Dublin Core: History, Key Concepts, and Evolving Context (part one)

DC-2010 INTERNATIONAL CONFERENCE ON

DUBLIN CORE AND METADATA APPLICATIONS

Pittsburgh, PA, October 20, 2010

Jane Greenberg, Professor, Director of the
SILS Metadata Research Center
janeg@email.unc.edu

Dublin Core Metadata Initiative®

Overview

- Introduction; Metadata defined
- Dublin Core (history & growth)
  - Dublin Core Metadata Element Set, Version 1.1
    - History workshop → *full* conference
    - Founding principles (characteristics)
- DC semantics and DCMI Metadata Terms
    - Dublin Core abstract model
    - Singapore Framework
      - *** An evolving context / *Semantic Web alignment***
- Dublin Core Metadata Initiative (DCMI)
- Q&A

Dublin Core Metadata Initiative®

Introduction

- Who is here?
  - √ librarians/archivist/museum professionals
  - √ scientists
  - √ government specialists
  - √ educators
  - √ business/corporate workers
  - – other?
- Dublin Core
  - – Using DC; planning to use DC; want to learn more

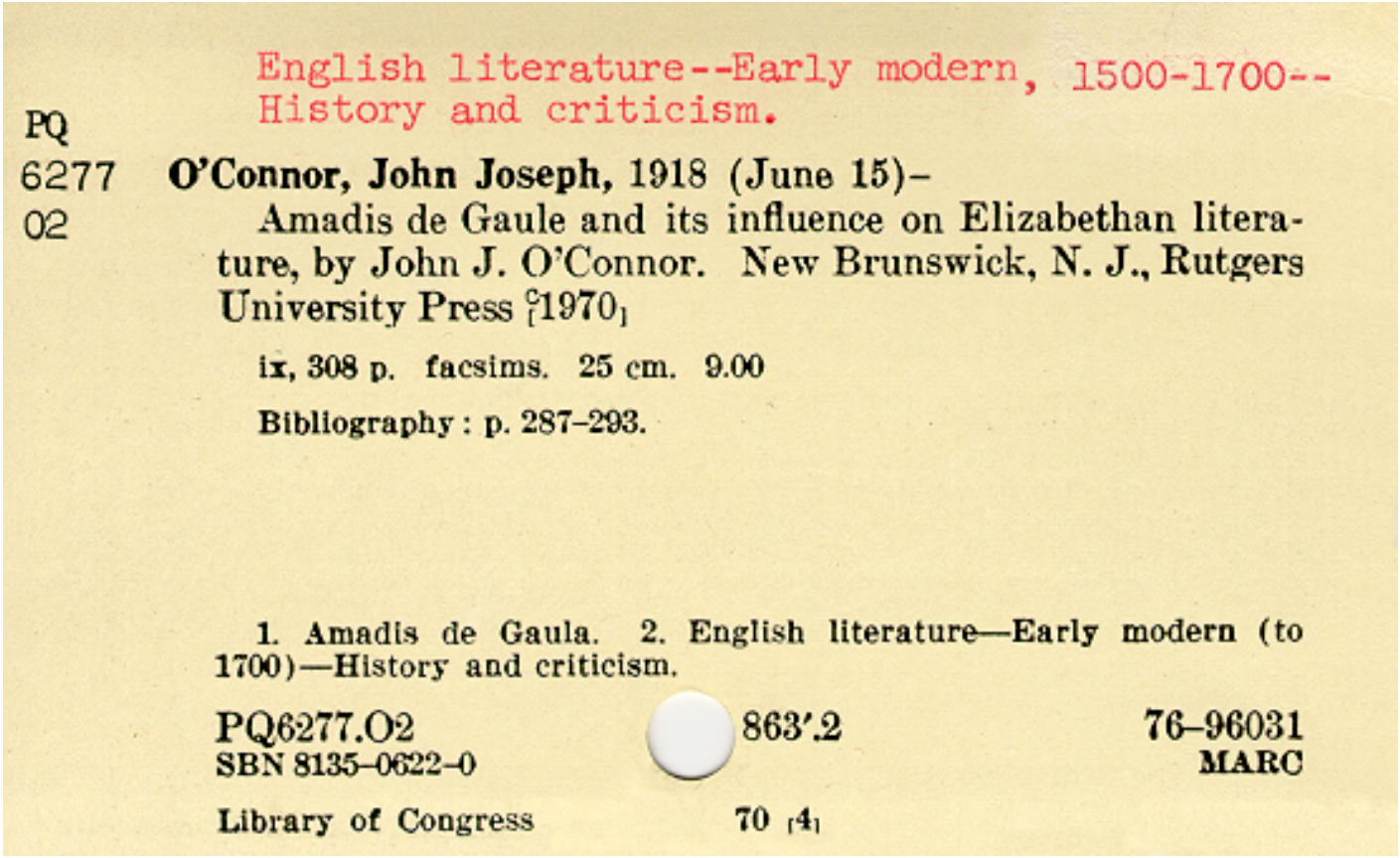
Dublin Core Metadata Initiative®

Metadata defined

…….data about data

…….information about data

Library catalog card



English literature--Early modern, 1500-1700--
History and criticism.

PQ
6277
02

O'Connor, John Joseph, 1918 (June 15)–

Amadis de Gaule and its influence on Elizabethan literature, by John J. O'Connor. New Brunswick, N. J., Rutgers University Press [©1970]

ix, 308 p. facsims. 25 cm. 9.00

Bibliography: p. 287–293.

