

DCMI Metadata Schema Registry for Sharing Authoritative Information about Metadata Schemas

by Mitsuharu Nagamori and Shigeo Sugimoto

Information Standards

Since the emergence of the World Wide Web in the mid-1990s, metadata has been recognized as a key technology for digital libraries. Metadata is typically defined as “data about other data” – a simple definition that embraces a broad range of resources from library catalogues and indexes to thesauri, ratings, reviews and terms and conditions for use. On the Internet, metadata is designed for tasks ranging from resource description and discovery to archiving, trading, content filtering, resource syndication and information management. This diversity of purpose reflects the variety of information resources available on the Internet, ranging from personal web pages to huge portals for government information, digital libraries and shopping catalogues, as well as the variety of users ranging from young children to businesses and professionals.

A metadata schema registry (or simply, metadata registry) is a formal system that provides services over metadata vocabularies to users and machines. A metadata schema registry is widely recognized as an important tool not only to share information about metadata vocabularies but also to enhance reusability of metadata vocabularies. A registry also plays important roles in facilitating semantic metadata interoperability among communities speaking different languages and over time [1]. Achieving metadata interoperability is fundamental to making information resources shareable and discoverable.

The Dublin Core Metadata Initiative (DCMI) [2] is running the DCMI metadata schema registry that currently provides the reference descriptions of the metadata terms defined in RDF (Resource Description Framework)

Mitsuharu Nagamori and Shigeo Sugimoto are with the Graduate School of Library, Information and Media Studies at the University of Tsukuba, Japan. Their email addresses are nagamori@slis.tsukuba.ac.jp and sugimoto@slis.tsukuba.ac.jp

schemas. The authors have been developing the DCMI metadata schema registry since 1998, as well as developing some software tools for using metadata schema registries. From our experience, this article describes the basic requirements and functions for the DCMI metadata schema registry.

This article is based on our previous paper about a metadata schema registry as a tool to enhance metadata interoperability [3]. Our previous paper describes a general model of metadata schemas and requirements for metadata schema registries. This article describes the basic requirements and functions for a metadata schema registry of the DCMI.

DCMI Metadata Schema Definitions

Metadata schema refers to both the semantic and syntactic features of metadata. Very roughly speaking, a metadata schema is composed of a set of terms, a set of structural definitions of metadata instances and a binding scheme for implementation. DCMI defines metadata schemas, including properties, vocabulary encoding schemes, syntax encoding schemes and classes [4]. RDF schema versions of the DCMI term declarations are available at [5]. Figure 1 shows the definition of *Creator* from the *DCMI Metadata Terms* [4].

Each term is specified with the following minimal set of attributes:

Name: A token assigned to the term, unique within the term’s DCMI namespace

Label: The human-readable label assigned to the term

URI: The uniform resource identifier used to uniquely identify a term

Definition: A statement that represents the concept and essential nature of the term

Type of Term: The type of term as described in the DCMI Abstract Model [6]

NAGAMORI and SUGIMOTO, continued

FIGURE 1. Definition of **Creator** from *DCMI Metadata Terms* [4]**Term Name:** creator

URI:	http://purl.org/dc/elements/1.1/creator
Label:	Creator
Definition:	An entity primarily responsible for making the resource.
Comment:	Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.
Type of Term:	Property
Refines:	http://purl.org/dc/elements/1.1/creator
Refines:	http://purl.org/dc/terms/contributor
Has Range:	http://purl.org/dc/terms/Agent
Version:	http://dublincore.org/usage/terms/history/#creatorT-001

Where applicable, the following attributes provide additional information about a term:

Comment: Additional information about the term or its application

See: Authoritative documentation related to the term

References: A resource referenced in the Definition or Comment

Refines: A property of which the described term is a sub-property

Broader Than: A class of which the described term is a super-class

Narrower Than: A class of which the described term is a sub-class

Has Domain: A class of which a resource described by the term is an instance

Has Range: A class of which a value described by the term is an instance

Member Of: An enumerated set of resources (Vocabulary Encoding Scheme) of which the term is a member

