

Dublin Core in Practice

Marcia L. Zeng

mzeng@kent.edu

DC2008, Berlin, 2008-09-22

Implementation

= the realization of an application
= execution of a plan, idea, model,
design, specification, standard ...



Tangram = 七巧板 = qī qiǎo bǎn = "seven boards of skill"

Understanding the Pieces of the Puzzle

Outline

1. Standards
2. Metadata types and functions
3. Crosswalking and mapping
4. Sources of metadata
5. Quality control
6. Workflows
7. Reference sources

Understanding the Pieces of the Puzzle (1)

1. Standards

- Standards for ***data structures and semantics***
 - metadata element sets
 - e.g., Dublin Core Metadata Element Set ([DCMES](#))
- Standards and Guides for ***data content***
 - created to guide the practices of metadata generation
 - e.g., [CCO](#), [Using DC](#)
- Standards for ***data values***
 - value encoding schemes ([link to list](#))
 - e.g., [DCMI Type Vocabulary](#) | [RFC 4646 Tags for Identifying Languages](#)
- Standards for ***data exchange***

Term Name:	type
URI:	http://purl.org/dc/elements/1.1/type
Label:	Type
Definition:	The nature or genre of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary [DCMITYPE]. To describe the file format, physical medium, or dimensions of the resource, use the Format element.
References:	[DCMITYPE] http://dublincore.org/documents/dcmi-type-vocabulary/

Standards for ***data structures and semantics*** –element sets
 Standards and Guides for ***data content*** – rules and guides
 Standards for ***data values*** – value encoding schemes

-- from *Dublin Core Metadata Element Set, Version 1.1*

Term Name: date	
URI:	http://purl.org/dc/elements/1.1/date
Label:	Date
Definition:	A point or period of time associated with an event in the lifecycle of the resource.
Comment:	Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF].
References:	[W3CDTF] http://www.w3.org/TR/NOTE-datetime

Examples from values associated with *DATE* element

- ◆ 1979
- ◆ 2000-03
- ◆ 2000-03-01
- ◆ 2001-01-02T21:48.00Z
- ◆ 200003
- ◆ C1999, 2000
- ◆ January, 1919
- ◆ May, 1919
- ◆ 1987, c2000
- ◆ ?1999
- ◆ 1952 (issued)
- ◆ (1982)
- ◆ 1930?]
- ◆ 1823-1845
- ◆ Between 1680 and 1896?

Value space that should follow *syntax encoding schemes*

Term Name: language	
URI:	http://purl.org/dc/elements/1.1/language
Label:	Language
Definition:	A language of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as RFC 4646 [RFC4646].
References:	[RFC4646] http://www.ietf.org/rfc/rfc4646.txt

Examples from values associated with *LANGUAGE* element

- ◆ en
- ◆ eng
- ◆ en-GB
- ◆ en-US
- ◆ English
- ◆ engfre
- ◆ new
- ◆ Korean
- ◆ Deutsch
- ◆ German LOCLANGUAGE:: German

**Value space that should apply
controlled vocabularies / value encoding schemes**

Examples from values associated with **FORMAT** element

Term Name: format	
URI:	http://purl.org/dc/elements/1.1/format
Label:	Format
Definition:	The file format, physical medium, or dimensions of the resource.
Comment:	Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types [MIME].

- ◆ text
- ◆ text/html
- ◆ text/plain
- ◆ plain
- ◆ digital TIFF
- ◆ image/tiff
- ◆ other http://...../.../postscript.pdf
- ◆ application/msword
- ◆ application/Flash (animation)
- ◆ ascii http:///.../sample.txt
- ◆ pdf http:///.../sample.pdf
- ◆ ps http:///.../sample.ps
- ◆ 3.6 megabytes
- ◆ 1000149 bytes
- ◆ language/java
- ◆ Application/JAVA applet
- ◆ Java
- ◆ CLASS
- ◆ Model/VRML
- ◆ AVI, MOV, QTM
- ◆ 1 v. (various pagings)
- ◆ 10 p., [6] p. of plates
- ◆ p.461-470
- ◆ viii, 82 p.
- ◆ MPEG-4

**Value space that should apply
controlled vocabularies / value encoding schemes**

Implement standards and follow best practices to ensure consistency

- ◆ In shared repositories, consistency may be an issue in the following areas:

- Data Recording
- Source Links
- Identification and Identifiers
- Description of Sources
- Metadata Representation (in the search results display)
- Data Syntax
-

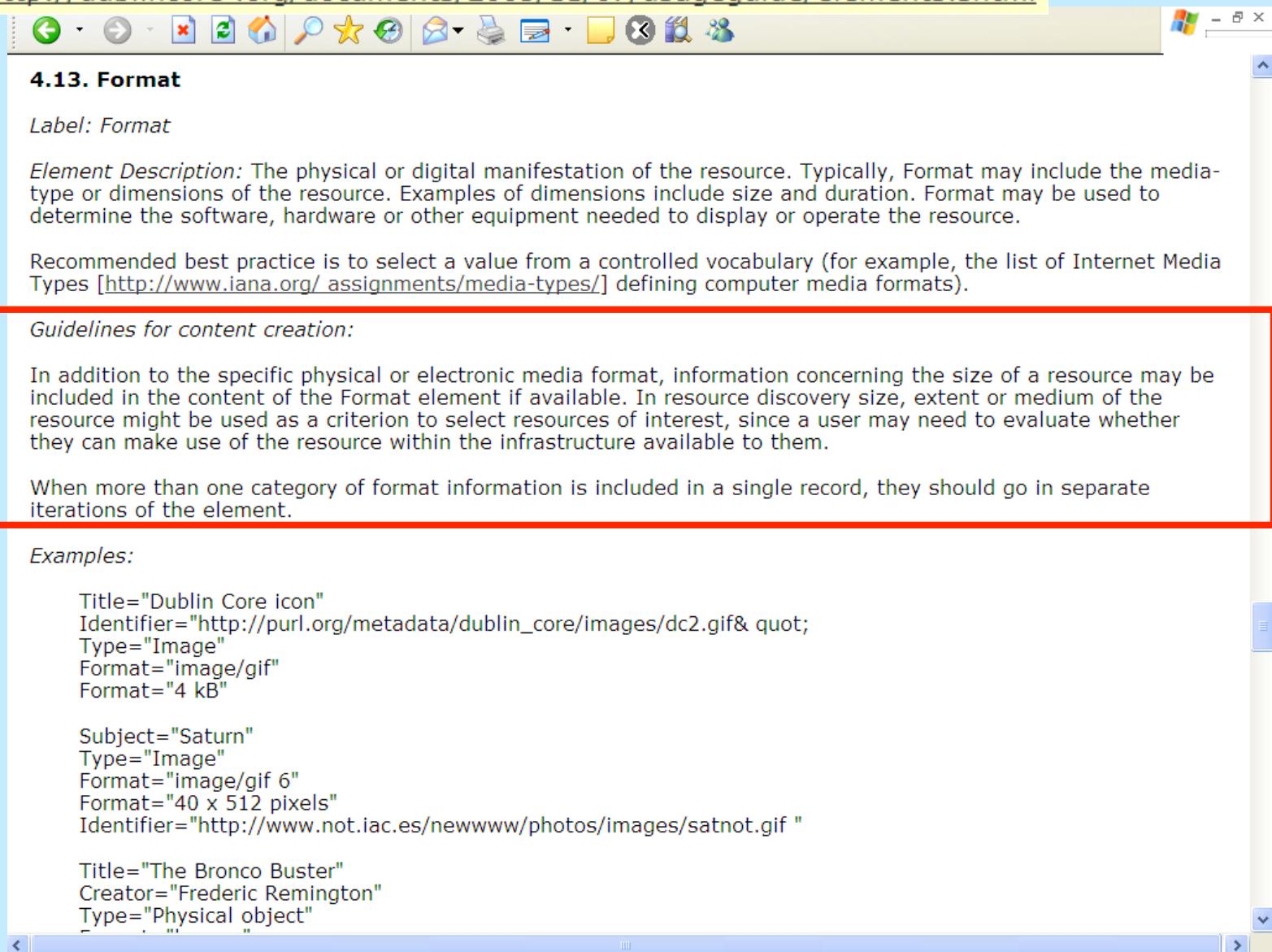
Using Dublin Core - The Elements

Creator: Diane Hillmann

Date Issued: 2005-11-07

Identifier: <http://dublincore.org/documents/2005/11/07/usaguide/elements.shtml>

Implementing Content Standards and Guidelines



The screenshot shows a Microsoft Internet Explorer window displaying the '4.13. Format' section of the Dublin Core Usage Guide. The page includes a toolbar at the top, a header with the section title, a label definition, an element description, recommended best practices, guidelines for content creation, examples of XML code, and a sidebar on the right.

4.13. Format

Label: Format

Element Description: The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Examples of dimensions include size and duration. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource.

Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [[http://www.iana.org/ assignments/media-types/](http://www.iana.org/assignments/media-types/)] defining computer media formats).

Guidelines for content creation:

In addition to the specific physical or electronic media format, information concerning the size of a resource may be included in the content of the Format element if available. In resource discovery size, extent or medium of the resource might be used as a criterion to select resources of interest, since a user may need to evaluate whether they can make use of the resource within the infrastructure available to them.

When more than one category of format information is included in a single record, they should go in separate iterations of the element.

Examples:

```
Title="Dublin Core icon"
Identifier="http://purl.org/metadata/dublin_core/images/dc2.gif"
Type="Image"
Format="image/gif"
Format="4 kB"

Subject="Saturn"
Type="Image"
Format="image/gif 6"
Format="40 x 512 pixels"
Identifier="http://www.not.iac.es/newww/photos/images/satnot.gif"

Title="The Bronco Buster"
Creator="Frederic Remington"
Type="Physical object"
```

Implementing Content Standards and Guidelines

Implementing Content Standards and Guidelines

ShareableMetadataPublic - oaibp - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Favorites

Address <https://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/ShareableMetadataPublic> Go Google DCMI Settings

article discussion edit history

ShareableMetadataPublic

Main Page

Contents [hide]

1 Best Practices for Shareable Metadata

- 1.1 Introduction
- 1.2 General Recommendations
- 1.3 Recommendations for Classes of Data Elements
- 1.4 Best Practices for Technical Aspects of Metadata
- 1.5 Final Preparations

Best Practices for Shareable Metadata

[edit]

A MS Word version of this document is available at:

<https://webservices.itcs.umich.edu/mediawiki/oaibp/sites/oaibp/uploads/f/f4/ShareableMetadataPublic.doc> (August 2005)

Digital Library Federation

navigation

- Main Page
- Community portal
- Current events
- Recent changes
- Random page
- Help
- Donations

search

Go Search

toolbox

- What links here
- Related changes

http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/ShareableMetadataPublic

Understanding the Pieces of the Puzzle (2)

2. Metadata Types and Functions

The DCMI defines metadata as ...

