, we started this project to build a schema registry for application profiles and metadata vocabularies in order to enhance metadata interoperability. In this project, we built a metadata schema registry named Meta-Bridge and a set of guidelines to create metadata schemas based on the Singapore Framework (2008). In this project, we used the existing standards of W3C and the DCMI as far as possible in order to make Meta-Bridge and related components interoperable with other systems and services.

This project was funded by the Ministry of Internal Affairs and Communications as one of the ten projects funded under the program to promote usage of the evolving ICT (Information and Communication Technology) environment, especially digital resources and books, and digital

archives. This project was primarily hosted at the Research Center for Knowledge Communities (RCKC), University of Tsukuba, Japan and carried out by Infocom and the collaborating companies (2010). The RCKC collaborated with the partners of the project, which are National Diet Library (NDL), National Institute of Informatics (NII), National Archives of Japan (NAJ), Tokyo National Museum, National Museum of Modern Arts, and the Toppan Printing Co. Ltd.

2. Objectives

2.1. Project Outline

Metadata as content available on the Internet has a lifecycle just like published content and Web resources – planning, production, use, maintenance and preservation. As shown in Figure 1, we define the lifecycle as composed of four stages – metadata schema design, metadata creation, metadata use, and maintenance and preservation of metadata. Each of these lifecycle phases is carried out in accordance with requirements of the domain of the metadata, whereas the lifecycle model is domain-neutral. Because metadata records are used for many purposes over a long period, the metadata schemas must be well maintained over time.

The paragraphs below show the sages of this project

- Develop a metadata schema registry which stores metadata vocabularies and application profiles. RDF Schema is used as the base framework to define the metadata vocabularies. The Singapore Framework and RDF are used to express an application profile, i.e. metadata schema defined for a particular application.
- Develop a set of guidelines for defining metadata schemas for those who are wanting to define a new schema. In this guideline, reuse of existing schema is strongly recommended in order to enhance metadata interoperability.
- Collect schemas from partners and other participants, and make them available on line in the standard form defined in the project.

(Note: In this report, both metadata vocabularies and application profiles are called metadata schemas. In this sense, *metadata schema* is a broader concept of a metadata vocabulary and an application profile.)

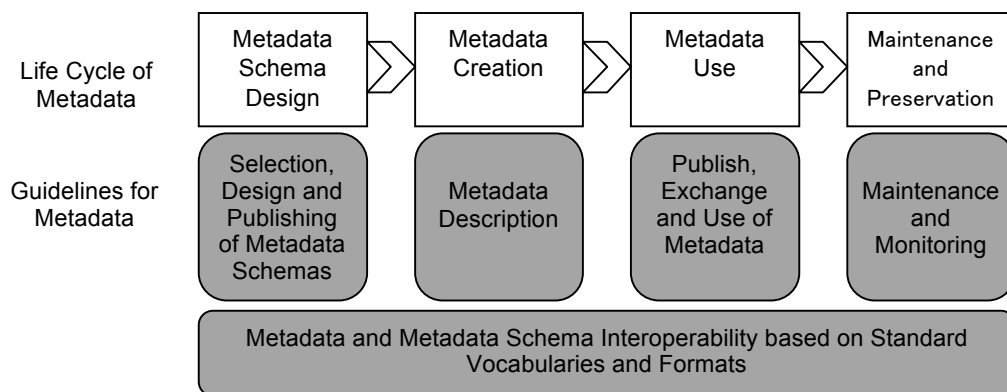


FIG. 1. Life Cycle of Metadata Schemas and Metadata

2.2. Registry Use Cases and Functional Requirements

The following paragraphs show use cases prepared to clarify functional requirements of the metadata schema registry.

Case 1: express and publish existing schemas in a canonical form.

