Metadata and Digital Libraries

Marty Kurth
UAEU Libraries
October 4-8, 2009

Participant Manual
<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mon 10/5</td>
<td>10-11:30</td>
<td>Session 1: Introduction to digital library system objectives, functionality, and metadata</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 2: Building digital collections</td>
</tr>
<tr>
<td></td>
<td>11:30-1</td>
<td>Session 3: More about metadata and introduction to Dublin Core</td>
</tr>
<tr>
<td></td>
<td>2-3:30</td>
<td>Session 4: Understanding functional requirements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 5: Metadata and functionality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 6: Four characteristics of metadata practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 7: Metadata conversion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 8: Metadata workflows</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 9: Digital library development project (exercise)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 10: Business planning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 11: Digital preservation planning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 12: Metadata and ensuring access over time (collections)</td>
</tr>
<tr>
<td>Wed 10/7</td>
<td></td>
<td>Session 13: Metadata and ensuring access over time (objects)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 14: Project management--team building and work planning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 15: Project management--proposal writing and assessment</td>
</tr>
<tr>
<td>Thurs 10/8</td>
<td></td>
<td>Daily debriefing</td>
</tr>
</tbody>
</table>
Metadata and Digital Libraries

Marty Kurth
UAEU Libraries October 4-8, 2009

Workshop approach

This workshop is weighted more toward group exercises than presentation of information, with time for conversation regarding UAEU Libraries’ digital library plans.
Disclaimer

Much of what I will present is illustrative rather than definitive. Nothing can substitute for thoughtful inquiry guided by your own circumstances and experience.

Session 1: Introduction to digital library system objectives, functionality, and metadata

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying content for this session)
Session goals

- Understand the relationship between system objectives and metadata
- Examine the objectives of the library bibliographic system and how those objectives impact system metadata
- Explore the connection between digital library systems and digital library metadata
- Underscore the importance of system objectives when working with metadata

The library catalog

- Why do we describe library materials in the way we do?
  - Why do we catalog in the way that we do?
  - Why do we assemble certain information (metadata) about library materials, and record this metadata in such a highly defined way?
Cutter (1876)

Objectives of a bibliographic system
- To enable a user to find a book if the author, title, or subject is known
- To show what the library has by a given author, on a given subject, or of a given kind
- To assist in the choice of a book based on its edition (bibliographically) or its character (literary or topical)

IFLA (1998)

- To find entities that correspond to the user’s stated search criteria
- To identify an entity
- To select an entity that is appropriate to the user’s needs
- To acquire or obtain access to an entity described
To locate
- Known entity
- Set of entities

To identify an entity

To select an appropriate entity

To acquire or obtain access to an entity

To navigate a bibliographic database

Exercise 1a: Library bibliographic system metadata
- How does the MARC metadata support the objectives of the library system? (For example, to find, identify, select, obtain)
- What other system objectives can we detect from the system’s metadata?
The library bibliographic system

- System objectives have led to specific practices in bibliographic description
  - Standards such as AACR2
- Uniform record creation is required by global bibliographic databases
  - Standard record formats such as MARC21
- Desired functionality requires precise cataloging rules and conventions

Exercise 1b: Digital library system metadata
- XML encoded metadata used by some type of digital information system
- What system objectives can we detect by examining this system’s metadata?
Digital library systems

- No agreed upon definition or objectives
- No agreed upon standards or formats
- Very little interoperability
- A huge number of players, many of whom are not librarians
- What is a “digital library,” anyway?
  - Digital (electronic) information systems?

A different world from the library bibliographic system, but not an alternate universe

Digital library system development...
- Still requires the articulation of objectives (desired system functionality)
- And those objectives will rely upon certain characteristics of available or generated metadata
Digital library system objectives

- To support...
  - Discovery
  - Navigation
  - Presentation, display
  - Access control
  - Administration, management
  - Preservation
  - Others?

System objectives?

- Who decides on the objectives of the digital library system?
- Who decides what functionality to support?
- Who are the players or stakeholders on digital library projects?
Digital library projects

- Digital library stakeholders:
  - Project sponsor
  - Project director
  - Project manager
  - Subject specialist
  - System developer/programmer
  - Metadata specialist
  - Library administrator/manager
  - End-users
  - Others?

Digital library system objectives?