1. Amadis de Gaula. 2. English literature—Early modern (to 1700)—History and criticism.

PQ6277.O2 863'.2 76–96031
SBN 8135–0622–0 MARC

Library of Congress 70 [4]

Dublin Core Metadata Initiative®

Metadata Example for a specimen

Family: Pinaceae

Species: *Pinus serotina*

Date identified: 1958-05-10

County: Pasquotank County

Location collected: Woodland Border, 2.3 miles north east of Nisonton

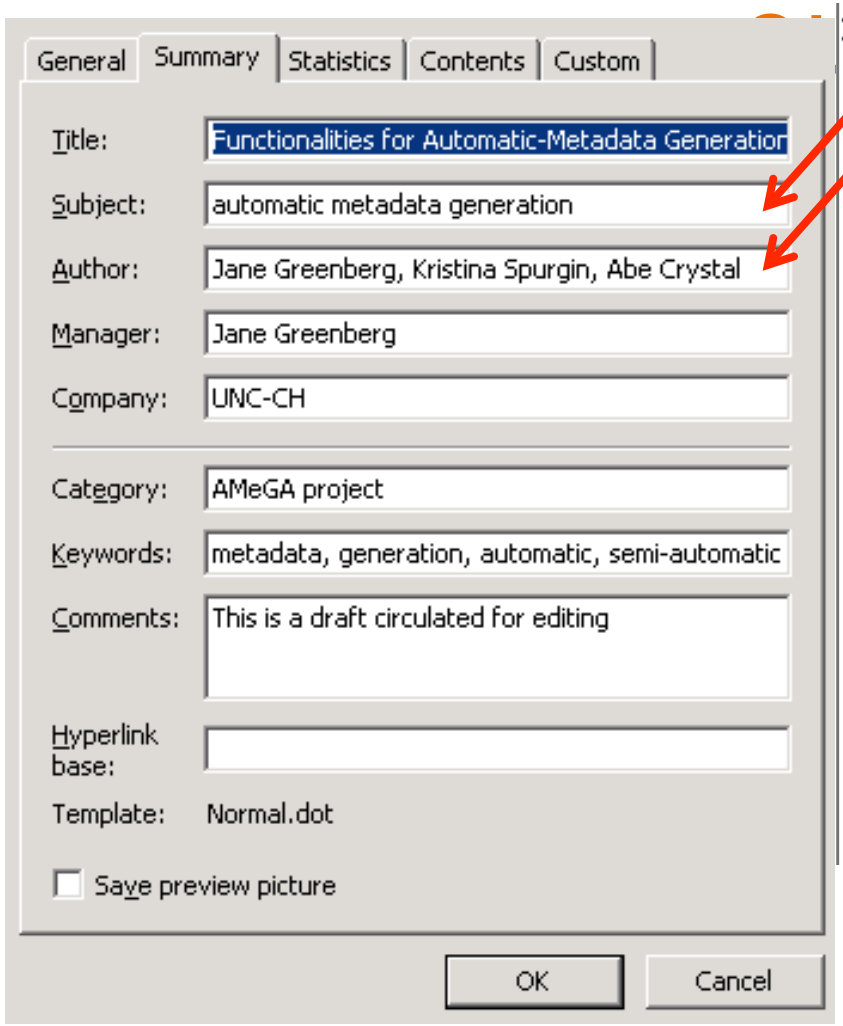
Collected by: Harry E. Ahles

<Species> Pinus serotina </ Species>

<Date.ID><scheme=SPEC.W3C DTF“>1958-05-10

</Date.ID>

MS WORD dialog box and




```
xmlns:st1="urn:schemas-microsoft-com:office:smarttags"><o:DocumentProperties>
  <o:Subject>automatic metadata generation</o:Subject>
  <o:Author>Jane Greenberg, Kristina Spurgin, Abe Crystal</o:Author>
  <o:Keywords>metadata, generation, automatic, semi-automatic</o:Keywords>
  <o:Description>This is a draft circulated for editing</o:Description>
  <o:LastAuthor>Kristina M Spurgin</o:LastAuthor>
  <o:Revision>2</o:Revision>
  <o:TotalTime>13</o:TotalTime>
  <o:LastPrinted>2004-11-06T18:41:00Z</o:LastPrinted>
  <o:Created>2004-11-12T16:50:00Z</o:Created>
  <o:LastSaved>2004-11-12T16:50:00Z</o:LastSaved>
  <o:Pages>1</o:Pages>
  <o:Words>9160</o:Words>
  <o:Characters>54690</o:Characters>
  <o:Category>AMeGA project</o:Category>
  <o:Manager>Jane Greenberg</o:Manager>
  <o:Company>UNC-CH</o:Company>
```


Dublin Core Metadata Initiative®

Dublin Core metadata

<dc:title>Godiva Chocolatier</dc:title>

<dcterms:alternative>Godiva store**</dcterms:alternative>**

<dc:creator>Nancy Confection</dc:creator>

<dc:creator>Confection, Nancy</dc:creator>

<dc:subject>Chocolate</dc:subject>

<dc:subject **xsi:type="dcterms:lcsh"**>Truffles (Confectionery)</dc:subject>

<dcterms:created xsi:type="dcterms.W3CDTF">
2008--6--28**</dcterms:created>**

<dc:identifier>http://www.godiva.com</dc:identifier>

<dcterms:abstract>Provides access to chocolate collections, gifts, ….**</dcterms:abstract>**

Dublin Core Metadata Initiative®

Metadata

- Structured, descriptive information about a **resource** (DCMI Glossary; Weibel, 1995)
- Data about the content, quality, condition, and other characteristics of data (FGDC Glossary, 1992)
- Additional information for the data it describes to support “use” (Wikipedia, 2008)
- Structured data about an object that supports **functions** associated with the designated object (Greenberg, 2003)

Dublin Core Metadata Initiative®

Some typical metadata functions

Discover resources	Manage documents	Control IP rights
Identify versions	Certify authenticity	Indicate status
Mark content structure	Situate geospatially	Describe processes