Requirements:

- Functions to register, store and provide access to schemas

- Functions to help novice users create and register schemas – a user-friendly interface to help users express schemas in a canonical, formal form

Case 2: Search metadata schema and/or schema components in a set of existing metadata schemas for reuse in a new schema

Requirements:

- Function to search and browse schemas and their components registered in the registry
- Function to support advanced search – search by objective domains and target functions, similarity search, and user support to select a best-fit schema from retrieved schemas

Case 3: Define a new schema based on existing schemas registered in the registry

Requirements:

- Functions to find and select appropriate schemas to help design the new schema
- Functions to define a new schema by augmenting existing schema
- Functions to find, select, tailor and combine existing schemas to design new schema

Case 4: Interoperation among metadata and their services of different schemas

Requirements:

- Function to map a metadata schema into existing metadata vocabularies and/or into metadata formats
- Function to advise users to enhance metadata interoperability

Case 5: Create a new schema by mashing-up existing metadata schemas

Requirements:

- Function for a third-party user without deep knowledge about metadata to check definition of a metadata schema and its components in order to define a new schema
- Function to use registered schemas through APIs
- Function to convert registered schemas into a simple and general form

2.3. Meta-Bridge: A Registry for Metadata Schemas and Application Profiles

The Meta-Bridge design is based on the requirements analysis shown above and on the following requirements specified from the user perspective.

1. Functions for metadata schema designers to collect, register, and update information about metadata schemas
2. Functions to help content providers attach metadata to their content and define metadata schemas for their content
3. Functions to help service providers create value-added services using metadata schemas

In this project, we used RDF to define metadata schemas. We defined a scheme to define description set profiles (DSP) using OWL for description of an application profile. All of the schemas stored in this system are represented in RDF and OWL. However, because not all schema designers are familiar with RDF or OWL, this system provides those users with simplified schemes to define DSPs based on a Tab Separated Value format. This simplified DSP enables those novice users to define metadata schemas without learning formal description schemes. With Meta-Bridge, users can not only search metadata schemas and the related information, but also check relationships between registered schemas. Meta-Bridge helps users browse schemas on a Web browser and download the schemas in such machine-friendly formats as RDF/XML, Turtle, and Topic Maps.

After discussions with our partners, we agreed that metadata schemas should be recorded and maintained in the form of Excel file. Based on this finding, Meta-Bridge is given a function to upload a schema stored in an Excel files and converted into RDF. So Meta-bridge provides

metadata schemas in a human-readable format with its own user interface on a browser for human users and in RDF via APIs for software.

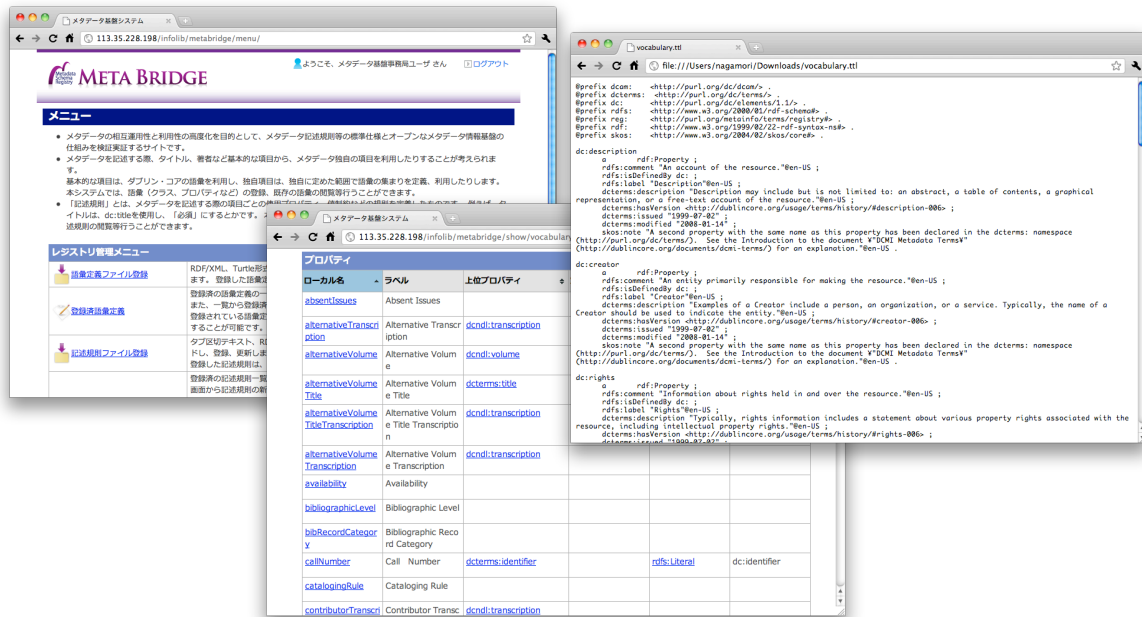


FIG. 2. Screenshots of Meta-Bridge

2.3. Guidelines for Sharing Information about Metadata

We developed a set of guidelines to create and share metadata schemas as a major product as well as in the Meta-Bridge system. The guidelines provide recommendations about tasks performed and resources used in each stage of the lifecycle of metadata. The paragraphs below show a list of guideline items with priority (High A – Low C).