- How do the stakeholders decide on system objectives?
- How is system functionality developed?
- What are some processes by which decisions are reached?
Session 2: Building digital collections

Reflection

“Digital collections must now intersect with the user’s own context—within the course, within the research process, within the leisure time activities, and within the social networks that are important to the end user.”—A Framework of Guidance for Building Good Digital Collections, 3rd ed., 2007
A good digital collection . . .

- Is built following a collection policy
- Is described so a user can discover its characteristics
- Contains actively managed resources
- Is broadly available
- Respects intellectual property rights
- Supplies use data
- Is interoperable
- Integrates into the user’s workflow
- Is sustainable

A good digital object . . .

- Is in a format that supports its use
- Is preservable
- Is meaningful and useful outside local context
- Is named with a persistent, unique, resolvable identifier
- Can be authenticated
- Has associated metadata
Good metadata . . .

- Conforms to community standards
- Supports interoperability
- Uses authority control and content standards
- Includes terms of use
- Supports long-term curation and preservation
- Has the qualities of a good digital object

A good digital initiative . . .

- Has substantial design and planning
- Has staff with expertise
- Follows project best practices
- Has an evaluation component
- Markets itself and shares process and outcomes
- Considers the digital life cycle
Digital collection selection criteria

- Legal rights and restrictions
- Increased or transformed access
- Content
  - Virtual collection building, scholarship driven, local utility
- Preservation

Reflection

“There are only local collections, built with local funding, in support of local needs.”—Ross Atkinson
Exercise 2: Selection for digitization

For further study:

And just out:


---

Session 3: More about metadata and introduction to Dublin Core

(Many thanks to Diane Hillmann for sharing her content for this session)
Reflection (point)

“Metadata consists of statements we make about resources to help us find, identify, use, manage, evaluate, and preserve them.”—Me?

Reflection (counterpoint)

“Metadata is what we know and data is what we're looking for.”—David Weinberger
Some typical metadata functions

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

Metadata building blocks (in words)

1. The basic unit of metadata is a statement.
2. A statement consists of a property (aka, element) and a value.
3. Metadata statements describe resources.
Metadata building blocks (in pictures)

(An oversimplification of the DCMI abstract model for resources)

What are the *properties* and *values* in these metadata statements?

245  00  $a  Amores perros $h  [videorecording]

<title>Nueve reinas</title>
<Type>MovingImage</Type>
Who cares about metadata?

The term “metadata” has meaning in contexts such as:

- Data modeling
- Library cataloging
- Internet/World Wide Web resource discovery
  - Led to a convergence between the first two
  - Formed the context in which Dublin Core arose

Introduction to the Dublin Core
How and why did the Dublin Core come to be in 1995?

- Dramatic increase in the number of document-like resources on the net
- Slow improvement in indexing services made resources hard to discover
- Belief that descriptive metadata would improve discovery
- Perceived need for a descriptive standard that was simple to apply (by non-professionals)

Dublin Core Metadata Element Set

Creator  Title  Subject
Contributor  Date  Description
Publisher  Type  Format
Coverage  Rights  Relation
Source  Language  Identifier
Characteristics of the Dublin Core

- A flat element structure, with:
  - All elements optional
  - All elements repeatable
- Elements displayed in any order
- Extensible (elements, qualifiers)
- Syntax independent
- International
- Subject independent

Resources for which DC is often used

<table>
<thead>
<tr>
<th>DCMI Type Vocabulary</th>
<th>Collection</th>
<th>Dataset</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image</td>
<td>Interactive Resource</td>
<td>Moving Image</td>
<td></td>
</tr>
<tr>
<td>Physical Object</td>
<td>Service</td>
<td>Software</td>
<td></td>
</tr>
<tr>
<td>Sound</td>
<td>Still Image</td>
<td>Text</td>
<td></td>
</tr>
</tbody>
</table>
Dublin Core principles

- Dumb-down
- The one-to-one principle
- Appropriate values

Dumb-down

- Simple DC does not use *element refinements* or *encoding schemes* and statements only contain *value strings*
- Qualified DC uses features of the DCMI Abstract Model, particularly *element refinements* and *encoding schemes*
- *Dumbing down* is translating qualified DC to simple DC (*property dumb-down* and *value dumb-down*)
- For more info, see the DCMI Abstract Model
Element refinements

- Element refinements narrow the meaning of DC elements
  - `hasVersion` and `isVersionof` refine `relation`
  - `bibliographicCitation` refines `identifier`

- Element refinements are *properties*, so we typically render them independently
  - `<dcterms:alternative>Nine queens</dcterms:alternative>`