- ◆ data associated with either an information system or an information object for purposes of:
 - description
 - administration
 - legal requirements
 - technical functionality
 - use and usage
 - preservation

--DCMI Glossary, 2005

<http://dublincore.org/documents/usageguide/glossary.shtml>

Types of Metadata

Type	Definition	Examples
Administrative	Metadata used in managing and administering information resources	<ul style="list-style-type: none"> - Acquisition information - Rights and reproduction tracking - Documentation of legal access requirements - Location information - Selection criteria for digitization - Version control and differentiation between similar information objects - Audit trails created by record keeping systems
Descriptive	Metadata used to describe or identify information resources	<ul style="list-style-type: none"> - Cataloging records - Finding aids - Specialized indexes - Hyperlinked relationships between resources - Annotations by users - Metadata for record keeping systems generated by records creators
Preservation	Metadata related to the preservation management of information resources	<ul style="list-style-type: none"> - Documentation of physical condition of resources - Documentation of actions taken to preserve physical and digital versions of resources, e.g., data refreshing and migration
Technical	Metadata related to how a system functions or metadata behave	<ul style="list-style-type: none"> - Hardware and software documentation - Digitization information, e.g., formats, compression ratios, scaling routines - Tracking of system response times - Authentication and security data, e.g., encryption keys, passwords
Use	Metadata related to the level and type of use of information resources	<ul style="list-style-type: none"> - Exhibit records - Use and user tracking - Content re-use and multi-versioning information

-- Baca, M. ed. *Introduction to Metadata: Pathways to Digital Information*. V.3. Anne J. Gilliland: Setting the Stage, Table 2.

Greenberg: Mapping “Types” to “Functions”

Typology of 7 types of metadata	Metadata Functions <i>This type of metadata satisfies:</i>	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 by Greenberg, J. "Metadata Quality: A Layer Cake of Criteria", presentation at ASIST 2005 Conference

A wish list of Functions

We would like users to be able to:

Search records by: Title, Author name, Keyword,
Type of document, Publication, Conference
name, and Year

Browse records by: Year, Department,
Classification, Object type, Subject matter, etc.

View latest additions to the archive

We would like to be able to:

Link together records from the same Conference,
Publication

Filter by: Year and Language

- Based on Guy, Powell, and Day. "Improving the Quality of Metadata in Eprint Archives." *Ariadne*, no. 38. 2004.

An operational system

The screenshot shows a search interface with four main tabs at the top: General Search (selected), K-12, Higher Ed & Research, and Specialized Search. Below the tabs is a search bar with a green "Search" button. To the right of the search bar are links for "Reset Search Options" and "Search Help". A red box highlights a section containing two rows of search filters. The first row is for "Search by Grade Level" with checkboxes for Grades preK-2, Grades 3-5, Grades 6-8, and Grades 9-12. The second row is for "Search by Format" with checkboxes for Audio, Data, Image, Interactive resource, Text, and Video.

General Search K-12 Higher Ed & Research Specialized Search

Search

Reset Search Options Search Help

Search by Grade Level: Grades preK-2 Grades 3-5 Grades 6-8 Grades 9-12

Search by Format: Audio Data Image Interactive resource Text Video

Will the metadata support these functions?

title

[Mars](#)
This site focuses on the planet Mars. Some topics discussed are Mars interior, surface, atmosphere,...
[View all related information](#) - [Include/Exclude results like this](#)

abstract

[Trip to Mars](#)
Trip to Mars
[View all related information](#) - [Include/Exclude results like this](#)

[Mars](#)
This National Aeronautics and Space Administration (NASA) Observatorium site contains a tutorial ab...
[View all related information](#) - [Include/Exclude results like this](#)

media type

[text]

[text] [image]

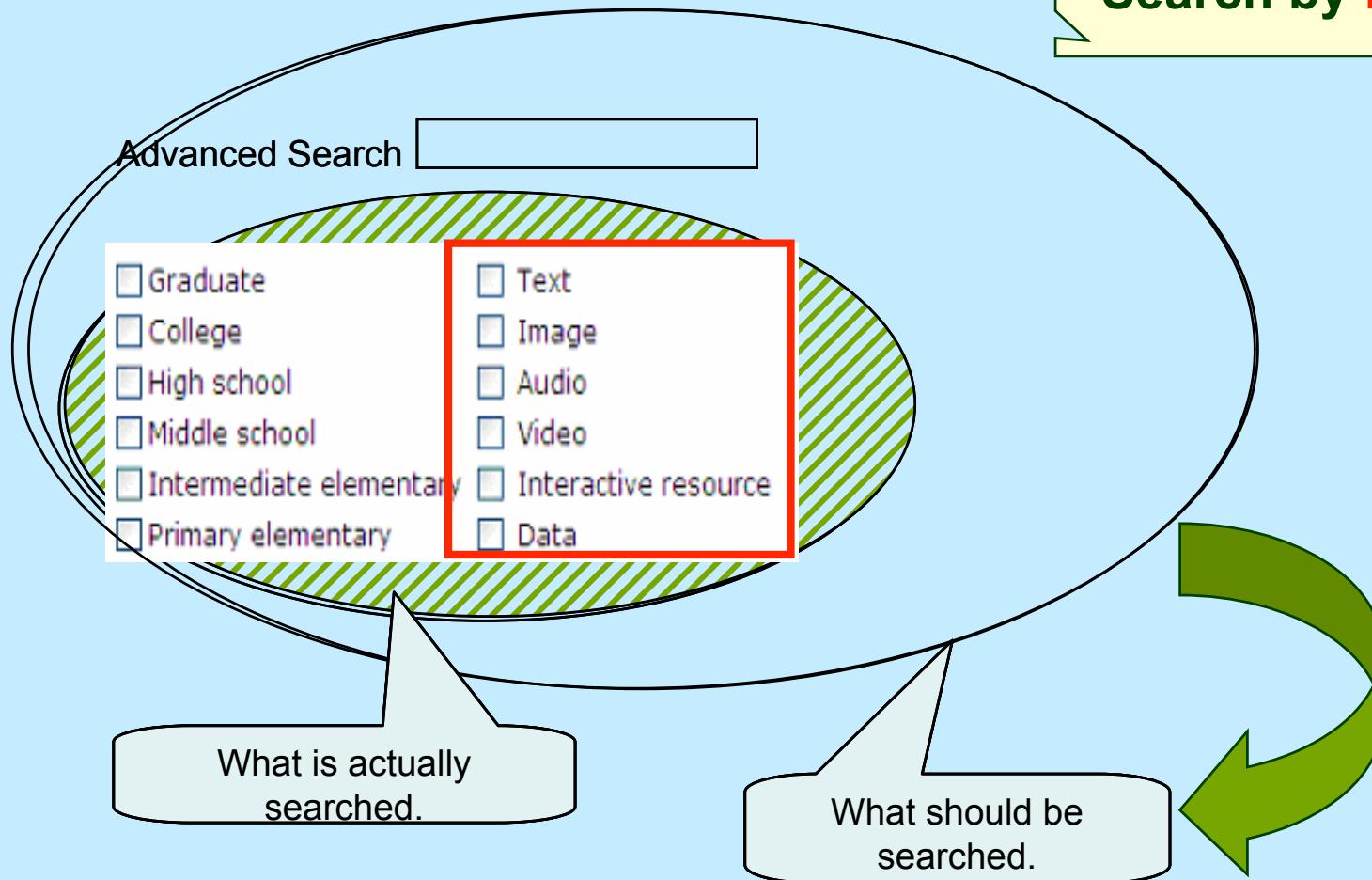
Video?
Image?
Text?

No FORMAT information

The diagram illustrates the process of extracting metadata from web pages. It starts with two boxes labeled 'abstract' and 'title' pointing to specific fields ('[text]' and '[text] [image]') on three separate web pages. Arrows then point from these fields to a thought bubble containing 'Video?', 'Image?', and 'Text?'. Finally, an arrow points from the thought bubble to a large green text box stating 'No FORMAT information'.