Dublin Core Metadata Initiative®

Typology of 7 types of metadata	Metadata Functions "This Type Of Metadata Facilitates:"	Element examples*
Identification / description metadata	Resource Discovery / Information Retrieval	**Creator** (Author), Title, Subject
Administrative metadata	Resource Management	Price, Condition
Terms and conditions metadata	Resource Usage	Rights, Reproduction restrictions
Content ratings metadata	Resource Use By Appropriate Audiences	Audience
Provenance metadata	Resource Authentication And Other Provenance-Related Activities	**Creator**, **Source**
Linkage / relationship metadata	Resource Linking With Related Resources	Relation, **Source**
Structural metadata	Resource Hardware And Software Needs	Compression ratio

(Lagoze, 1996; mapped in Greenberg, 2005)

Dublin Core Metadata Initiative®

Dublin Core history and growth

Internet / WWW

- **1970 ARPANET begin operation**
- 1991 Gopher released by Univ. of Minnesota
- **1991 WWW released by CERN**
- 1993, Lycos (first engine to achieve commercial success)
- 1994, WebCrawler first full-text Web search engine
- Late 1990s Web directories become popular
- **1998 Google, relevancy ranking**
- **2009++—Linked data**

~~~ Cataloging/Metadata

- ***c.1450 printing press***
- 18th c. French cataloging code
- 1876 Cutter's catalog objects
- 1960/61 Lubetsky/Paris principles
- 1988 Anglo-American Cataloigng Rules, 2nd ed. *rev*.
- **1994, 2nd WWW Conference in Chicago**
  - need for a metadata "core": a small, common set of metadata elements to describe Web content
- **1995, NCSA/OCLC workshop in Dublin, Ohio**
- **1998 FRBR**
- **2009 RDA**

Dublin Core Metadata Initiative®

Dublin Core (DC) history

DC 1: March, 1995, Dublin, Ohio (OCLC/NCSA workshop)

Identified 13 core elements (properties) essential for resource discovery on the Internet

- Title
- Author
- Subject
- Publisher
- Other Agent
- Date
- Object Type
- Form
- Identifier
- Relation
- Source
- Coverage
- Rights Management

1994, informal discussion at the 2nd WWW conf., Chicago

Dublin Core Metadata Initiative®

Dublin Core (DC) history

DC 2: April, 1996, Warwick, England

- Established the Warwick Framework
- Establish a syntax for the DC…HTML tags for embedding in web documents (HTML 2.0)

DC 3: September, 1996, Dublin, Ohio