Guidelines for selection, design and publishing of metadata schemas:

1. Interoperability is a prime consideration in the design and selection of schemas. (A)
2. Take care to preserve the original meaning of terms and schemes in any newly designed scheme. (A)
3. Prepare information to help others understand a newly designed metadata schema for a particular domain. (B)
4. Describe definition of metadata schema in a standard form interpretable by computers. (C)

Guidelines for metadata description:

1. Assign a URL to every resource. (A)
2. Assign a human-readable label to a resource in addition to a machine identifier. (A)
3. Describe a creator name(s) in a standard and reusable form. (B)
4. Describe date-time and location in formation of metadata in an unambiguous format.(B)
5. Assign keywords in a controlled vocabulary if applicable. (B)
6. Use a language tag or attach a sub-structure to assign “yomi” to a label text.(C)
(Note: “Yomi” in Japanese is a script attached to a target script to show a proper pronunciation of the target script. Yomi is usually given in syllabic characters, i.e.

Hirakana or Katakana, and the target text is not limited but usually Kanji (Chinese characters). (C)

7. Literal values for a particular data type or in a particular language should be used consistently and declared as part of metadata schema. (C)

Guidelines to publish, exchange and use metadata:

1. Use RDF to publish metadata as a standard on the Web in a standardized way.(A)
2. Declare metadata schema(s) to assure that metadata is assigned consistently and with precision. Revise metadata or metadata schemas if inconsistency or any critical problem is found. (B)
3. Do not lose consistency of a metadata schema or functionality of metadata when the schema is published. Dumbing-down should be done by users in accordance with their domain needs. It is a good idea to provide simple sample values for principal properties. (B)

Guidelines for maintenance and monitoring:

1. Provide administrative information for metadata schema maintenance and keep versions of schemas appropriately. (A)
2. Attach administrative information to metadata instances, e.g. metadata creator, date, and base schema. (A)
3. Maintain provenance information if a set of metadata is gathered and stored. (B)
4. Register metadata schemas in an open metadata schema registry to help users find schemas and also to properly maintain versions of the schema. (B)
5. Check consistency between metadata and relevant schemas when creating or making metadata open to the public. (C)

3. Conclusion

This is a brief report of the Metadata Information Infrastructure project funded in the Japanese fiscal year of 2010. Meta-Bridge provides a rich set of functions for organizations which design and use metadata schemas. Meta-Bridge is designed to store not only metadata vocabularies but also application profiles. Meta-Bridge is useful for Linked Open Data activities on the Web, which will benefit from registration and maintenance of the terms. It is also useful for organizations which are collaborating with other organizations of different types, e.g. Museum-Library-Archives collaboration. The authors hope that Meta-Bridge will help augment interoperability of metadata on the Web.

Our next step is to collect more schemas to enrich Meta-Bridge. Collaboration with memory organizations is inevitable to collect schemas. Meta-Bridge currently has only a Japanese interface. A user interface in English is essential for its promotion in the international community but it is left as a future task.

Acknowledgements

We are grateful to the Ministry of Internal Affairs and Communications for their support for this project. We thank all of members of the advisory panel for this project for their valuable comments and suggestions. We express our thanks to our partners who provided lots of metadata resources to this project.

References

- DCMI.(2007). DCMI Registry at Tsukuba. Retrieved April 30, 2011, from <http://demi.kc.tsukuba.ac.jp/>.
- Hillmann, Diane, Stuart A. Sutton, Jon Phipps, Ryan Laundry. (2006), A Metadata Registry from Vocabularies Up: The NDSL Registry Project. Proceedings of International Conference on Dublin Core and Metadata Applications, 2006, 65-75.
- Nilsson, Mikael, Thomas Baker, Pete Johnston. (2008). Singapore Framework for Dublin Core Application Profiles. Retrieved, April 30, 2011, from <http://dublincore.org/documents/singapore-framework/>.
- Metadata Information Infrastructure Project. (2010). Web Site of Metadata Information Infrastructure Project. Retrieved, April 30, 2011, from <http://meta-proj.jp/>.