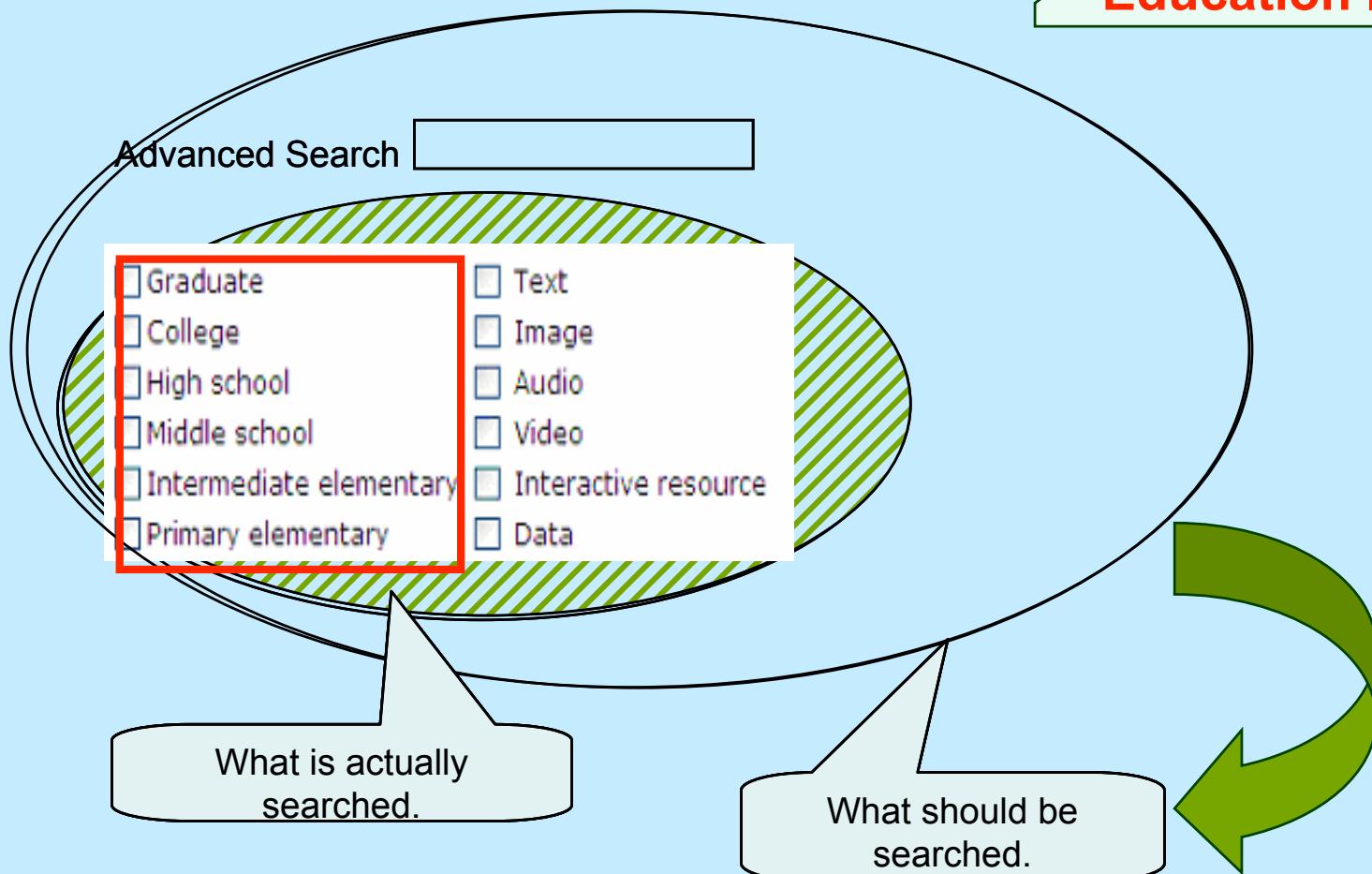
- Collections which did not provide FORMAT information are excluded from being searched

Search by format



- Collections which did not provide Education Level information are excluded from being searched

Search by
Education level

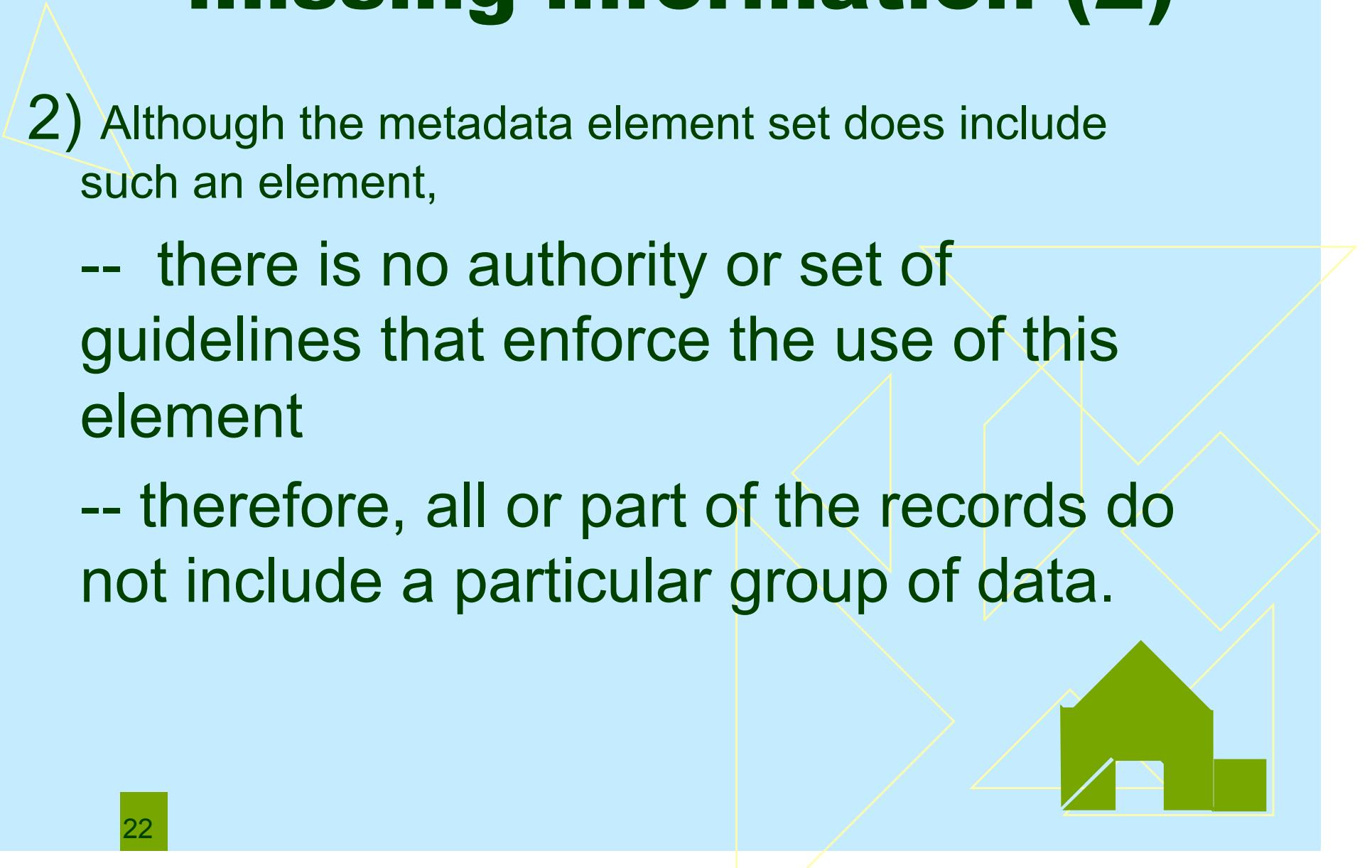


The causes of missing information (1)

1) The metadata element set might not include such an element.

e.g., the legacy DC 15 elements does not include EDUCATION LEVEL

The causes of missing information (2)

- 
- 2) Although the metadata element set does include such an element,
 - there is no authority or set of guidelines that enforce the use of this element
 - therefore, all or part of the records do not include a particular group of data.

File Edit View D

Back

Google G

For example, for the same source,
records quality could be very different --

NATIONAL AERONAUTICS
AND SPACE ADMINISTRATION[+ NASA Homepage](#)
[+ NASA en Español](#)
[+ Marte en Español](#)

SEARCH MARS



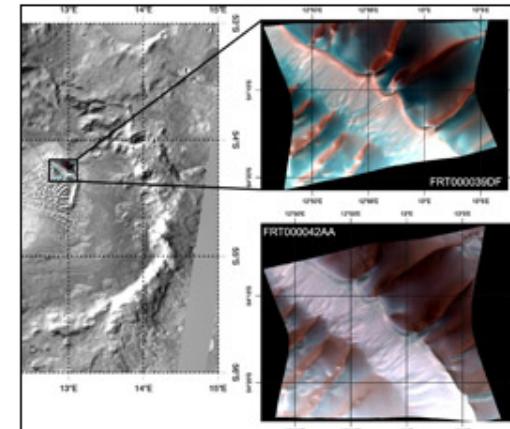
NASA's Mars Exploration Program

[OVERVIEW](#)[SCIENCE](#)[TECHNOLOGY](#)[MISSIONS](#)[PEOPLE](#)[FEATURES](#)[EVENTS](#)[MULTIMEDIA](#)[ALL ABOUT MARS](#)[Mars for Kids](#)[Mars for Students](#)[Mars for Educators](#)[Mars for Press](#)[+ CURRENT MISSIONS](#)[MARS GLOBAL SURVEYOR](#)[2001 MARS ODYSSEY](#)[MARS EXPLORATION ROVER](#)[MARS EXPRESS](#)[MARS RECONNAISSANCE ORBITER](#)[+ ALL MARS MISSIONS](#)

Latest Rover/Lander Update

[Silica-Rich Soil Found by Spirit >>](#)

Latest Orbiter Update

[Making Quick Work of its Science Goals >>](#)Latest: [Mars Global Surveyor](#)[Mars Exploration Rover](#)

Internet

start



>>

yizha...

2 Mi...

3 In...

ppt

Macro...

OAI

EN



12:07 PM

Resource Information:

Description All sorts of information about and from the Mars Global Surveyor, Mars Pathfinder, Mars Surveyor 98, Mars Surveyor 2001, Mars Exploration Rover and other missions is presented. Status reports, science results, photographs, animations, images, and specifications of the spacecraft are among the documents included.

Subject Mars

Keyword(s) Outer space – Exploration

Animations

G

Images

Lobal surveyor

MGS

Mars Climate Orbiter

Mars Exploration Rover

Mars Global Surveyor

Mars Lander

Mars Orbiter

Mars Polar Lander

Mars Surveyor

Mars missions

Mars pathfinder

Marx Exploration Rovers

Outer space exploration

Pathfinder

Photographs

Photos

Pictorial works

Space exploration

Surveyor

Video

Videos

Web

Title

Mars Exploration

Record #1

No date.

No format, type, grade level, language, rights.

Resource Information:

Description	The Jet Propulsion Laboratory (JPL) is the leader in robotic Mars Exploration. Explore the wonders of Mars with JPL's help, by visiting the Mars Exploration home page. You will find fascinating information in sections called: Extreme Planet, Follow the Water, Mars Rocks, Life on Mars?, The Martian Mystique and Just for Kids.
Resource Format	text/html
Language	en
Subject Keyword(s)	Mars exploration, red planet, Mars odyssey, 2001 Mars, Mars launch, Mars climate, spacecraft assembly, Jet Propulsion Laboratory, Mars program, robotic Mars, JPL Mars, Pathfinder, Mars ice, life on Mars, NASA Mars, Mars probe, Mars Home, Mars education, Mars teachers, space robot, sojourner, Martians, Mars fun, Mars kids, children Mars, learn about Mars, explore Mars, men on Mars, Mars Global Surveyor, Mars orbiter, Mars lander, Mars Polar Lander, Mars landing
Title	Mars Exploration Homepage
Resource Type	Text

Record #2

No date.

format missed images.