- DC elements extended to include digital images, 15 properties

DC 4: March 1997, Canberra, Australia

- *Minimalists* and *structuralists*
- Canberrra qualifiers (meta-meta)
  - Refinement rather than redefinition
    - Coverage: spatial and temporal
  - Schemes
    - ISO 8601 for date encoding (syntax encoding)
    - LCSH or AAT for subject metadata (vocabulary encoding)

Dublin Core Metadata Initiative®

DC-5: 1997, Helsinki, Finland
DC-6: 1998, Washington, D.C.
DC-7: 1999, Frankfurt, Germany
DC-8: 2000, Ottawa Canada
2001, Tokyo, Japan (Conference w/proceedings)
2002, Florence, Italy (Application profiles)
2003, Seattle, WA
2004, Shanghai, China
2005, Madrid, Spain
2006, Mexico
2007, Singapore (Singapore Framework)
2008, Berlin, Germany
2009, Seoul, Korea
2010, Pittsburgh, PA

Dublin Core Metadata Initiative®

Dublin Core:
the original idea and guiding principles

- Simple
- Semantic interoperability

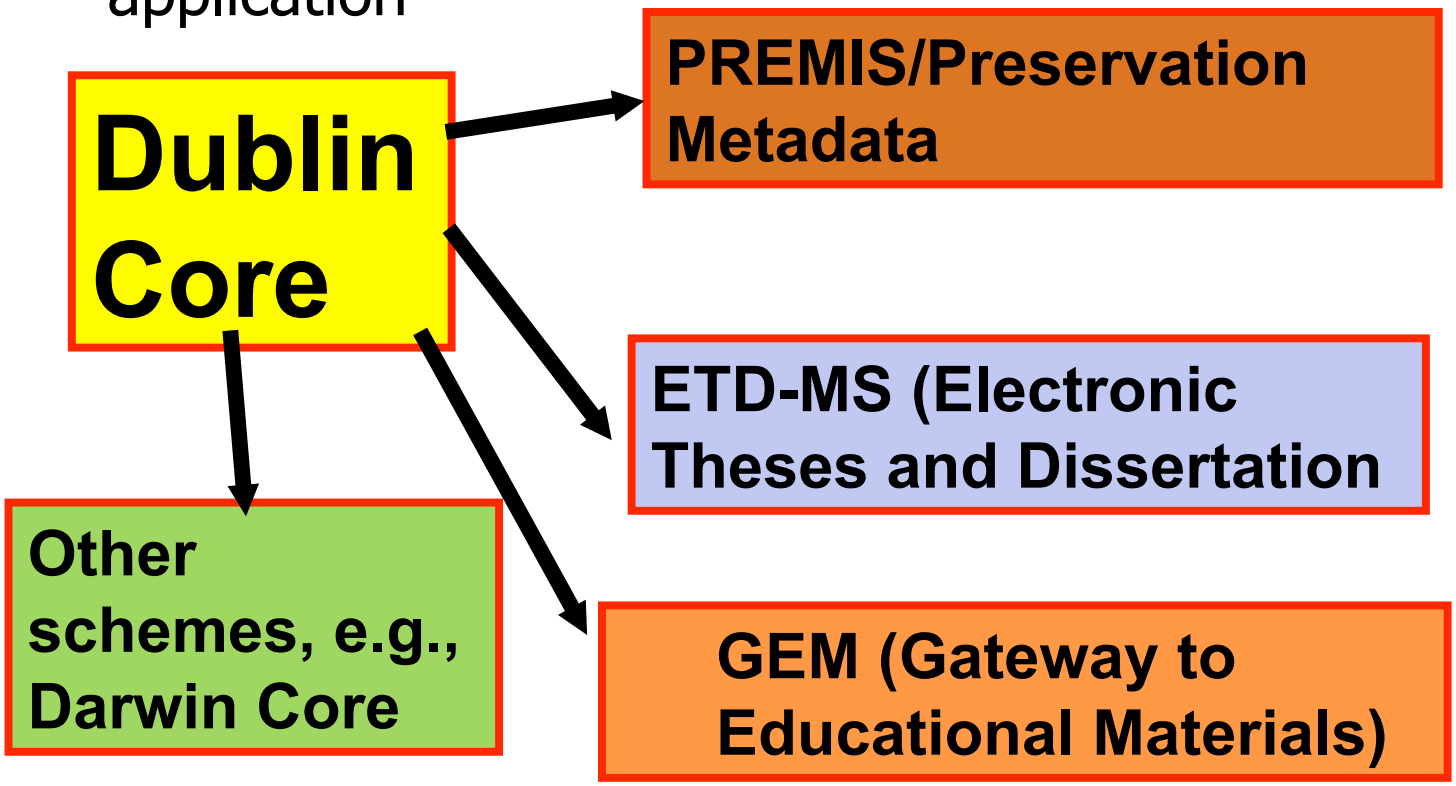
  Title = title = title
- International consensus
- Interdisciplinary (any domain)
- For any format
- Extensible
- Modular

Making it easier to find information on the Web as it develops

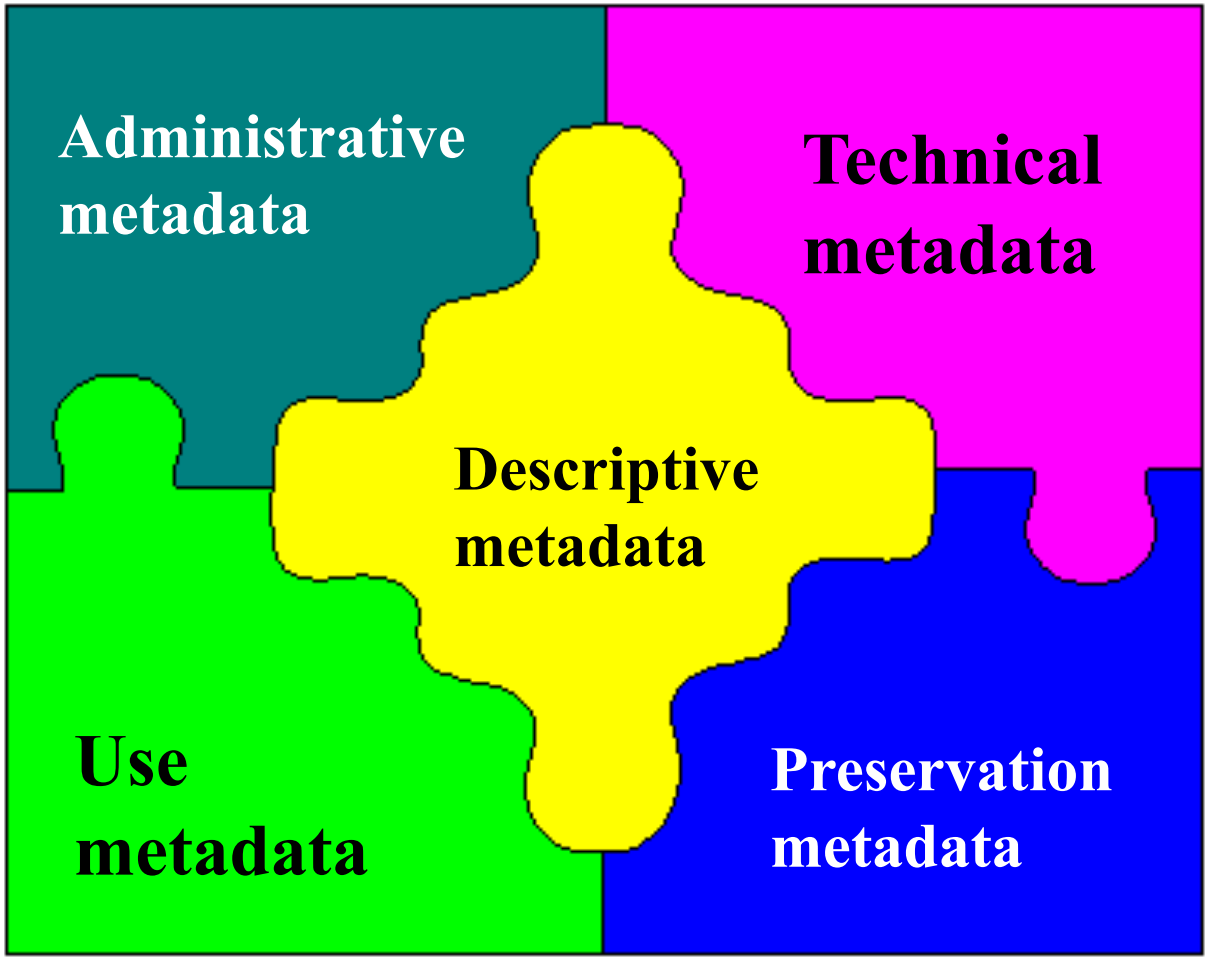
Dublin Core Metadata Initiative®

Extensibility (Zeng, M., 2004, JG *rev.*, 2006)

Allows for extensions to metadata schema so as to accommodate the particular needs of any given application



Modularity (Weibel's notion of Legos)



Dublin Core Metadata Element Set (DCMES)
(version 1.1)

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

•Dublin Core Element Set: http://dublincore.org/documents/dces/
•DCMI Metadata Terms: http://dublincore.org/documents/dcmi-terms

Finalized 1998 RFC 2413; CEN CWA 13874 (2000); ANSI/NISO Z39.85 (2001, rev. 2008); ISO 15386 (2003, rev. 2009); Internet RFC 5013 (2007)

Legacy/common usage

- Modest level above HTML <title> and <keywords> tags, and local XML elements like <name> or <author>
- Adding standard meaning
- Widely deployed for basic description and basic exchange (e.g. OAI-PMH, DL's, archives)
- "Flat" metadata model

Dublin Core Metadata Initiative®

DC semantics, thinking outside the box

- Harper, C. (2010). Dublin Core Metadata Initiative: Beyond the Element Set (www.niso.org/publications/isq/free/FE_DCMI_Harper_isqv22no1.pdf)
