Introduction to Information Organization

Overview Series #3

Information Representation and Metadata

Objectives of this lecture

- Define terms related to information representation
- Explain bibliographic control in terms of representation and metadata
- Explore concept and power of metadata
- Link lecture’s concepts and terms to IOP Section 2

Model of information retrieval

User's info need

Represented as query

Information retrieval system

matches representations

Subset of metadata representations

Representation

- Many types and levels of representation:
  - Representation of users’ information needs
  - Representations of information objects
  - Representations of relationships among objects
  - Representations manipulated by systems
- Ultimately we create and interact with multiple representations to connect users to information
  - Do users understand these representations?
  - Are some representations better than others?
- Representation is foundation of bibliographic control

Bibliographic control

- We attempt to exert bibliographic control over the bibliographic universe
  - Describing or representing information objects as metadata (name, title, subject, etc.)
  - Entering data in metadata records that serve as surrogates for information objects
  - Putting metadata records into information retrieval systems where they can point to actual information objects

Bibliographic universe

The Bibliographic Universe

- books
- journals
- videos
- sound recordings
- photos
- websites
- computer files
- maps
- globes
Two primary challenges

- How to individually represent information objects as concrete entities so they can be found when needed?
  - Container-oriented
  - Uniqueness-oriented
- How to represent and show relationships among objects as sources of information on various subjects so they can be found by those in search of the information?
  - Content-oriented
  - Relationship-oriented

Representation based on object

Information Object

represented by
- Attributes related to its Creator
- Attributes related to its Name
- Attributes related to its Content

Capturing relationships among items

Training Your New Pet
By William Jones
Dogs
Cats
Pets
Training

Kids and Pets: A guide to training your child
By Judy Blakely
Dogs
Cats
Pets
Training

101 Tips to better pet ownership
By Judy Blakely
Dogs
Cats
Birds
Pets

Representation responsive to users

- What do users want to do?
  - Search?
  - Evaluate?
- How do our representations support their tasks?
- Library catalog representations support four generic user tasks
  - Find
  - Identify
  - Select
  - Acquire or obtain

Meeting in the middle

Illustrations
Visual Appeal
Features
Title
Subject
Author
Length
Complexity
Publisher
ISBN
Edition

The FRBR user tasks

- The metadata record needs to support user’s tasks:
  - Find: Discovering if something exists by searching one or more attributes
  - Identify: Examine retrieved records to determine the items that met user’s search request
  - Select: Examine retrieved records for those that meet other user needs/requirements
  - Obtain: Using data in retrieved records to gain physical access to the described object
The goal....

- To have enough attributes to be sure to get the user to the shelf
  - How do they want to search the database?
    - Subject? Title? Author?
  - How do they want to pick from the screen?
    - Subject? Length? Features?
  - How can we help them find it on the shelf
    - Call number?

Library paradigm for representation

Representation answers two questions

- What is the item?
  - Represent the container
  - Bibliographic representation
  - Descriptive, not evaluative
  - Uniquely identify/describe an object
  - Distinguish one object from another through metadata

- What is the item about?
  - Represent the content
  - Subject representation
  - Identify topics, concepts, perspectives, etc.
  - Adequately represent object's content so user can find the object based on search topic

Other questions... How good is it? How will it be useful to me?

Is and About

It "is" a book
It is "about" sailboats

Representing information objects

- Identify attributes of objects
- Choose those that are important to users
  - Name of object?
  - Creator?
  - Subject?
  - Date of creation?
  - Location?
- Think how users will use this information (users' tasks)
  - Find
  - Identify
  - Select
  - Acquire or obtain

What to represent: entity level

- Entity: discrete and unique representation of a physical object or concept
- Entity level: the object chunk represented
  - Discrete work or text that is represented in a metadata record
- Entity level = unit of analysis = unit of representation
  - One entity is represented by one metadata record
- Examples of entity level
  - Book: whole book, section, or chapter
  - Conference proceedings: whole or separate papers
  - Journal: Whole title, specific issue, or articles
  - Web resources: web site, web page
  - Sound recording: complete CD, individual song

