Introduction to Metadata Analysis Using the DLF Metadata Framework

Kate Flynn, Chicago Collections, University of Illinois, Chicago
Hannah E. V. Stitzlein, University of Illinois,
Urbana-Champaign

The Basics

- Baseline
- Glossary: http://bit.ly/2glzY6W
- Types of metadata
 - Descriptive metadata
 - Technical metadata
 - Preservation metadata
- What people might be working with when assessing
 - Databases
 - Character delimited value formats
 - o XML
 - JSON
 - Spreadsheets (XLS, XSLX)

The Basics

Metadata Assessment Framework

http://dlfmetadataassessment.github.io/Framework

- A high-level framework for assessing descriptive metadata in digital collections
- Many thanks to:
 - Sara Rubinow
 - Maggie Dickson
 - Kathryn Gronsbell
 - Anna Neatrour
 - Bethany Scott

Migrations or Data Sharing

- Migrations
 - Before or after (or both!)
- Data sharing
- Aggregation (exposing data)
 - Shareability
 - Interoperability

Validation

- Standards
- Application profile or guidelines
- Schema
- Other pre-defined structures

Targeted Enhancement

- What change would have the most impact?
- Consider resources available

Profile Generation

 Lets users (internal & external) know what is expected for metadata

Standards & Best Practices

- Staff changeovers may mean internal knowledge loss
- Systems & contexts can change over time

Handling Emerging Object
Types

Digital humanity use

Impact of Metadata Work

- Usability by users
- Staff changeovers and internal knowledge
- External access and aggregation

Documenting assessment

Why?

- Advocacy
- Demystification
- Sustainability
- Compliance

How / Potential outputs

- Best Practices / Guidelines
- Clean-up plans
- Application profile
- Reports
- Other

Building a scope + Parameters

- Use practices and policies for efficiency & effectiveness
 - Guidelines
 - Data dictionaries
 - Application profiles

Building a scope + Parameters

- Prioritizing assessment
 - Baseline compliance
 - Vocabulary compliance
 - Repository migration
 - Usability

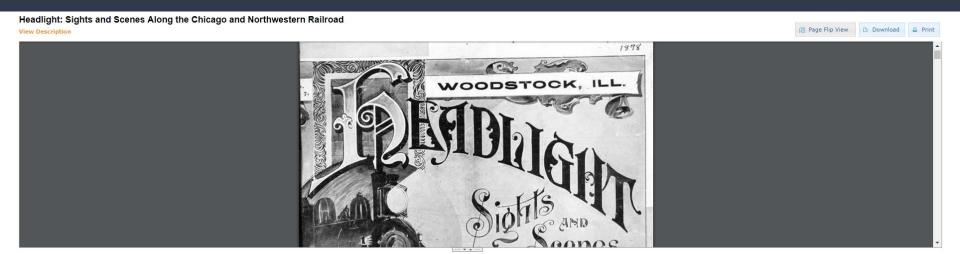
Building a scope + Parameters

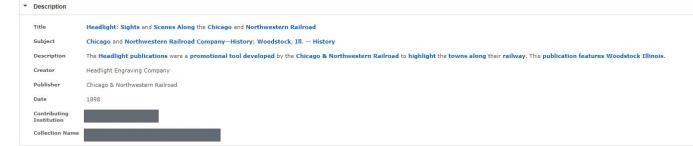
- Keep it simple
 - Use an approach that allows you to meet your objectives

Completeness

The element/property/attribute is present

Completeness





Missing rights, type fields

Consistency

Values are consistent within your domain. Semantic and structural values and elements are represented in a consistent manner across records

Example

- Controlled vocabularies
- + Consistent field use

Consistency



Accuracy

Information is correct and factual both semantically and syntactically

Accuracy



Example

Project Location: Beverly (Chicago, Ill.)

Project Location: Loop (Chicago, III.)

Accessibility

Metadata can be read by a human and/or a machine

Accessibility

Type Photograph. Horizontal; Image

Type Image

| Туре | ** |
|-----------------|--------|
| image | 210521 |
| text | 94711 |
| sound | 1388 |
| moving image | 747 |
| physical object | 244 |

Conformance to Expectations

Values adhere to the expectations of your user communities (both internal and external)

Example

- + Consortium requirements
- + Aggregator requirements
- + Best practices

Conformance to Expectations



Mountain West Digital Library Dublin Core Application Profile

Version 2.0 July 20, 2011

I. Best Practices for All Fields

Use of semicolons

Data in one field may be combined with data from another field during the harvesting/aggregating process. To retain integrity of data, put a semicolon after the last word in each field. Likewise, separate multiple entries within a single field by inserting a semicolon and a space between each two entries.

MARC mapping

MARC tags are provided in each table to clarify the type of data contained in a field as well as to facilitate mapping data from Dublin Core records to MARC records.

Requirement status

- Required: Field must be included in record. There are eight required fields:
 - o date
 - description

1

MWDL Dublin Core Application Profile

Version 2.0 (July 20, 2011)

- o format
- o <u>identifier</u>
- o rights
- o subject
- o title
- o type

Timeliness

When the resource changes, the metadata is updated accordingly. When additional metadata becomes available or when metadata standards change, the metadata associated with the resource also changes.

Timeliness

TITLE: Evanston's Central Business District

Before Metadata
Guidelines update

CREATOR: Perkins & Will; Brubaker, C. William

Title Evanston's Central Business District

Creator Brubaker, C. William

Contributor Perkins & Will

After Metadata Guidelines update

Provenance

You can track metadata transformations back to the original form. You have information about the source of the original metadata.

Provenance

Pittsburgh



Get full image from University of Michigan. Libraries

| Creator | Etienne Turpin |
|--------------------------|---|
| Created Date | 2013 |
| Partner | Michigan Service Hub |
| Contributing Institution | University of Michigan. Libraries |
| Location | Ann Arbor United States |
| Format | Image |
| Туре | image |
| Subject | Labor |
| Language | Undetermined |
| Rights | These digitized materials are made available for education and research purposes. Where possible, we indicate what we know about copyright and rights of privacy, publicity, or trademark in the record for each item. Due to the nature of library and archival collections, we do not always know if there is a copyright holder or how to find them in any given case. We are eager to hear from any rights owners to learn more about our collections and obtain accurate information (see more at http://quod.lib.umich.edu/l/lbc2ic?page=rights-permissions). Http://quod.lib.umich.edu/t/text/accesspolicy.html |

Resources

DLF Assessment Interest Group Metadata Working Group Toolkit

http://dlfmetadataassessment.github.io

Metadata Application Profile Clearinghouse https://dlfmetadataassessment.github.io/Metadata
SpecsClearinghouse/

Glossary of Terms (draft)
https://docs.google.com/document/d/1ejPUZs0Qq
z_HHx7gWPIUEuP8Qdg0EXZRop7bEdjVgq4/edit?us
p=sharing

Metadata Assessment Framework http://dlfmetadataassessment.github.io/Framework http://dlfmetadataassessment.github.io/Framework