Encoding schemes

- **Vocabulary encoding schemes**
  - Indicate that a value comes from a controlled vocabulary (e.g., that “Spanish American literature” is an LCSH term)

- **Syntax encoding schemes**
  - Indicate that a string is formatted in a standard way
    (e.g., that “1956-11-12” follows ISO 8601)

- DCMI recommends using encoding schemes with *coverage, date, format, language, subject,* and *type*
The one-to-one principle

- Create one metadata *description* for one and only one resource
  - E.g., do not describe a digital image of the Mona Lisa as if it were the original painting
- Group related *descriptions* into *description sets*
  - I.e., describe an artist and his/her work separately, not in a single description

Appropriate values

- Use elements and qualifiers to meet the needs of your local context, but . . .
- Remember that your metadata may be interpreted by machines and people, so . . .
- Consider whether the values you use will aid discovery outside your local context and . . .
- Make decisions about your local practices accordingly
Metadata creation and distribution models

- **Federation**
  - Extensive specifications, standards, protocols, training

- **Harvesting**
  - Basic agreements, reliance on best practices

- **Gathering**
  - Automated indexing of content, algorithms yield results from search terms, less likely to use descriptive metadata per se

---

**Harvesting model** key features

- Integrating metadata from many sources calls for common element sets, record structures, and harvesting protocols
- Open Archives Initiative Protocol for Metadata Harvesting serves as a framework for sharing metadata and mandates ‘simple DC’ as a common metadata format
- Harvesting promotes metadata reuse
- Best practices balance cost and interoperability
- Communities add value to basic infrastructure (more complex metadata, new uses for protocol)
Reflection

“Good metadata should be coherent, meaningful, and useful in global contexts beyond those in which it was created.”—A Framework of Guidance for Building Good Digital Collections, 3rd ed., 2007

Exercise 3: Creating Dublin Core metadata for digital objects
Day 1 debriefing

Session 4: Understanding functional requirements

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying content for this session)
Goals of session

- Understand functional requirements and their usefulness
- Recognize how functional requirements inform system metadata decisions
- Understand “use cases” and how they define and record functional requirements
- Learn how a use case should be “read” by a metadata specialist

Functional requirements, pt. 1

- What are functional requirements?
  - In this context, functional requirements are those of an information system, not of bibliographic records (FRBR)
  - A more specific and detailed description of system objectives
  - They describe and define specific, required system behaviors
  - Ideally, they are developed through a requirements analysis process
  - They guide system implementation and programming work
Functional requirements, pt. 2

- How do project stakeholders develop functional requirements?
- Ideally, system designers use some reasonably formal design process
- Examples of design processes:
  - Rational Unified Process (RUP)
  - User centered design
  - Agile software development

Software design processes

- Systematic methods for generating and defining functional requirements
- Different design processes emphasize different methodologies, but there are often many similarities among them
- Most processes employ “use cases,” though they may exploit different methods to generate and develop them
Use cases

- Each use case describes a single function of the system
- Each function is an interaction between the system and an external USER
- Each use case describes functionality, but not how that functionality will be accomplished
- The entire system may have dozens or hundreds of use cases
- Taken altogether, the use cases define the system’s functional requirements

The USER in a use case

- USERs are anything external to the system that will interact with it
- A USER may represent a class of users
  - Data entry staff
  - System administrators
  - General public users
- A USER may represent another system
  - An OAI harvester
Sample use case

- Exercise 4: Sample use case
- Typical use case components:
  - Priority
  - Preconditions
  - Flow of Events (scenario)
  - Alternative Events (exceptions)
- What in this use case will depend on or impact system metadata?

Generating use cases

- The design process used will likely guide how use cases are generated
- A typical approach is to enumerate all the possible USERs of the system (everyone and everything that will interact with it), and then list every interaction
- Each of these interactions will become a use case
A complete set of use cases

- Together, they define the functional requirements of the proposed system
- Documented, they form a contract among stakeholders about what the system will do and not do
- Requirements help in the inevitable “panic phase” of a project
- Requirements inform our decisions about metadata, standards, software, vendors...

Build or buy?

