

Friday, February 4, 2022

## Class 4: Content Standard Principles

---

### LIBR 509: Foundations of Resource Description and Knowledge Organization

#### - **Intro:**

- Content standards = agreed upon sets of attributes to provide a useful description of resources in a collection. They include attributes like those we've covered so far (like classifications and thesaurus terms) but also information about the resource's physical properties, its identification, and relationships to people and other resources

- What is a content standard?

- What should we describe about a resource?

- What are the minimal agreed-upon attributes to include in a record to make them available to others and search ability?

#### - **Key Qs:**

- Which *aspects* of a resource to represent?
- What are the *constraints* placed on those representations?
- How much *time and expertise* does the person doing the description have?

#### - **Qs we will not cover in this course:**

- Is this bibliographic description / cataloguing or is this metadata?
- Where am I going to write all this down?
- Will there be paper cuts?

#### - **Central Concepts**

- Title
- Creator (the person, group, or event who is “responsible” for the resource)
- Vision

## - Other Concepts

- What the resource looks like
- What type of resource it is (text, sound recording, etc)
- Who published it and where
- How to identify the resource (ISBN, ISSN, DOI)
- What the resource is about (Thesaurus terms, Class number)

## - Important terms

- Content schema / Content standard are used interchangeably in this course. However, they are not exactly interchangeable:
  - **Schemas:** typically become standards once they are regulated, managed and shared between institutions. A schema is an outline for what a record should contain.
  - **Standards:** are typically regulated, validated, and may have more rules

## - Why is it like this?

### - (Authority) Control

- When we're working with everything you might include to describe a record, we ensure it is reliable through control - universal
- Language and consistency challenges - minor differences in language unique, repeatable way
- Bibliographic relationships - items have different relationships to each other - how these relationships be networked better?

### - Abstraction and Specificity

- How important is it to make similar resources appear the same?
- How important is it to distinguish similar resources from each other?

### - Access

- In a catalogue, typically "Access Points"
  - Creating a surrogate (stand in) in a catalogue, what items need to be:

- Searchable?
  - Filterable?
  - What will users need to find, sort, identify, and evaluate a record?
- **IFLA's General Principles**
- Think about these when making catalogue records
  - Exp: Convenience of the user ; Common usage ; Representation ; Accuracy ; Economy ; Integration ' Interoperability
- **Dublin Core (exp of a content schema with minimal rules)**
- Started at a metadata semantics workshop in 1995
  - Sets up a common "core" of metadata elements for web-based resources
  - Main goal is easier search and retrieval mechanisms
  - Split into 15 elements:
    - *General Info*: Title, Date, Creator, Contributor, Publisher, Rights
    - *Physical Description*: Identifier (ISBN, URI), Type (moving image), Format (mpeg), Language
    - *Intellectual Content*: Description, Subject (LCSH, DDC), Coverage (place, time), Relation, Source
      - Needs to be broad enough to be universal but understandable
- **Dublin Core Qualifiers (modifications)**
- More specific version of one of the elements
    - Instead of date: Date Digitized, Date Modified, Date Available
  - If the user's software doesn't understand the qualifier, it reverts back to the main element
- **Metadata & Cataloguing**
- They are both about saying something about a potentially informative resource
  - Function as a *surrogate* of a record

- Cataloguing is a kind of metadata work (and cataloging standards are types of metadata standards)

#### - **Metadata vs Cataloguing**

- Most metadata is created specifically to describe digital content
- Digital content often constantly changes
  - The “Extent” of a website or web page can be difficult to measure
  - Web content often has nothing resembling “editions”
- Digital content is harder to control
  - Metadata is specifically required to provide and constrain access
- Metadata is used to control and document

#### - **Choosing / Creating a Content Standard**

##### • **What you need to consider:**

- Balance between functionality and simplicity
  - How do you know if you have too many elements or too few?
- Support human and machine use
  - What is easily understood to a human is not always understood by computers
- Support interoperability and extensibility
  - *Interoperability*: That metadata from one institution should be readable by another (standardization). That metadata from one schema can be readable in another (enabled by translations, crosswalks)
  - *Extensibility*: That the metadata schema is adaptable for local needs. That an extended schema can be simplified for global needs (

#### - **Crosswalk between Dublin Core and EAD**

- Only go one way
- Dublin Core (dc.title) => EAD (Header) (titleproper)

#### - **Summary of lecture:**

- Each schema / standard has:

- A set of values
- Some instructions on which elements are necessary
- Some instructions on how to modify elements
- Some instructions on how to fill out the values
  - Possibly a reference to controlled vocabs for particular fields