No grade level, rights.

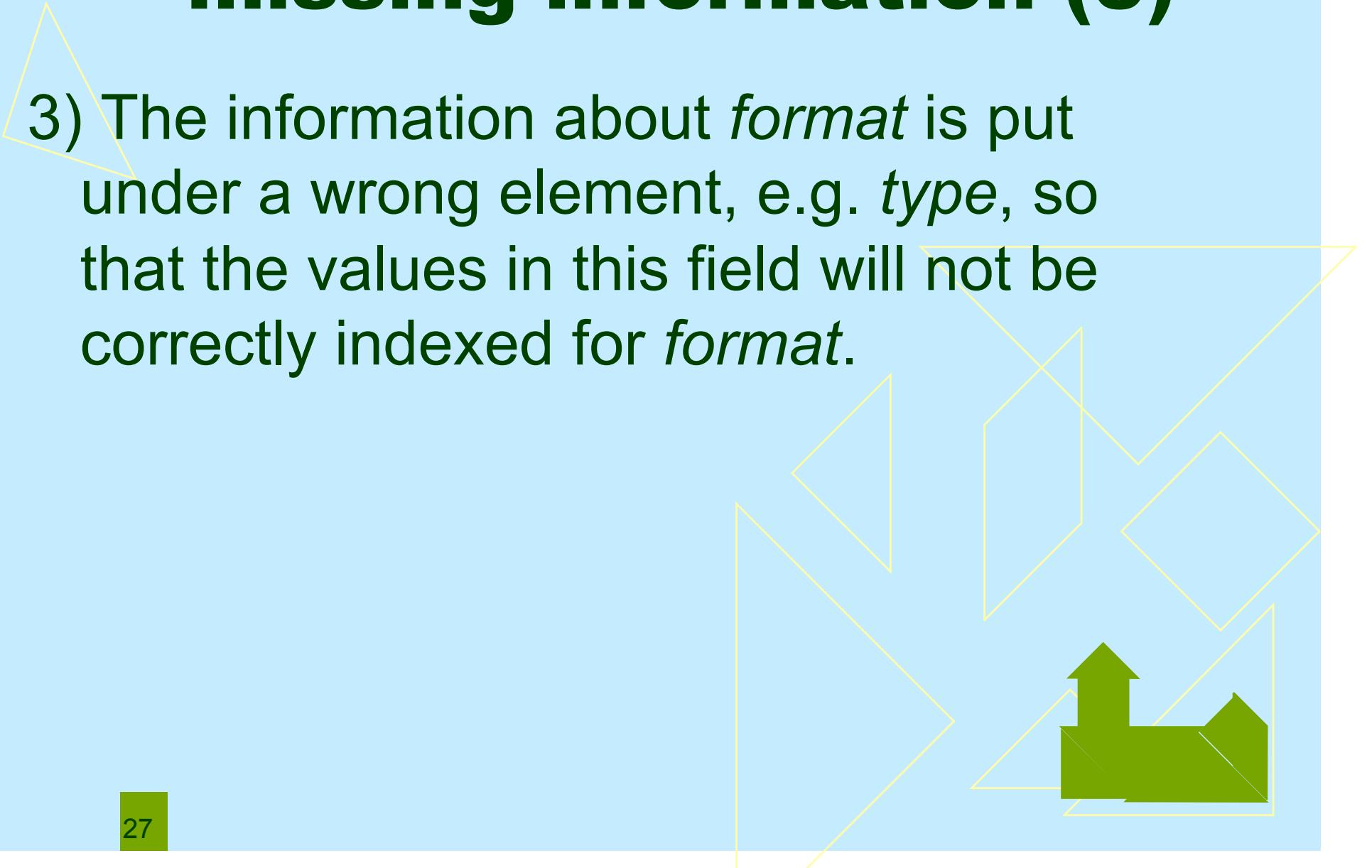
Record #3

Recorded grade level, format, type +
No date, language, rights.

Resource Information:

Contributor(s)	National Aeronautics and Space Administration (NASA)
Description	The Mars Exploration Program is a science-driven program that seeks to understand whether Mars was, is, or can be, a habitable world. To find out, we need to understand how geologic, climatic, and other processes have worked to shape Mars and its environment over time, as well as how they interact today.
Resource Format	image text/html
Subject Keyword(s)	Astronomy -- Study and teaching Astronomy Outer space -- Exploration Solar system Space sciences -- Study and teaching Astronomy Geology mars exploration; mars exploration rover; mars exploration rover spirit; mars rover; mars rover opportunity; NASA's mars exploration rover spirit; NASA's opportunity; NASA's spirit; opportunity rover; rover navigation
Title	Mars Exploration
Resource Type	Collection Image Text
Grade Level	Grades Pre-K to 12

The causes of missing information (3)

- 
- 3) The information about *format* is put under a wrong element, e.g. *type*, so that the values in this field will not be correctly indexed for *format*.

The causes of missing information (4)

- 4) The values are lost when the metadata records are converted from one database to another (due to incorrect mapping).

Original Record

Title	Game Theory Simulation
Author	Stephan Waner
Category	Game Theory
Subject	Game Theory
Type	Interactive Tutorial
Description	Here is a little on-line Javascript utility for game theory (row and column player). It is also designed to play against mixed strategy most of the time...) Notes: This will only work in Internet Explorer, version 3 or later. You need only enter the non-zero probabilities for your row strategies. (The computer will set the others to zero. To play against the computer, enter the row strategies. (The computer does not know your mixed strategy.) To "Play" and click on row strategies. (The computer will do the same for column strategies.)
URL	http://people.hofstra.edu/faculty/Stefan_Waner/RealWorld/GTSim.html

Converted Record

DC: Elements: 8	
dc:title:	Game Theory Simulation
dc:creator:	Stephan Waner
dc:subject:	Game Theory
dc:description:	Here is a little on-line Javascript utility for game theory (row and column player). It is also designed to play against mixed strategy most of the time...) Notes: This will only work in Internet Explorer, version 3 or later. You need only enter the non-zero probabilities for your row strategies. (The computer will set the others to zero. To play against the computer, enter the row strategies. (The computer does not know your mixed strategy.) To "Play" and click on row strategies. (The computer will do the same for column strategies.)
dc:date:	2003-01-01
dc:identifier:	http://www.econport.org:8080/econport/request?page=web_or_summary&contentMetadataID=116
dc:source:	http://people.hofstra.edu/faculty/Stefan_Waner/RealWorld/GTSim.html
dc:language:	en

- ◆ These records' TYPE values were not converted to the repository.
- ◆ These resources are excluded when users search by resource type.

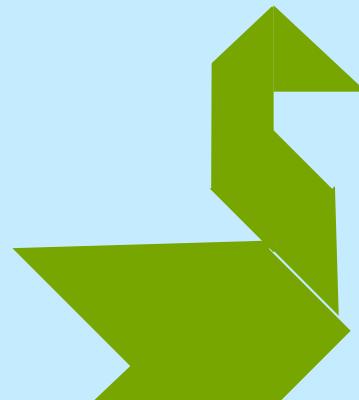
SEARCH

Search By Resource Format

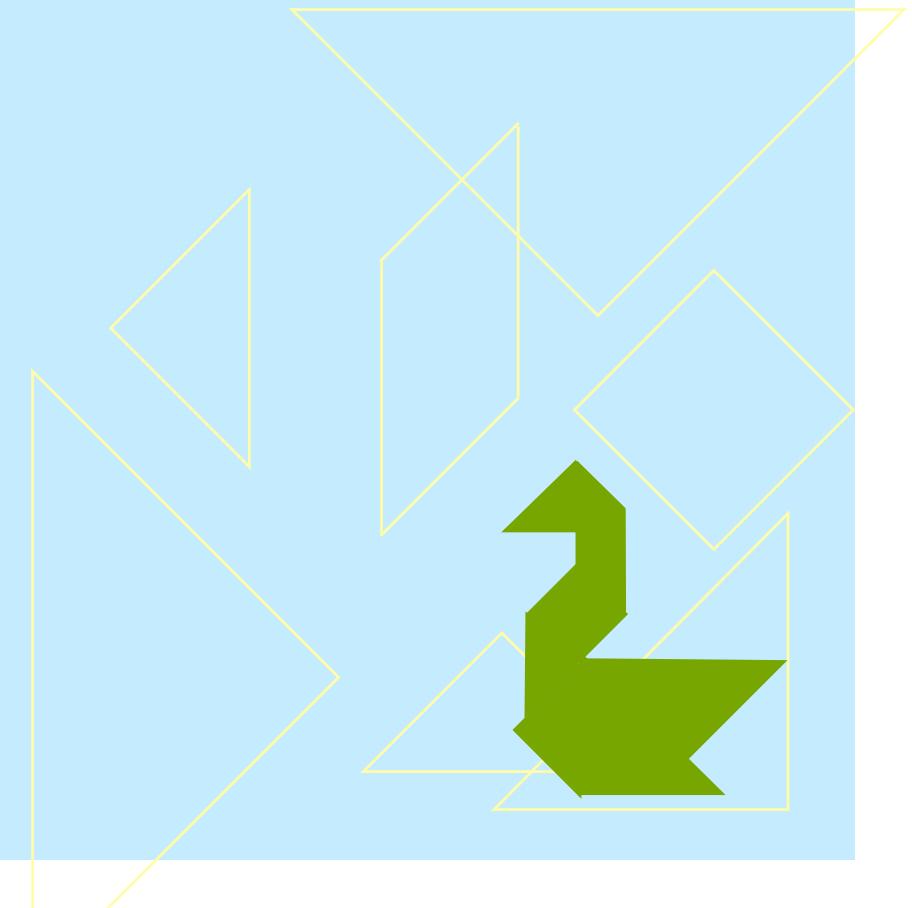
Text
 Image
 Audio
 Video
 Interactive resource
 Data

Understanding the Pieces of the Puzzle (3)

3. Crosswalking and Mapping



30



... ...

