Chinese-European Workshop on Digital Preservation

Beijing (China), July 14 – 16, 2004

Metadata-Registries

Dr. Heike Neuroth

Research & Development

Göttingen State and University Library (SUB)

neuroth@mail.sub.uni-goettingen.de

Table of Contents

- ➢ Metadata

- ➢ Application Profile

- ➢ Interoperability

- ➢ Metadata Registry

- ➢ Examples

- ➢ Semantic Web

- ➢ Metadata Registry & Semantic Web

Metadata

➢ Metadata are (structured) data about data

➢ They describe the content, quality, conditions (like rights, access), and other characteristics of (digital) objects

➢ For discovery, description, archiving, etc.

➢ Each metadata element has a detailed definition (semantic, syntax)

➢ Set of metadata elements = metadata element set (or schema)

Metadata is a language

➢ A metadata "sentence" might say:

- "This book has **Author** *Johann Wolfgang von Goethe*, **Title** *Faust I*, and **Date Published** 1808

➢ Dublin Core was designed as a simple metadata language for the discovery of digital objects (Tom Baker: "pidgin language for a digital tourist").

➢ In unqualified Dublin Core, the sentence above would say:

- "This <u>resource</u> has **Author** *Johann Wolfgang von Goethe*, **Title** *Faust I*, and **Date** *1808*

Application Profile

➢ Dublin Core Metadata Element Set

- Too simple for most local application

- Need to extend this set locally

➢ An application profile is a schema which consists of terms (metadata elements) drawn from one or more metadata schema optimised for a particular local application

➢ Examples

- Dublin Core Library Application Profile

- Renardus Application Profile

- ...

Application Profiles and Interoperability

Contents:

- Keep precision in the meaning (semantics) of the terms used

- Make the sources of the terms explicit

- Spell out the local restrictions of use (value range, syntax, repeatability, obligation etc.)

Goal:

- Exchange of data with other providers:
  Interoperability!

Interoperability

➢ *... enables* **different communities***, with different types of information and technologies, to achieve a general level of* **information sharing** *...*

➢ *... is defined as the ability of digital library components or services to be functionally and logically* **interchangeable** *...*

➢ *... different services and components can* **communicate with each other** *through open interfaces ...*

Eric Miller, W3C

...

	Attribute Space	Value Space
Level 3	e.g. DCMES, LOM	e.g. DDC, LCSH
Level 2	Representation & Exchange	
	e.g. METS, XML, RDF, SOAP, OAI – PMH	
Level 1	Transport	
	e.g. Z39.50, LDAP, HTTP	

© **based on White Paper of the DELOS Working Group on Registries**

Metadata Registry

Need of registration of Application Profiles

- Clear declaration of rights holder

- Somebody is responsible for maintenance and updates

- Guarantees interoperability between metadata implementers

- Increase cooperation (metadata exchange, metadata sharing)

- Avoid double work by re-using already defined metadata elements

- …

Examples

MetaForm (SUB Göttingen)

- Local "manifestations" of Dublin Core for specific projects introduce variations -- like "dialects"

- "Crosscuts": how are elements used in different implementations?

- Provides "mappings" and "crosswalks" between Dublin Core and other schemas of similar scope

- Demonstrates the sort of output one would want from queries to a distributed registry

http://www2.sub.uni-goettingen.de/

MetaForm
CROSSWALKS, CROSSCUTS & MAPPINGS

Crosswalks

Format	Project description	Version	Mappings
Dublin Core and its Dialects			
AGLS	Australian Government Locator Service		Mapping
BIBLINK	BIBLINK Project	Stand: 28.07.1998	Mapping
BSZ	Dublin Core in der Interpretation des BSZ (Bibliotheksservice-Zentrum Baden-Württemberg)	Draft Stand: 27.09.2000	Mapping
CEDARS	Metadata for digital preservation	Stand 2002	Mapping
CIMI - Z39.50	CIMI Z.39.50 Application Profile for use in project CHIO	Draft Version 3 15.07.1996	Mapping
CIC	Gesellschaft Deutscher Chemiker Keine Formatbeschreibung vorhanden		Mapping
DEF	Denmarks Elektroniske	Stand: 2000	Mapping