- Build or buy decisions are typical in digital library development projects
  - Building a digital library system
    - Defining one's own functional requirements
    - Hiring programmers to build the system
    - Testing, evaluation, maintenance, updates
  - Acquiring a pre-built digital library system
    - Finding a system with functionality that meets your requirements as nearly as possible
Build or buy

- Both cases require articulating and documenting desired objectives and functionality
- If *build*, these will develop into complete use cases
- If *buy*, they can be used in an RFP process, and later to evaluate competing systems

Requirements and metadata

- Certain functional requirements will depend upon or impact system metadata
  - What data elements are required
  - What content value practices need to be adopted
  - Whether metadata standards can or should be used
- If we have existing metadata, requirements will inform our analysis and conversion of it
Exercise 4: Sample use case

Session 5: Metadata and functionality

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying content for this session)
Session goals

- Review or familiarize ourselves with concepts and vocabulary of metadata assessment and analysis
- Explore the connection between metadata and functionality

Metadata specialist scenario

- The typical digital library development situation facing the metadata specialist:
  - We have some functional requirements to meet, AND we have some metadata
  - BUT the metadata must be altered in some way (cleaned-up, augmented, enhanced, mapped...) so that it will meet our requirements
Metadata and functionality

- In order to match metadata with functionality, we need first to assess, or analyze, our existing metadata
- Then we can begin to evaluate whether our metadata will or will not support particular functionality and how it will need to be converted

Metadata assessment

- If we look at existing metadata, how do we describe what we observe?
  - File format
  - Type of metadata
  - Semantics
  - Content values
  - Structure
  - Use
  - Status
Metadata analysis: File format

- File, or data exchange, formats:
  - SGML / HTML
  - XML / XHTML
  - MARC
  - “Delimited” plain-text file
  - Binary (not plain-text) formats, either open or proprietary
Metadata analysis: Type

- Types of metadata
  - Descriptive
  - Structural
  - Administrative
  - Technical
  - Preservation
  - Access/rights
Metadata analysis: Semantics

- Metadata element sets ("schemes")
  - MARC21
  - Dublin Core (DC)
  - EAD
  - MODS
  - VRA Core
  - METS

```xml
<info>
  <title>Arithmetic</title>
</info>

<name type="personal">
  <namePart>Rand, Fed</namePart>
</name>

<name type="personal">
  <namePart>Sandburg, Carl</namePart>
</name>

<dates>
  <date-type>20th-Century</date-type>
  <date-start>1878-1967</date-start>
</dates>

<dates>
  <date-type>20th-Century</date-type>
  <date-start>1912-1967</date-start>
</dates>

<place>
  <place-abbr>San Diego</place-abbr>
</place>

<publisher>Harcourt Brace Jovanovich</publisher>

<date-issued>1991</date-issued>

<date-created>1991</date-created>

<edition>1st ed.</edition>

<isn>monographic</isn>

<issn>1076-0656</issn>

<language>en</language>

<abstract>
  A poem about numbers and their characteristics. Features anamorphic, or distorted,
  drawings which can be restored to normal by viewing from a particular angle or by viewing
  the image's reflection in the provided Mylar cone.
</abstract>

<target-audience>
  <category>adults</category>
  <category>students</category>
</target-audience>

<statement-of-responsibility>C. A. H. J. Sandburg; illustrated as an anamorphic
adventure by Ted Rand.</statement-of-responsibility>
```
<html xmlns="http://www.w3.org/1999/xhtml" version="1.0" encoding="UTF-8">
	<title>Arithmetic</title>
	<meta name="DC:creator" content="Sandburg, Carl, 1878-1967."/>
	<meta name="DC:subject" content="Children's poetry, American."/>
	<meta name="DC:subject" content="Arithmetic"/>
	<meta name="DC:subject" content="American poetry"/>
	<meta name="DC:subject" content="Visual perception"/>
	<description>A poem about numbers and their characteristics. Features anamorphic, or distorted, drawings which can be restored to normal by viewing from a particular angle or by viewing the image's reflection in the provided Mylar cone.
</description>
</head>

<script type="text/javascript">/* <![CDATA[
    function preload_images(args) {
        myImages = new Array();
        var i = 1;
        args = args;
        while (i < args.length) {
            myImages[i] = new Image();
            myImages[i].src = args[0];
            i++;
        }
    } /* ]]></script>
Metadata analysis: Content

- Does the metadata...
  - Adhere to any published content standards or best practices?
    - AACR2/RDA, EAD Best Practice (RLG), CCO
  - Use any known and shared vocabularies?
    - LCSH, AAT, TGN
  - Adhere to any application profiles?

- Degree of conformance to any employed standards, practices, or vocabularies?

Metadata analysis: Structure

- Structure
  - What is the record structure?
    - Flat or hierarchical (nesting)
  - What relationships are possible? How complex can they be?
  - Is element qualification allowed?
  - Degree of ambiguity within data?