Instance Of: A class of which the described term is an instance

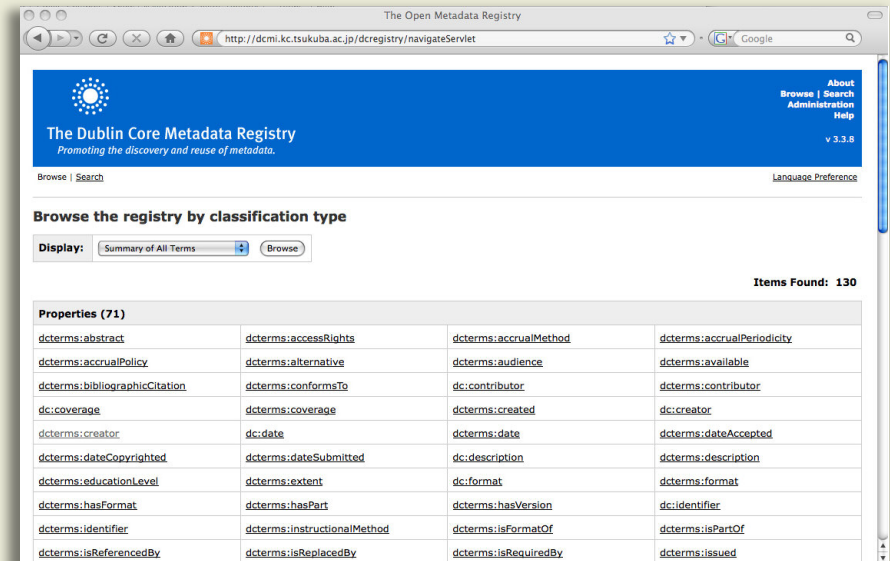
Version: A specific historical description of a term

DCMI Metadata Schema Registry

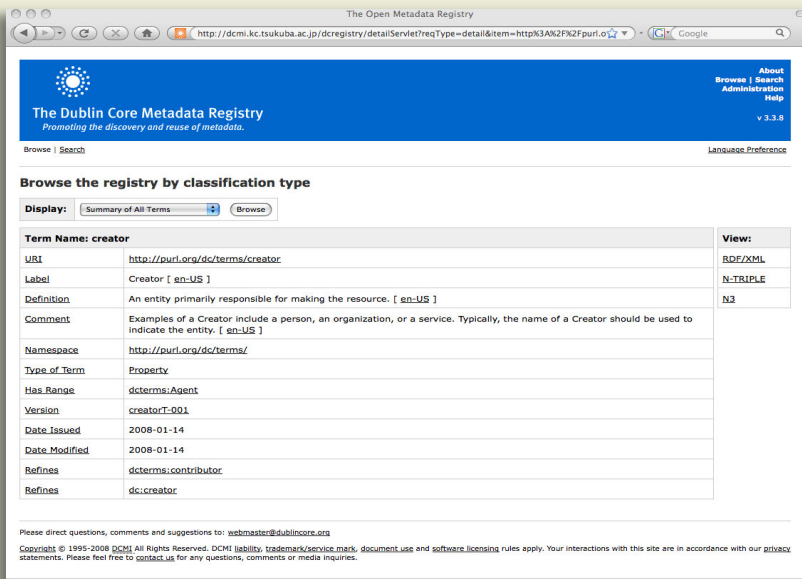
The DCMI Registry Working Group, which was established in December 1999, is an international forum for people who have interests in

the registry [7]. The authors have been involved in the working group since 1998. The DCMI metadata schema registry [8] provides authoritative reference descriptions of metadata schemas (see Figure 2). The reference descriptions are internally encoded in RDF schemas, and volunteers have translated them into 24 different languages. The reference descriptions are presented in a user-friendly form for human users and in RDF schemas for machines.

FIGURE 2. The DCMI Metadata Schema Registry [8]



The application program interface is based on web services protocols, specifically both REST (Representational State Transfer) and SOAP. Description of each metadata term includes a unique name of the term, language-dependent labels, definition statement of the term, date(s) of issue, type of the term and other essential information. The DCMI metadata schema registry is available as open source software for use by other communities.

FIGURE 3. The display for human users of `dcterm:creator` in the DCMI Metadata Schema Registry [8]

DCMI Metadata Schema Registry Services

A metadata schema defines a framework for representing metadata. In order to make metadata shareable and discoverable across user communities and languages, it is necessary to improve interoperability of metadata schemas. The DCMI metadata schema registry stores metadata schemas and serves information about them, such as definitions of the metadata schemas and relationships between metadata terms, to users. DCMI metadata schema registry provides its services not only for human users but also for machines.

For human users. The DCMI metadata schema registry provides human users with functions to find and browse reference descriptions of metadata schemas, such as that of a metadata element like `dcterm:creator` shown in Figure 3. A person who is engaged in the design of a new metadata schema can use a metadata schema registry as a dictionary to find existing metadata schemas. A person who writes and/or edits metadata, such as a cataloguer, can use a metadata schema registry as a reference source to understand metadata schemas.

