Metadata Standards and Applications

5. Applying Metadata Standards: Application Profiles
Goals of Session

- Learn how metadata standards are applied, used and documented:
- Learn about the concept and use of application profiles
- Explore how different metadata standards may be used together in digital library applications
Why Application Profiles?

- Describes the set of metadata elements, policies, guidelines and vocabularies defined for a particular domain, implementation, or object type
  - Declares the metadata terms an organization, information resource, application, or community uses in its metadata
  - Documents metadata standards used in instance data, including schemas and vocabularies, policies, required elements, etc.
  - Called “application profile” or just “profile”
Benefits of Documenting Terms We Use

- To provide authoritative specification of term usage
- To facilitate interoperability by informing potential users of domain consensus
- To support evolution of vocabularies
- To encourage alignment of practice
- To enable interpretation of legacy metadata
Less Flexibility, More Predictability

Many metadata standards are sufficiently flexible that they need a mechanism to impose some constraints

- Profiles allow expression of the decisions made for a project in machine-readable form (XML or RDF)

Refining

- Allow a narrower interpretation of a standard to suit your project

Combining

- Enable mixing elements from various different standards (there are limits to this!)
Components of an AP

- **Human readable documentation**
  - Property descriptions and relationships
  - Domain or project specific instruction
  - Obligation and constraints

- **Machine-readable versions may contain:**
  - Specific encoding decisions and XML or RDF schemas
  - Models of data relationships specific to the AP represented in the schemas
  - Functional requirements and use cases supporting decisions
Using Properties from other Schemas

- DC APs set stringent requirements for determining reusability of terms:
  - Is the term a real “property” and defined as such within the source schema?
  - Is the term declared properly, with a URI and adequate documentation and support?
  - In general, properties whose meaning is partly or wholly determined by its place in a hierarchy are not appropriate for reuse in DC APs without reference to the hierarchy.

- Other styles of profiles have different requirements and strategies for developing machine-readability and validation.
Documenting new properties

- Minimum: a web page, with the relevant information available to other implementations
- Better: a web page and an accessible schema using your terms as part of your application profile
- Best: all terms available on a distributed registry
Singapore Framework

- A Framework for designing metadata applications for maximum interoperability
  - Defines a set of descriptive components that are necessary for documenting an Application Profile
  - Forms a basis for reviewing Dublin Core application profiles
  - Relates APs to standard domain models and Semantic Web standards
An RDA Application Profile

- A DCMI/RDA Task Group has been defining RDA properties and value vocabularies as formal RDF vocabularies (with URIs)
  - IFLA has stated an intention to declare FRBR entities and attributes as well
  - Next step is a DC application profile of RDA according to the Singapore Framework
  - See http://metadataregistry.org for the provisionally registered properties/vocabularies
METS Profiles

- Description of a class of METS documents provides document authors and programmers guidance to create and process conformant METS documents
  - XML document using a schema
  - Expresses the requirements that a METS document must satisfy

- METS Profiles are output in human-readable prose and not intended to be “machine actionable” (but they use a standard XML schema)
Components of a METS Profile

- 1. Unique URI
- 2. Short Title
- 3. Abstract
- 4. Date and time of creation
- 5. Contact Information
- 6. Related profiles
- 7. Extension schemas
- 8. Rules of description
- 9. Controlled vocabularies
- 10. Structural requirements
- 11. Technical requirements
- 12. Tools and applications
- 13. Sample document
MODS Profiles

- Some applications are establishing MODS profiles to document usage, required elements, controlled vocabularies used, etc.

- Some examples:
  - DLF Aquifer MODS profile: to establish implementation guidelines for rich shared metadata for cultural heritage materials
  - British Library electronic journal MODS profile
METS & MODS Together

- METS can be used to package together the metadata with the objects
  - METS allows for use of any XML metadata schema in its extensions
  - MODS can be associated with any level of the description
  - Technical metadata can be inserted and associated with specific files
Many metadata standards are sufficiently flexible that profiling is necessary

- Documenting what is used in an application will simplify and enhance data presentation, conversion from other sources, ability to provide different outputs
- Constraining a metadata standard by specifying what is used and how facilitates data exchange and general interoperability

Documentation is always a good value!
DC Application Profile Examples

- **Collections AP**

- **Scholarly Works Application Profile (SWAP)**
  - [http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile](http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile)

- Both these have been reviewed by the DC Usage Board and are deemed compliant with the DC Abstract Model
METS/MODS AP Examples

- University of Maryland Descriptive Metadata

- UVa DescMeta

- Texas Digital Library profile for electronic theses and dissertations
Vocabulary Development in an AP: a Case Study of KMODDL

- http://kmoddl.library.cornell.edu/aboutmetadata2.php

- Needed to describe all kinematic models, plus materials related to the models

- Developed several special vocabularies:
  - Voigt1 and Voigt2 (plus the IDs)
  - KMODDL Type (plus ID)
  - An AAT subset for Medium

- Adapted:
  - DLESE GradeRange for Audience
  - MARC Organization List for MODS:physicalLocation
Voigt1 and Voigt2

- Based on a 19th century treatise on the Reuleaux models, which classified them based on mechanical principles (http://kmoddl.library.cornell.edu/model.php)
- Names of the mechanisms used as subject terms
- Classification numbers used as IDs to tie the related materials together
KMODDL Type Vocabulary

- [http://kmoddl.library.cornell.edu/aboutmeta3.php](http://kmoddl.library.cornell.edu/aboutmeta3.php)

- Used to differentiate a complex array of versions and related materials from one another

- Allows distinctions between print and digital (of the same resources) and identifies granular levels within those resources

- Enables creation of organized web pages presenting the information to users
KMODDL Example

- [http://kmoddl.library.cornell.edu/model.php?m=244](http://kmoddl.library.cornell.edu/model.php?m=244)

Note:

- Browse tree tab on left
- Lists of linked References and Resources at the bottom of the page
- Attributed description
- Tabs for Image and Movie on the top right
Exercise

- Critique an Application Profile for a community or project, e.g., the Open Language Archives Community Metadata Set (OLAC-MS)

  - OLAC Metadata (DC-based)
    http://www.language-archives.org/OLAC/metadata.html

  - University of Maryland Descriptive Metadata
    http://www.lib.umd.edu/dcr/publications/taglibrary/umdmdm.html

  - UVa DescMeta
    http://lib.virginia.edu/digital/metadata/descriptive.html
Exercise: Questions to address

- Does the profile define its user community and expected uses?
- How usable would the profile be for a potential implementer?
- How (well) does the profile specify term usage?
- How (well) does the profile define and manage vocabularies?
- Are there key anomalies, omissions, or implementation concerns?