- General character and complexity
  - Simple unstructured
  - Simple structured
  - Richly structured
Metadata analysis: Use

- What is, or was, the intended or potential use of this metadata?
  - Understanding why metadata was created and how it was used can help tell you what you can expect from it, in terms of consistency, reliability, interoperability...

Metadata analysis: Status

- Static vs. dynamic
  - Static metadata will not be updated, augmented, etc.—it is essentially “dead”
  - Dynamic metadata is “living,” maintained by someone, updated when needed, perhaps regularly supplemented

- This distinction will have an impact on conversion strategies and workflows
## A typology of data standards, pt. 1

<table>
<thead>
<tr>
<th>Type of Data Standard</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data structure standards (metadata element sets, schemas). These are &quot;categories&quot; or &quot;containers&quot; of data that make up a record or other information object.</td>
<td>the set of MARC (Machine-Readable Cataloging format) fields, Encoded Archival Description (EAD), Dublin Core Metadata Element Set (DCMES), Categories for the Description of Works of Art (CDWA), VRA Core Categories</td>
</tr>
</tbody>
</table>


## A typology of data standards, pt. 2

<table>
<thead>
<tr>
<th>Type of Data Standard</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data content standards (cataloging rules and codes). These are guidelines for the format and syntax of the data values that are used to populate metadata elements.</td>
<td>Anglo-American Cataloguing Rules (AACR), Resource Description and Access (RDA), International Standard Bibliographic Description (ISBD), Cataloging Cultural Objects (CCO), Describing Archives: A Content Standard (DACS)</td>
</tr>
</tbody>
</table>

### A typology of data standards, pt. 3

<table>
<thead>
<tr>
<th>Type of Data Standard</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data format/technical interchange standards (metadata standards expressed in machine-readable form). This type of standard is often a manifestation of a particular data structure standard (type 1 above), encoded or marked up for machine processing.</td>
<td>MARC21, MARCXML, EAD XML DTD, METS, MODS, CDWA Lite XML schema, Simple Dublin Core XML schema, Qualified Dublin Core XML schema, VRA Core 4.0 XML schema</td>
</tr>
</tbody>
</table>


---

### Exercise 5: Metadata analysis
Session 6: Four characteristics of metadata practice


Purpose of the session

- To identify what metadata practitioners contribute to facilitating the use of networked information
- To help relate metadata practice to cataloging practice
To cover

- Typical approaches to metadata work
- Metadata practitioners' responsibilities
- Primary role that practitioners perform
- Central contribution of practitioners to scholarly communication and collaboration

Background context

Because what I will say has been shaped by my experience of working with metadata for digital resources in libraries, I will briefly give the historical context of that experience.
From "CUL Goals & Objectives 2002-2007"

II.3. Establish and operate a "consulting to production" metadata service capable of producing metadata in a variety of formats to organize, manage, and preserve collections over time and to enable effective discovery and use.

Cornell Library Technical Services formed Metadata Services by reallocation in 2002
The DCAPS service model

DCAPS: Digital Consulting & Production Services

The Metadata Services mission

Metadata Services provides metadata consulting, design, development, production, and conversion services to Cornell's faculty, staff, and community partners to increase the value of their digital resources.
Metadata defined for clients

Metadata organizes information about digital resources, including titles, authors, keywords, format, versions, and rights. It increases the value of digital resources by making them easier to access, use, share, and re-purpose.

Given this context . . .

We will discuss metadata practice in terms of the metadata that practitioners design, develop, produce, and convert to other formats in order to manage digital resources and make them accessible to end users.
1. Metadata practice approaches metadata in aggregates

Practitioners work in the context of

- Projects
- Collections
- Services
- Communities of practice
NISO Framework of Guidance for Building Good Digital Collections

"Objects, metadata, and collections must now be viewed not only within the context of the projects that created them but as building blocks that others can reuse, repackage, and build services upon."

Metadata work scenarios first consider

- "Project" goals, requirements, user needs
- Scholarly communities the effort will serve
- Other initiatives serving those communities
- Interoperability mechanisms that may apply
Which means that . . .

Practitioners seek to understand the big picture before they design the parameters for the structure and content of individual metadata records

This approach differs from most cataloging, where . . .