4- TECHNICAL

- o 4.1 Format
- o 4.2 Size (bytes)
- o 4.3 Location
<http://lcs.www.media.mit.edu/people/fredm/papers/mb/>
- o 4.4 Requirements

Composite

Type : Operating System

Name : Min/Max Operating System Requirement

- o 4.5 Installation remarks
- o 4.6 Other Platform Requirements
- o 4.6 Duration

N/A

5- EDUCATIONAL

- o 5.1 Interactivity Type
Unknown
- o 5.2 Learning Resource Type
Technical Reference
- o 5.3 Interactivity Level
Lower Division
- o 5.4 Semantic Density
Unknown
- o 5.5 Intended End User Role
- o 5.6 Learning Context
- o 5.7 Typical Age Range
- o 5.8 Difficulty
Unknown
- o 5.9 Typical Learning Time
N/A
- o 5.10 Description
No description
- o 5.11 Language
English

6- RIGHTS

- o 6.1 Cost
0.00USD
- o 6.2 Copyright and Other Restrictions
- o 6.3 Description

7- RELATION

8- ANNOTATION

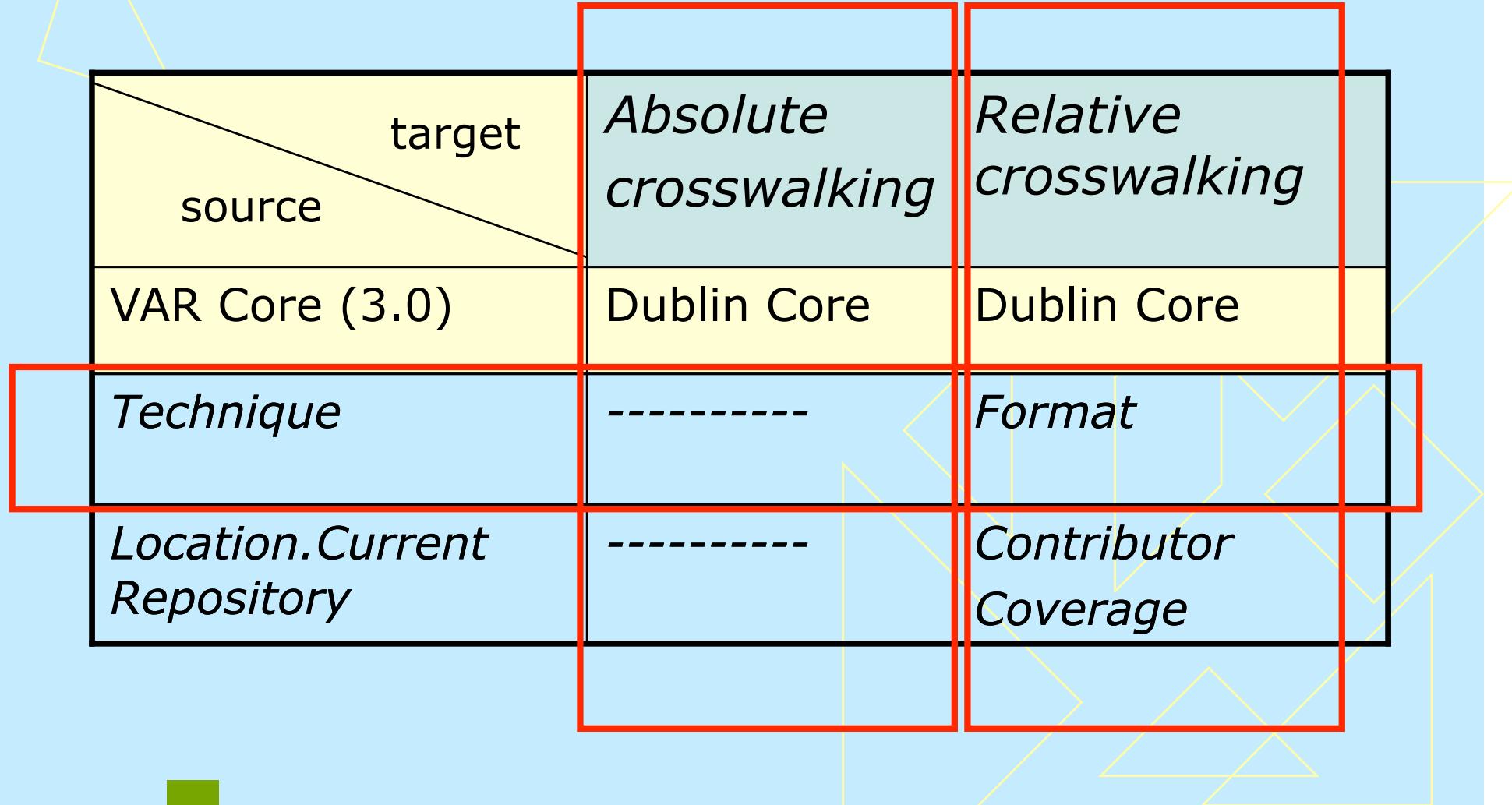
9- CLASSIFICATION

A LOM record (partial)

- hierarchical structure (not flat)
- many more elements than DC
- several elements correspond to one DC element

7

Common Crosswalking Approaches



Pros and Cons

-- Absolute Crosswalking

vra:title = dc:title

vra:technique = 

- ◆ ensures the equivalency (or closely-equivalent matches) of elements
- ◆ does not work well for data conversion
- ◆ data values in non-mappable space will be left out, especially when a source schema has a richer structure than that of the target schema.
 - E.g., from LOM records to DC records

Pros and Cons

-- Relative Crosswalking

vra:title = dc:title

vra:technique = dc:format

- ◆ When data conversion is conducted, at least the values in the fields of a source database will find a place to reside in the target database.
- ◆ Appears to work better when mapping from complex to simpler schema, e.g., from MARC to DC, but not vice versa. <http://www.loc.gov/marc/marc2dc.html>
- ◆ The problem is that mapped elements are not really equivalent.

before ...

Incorrect element mapping

OPTIONS mapped to **SUBJECT**, missing all **KEYWORDS**

Type of Document	Dissert										
Author	Pivovarov, Eugene										
URN	etd-05302002-130637										
Persistent URL	http://resolver.caltech.edu/CaltechETD:etd-05302002-130637										
Title	Aspects of non-Fermi-liquid metals										
Degree	PhD										
Option	Physics										
Advisory Committee	<table border="1"><tr><td>Advisor Name</td><td>Title</td></tr><tr><td>John Preskill</td><td>Committee C</td></tr><tr><td>Chetan Nayak</td><td>Committee N</td></tr><tr><td>Michael Cross</td><td>Committee N</td></tr><tr><td>Nai-Chang Yeh</td><td>Committee N</td></tr></table>	Advisor Name	Title	John Preskill	Committee C	Chetan Nayak	Committee N	Michael Cross	Committee N	Nai-Chang Yeh	Committee N
Advisor Name	Title										
John Preskill	Committee C										
Chetan Nayak	Committee N										
Michael Cross	Committee N										
Nai-Chang Yeh	Committee N										
Keywords	<ul style="list-style-type: none">phase transitionsHubbard modelstrongly correlated electronscharge-density wavesuperconductivityelectronic conductance										
Date of Defense	2002-05-22										
Availability	unrestricted										
Abstract	We consider several examples of metallic systems that exhibit non-Fermi liquid behavior. In the first model, we study the effect of a magnetic field on the electronic properties of a Fermi liquid. We find that the magnetic field can induce a transition to a non-Fermi liquid state. In the second model, we study the effect of a temperature on the electronic properties of a Fermi liquid. We find that the temperature can induce a transition to a non-Fermi liquid state. In the third model, we study the effect of a pressure on the electronic properties of a Fermi liquid. We find that the pressure can induce a transition to a non-Fermi liquid state. In the fourth model, we study the effect of a magnetic field on the electronic properties of a Fermi liquid. We find that the magnetic field can induce a transition to a non-Fermi liquid state. In the fifth model, we study the effect of a temperature on the electronic properties of a Fermi liquid. We find that the temperature can induce a transition to a non-Fermi liquid state. In the sixth model, we study the effect of a pressure on the electronic properties of a Fermi liquid. We find that the pressure can induce a transition to a non-Fermi liquid state.										

after ...

dc:title:	Aspects of non-Fermi-liquid metals
dc:creator:	Pivovarov, Eugene
dc:subject:	Physics
dc:description:	We consider several examples of metallic systems that exhibit non-Fermi liquid behavior. In the first model, we study the effect of a magnetic field on the electronic properties of a Fermi liquid. We find that the magnetic field can induce a transition to a non-Fermi liquid state. In the second model, we study the effect of a temperature on the electronic properties of a Fermi liquid. We find that the temperature can induce a transition to a non-Fermi liquid state. In the third model, we study the effect of a pressure on the electronic properties of a Fermi liquid. We find that the pressure can induce a transition to a non-Fermi liquid state. In the fourth model, we study the effect of a magnetic field on the electronic properties of a Fermi liquid. We find that the magnetic field can induce a transition to a non-Fermi liquid state. In the fifth model, we study the effect of a temperature on the electronic properties of a Fermi liquid. We find that the temperature can induce a transition to a non-Fermi liquid state. In the sixth model, we study the effect of a pressure on the electronic properties of a Fermi liquid. We find that the pressure can induce a transition to a non-Fermi liquid state.
dc:publisher:	California Institute of Technology
dc:contributor:	Nai-Chang Yeh
dc:contributor:	Chetan Nayak
dc:contributor:	John Preskill
dc:contributor:	Michael Cross
dc:date:	2002-06-03
dc:type:	Text
dc:format:	application/pdf
dc:identifier:	http://resolver.caltech.edu/CaltechETD:etd-05302002-130637
dc:source:	http://etd.caltech.edu/etd/available/etd-05302002-130637/
dc:language:	en
dc:rights:	unrestricted

missing keywords

AUTHOR mapped to DESCRIPTION

before ...