~

## - **Creating your own content schema (instructions)**

### • **Choose a set of resources**

- Who is searching / browsing through them - how many are there - what are the resources might someone want to search / filter by, or would need to know to identify the “right” item?
- What: Art Collection
- Who: Museum workers

### • **Create a List of Attributes**

- Bibliographic attributes usually include title / creator / version
- Consider static and dynamic, intrinsic and extrinsic properties => Read p.262, *DoO*

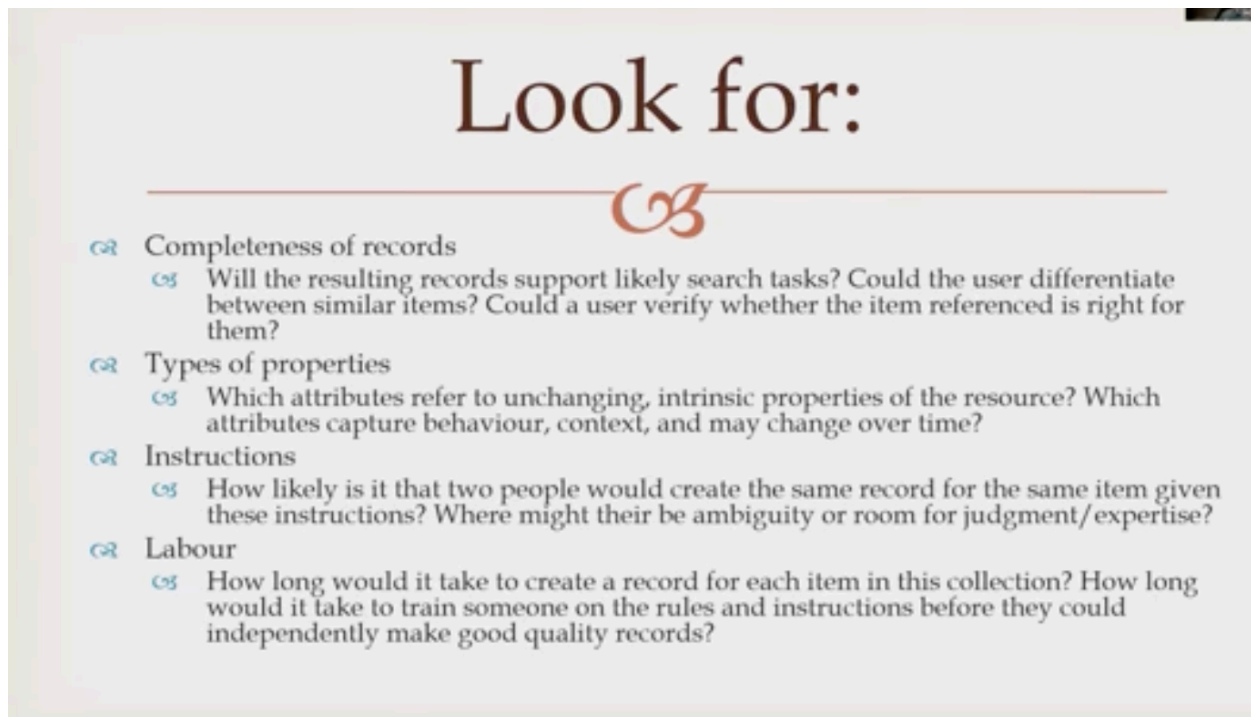
### • **Scope**

- At least 6 attributes
- Which ones are required, optional, repeatable (exp. Creator field)
- Instructions for anything that is not obvious or that might be interpreted differently across individuals, across time, or with peculiar items

### - **Types of Entry**

- Which attributes should be limited to values from a controlled list? Possibly connections to thesauri and classifications
  - If there are any values that should take from a particular type of thesauri or classification
- Which attributes should have values as free-text entry?

- Which attributes will need particular types of formatting for ordering / sorting?
- **Description**
  - You can make a couple sample records using the instruction you have applied
- **Peer Review Guidelines**
  - Should be about 2 hrs
  - To give feedback on:



# Look for:

---

☞ Completeness of records

- ☞ Will the resulting records support likely search tasks? Could the user differentiate between similar items? Could a user verify whether the item referenced is right for them?

☞ Types of properties

- ☞ Which attributes refer to unchanging, intrinsic properties of the resource? Which attributes capture behaviour, context, and may change over time?

☞ Instructions

- ☞ How likely is it that two people would create the same record for the same item given these instructions? Where might there be ambiguity or room for judgment/expertise?

☞ Labour

- ☞ How long would it take to create a record for each item in this collection? How long would it take to train someone on the rules and instructions before they could independently make good quality records?