Entity level and metadata record

- Choice of entity level
  - What chunk of information are users interested in?
  - What attributes are available for that chunk?
  - Is the chunk related to other chunks?
- Record holds data about the entity
- What is in your database? A whole book as each record?
Information Representation and Metadata

**Metadata**
- Data about an object
- Entity is the sum of that data
- Structured description of an information object
- Assertions made about an information object
- Metadata scheme consists of metadata elements
  - Library catalog metadata scheme contains metadata elements for title, author, subject, publisher, etc.
- Metadata element represents an attribute of the object

**Purpose and functions of metadata**
- **Purpose:** Provide information about an entity
- **Functions**
  - Discovering or finding that an object exists
  - Evaluating or selecting an object
  - Locating or accessing an object
- **Application areas**
  - Libraries
  - Museums
  - Websites
  - Government agencies
  - Your kitchen

**The chaos of information: No structure**

**The power of metadata: Structure**
Author: Taylor, Arlene G., 1941- Title: The organization of information
Edition: 2nd ed.
Descrip: xxvi, 417 p. : 27 cm.
Series: Library and information science text series
Note: Includes bibliographical references (p. 385-405) and index.
ISBN/ISSN: 1563089696 (pbk. : alk. paper)
1563089769 (hbk. : alk. paper)
Subject: Information organization. Metadata.

**Metadata scheme: Dublin Core**
- 15 elements to represent digital resources
- Also used for analog resources
  - Title
  - Creator
  - Subject
  - Description
  - Publisher
  - Contributors
  - Date
  - Type
  - Format
  - Identifier
  - Source
  - Language
  - Relation
  - Coverage
  - Rights

**Metadata elements and semantics**
- Each element needs a name
- Each element needs a definition (semantics)
- The semantics describe the information content associated with an element
  - **DC element: Title**
    - Semantics: A name given to the resource.
    - Comment: Typically, a Title will be a name by which the resource is formally known.
  - **DC element: Creator**
    - Semantics: An entity primarily responsible for making the content of the resource.
    - Comment: Creators can include a person, an organization, or a service.
Information Representation and Metadata

Database records as representations

- **Records**
  - Aggregates of data that are associated with entities (information objects) described in databases
  - Serve as surrogates for information objects
  - Stored in information retrieval (IR) systems; catalogs, indexes, bibliographies, etc.

- **Data in records support the function of the bibliographic tool/instrument**

The connection:

- **Attributes:**
  - Author
  - Title
  - Subject

Become Elements

Become Fields

Order through representation!

- **The bibliographic universe is not self-organizing**
- **Bibliographic control = information organization**
- We use representations, metadata, and bibliographic tools…
  - To bring order
  - To connect users with information…

*Knowledge of the power given by an instrument depends on the clarity of rules according to which it was constructed, and on knowledge of those rules.*

Representing objects for IOP

- **Determine the entity level**
  - Chunk of object or entire object represented
  - One metadata record for each chunk

- **Identify and select attributes of object to represent**
  - Examine object for attributes
  - Consider what users need to know about object
  - Focus first on the container attributes, then on the content attributes

- **Develop metadata scheme**
  - Each attribute will have an associated metadata element(s)
  - Metadata elements will label the selected attributes
  - Develop list of elements and their semantics

Concepts & terms

- Attributes
- Bibliographic control
- Bibliographic relationships
- Bibliographic representation
- Bibliographic tools
- Bibliographic universe
- Container
- Content
- Controlled vocabulary
- Database
- Entity
- Entity level
- Function of bibliographic tools
- Information needs
- Information objects
- Information retrieval system
- Metadata
- Metadata element
- Metadata scheme
- Name authority control
- Records
- Representation
- Semantics
- Structured representation
- Subject representation
- Unit of analysis
- Users' tasks
- Work and text