Title: Investigation of technical considerations in a proposed "Live from Lincoln Center" pay television service
Authors: Ward, John Erwin.
Issue Date: 1976
Publisher: Electronic Systems Laboratory, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology
Series/Report no.: Report (Massachusetts Institute of Technology. Electronic Systems Laboratory) ; ESL-R-650.
Description: M.I.T. Project DSR 82300. Research supported by Lincoln Center for the Performing Arts.
URI: <http://hdl.handle.net/1721.1/1259>
Appears in Collections: LTDS Technical Reports

after ...

DC: Elements: 14

dc:title	Investigation of technical considerations in a proposed "Live from Lincoln Center" pay televis
dc:subject	Subscription television
dc:subject	TK7855.M41 E386 no.650
dc:description	John E. Ward.
dc:description	M.I.T. Project DSR 82300. Research supported by Lincoln Center for the Performing Arts.
dc:publisher	Electronic Systems Laboratory, Dept. of Electrical Engineering and Computer Science, Mass
dc:contributor	Ward, John Erwin.
dc:date	2002-06-21T03:39:57Z
dc:date	1976
dc:format	v, 49, [31] leaves
dc:format	application/pdf
dc:identifier	http://hdl.handle.net/1721.1/1259
dc:language	eng
dc:relation	Report (Massachusetts Institute of Technology. Electronic Systems Laboratory) ; ESL-R-650.

no CREATOR

3357 records !

Theorems about Composition

Charpentier, Michel and Chandy, K. Mani (2000) Theorems about Composition.

Full text available as:
[Postscript](#) - Requires a viewer, such as GhostView
[Other \(Adobe PDF \(232KB\)\)](#)

Abstract

Compositional designs require component specifications. On the other hand, components specifications should be abstract enough to allow component reuse and to hide substantial parts of correctness proofs in components verifications. Part of the problem is that too abstract specifications do not contain enough information to be composed. Therefore, the right balance between abstraction and composability must be found. This paper explores the systematic construction of abstract specifications that can be composed through specific forms of composition called existential and universal.

Department: Computer Science
Institution: California Institute of Technology
Month: January
Year: 2000
Unique Identifier: CaltechCSTR:2000.cs-tr-00-02
PURL: <http://resolver.caltech.edu/CaltechCSTR:2000.cs-tr-00-02>

before ...

```
<!DOCTYPE html PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html>
  <head>
    <title>Caltech Computer Science Technical Reports - Theorems about Composition</title>
    <link rel="stylesheet" type="text/css" href="http://caltechcstr.library.caltech.edu/eprints.css" title="screen_stylesheet" media="screen" />
    <link rel="Top" href="http://caltechcstr.library.caltech.edu/" />
    <link rel="search" href="http://caltechcstr.library.caltech.edu/perl/search" />
    <link rel="Help" href="http://caltechcstr.library.caltech.edu/help" />
    <link rel="schema.DC" href="http://purl.org/DC/elements/1.0/" />
    <meta content="Theorems about Composition" name="DC.title" />
    <meta content="Charpentier, Michel" name="DC.creator" />
    <meta content="Chandy, K. Mani" name="DC.creator" />
    <meta content="All Records" name="DC.subject" />
    <meta content="Compositional designs require component specifications that can be composed: Designers have to be able to deduce system properties from components specifications. On the other hand, components specifications should be abstract enough to allow component reuse and to hide substantial parts of correctness proofs in components verifications. Part of the problem is that too abstract specifications do not contain enough information to be composed. Therefore, the right balance between abstraction and composability must be found. This paper explores the systematic construction of abstract specifications that can be composed through specific forms of composition called existential and universal." name="DC.description" />
    <meta content="California Institute of Technology" name="DC.publisher" />
    <meta content="2000-01-01" name="DC.date" />
    <meta content="Monograph" name="DC.type" />
    <meta content="NonPeerReviewed" name="DC.type" />
    <meta content="http://resolver.caltech.edu/caltechcstr/2000.cs-tr-00-02" name="DC.identifier" />
    <meta content="application/postscript" name="DC.format" />
    <meta content="http://caltechcstr.library.caltech.edu/173/00/postscript.ps" name="DC.relation" />
    <meta content="application/octet-stream" name="DC.format" />
    <meta content="http://caltechcstr.library.caltech.edu/173/01/postscript.pdf" name="DC.relation" />
    <meta content="http://resolver.caltech.edu/caltechcstr/2000.cs-tr-00-02" name="DC.relation" />
  </head>
  <body bgcolor="#ffffff" text="#000000" topmargin="0" rightmargin="0" leftmargin="0" marginwidth="0" marginheight="0">
```

after ...

DC Elements: 10

dc:title:	Theorems about Composition
dc:creator	Charpentier, Michel
dc:creator	Chandy, K. Mani
dc:subject	All Records
dc:description	Compositional designs require component specifications that can be composed: Designers have to be able to deduce system properties from components specifications. On the other hand, components specifications should be abstract enough to allow component reuse and to hide substantial parts of correctness proofs in components verifications. Part of the problem is that too abstract specifications do not contain enough information to be composed. Therefore, the right balance between abstraction and composability must be found. This paper explores the systematic construction of abstract specifications that can be composed through specific forms of composition called existential and universal.
dc:date	2000-01-01
dc:type	Main
dc:format	ps http://caltechcstr.library.caltech.edu/archive/00000173/00/postscript.ps
dc:format	other http://caltechcstr.library.caltech.edu/archive/00000173/01/postscript.pdf
dc:identifier	http://resolver.caltech.edu/CaltechCSTR:2000.cs-tr-00-02

Incorrect values

If re-generate records based on the embedded metadata, all can be corrected.

Inappropriate mapping

before ...

Id	Not Available
Title	Environmental Testing of MEMS for Space Applications
Author(s)	K. Man, J. Newell
E-mail(s)	N/A, N/A
Department	Jet Propulsion Laboratory, Jet Propulsion Laboratory
Organization	NASA, NASA
Description	Agenda: Environmental Verification Objectives; Test Sequencing;
Keywords	Thermal, Vacuum Pressure, Quasi-Static Accelerations, Sinusoidal → missed?
Classification	Aerospace Aeronautical Engineering / Structural Mechanics and Dynamics
Date of Creation	1998-8-4
Copyright Information	This resource is freely available for educational purposes.
Format	Portable Document Format
Size	854 KB
Requirements	Application
Interactivity Type	MIXED: Both
Resource Type	Slide Presentation
Interactivity Level	MEDIUM:
Intended Audience	Teachers, Researchers
Context of Use	Professional
Learning Duration	1 HOUR
Resources	

after ...

DDC Elements: 13	
dc:title:	Environmental Testing of MEMS for Space Applications
dc:creator	K. Man
dc:creator	J. Newell
dc:subject	Aerospace Aeronautical Engineering
dc:subject	Structural Mechanics and Dynamics
dc:subject	Aircraft Structure and Materials
dc:description	Agenda: Environmental Verification Objectives; Test Sequencing; En
dc:publisher	Digital Library Network for Engineering and Technology
dc:date	1998-08-04
dc:format	PDF
dc:identifier	http://www.dlnet.vt.edu/repository/previewRepository/AE000000/AE000000000000000000000000000000
dc:language	en-US
dc:rights	This resource has NO copyright restrictions; the contributor(s) agree

CLASSIFICATION mapped to **SUBJECT** and missed all the **KEYWORDS**.

Missed Data Values

- ◆ When a group of elements map to a single one, there could be missed data values.

RightsManagement

Role
Description
Contact
Conditions

Copyright Holder
No copyright notice. Copyright 2001 Rainer Dick presumed.
arXiv.org e-Print archive. <http://xxx.arxiv.cornell.edu/>



dc:rights



Missed Data Values which may cause in-accessible IDs

- WHY do some data sets have 100% inaccessible IDs in a repository after a few years?
 - The IDENTIFIER only provided the information of the record, such as a local access number ID.
 - When data were migrated or moved around in these collections' servers, these IDs became obsolete.

From a LOM record:

1- GENERAL

o 1.1 Identifier

-- Catalog : NEEDS

-- Entry : 49FEC5AD-7B27-4540-B0F7-EF20AB5126D9

-- Catalog : POND

-- Entry : lui236ADF369D304FCEB5BDA8A572F31B13

o 1.2 Title

The Mini Board Technical Reference

3- META-METADATA

o 3.1 Identifier

-- Catalog : NEEDS

-- Entry : 49FEC5AD-7B27-4540-B0F7-EF20AB5126D9

-- Catalog : x-ims-plirid-v0

-- Entry : urn:x-ims-plirid-v0:licefteluq.quebec.ca:sys:lorix:4

4- TECHNICAL

o 4.1 Format

o 4.2 Size (bytes)

o 4.3 Location

<http://lcs.www.media.mit.edu/people/fredm/papers/mb/>

o 4.4 Requirements

identifiers

- When crosswalking, there should be multiple elements mapped to dc:identifier element

Common properties in crosswalks

- ◆ definition of each metadata element
- ◆ mandatory, optional, or mandatory based on certain conditions
- ◆ may occur *multiple* times
- ◆ constraints due to the structure
 - ◆ e.g., hierarchical parent-child relationships
- ◆ constraints imposed on the *value*
 - ◆ e.g., free text, numeric range, date, or a controlled vocabulary
- ◆ *locally defined* metadata element (s)

- Based on St. Pierre and Jr. William. 1998. *Issues in Crosswalking Content Metadata Standards*. Bethesda, MD: NISO Press.