For machines. Software tools should be able to access metadata schemas stored in a metadata schema registry through application program interfaces over global or local networks. Software tools can use metadata registries as software components. For example, a metadata schema registry can function as an information data source for metadata schemas, provide relationships between metadata terms to realize crosswalking and transform metadata among other things. The DCMI metadata schema registry provides information about the registry contents in a format that is suitable for applications and for other registries. This function is accomplished with a collection of web services, collectively known as the Registry Application Interface. The two principal approaches to web services are SOAP and REST [9]. The Registry currently supports both approaches. Each of the services will return results in RDF/XML (Resource Description Framework/Extended Markup Language) format that can be parsed by applications (see Figure 4).

FIGURE 4. A definition of `dcterm:creator` in RDF/XML suitable for machine processing [8]

```
<rdf:RDF>
  <rdf:Description rdf:about="http://purl.org/dc/terms/creator">
    <dcterms:issued>2008-01-14</dcterms:issued>
    <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/terms/contributor"/>
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
    <rdfs:label xml:lang="en-US">Creator</rdfs:label>
    <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/elements/1.1/creator"/>
    <rdfs:range rdf:resource="http://purl.org/dc/terms/Agent"/>
    <rdfs:comment xml:lang="en-US">
      An entity primarily responsible for making the resource.
    </rdfs:comment>
    <dcterms:modified>2008-01-14</dcterms:modified>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/terms/">
    <dcterms:hasVersion rdf:resource="http://dublincore.org/usage/terms/history/#creatorT-001"/>
    <dcterms:description xml:lang="en-US">
      Examples of a Creator include a person, an organization, or a service. Typically, the name of a
      Creator should be used to indicate the entity.
    </dcterms:description>
  </rdf:Description>
</rdf:RDF>
```

The extended functional requirements of a metadata registry are summarized below.

(1) Searching

Human users should be able to search in various ways for information stored in a metadata schema registry, such as for the name of a metadata schema, for a definition of metadata terms or for related terms. Search results are a set of entities such as metadata schemas, metadata terms or value vocabularies.

(2) Browsing

A metadata schema registry has to provide user interfaces for human users to browse information stored in the metadata schema registry via a web browser. Users should be able to browse metadata schemas in different ways.

(3) Schema Mapping

A metadata schema registry should provide information about the relationships among metadata schemas, terms and vocabularies. This function is required for the searching and browsing function.

(4) Version Management

It is a crucial function to manage revision history of metadata schemas to enhance metadata interoperability over time. The revision history will support migration and transformation of metadata. This function enhances metadata usability over time.

(5) Multilanguage User Interfaces

A metadata schema registry stores metadata schemas in multiple languages in order to support broader usage of metadata in the global communities on the Internet. A variety of users may use a metadata schema registry across languages. Multilingual user interfaces are useful for users who speak languages other than English. This function enhances the usability of a metadata schema registry for the international community.

(6) API for software tools

Metadata schemas should be provided not only for human users but also for software tools. The application program interface provided should be based on web services (REST, SOAP).

Related Registry Activities

A goal of metadata schema registries is to make metadata schemas understandable by both humans and machines, as well as shareable among user communities. Metadata schema registries have captured the interest of broad metadata communities because of the critical requirements of interoperability and longevity of metadata schemas. ISO/IEC 11179 addresses the semantics of the data, the representations of data and the registration of the descriptions of the data [10]. ISO/IEC 11179 allows the creation of a shared data environment. The Universal Description Discovery and Integration (UDDI) registries act as reference points for web services that allow for common descriptions and discovery of those services [11]. UDDI is based on XML standards and is platform-independent. ISO IEC JTC1 SC32 WG2 has been organizing a series of workshops on metadata registries [12].

The white paper reported by the DELOS Working Group on Registries [13] describes basic concepts of metadata schemas, for instance, metadata vocabulary, layers for metadata interoperability, data model and so forth. The layered model discussed in the white paper gives a framework for metadata vocabularies.

Beginning in January 2004, the JISC IE (Joint Information Systems Committee Information Environment) Metadata Schema Registry (IEMSR) project started development of a metadata schema registry as a pilot, shared service within the JIS Information Environment [14]. The IEMSR will act as the primary source for authoritative information about metadata schemas recommended by the JISC IE Standards framework. Metadata within the JISC IE is based on the Dublin Core and IEEE LOM standards.

The Schema Web is a repository for RDF schemas expressed in the RDF Schema, OWL (Web Ontology Language) and DAML+OIL (DARPA Agent Mark-up Language+Ontology Interface Layer) schema languages [15]. It provides a simple directory of RDF schemas for both human users and machines to search and browse metadata schemas.

