PB138 — Metadata Describing XML Resources

Metadata Frameworks Describing the XML and Data Resources

RDF Framework

RDF Model and RDF Schema are W3C Recommendation.

link:http://www.w3.org/RDF [Specification and further W3C Working Group information]

RDF Model

RDF is general mechanism for metadata specification can be used with any (not only digital) resources.

- RDF model is based on triplets:
  - resource - http://www.fi.muni.cz/~tomp/xml for example
  - property - description for example
  - value - PB138 at FI MU Homepage for example.
- Triplets can be represented:
  - graphically,
  - like a triplets (r, p, v)
  - by an XML syntax See more

RDF Model (2)

See more:

- Good introduction on xml.com: What is RDF?
- Zvon RDF Tutorial
- RDF Tutorial - W3Schools RDF Tutorial
- RDF Tutorial

RDF Schema

- Specifies the set of properties and their definition domain and the range of their values.
- To model the RDF Schema is used the RDF.

RDF representation of used metadata schemes - Z39.50, Dublin Core etc.

- RDF is general framework for metadata modelling.
- There is a need of definition of allowed properties, their domains and (allowed) for particular use.
- This is the way to create RDF metadata schema representation.
- Representation may be in a form of RDF Schema.

Dublin Core - Example of RDF Schema

What is Dublin Core?

- Generally usable generic metadata schema.
- Based on initiative of librarians to create framework for bibliographic information description.
- Today is commonly used - metadata description of public service information (e-Government) for example.
- Created by 15 basic elements with partially defined semantics.
- Elements can be extended - by decomposition to (usually disjoint) subsets.
    - It always must be a subsets of originally designed elements.

Simple Dublin Core

"Simple" or "unqualified" Dublin Core ("Simple DC" in what follows) represents the basic set of 15 element designed and supported by:

- Dublin Core Metadata Initiative (DCMI, http://dublincore.org).
- Present version - Dublin Core 1.1.
- Accepted by IETF consortium as RFC (Request For Comment) 2431 as well as ISO Standard 15836-2003 since 2003.

Dublin Core - elements

- Nazev - Jméno dané zdroji
- Tvurce - Entita primárně odpovědná za vytvoření obsahu zdroje
- Predmet a klíčová slova - Téma obsahu zdroje
- Popis - Vysvětlení obsahu zdroje
- Vydavatel - Entita odpovědná za zpřístupnění zdroje
- Prispevatel - Entita, která přispěla k vytvoření obsahu zdroje

- Datum - Datum spojené s určitou událostí během existence zdroje
- Typ zdroje - Povaha nebo druh obsahu zdroje
- Format - Fyzická nebo digitalní reprezentace zdroje

Dublin Core - elements (2)

- Identifikator zdroje - Jednoznačný odkaz na zdroj v rámci daného kontextu
- Zdroj - Odkaz na zdroj, z něhož je popisovaný zdroj odvozen
- Jazyk - Jazyk intelektualního obsahu zdroje
- Vztah - Odkaz na příbuzný zdroj
- Pokryti - Rozsah nebo záběr obsahu zdroje
- Sprava autorskych prav - Informace o právech vztahujících se k popisovanému zdroji

DC - Metadata description example

```
Nazev: Zelena kniha o elektronickem obchodu
Tvurce: Urad pro verejne informacni systemy, Urad vlady
Predmet: Elektronicky obchod, elektronicky podpis, bezpecnost, sprava
Popis: Vladni navrh podpory elektronickeho obchodu v Ceske republice
Datum vytvoreni: 2001-09-20
Datum zverejneni: 2001-10-17
Identifikator: ISBN:?????
```

Qualified Dublin Core

- contains the same set of elements as the Simple DC and recommends more in depth specification and limitation on every element.
- Typically based on formal or de-facto international standard.
- The language must be specified according the ISO 639 standard for example.

XML coding of DC

DTD

XML Schema

RDF Schema - www.fi.muni.cz/~tomp/xml/rdf/dc-rdf-schema-cz.rdf

RDF Schema pro slovnik typu (Type Vocabulary) - /~tomp/xml/rdf/dc-tv-rdf-schema-cz.rdf

RDF Tools

Jena Java RDF API and toolkit

The ICS-FORTH RDFSuite

DC Creator on University of Bath

for more see http://www.w3.org/RDF

Example of Metadata Practical Usage - Public Service

Metadata Framework of ISVS CR

Construction steps

- Accept the DC recommendation and adopt it as Narodni metadatovy standard (NMS).
- Enhance the standard so it correspond the public service needs for easy searching of information and information resources management.
- Develop Application profile of NMS, that will specify required coding schemes and mandatory definition of all metadata elements.
- Prepare Public service thesaurus (Tezaurus verejne spravy).

Adaptation of Dublin Core for Public Service

Specific application profile of DC has been prepared for needs of public service in EU countries, Australia, Canada and New Zealand.

The objectives of MIReG (Metadata Information Resource for eGoverment) are:

- development of metadata framework
- corresponding reference software tools and best practice algorithm for implementation in the countries and sectors.

Adaptation of Dublin Core for Public Service (2)

MIReG cooperates with European Standardisation Authority (CEN), that is preposition of accepting the recommendation in entire Europe.

- process started by a series of workshops Managing information resources for e-government (MIReG) and become a part of the program Interchange of Data between Administrations (IDA) of EU.

- The next partner involved in a design of european metadata framework is the project ParlM.
  - The priorities of ParIM is to make the information of European Parliament accessible.
- The corresponding working group prepares the DC-Gov Application Profile recommendation.

NMS Application Profile

contains:

- Element refinement of metadata elements.
- specifies the elements semantics more precisely and divides them on more accurate subelements (general date is divide into creation date, publication date, date of validity, etc for example).
- Qualified element can be processed by tools for unqualified specification
- The tools work with qualified element as it is unqualified one (the date instead of publication date for example. This leads to a loss of some semantics but the element can be used for searching further for example.
- Encoding scheme - specifies the form used to store the value of the metadata element (date is always in the form of yyyy-mm-dd defined by ISO 8601).
- Besides the form of the value the qualification may specified the unit of the value for example as well.

Ontologies

What are ontologies?

Instrument how to describe a knowledge.

Set of terms and constructs and how they can be joint, derived etc.

Basic categories of ontologies are:

- Classes (general things) in the many domains of interest
- The relationships that can exist among things
- The properties (or attributes) those things may have

Uses the metadata frameworks (e.g.. RDF), but is better with more exact semantics.

There are general frameworks for creation of specific ontologies.

Use Cases

- Web portals - data integration on web.

- Multimedia collections.
- Large web sites management
- Design documentation
- Intelligent agents
- "Ubiquitous computation"

W3C Working Group

XML Topic Maps

The next proposal to WebOnt WG - http://www.topicmaps.org/xtm/1.0