Understanding the Pieces of the Puzzle (4)

4. Sources of Metadata

1) Manual creation

- ◆ Professional-created
- ◆ Authors-contributed
- ◆ User-tagged

2) Automatic metadata generation

- ◆ extraction
- ◆ harvesting

Sources of Metadata -- some tips

- ◆ Understanding DCMI conceptual model -- 'metadata statements' as the basic unit (rather than 'records')
- ◆ Using templates for better consistency and efficiency
- ◆ Utilizing automatically captured metadata
 - e.g., Adobe products, Microsoft Office documents
- ◆ Utilizing author-generated metadata
 - E.g. embedded metadata found at webpage's source codes, Word documents, Photoshop images
- ◆ Utilizing publisher-provided metadata

Document Metadata for instruction.pdf



Description
Camera Data 1
Camera Data 2
Categories
History
IPTC Contact
IPTC Content
IPTC Image
IPTC Status
Adobe Stock Photos
Advanced

Description

Document Title: Instruction in Metadata: Meeting the needs of different audiences

Author: Marica Zeng

Description: This talk shares experience of instruction in metadata at basic and advanced levels. Through different learning modules, exercises, and projects, the instruction helps participants to understand the concepts and structure of metadata, to explore

Description Writer: Marcia Zeng

Keywords: metadata decisions, metadata quality, functional requirements of metadata, metadata interoperability, maximizing metadata

Commas can be used to separate keywords

Copyright Status: Unknown

Copyright Notice: Prepared by Marcia Zeng for 2007 OCLC Western Digital Forum, San Diego, August 9-10. Use for educational and research purposes only.
Available at: <http://www.oclc.org/western/digitalforum/presentations/zeng.pdf>

Copyright Info URL: [Go To URL...](#)

Created: 8/11/2007
Modified: 8/11/2007
Application: PScript5.dll Version 5.2.2
Format: application/pdf

Description metadata partially captured automatically and partially entered manually are recorded for an Adobe Acrobat PDF file

Document Metadata for instruction.pdf



Description
Camera Data 1
Camera Data 2
Categories
History
IPTC Contact
IPTC Content
IPTC Image
IPTC Status
Adobe Stock Photos
Advanced

Advanced

- + PDF Properties (pdf, http://ns.adobe.com/pdf/1.3/)
- + Adobe Photoshop Properties (photoshop, http://ns.adobe.com/photoshop/1.0/)
- + TIFF Properties (tiff, http://ns.adobe.com/tiff/1.0/)
- + XMP Core Properties (xmp, http://ns.adobe.com/xap/1.0/)
- + XMP Media Management Properties (xmpMM, http://ns.adobe.com/xap/1.0/mm/)
- XMP Rights Management Properties (xmpRights, http://ns.adobe.com/xap/1.0/rights/)
 - Copyright (alt container)
 - [x-default]: Prepared by Marcia Zeng for 2007 OCLC Western Digital Forum, San Diego, August 9-10. Use for
- Dublin Core Properties (dc, http://purl.org/dc/elements/1.1/)
 - format: application/pdf
 - title (alt container)
 - [x-default]: Instruction in Metadata: Meeting the needs of different audiences
 - description (alt container)
 - [x-default]: This talk shares experience of instruction in metadata at basic and advanced levels. Through differ
 - creator (seq container)
 - [1]: Marcia Zeng
 - subject (bag container)
 - [1]: metadata decisions
 - [2]: metadata quality
 - [3]: functional requirements of metadata
 - [4]: metadata interoperability
 - [5]: maximizing metadata
 - rights (alt container)
 - [x-default]: Prepared by Marcia Zeng for 2007 OCLC Western Digital Forum, San Diego, August 9-10. Use for

Powered By
xmp

Replace All...

Append All...

Save All...

Delete

OK

Reset

Metadata statements showing an integrated record

Create Shareable Metadata

Avoid semantic conflicts

-- Always indicating which scheme that your values come from

Example for html encoding:

```
<meta name="dc.element"
      scheme="dcterms.scheme" content="value" />
```

Examples:

```
<meta name="subject" scheme="dcterms.MeSH"
      content="pressure" />
```

```
<meta name="subject" scheme="INSPEC Thesaurus"
      content="pressure" />
```

See also *Tutorial 2. Page 22 RDF example under <dcterms:subject>*

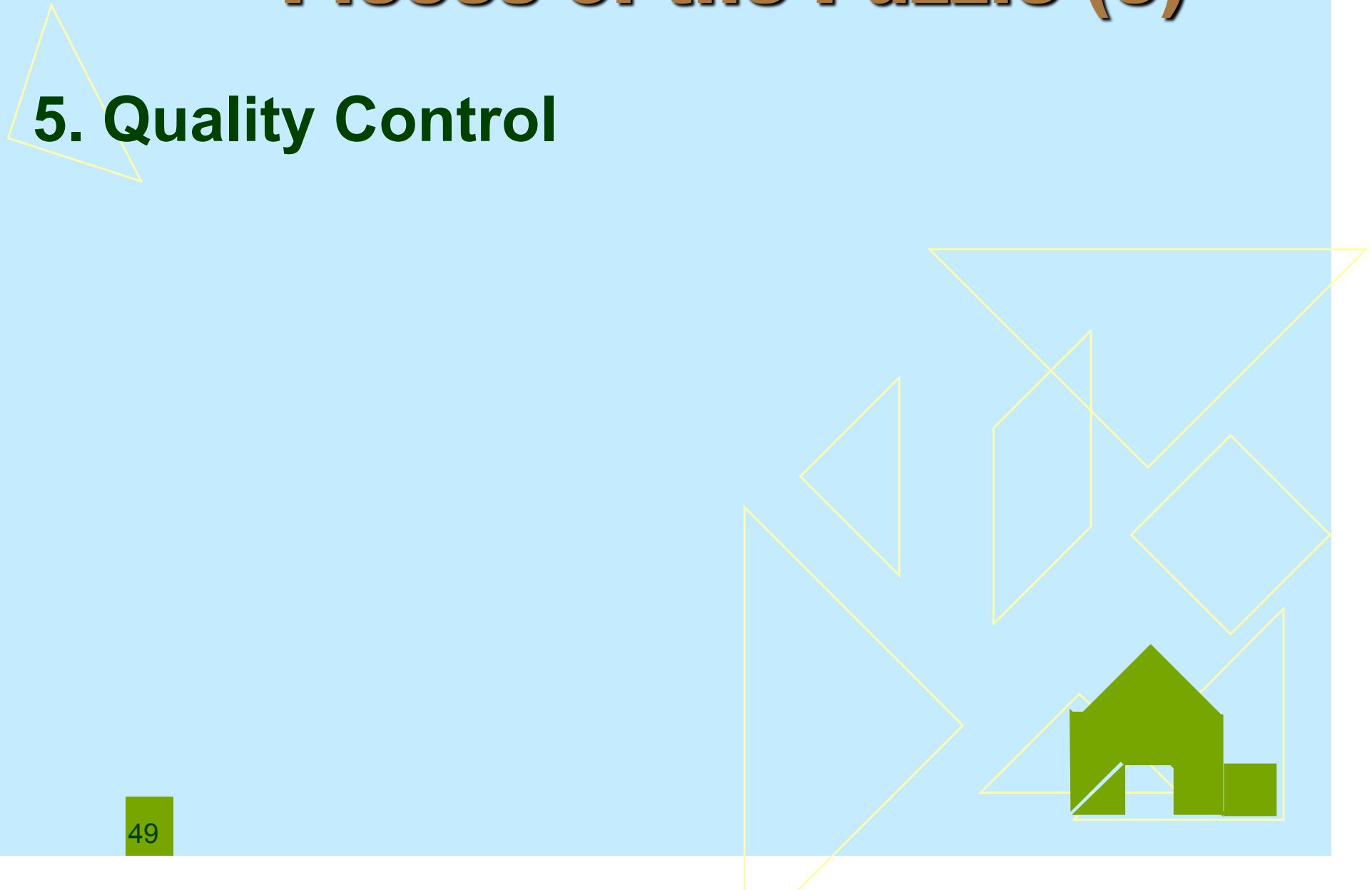
Reuse Metadata

Combining and recombining (the twin aspects of reuse) metadata descriptions

- 1) Creating basic metadata descriptions can be a combined effort of machine and human processing
- 2) Existing metadata descriptions can be reused for any appropriate projects
- 3) Quality of metadata can be enhanced through recombinant metadata
- 4) Integrated records can be generated for better access and sharing.

Understanding the Pieces of the Puzzle (5)

5. Quality Control



Quality Measurement Aspects

- ◆ Completeness
- ◆ Correctness
 - content-info
 - format
 - input
 - browser interpretation
 - mapping/integration
 - redundancy

- ◆ Consistency
- ◆ Duplication
 - intra-collection
 - inter-collection

Can any quality measurement be done automatically?

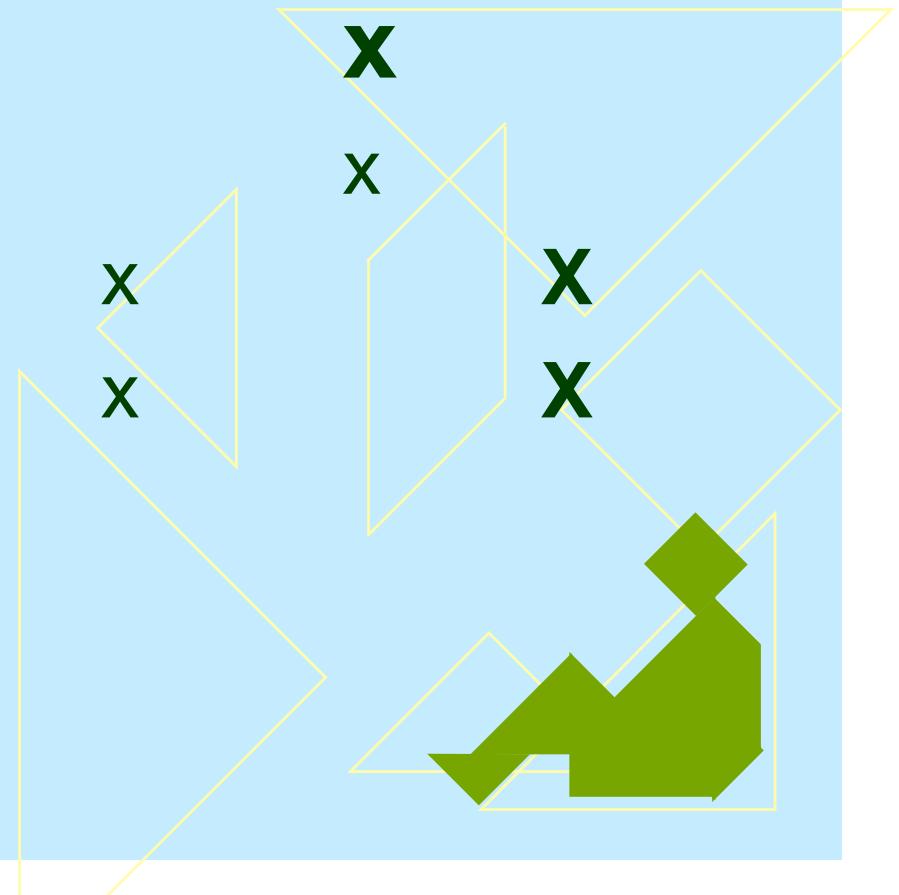


Manual

Machine-
assisted

- ◆ (a) completeness
- ◆ (b) correctness
- ◆ (c) consistency
- ◆ (d) duplication

X



How Can We Ensure a Better Quality?