The National Science Digital Library (NSDL) Metadata Registry is designed to provide humans and machines with the means to discover, create, access and manage metadata schemas, application profiles, crosswalks and concept mappings [16]. The NSDL Registry will complement the existing NSDL Metadata Repository by providing a means for NSDL and its affiliated

projects to identify, declare and publish through registration of their metadata schemas and controlled vocabularies.

Conclusion

We have been working on the DCMI metadata schema registry since

1998. We believe that the metadata schema registry has tremendous potential to enhance usability, reusability and interoperability of metadata schemas. More effort is needed to integrate software tools with metadata schema registries, as well as to collect and organize more schemas, so as to enhance the usability and reusability of metadata schemas and their registries. ■

Resources Mentioned in the Article

- [1] Sugimoto, S., Nagamori, M., Sakaguchi, T., Tabata, K., & Baker, T. (2001) Versioning the Dublin Core across multiple languages and over time. In *SAINT-W '01: Proceedings of the 2001 Symposium on Applications and the Internet – Workshops [SAINT 2001 Workshops]* Washington, DC: IEEE Computer Society, 151-156.
- [2] *Dublin Core Metadata Initiative*. (2008, August 4). Dublin, OH: Dublin Core Metadata Initiative. Retrieved August 28, 2008, from <http://dublincore.org/>.
- [3] Nagamori, M., & Sugimoto, S. (2006). A metadata schema registry as a tool to enhance metadata interoperability. *TCDL Bulletin*, 3(1). Retrieved August 28, 2008, from www.ieee-tcdl.org/Bulletin/v3n1/nagamori/nagamori.html.
- [4] *DCMI metadata terms*. (2008, January 14). Dublin, OH: Dublin Core Metadata Initiative. Retrieved August 28, 2008, from <http://dublincore.org/documents/dcmi-terms/>.
- [5] *DCMI term declarations represented in RDF schema language*. (2008). Dublin, OH: Dublin Core Metadata Initiative. Retrieved August 28, 2008, from <http://dublincore.org/schemas/rdfs/>.
- [6] Powell, A., Nilsson, M., Naeve, A., Johnston, P., & Baker, T. (2007, June 4). *DCMI abstract model*. Dublin, OH: Dublin Core Metadata Initiative. Retrieved August 28, 2008, from <http://dublincore.org/documents/abstract-model/>.
- [7] *DCMI Registry Community*. (2008, June 9). Dublin, OH: Dublin Core Metadata Initiative. Retrieved August 28, 2008, from <http://dublincore.org/groups/registry/>.
- [8] *DCMI Registry*. Retrieved August 28, 2008, from <http://purl.org/dcregistry/>.
- [9] [DCMI Registry] Application Interface. (2008). Section V of *DCMI Metadata Registry: Promoting the discovery and reuse of metadata*. Dublin, OH: Dublin Core Metadata Initiative. Retrieved August 28, 2008, from http://dcmi.kc.tsukuba.ac.jp/dcregistry/pageDisplayServlet?page=help_en-US.xsl.
- [10] ISO/IEC JTC1 SC32 WG2. (2004, September 15). ISO/IEC 11179, information technologies – metadata registries (MDR). 11179-1: Framework. (2nd ed.) Retrieved August 28, 2008, from <http://metadata-standards.org/11179/#11179-1>.
- [11] *UDDI XML.org: Online community for the universal description, discovery, and integration OASIS Standard*. (2008). Retrieved August 28, 2008, from <http://uddi.xml.org/>.
- [12] ISO/IEC JTC1 SC32 WG2. (2008, June 27). *SC32/WG2 (Metadata) home page*. Retrieved August 28, 2008, from <http://metadata-standards.org/>.
- [13] Baker, T., et al. (n.d.) *Principles of metadata registries: A white paper of the DELOS Working Group on Metadata Registries*. [n.p.]: DELOS Network of Excellence on Digital Libraries. Retrieved August 28, 2008, from <http://delos-noe.iei.pi.cnr.it/activities/standardizationforum/Registries.pdf>.
- [14] Johnston, P. (2007, May 25). *JISC IE Metadata Schema Registry*. Joint Information Systems Committee (JISC). Retrieved August 28, 2008, from <http://www.ukoln.ac.uk/projects/iemrs/>.
- [15] *SchemaWeb*. (n.d.) Retrieved August 28, 2008, from <http://www.schemaweb.info/>.
- [16] *NSDL Registry*. (2008). National Science Digital Library. Retrieved August 28, 2008, from <http://metadataregistry.org/>.