CEDARS

CEDARS	Definition CEDARS	DC-Element	DC Definition
Preservation Description Information.Provenance Information.Rights Management.Rights Information.Copyright Statement.Name of Publisher	This sub-element contains the name of the publisher of the digital object.	DC.Publisher	An entity responsible for making the resource available
Preservation Description Information.Provenance Information.Rights Management.Rights Information.Copyright Statement.Publication-Place	This sub-element contains the place of publication of this version of this digital object.		
Preservation Description Information.Provenance Information.Rights Management.Rights Information.Copyright Statement.Date of Publication	This sub-element contains the date of publication of this version of this digital object.	DC.Date.Created	Date of creation of the resource
Preservation Description Information.Context Information.Related Information Objects	This element specifies any other information objects which were judged, at the time of ingest, to be significantly related to the ingested digital object.	DC.Relation	A reference to a related resource
Preservation Description Information.Provenance Information.Rights Management	This metadata section contains information relating to the intellectual property rights relevant to the digital object.	DC.Rights	Information about rights held in and over the resource

Examples

Schemas/CORES

- Registry of "all" Application Profiles and Metadata Element Sets

- "Rights holder", maintainer

- Updates

- Search for Aps, single metadata elements etc.

- …
  http://www.cores-eu.net/registry/

CORES Registry

Download schema creation tool
Help on using the registry

you are not logged in

Index

Agencies: Browse -Search

Element Sets: Browse -Search

Elements: Browse -Search

Encoding Schemes: Browse -Search

Application Profiles: Browse -Search

Element Usages: Browse -Search

Sandbox registry - Index - Agencies - Element Sets - Elements - Encoding Schemes - Application Profiles - Element Usages - login

If you have any problems please contact the administrator: cores@dsd.sztaki.hu

© 2002 MEG Registry Project, ILRT and UKOLN
© 2002, CORES Project, MTA SZTAKI DSD

CORES Registry

Download schema creation tool
Help on using the registry

you are not logged in

Application Profiles

Name	Version	Agency	
		Food and Agriculture Organization of the United Nations	Detail
Omnipaper Application Profile	1.0	Universidade do Minho	Detail
Renardus Application Profile		Renardus Consortium	Detail
The British Library Application Profile	1.0	The British Library	Detail
The IEEE LOM Application Profile	1.0	Metadata for Education Group	Detail
The IEEE LOM Metametadata Application Profile	1.0	Metadata for Education Group	Detail
The Qualified Dublin Core Application Profile		The Dublin Core Metadata Initiative	Detail
The RDN Record Sharing (rdn_dc) Application Profile		Resource Discovery Network	Detail
The RSLP Collection Description Application Profile	1.0	Research Support Libraries Programme	Detail
The Simple Dublin Core Application Profile		The Dublin Core Metadata Initiative	Detail

Sandbox registry - Index - Agencies - Element Sets - Elements - Encoding Schemes - Application Profiles - Element Usages - login

If you have any problems please contact the administrator: cores@dsd.sztaki.hu

Examples

Dublin Core Registry

- Defines Dublin Core schemata (terms, controlled vocabularies etc.)

- And relation between them

- Translations of elements, refinements and their definitions in different languages

  http://dublincore.org/dcregistry/index.html

The Dublin Core Metadata Registry

The Dublin Core Metadata Initiative's (DCMI) Metadata Registry is designed to promote the discovery and reuse of exiting metadata definitions. It provides users, and applications, with an authoritative source of information about the Dublin Core element set and related vocabularies. This simplifies the discovery of terms and related definitions, and illustrates the relationship between terms.

The reuse of existing metadata terms is essential to standardization, and promotes greater interoperability between metadata element sets. The discovery of existing terms is an essential, and prerequisite, step in this process. This application promotes the wider adoption, standardization and interoperability of metadata by facilitating its discovery, and reuse, across diverse disciplines and communities of practice.

This application was developed by the OCLC Office of Research, in cooperation with the Dublin Core Metadata Initiative Registry Working Group. It was developed, and is distributed, as an open-source project, built entirely upon open-source/open-standards software.

Please select from one of the following supported languages or click on the Preferences link above for additional options.

Information about the **application interface** can be found here.

Having trouble displaying the international fonts? Click here for help.

* Language support for those languages marked with an '*' are limited to terms only and do not currently include user-interface support.

