Dublin Core: The Road from Metadata Formats to Linked Data

Joint NISO/DCMI Webinar
25 August 2010
Makx Dekkers and Thomas Baker
Dublin Core in the
Early Web Revolution
First steps

- In October 1994, informal discussion at second Web Conference, Chicago
- Identified a need for a “core” set of descriptors to help discover content on the Web
- 1-3 March 1995, OCLC/NCSA workshop in Dublin, Ohio at OCLC Headquarters
Dublin Core: the original idea

- A basic description mechanism for digital information that:
  - can be used in all domains
  - can be used for any type of resource
  - is simple, yet powerful

- Making it easier to find information on the Web as it develops (1995!)
1995: The Dublin Core

- “Core” set, simple enough for non-experts to understand and create
- A “library catalog card” for Web objects
- Based on consensus across domains
- Standardized:
  - ISO 15836:2003, revised 2009

Elements
1. Identifier
2. Title
3. Creator
4. Contributor
5. Publisher
6. Subject
7. Description
8. Coverage
9. Format
10. Type
11. Date
12. Relation
13. Source
14. Rights
15. Language
1996: Modular metadata

- Not: “One size fits all”
- Different ways to describe one object:
  - MARC records for library catalogs
  - Dublin Core for simpler descriptions
  - Specialized metadata for terms and conditions of use
- Recognized need for a general framework for different types of metadata
1996: Towards metadata “frameworks”

- Warwick Framework: “packages” in “containers”
  - Descriptive Dublin Core
  - Administrative Admin Core
  - Rights Creative Commons
  - Creative Commons Licence

- W3C Resource Description Framework (RDF)
  - Generic, interoperable expression of metadata
1997: Qualification to add precision

- Not just any Date, but a Date the resource was *Created*
- Not just any Subject, but a *Library of Congress Subject Heading*
- Dumb-down: ignore extra details to see just a “core” description
2000: Application Profiles

- Customized implementations
  - Use “the Dublin Core” with other vocabularies
  - Local rules and guidelines
- Application profile provides documentation so that others can follow
- Not “take it or leave it”, but “take what you want, create what you need”
Dublin Core usage

- The Web exploded (now over a trillion pages!)
- Search engines took care of the open Web
- Dublin Core metadata came to be used widely in “controlled environments”
  - As a basic description mechanism
  - As a basic exchange format
- But never intended to be a “complete” solution
Interoperability in the early Web

- No interoperability of **data** in the Open Web
  - Mostly pages with links for human navigation
- Controlled environments: “intra-operability” (cooperation between known partners)
  - But “intra-net” can be quite large, e.g. OAI-PMH
- Semantic Web/Linked Data intends to “open up” controlled environments (“silos”)
  - Global interoperability across silos through typed links
Linked Data basic principle

This presentation Has subject Dublin Core

Object (Resource) — Predicate (property) — Subject (Resource)
Dublin Core principles and Linked Data

- Dublin Core principles (1995-2005)
  - One-to-one (describe one and one thing only)
  - Dumb-down
  - Appropriate values

- Corresponding Linked Data design principle
  - A statement is “about” a named resource
  - Sub-property relations
  - Choice between text strings and links to other resources

Dublin Core was one of the inspirations for RDF
Dublin Core development

- Started from a vision for the open Web (HTML)
- Came to be widely deployed in controlled environments (XML)
- Further developed since 2000 in conjunction with Semantic Web and Linked Data (RDF)
- From a “Core Metadata Element Set” for the Web to a “core vocabulary” for Linked Data