  - Shift the focus of the conversation from metadata *format* to metadata *vocabulary*
  - Link with RDF and the Semantic Web / Linked data movement
    - RDF uses URIs to identify a resource (subject) and properties (predicate)
- Why?

Dublin Core Metadata Initiative®

Why a shift in thinking?

- maybe not a shift, but the time is now

- Share and reuse metadata
- Generating metadata **costs money**
- There's a lot of duplication
- Semantic interoperability, linking, discovery, intelligence…
- Machine processability (take advantage, realize new opportunities)

★Finally, an information infrastructure, and enabling technology, together supporting longstanding ideas/ ideals

  – Turing machine; Bush's memex;

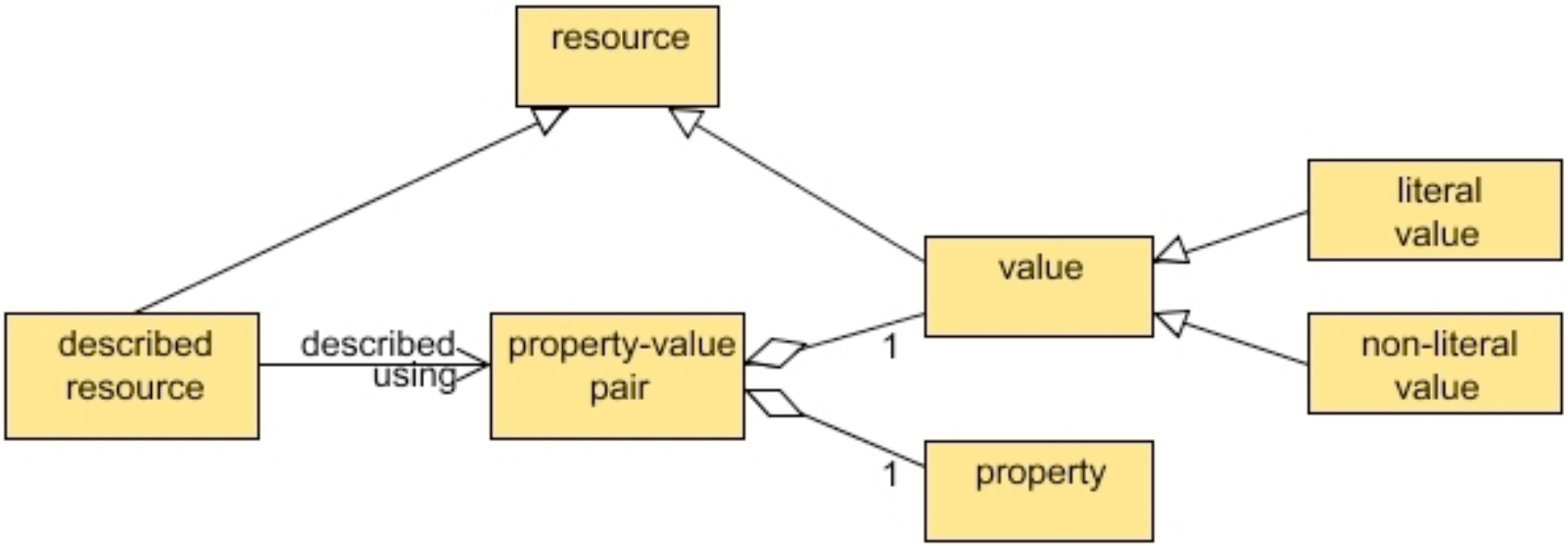
Dublin Core Metadata Initiative®

Dublin Core abstract model

- Describes how metadata statements can be packaged into records or “descriptions”
- Basic building block
  - statement that says one thing about one resource
    Subject-Predicate-Object (and RDF statement)
- A description contains one or more statements about one and only one resource (using DC and other vocabularies)

Dublin Core Metadata Initiative®

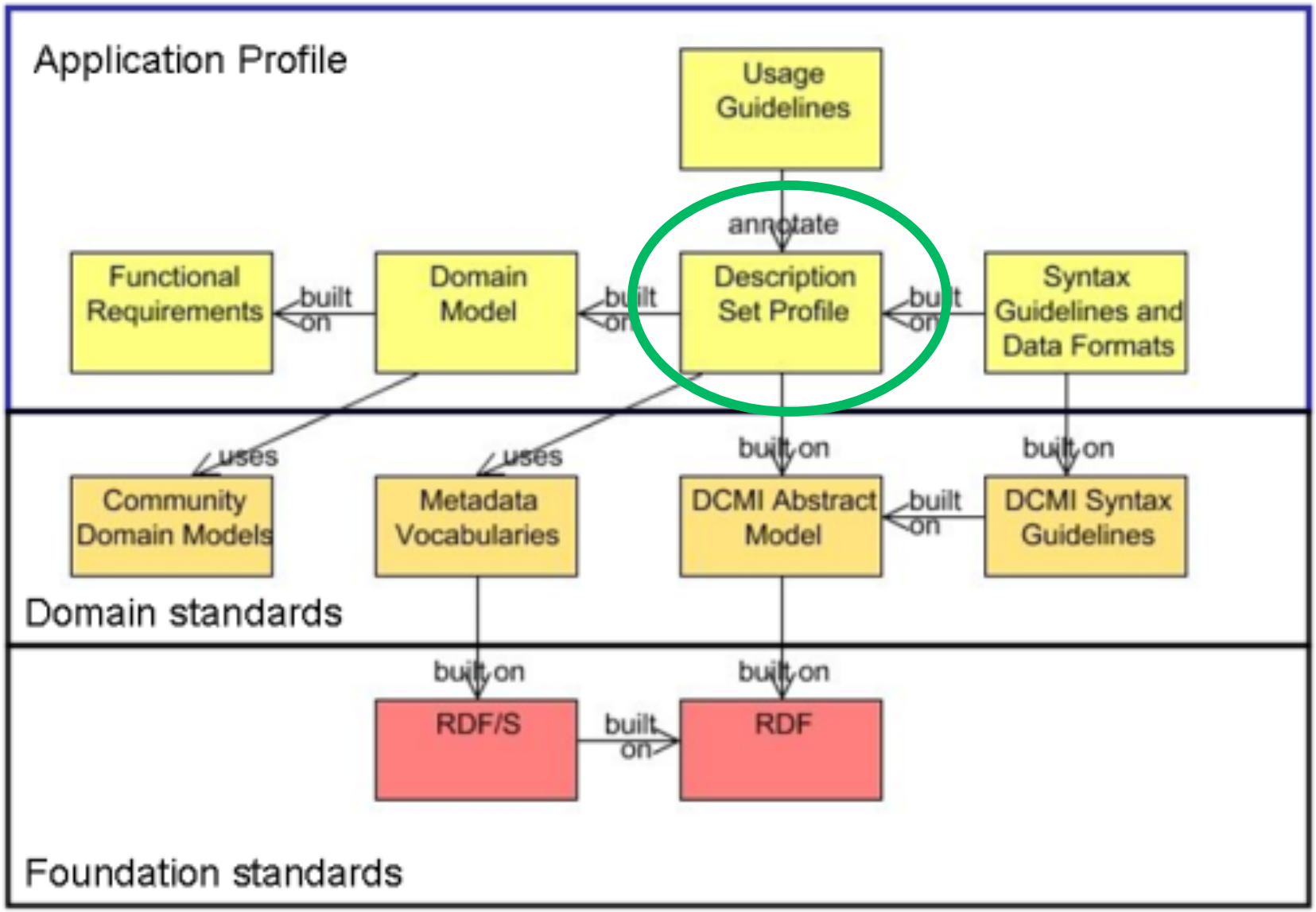
DCMI resource model



Dublin Core Metadata Initiative®

Singapore Framework

(machine processing, long term quality control Semantic web/linked data)



- **A loose standard for DC endorsed application profiles**

DCMI Metadata Terms (an evolving approach/view)

Creator: DCMI Usage Board
Date Issued: 2010-10-11
Document Status: This is a DCMI Recommendation.
Description: …an up-to-date specification of all metadata terms maintained by the Dublin Core Metadata Initiative, including properties, vocabulary encoding schemes, syntax encoding schemes, and classes.

Table of Contents

http://dublincore.org/documents/dcmi-terms/

abstract
accessRights
accrualMethod
accrualPeriodicity
accrualPolicy
alternative
audience
available
bibliographicCitation
conformsTo
contributor
coverage
created
creator
date
dateAccepted
dateCopyrighted
dateSubmitted
description
educationLevel
extent
format
hasFormat
hasPart
hasVersion
identifier
instructionalMethod
isFormatOf
isPartOf
isReferencedBy
isReplacedBy
isRequiredBy
issued
isVersionOf
language
license
mediator
medium
modified
provenance
publisher
references
relation
replaces