◆ Make Policies on:

- minimum quality requirements
- quality measurement instruments
- quality enforcement policies
- quality enhancement actions
- the training of metadata creators

◆ Training!

- A 2-hour training may eliminate hundreds of errors

◆ IT team should talk with content team

- A test of crosswalk for OAI harvest may prevent thousands of mis-matched or missed values

◆ Use Tools:

- Provide instructions on best practices
- Use template for inputting records, with suggested syntax, vocabularies, and build-in values
- Use validators
- Implement duplicate-checking algorithm

Understanding the Pieces of the Puzzle (6)

6. Workflows

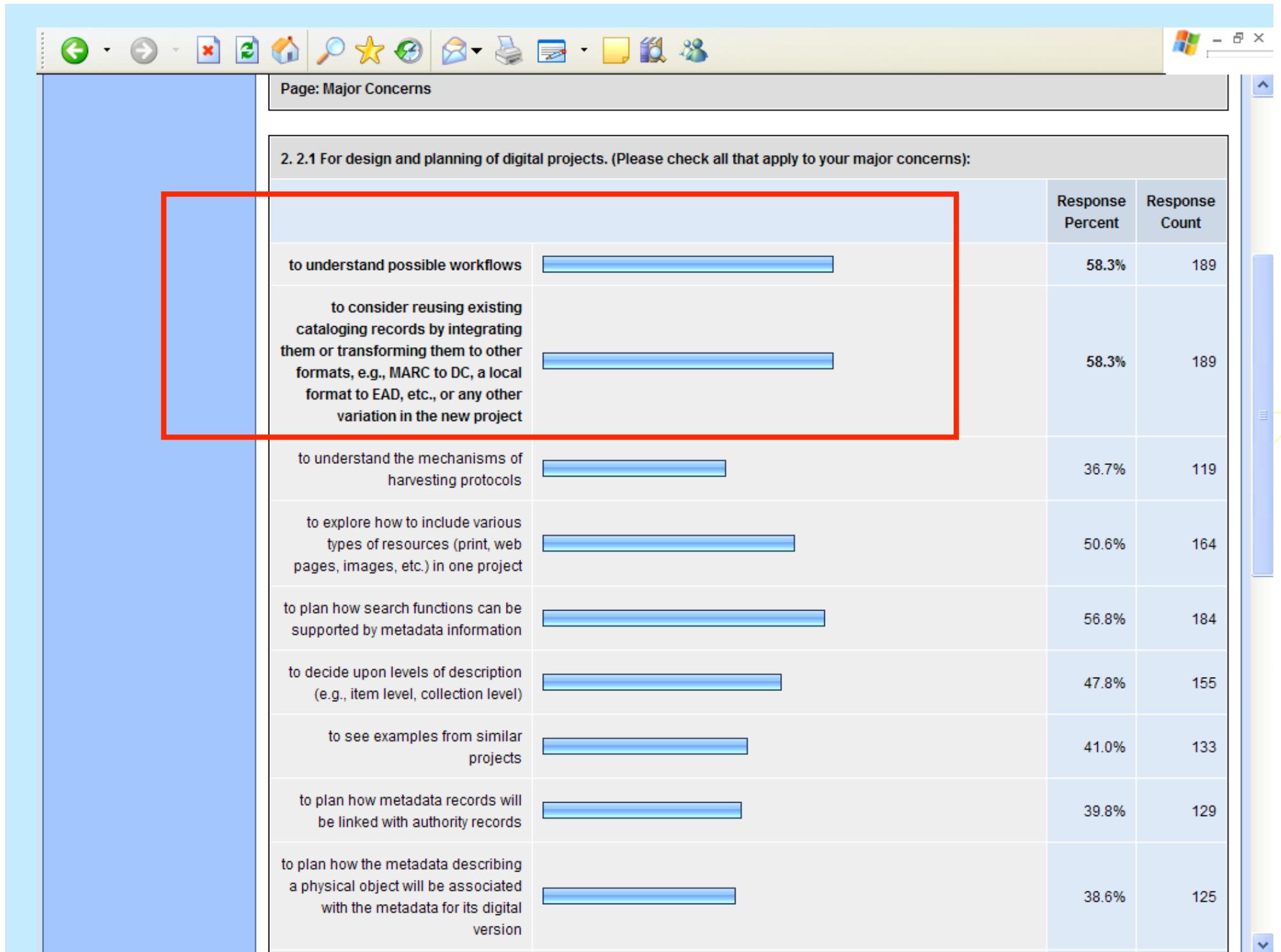
Metadata Decisions

Major concerns in DL projects that relate to metadata include making decisions about:

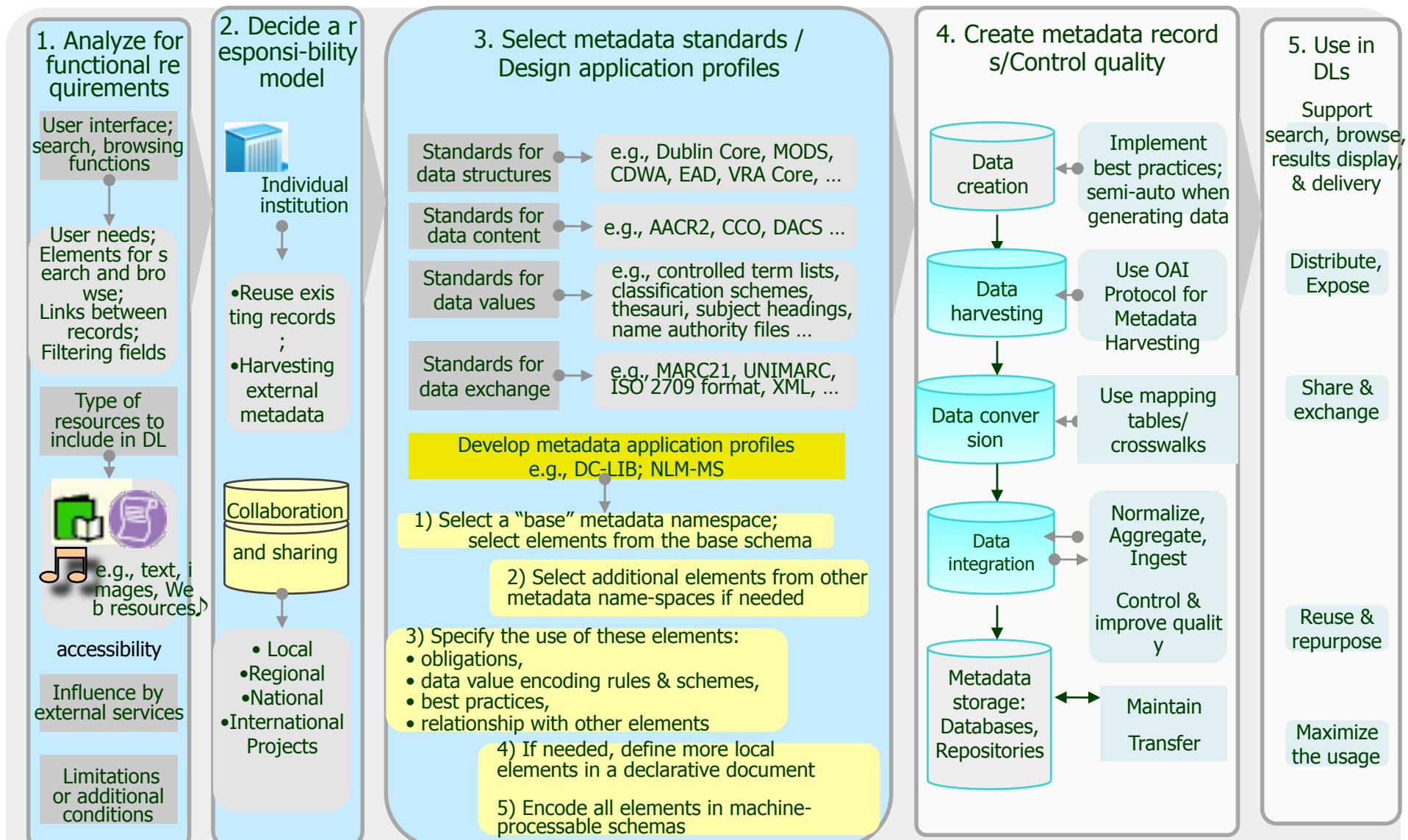
- design and planning of digital projects
- element set standards (data structure decision)
- data contents in a record (data content decision)
- authority files and controlled vocabularies (data value decision)
- metadata encoding (data format/technical interchange decision)

A survey on Metadata Decisions for Digital Libraries

– 426 responses from all over the world, (Zeng, Lee, Hayes, 2008)



Metadata Workflows



Created by M Zeng, J Lee, and A Hayes. ©IFLA GDL-WG, 2008

Recommendations for quality metadata

- ◆ Use standards or create application profiles based on existing standards to optimize the interoperability of metadata records.
- ◆ Employ controlled vocabularies and authority files for data values.
- ◆ Create quality, shareable metadata records.
- ◆ Reuse existing metadata whenever possible.
- ◆ Consider collaboration and partnership for metadata exchange.
- ◆ Consider preservation of objects in the collection.

Understanding the Pieces of the Puzzle (7)

7. Reference sources

- ◆ NISO Framework Advisory Group. 2007. *A Framework of Guidance for Building Good Digital Collections.* 3rd ed.
<http://www.niso.org/publications/rp/framework3.pdf>
- ◆ *Using Dublin Core.* Hillmann, Diane. 2005.
<http://dublincore.org/documents/2005/11/07/usageguide/>
- ◆ *Best Practices for Shareable Metadata.* NSDL. 2005-
<http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/ShareableMetadataPublic>
- ◆ Lists of standards and best practices can be found at
<http://www.metadataetc.org/book-website/index.html>