العربية النسخة/عربي [ar-SA]	Castellano [ca-ES]	Česky [cs-CZ]
Cymraeg [cy-GB]*	Dansk [da-DK]*	Deutsch [de-DE]
Ελληνικά [el-GR]	English [en-US]	Español [es-ES]
Suomeksi [fi-FI]	Français [fr-FR]	Italiano [it-IT]
日本語 [ja-JP]	한국어 [ko-KR]	मराठी [mr-IN]
Norsk [no-NO]*	Polski [pl-PL]	Português [pt-PT]
Русский [ru-RU]	Svenska [sv-SE]	ไทย [th-TH]
українська [uk-UA]	简体中文 [zh-CN]	繁體中文 [zh-TW]

The Dublin Core Metadata Registry

Help
Preferences
Search
Administration

v 3.0.0

Search / Browse

Display: [Controlled Vocabulary Terms ▼]   [Submit]

Summary of DCMI Controlled Vocabulary Terms   Items Found: 12

Collection
Dataset
Event
Image
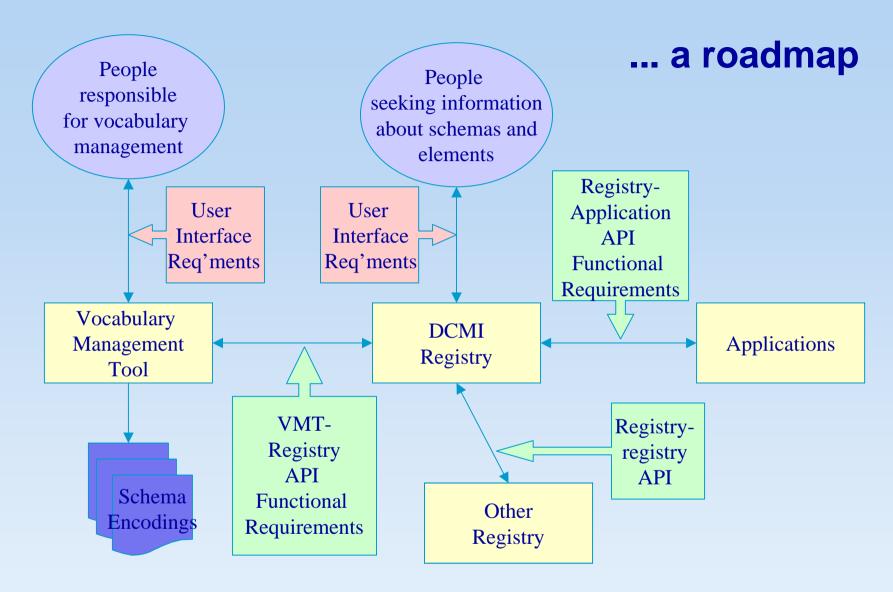InteractiveResource
MovingImage
PhysicalObject
Service
Software
Sound
StillImage
Text

... a roadmap

People responsible for vocabulary management

People seeking information about schemas and elements

User Interface Req'ments

User Interface Req'ments

Registry-Application API Functional Requirements

Vocabulary Management Tool

DCMI Registry

Applications

Schema Encodings

VMT-Registry API Functional Requirements

Other Registry

Registry-registry API

DCMI Roadmap for Development of Vocabulary Management and Registry Systems
(H. Wagner, R. Heery, S. Weibel)

Metadata Registries: Services

Advanced Registries will provide services for metadata
exchange and enhancements:

➢ Map from one data element to another

➢ Adjust value range (e.g. notation -> caption)

➢ Normalize data (e.g. date format)

➢ Support standardization (e.g by the use of authority
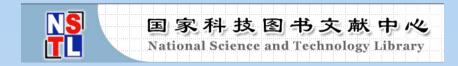files)

Semantic Web

A vision

"The Semantic Web is an extension of the current web in which *information is given well-defined meaning*, better enabling computers and people to work in cooperation."

Tim Berners-Lee, James Hendler and Ora Lassila, The Semantic Web, Scientific American, May 2001

Metadata Registries & Semantic Web

- Metadata provides well-defined meanings to information

- Registries help to maintain this well-defined meanings

- Increase interoperability, also in (semi-) automatically ways (metadata exchange, metadata sharing, metadata re-use)

- **But**: all registries so far are very complicated, maintenance is very time-consuming (depending on metadata developers)

- ?<u>Vision</u>: more simple, concentrating on core set

Thank you very much for your attention

Dr. Heike Neuroth
Research & Development
Göttingen State and University Library (SUB)
neuroth@mail.sub.uni-goettingen.de