requires
rights
rightsHolder
source
spatial
subject
tableOfContents

Properties in the /terms/ namespace

DCMI Metadata Terms

(Minimal attributes)

Attribute	Description
Name:	A token appended to the URI of a DCMI namespace to create the URI of the term.
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 [DCAM].

Dublin Core Metadata Initiative®

DCMI Terms

Attribute	Description
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.
Equivalent Property:	A Property to which the described term is equivalent

**__Where applicable__, the above attributes provide additional information about a term.

Dublin Core Metadata Initiative®

DCMI Metadata Terms

(Minimal attributes/example)

Term Name: subject	
URI:	http://purl.org/dc/terms/subject
Label:	Subject
Definition:	The topic of the resource.
Comment:	Typically, the subject will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary. To describe the spatial or temporal topic of the resource, use the Coverage element.
Type of Term:	Property
Refines:	http://purl.org/dc/elements/1.1/subject
Version:	http://dublincore.org/usage/terms/history/#subjectT-001
Note:	This term is intended to be used with non-literal values as defined in the DCMI Abstract Model (http://dublincore.org/documents/abstract-model/). As of December 2007, the DCMI Usage Board is seeking a way to express this intention with a formal range declaration.

Dublin Core Metadata Initiative®

Dublin Core metadata

<dcterms:title>Godiva Chocolatier</dcterms:title>

<dcterms:alternative>Godiva store</dcterms:alternative>

<dcterms:creator>Confection, Nancy</dcterms:creator>

<dcterms:subject>Chocolate</dcterms:subject>

<dcterms:subject xsi:type="dcterms:lcsh">Truffles (Confectionery)</dcterms:subject>

<dcterms:created xsi:type="dcterms.W3CDTF">
2008--6--28</dcterms:created>

<dcterms:identifier>http://www.godiva.com</ dcterms:identifier>

<dcterms:abstract>Provides access to chocolate collections, gifts, ….</dcterms:abstract>

DCMI Metadata Terms

Table of Contents

Dublin Core Metadata Initiative®

Encoding

Vocabulary Encoding Schemes	DCMIType, DDC, IMT, LCC, LCSH, MESH, NLM, TGN, UDC
Syntax Encoding Schemes	ISO3166, ISO639-2, ISO639-3, Period, Point, RFC1766, RFC3066, RFC4646, RFC5646, URI, W3CDTF

Dublin Core Metadata Initiative®

Dublin Core metadata

```
<dcterms:title>Godiva Chocolatier</dcterms:title>
<dcterms:alternative>Godiva store</dcterms:alternative>
<dcterms:creator>Confection, Nancy</dcterms:creator>
<dcterms:subject>Chocolate</dcterms:subject>
<dcterms:subject xsi:type="dcterms:lcsh">Truffles
   (Confectionery)</dcterms:subject>
<dcterms:created xsi:type="dcterms.W3CDTF">
    2008--6--28</dcterms:created>
<dcterms:identifier>http://www.godiva.com</
   dcterms:identifier>
<dcterms:abstract>Provides access to chocolate
   collections, gifts, ….</dcterms:abstract>
```

DCMI Metadata Terms

Table of Contents

Dublin Core Metadata Initiative®

DC Metadata Terms

- Modern Dublin Core (the dcterms:namespace) defines domains and ranges; e.g. the value of dcterms:subject is the concept not the string – could be a SKOS:Concept:

  <dcterms:subject rdf:resource="http://example.org/taxonomy/D003.53"/>

- This usage is in line with Linked Data approach

Dublin Core Metadata Initiative®

DCMI Terms

Attribute	Description
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.
Equivalent Property:	A Property to which the described term is equivalent

**__Where applicable__, the above attributes provide additional information about a term.

Dublin Core Metadata Initiative®

Classes

Agent
AgentClass
BibliographicResource
FileFormat
Frequency
Jurisdiction
LicenseDocument
LinguisticSystem
Location
LocationPeriodOrJurisdiction
MediaType
MediaTypeOrExtent
MethodOfAccrual
MethodOfInstruction
PeriodOfTime
PhysicalMedium
PhysicalResource
Policy
ProvenanceStatement
RightsStatement
SizeOrDuration
Standard

Dublin Core
Metadata Initiative®

Example of a class

Term Name: Agent	
URI:	http://purl.org/dc/terms/Agent
Label:	Agent
Definition:	A resource that acts or has the power to act.
Comment:	Examples of Agent include person, organization, and software agent.
Type of Term:	Class
Instance Of:	http://purl.org/dc/terms/AgentClass
Version:	http://dublincore.org/usage/terms/history/#Agent-001

Example of a class

Term Name: BibliographicResource	
URI:	http://purl.org/dc/terms/BibliographicResource
Label:	Bibliographic Resource
Definition:	A book, article, or other documentary resource.
Type of Term:	Class
Version:	http://dublincore.org/usage/terms/history/#BibliographicResource-001

Dublin Core Metadata Initiative®

Syntax Guidelines

and more @: http://dublincore.org/specifications/

- DC-TEXT [DCMI Recommended Resource]. "Expressing Dublin Core metadata using the DC-Text format"
- DC-HTML [DCMI Recommendation]. "Expressing Dublin Core metadata using HTML/XHTML meta and link elements"
- DC-DS-XML [Proposed Recommendation]. "Expressing Dublin Core Description Sets using XML (DC-DS-XML)"
- DC-XML-GUIDELINES [DCMI Recommendation].
- DC-RDF [DCMI Recommendation]. "Expressing Dublin Core metadata using the Resource Description Framework (RDF)"

Dublin Core Metadata Initiative®

Principles

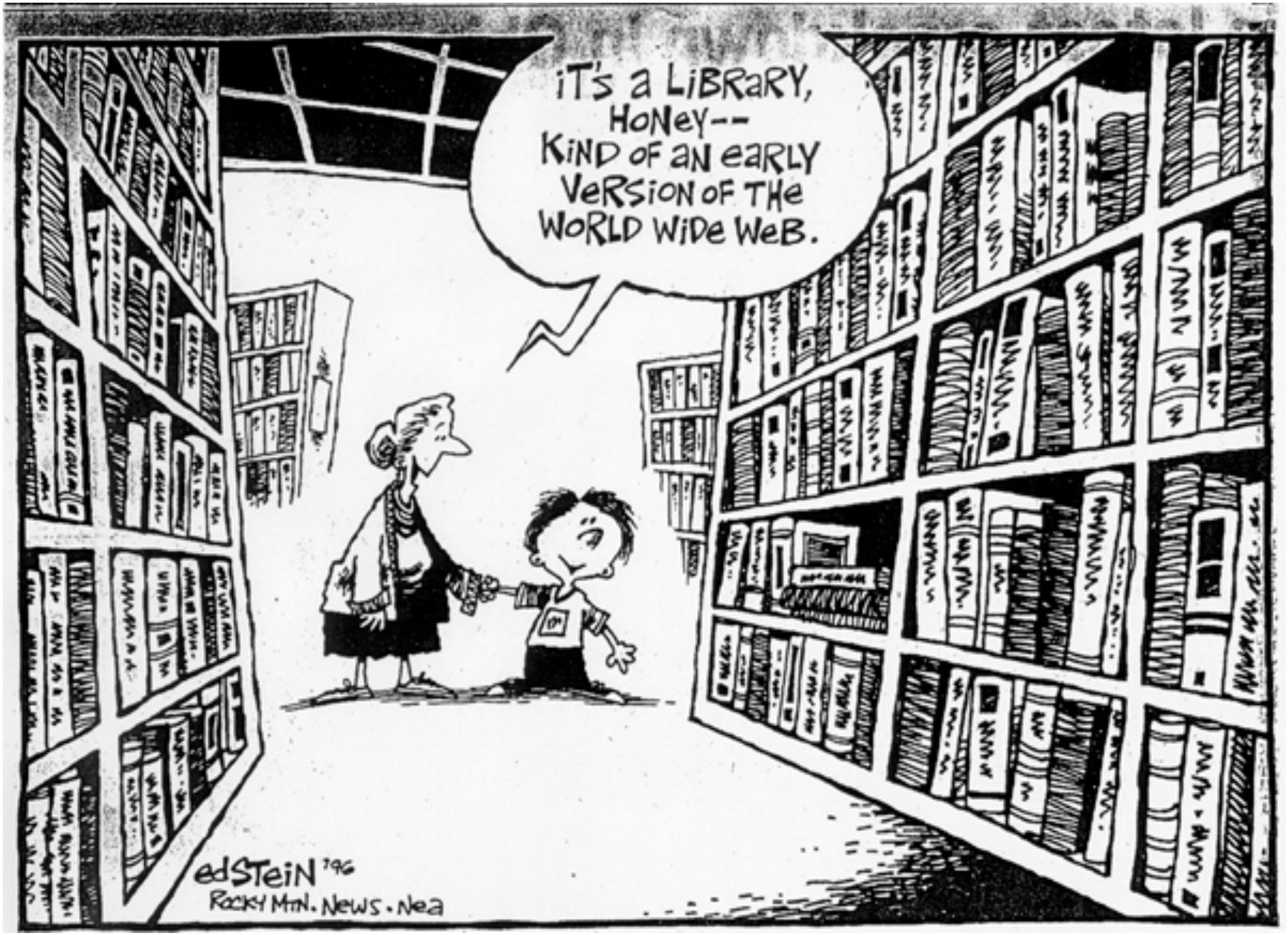
Original ideas

- Simple
- Semantic interoperability
- International consensus
- Interdisciplinary (any domain)
- For any format
- Extensible
- Modular

Today…(Makx Dekkers)

- Open consensus building:
- International scope
- Neutrality of purposes and business models
- Neutrality of technology
- Cross-disciplinary focus

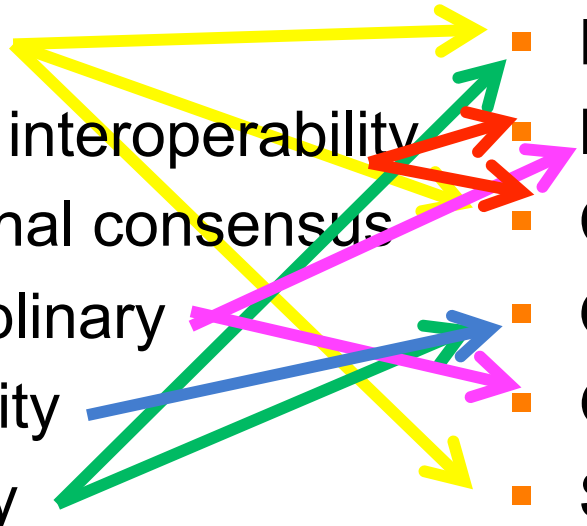
Dublin Core Metadata Initiative®


iT's a LiBRaRY, HONey-- KiND OF aN eaRLY VeRSiON OF THe WORLD WiDe WeB.
edSTeiN '96
Rocky Mtn. News · Nea

Toward a Semantic Web

Founding principles

- Simplicity
- Semantic interoperability
- International consensus
- Interdisciplinary
- Extensibility
- Modularity



Key Ideas and approaches

- RDF (very simple)
- Metadata/data reuse
- Consistency
- One size does not fit all
- Global context
- Smarter web, smarter data
  - Machine processable
- Application profiles

Dublin Core Metadata Initiative®

DCMI

- DCMI
  - *An initiative, a community*
  - Committed to **open standards**, support **interoperability**
- DCMI Community emphasizes open participation
  - Conferences, communities, discussion lists/wikis
- DC structure
  - Incorporated as non-profit organization in Singapore
  - Governance:
    - DCMI Executive (Managing Director/Chief Executive Officer and a Chief Information Officer; day-to-day)
    - Oversight committee (liaisons and promotion)
    - Advisory Board: technical and operational advice
    - Usage Board: maintenance and review of proposals
    - Partners: financial support
  - Work by the architecture forum, communities, and task groups

Dublin Core Metadata Initiative®

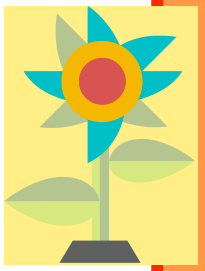
Summary

- Dublin Core (history & growth)
  - -History workshop → *full* conference
  - -Founding principles (characteristics)
- Toward a *more* Semantic Web… more guidelines
  - -DCMI Metadata Terms
- Dublin Core Metadata Initiative (DCMI)

Q&A

Thank you!

janeg@emal.unc.edu



-Marcia Zeng
→
←Makx Dekkers

-Tom Baker
-Corey Harper
-Diane Hillmann