- a cataloger considers the item in hand first
- content and encoding standards are clear
- creation tools and delivery mechanisms are pre-determined
- documentation is widely shared
- there is an established community that shares theory and practice
The two approaches contrasted

Content object
Unit record

cataloging workflows

Collection
Catalog
Delivery system
Online service

metadata workflows

Content object
Unit record

Collection
Catalog
Delivery system
Online service

Why is this important?

Cataloging skills still apply in a metadata environment because cataloging and metadata workflows are mirror images of one another.
2. Metadata practice comprises interpersonal, informational, and operational layers

Social aspects of metadata work

- Practitioners serve on teams that include "scholars, information professionals, and technologists" (Greenstein and Thorin)
- Metadata design and development are highly consultative
- Interactions involve advocacy, negotiation, and committing resources
Because of these social aspects, metadata practitioners perform multiple roles in project teams . . .

. . . and metadata practitioners' roles are similar to managerial roles: **interpersonal, informational, decisional**

---

**The metadata practitioner's organizational responsibilities**

(from Mintzberg by way of Choo)

- **Interpersonal Roles**
  - Liaison
  - Proxy
  - Negotiator

- **Informational Roles**
  - Investigator
  - Disseminator
  - Spokesperson
  - Strategist

- **Operational Roles**
  - Element set designer
  - Profile developer
  - Vocabulary creator
  - Resource describer
Is it not true that everyone on a team fulfills these roles?

What is unique about metadata work?

3. Metadata practice specializes in cross-community translation
How "communities of practice" function

- Members consult community history to make meaning
- Group knowledge yields competent members
- Group requires that members share information
- Group bestows identities and status on members

In libraries we know how this works

Policies, procedures, standards, rules, codes, reference sources, meetings, workshops, associations, conferences, awards, and . . .

. . . acronyms!
But . . .

- Shared culture and language make it hard to discuss group work outside the group
- Group’s conceptual framework and language create communication boundaries
- Communicating outside the group requires recoding
- Communicating across boundaries requires learning the language of the target group

Boundary spanning  (Tushman and Scanlan)

- Gather information from one side and translate it to match the culture and language of the other
- Develop formal and informal information sources inside and outside the community
- Use internal and external sources to support translation work
Sound familiar?

- Reconcile searches that subject experts want with system limitations and interoperability requirements
- Investigate controlled vocabularies that relate to natural language terms
- Develop element-set profiles and local vocabularies that meet user needs and delivery system constraints
- Map and transform local metadata for harvesting and reuse

In other words, metadata practice ...

- relays messages among communities to build systems that support community work
- actively engages with the languages of collaborators' communities
- helps communities make meaning
- regularizes community terminology (intra-)
- map community terminology to other communities (inter-)
4. Metadata practice’s semantic and syntactic translations support interoperability

Metadata is modular (in creation and use)

- ISBN, AACR, LCCN, LCSH
- Descriptive, technical, preservation, rights
Metadata work's central operations
("It is all translation")

- Mapping – establishing relationships between equivalent elements in different schemes
- Transformation – designing and implementing scripts/tools to move mapped metadata between schemes
- as in translating the language of a resource into ISBD, MARC 21, and subject vocabs

Mapping and transformation operate on

- Semantics – the meaningful content that metadata conveys
- Syntax – the structure that expresses that content

650  0 $a Veterinary therapeutics $z Tropics $v Congresses.
<term>Veterinary therapeutics--Tropics--Congresses </term>
What do mapping and transformation have to do with metadata aggregates, layers of metadata responsibilities, and cross-community translations?

NISO Framework of Guidance for Building Good Digital Collections

Digital objects, metadata, and collections are building blocks for reuse and integration
We create "boundary objects" to connect the building blocks (Bowker & Star)

We actively engage with user communities to build tools.

The tools we create influence the work that user communities perform.

Or, globally, on the surface of the Web . . .

We map and transform metadata to facilitate multidisciplinary research and instruction.

We create tools that support the semantic and syntactic interoperability of Web resources.
Conclusion: Implications for libraries

- Wholly manual processes do not scale
- Metadata workflows benefit from practitioners with complementary skills
- Opportunities lie in integrating automated and manual operations

Exercise 6: Metadata analysis scenarios

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying exercise content for this session)
Day 2 debriefing

Session 7: Metadata conversion

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying content for this session)
Session goals

- Explore the reasons for converting metadata
- Discuss measures for assessing and ensuring the quality of metadata
- Examine metadata mapping and its purposes
- Learn how to create a metadata map

Metadata conversion

- Two broad categories or types of metadata conversion work:
  - Enhancement: cleaning up, adding, expanding, disambiguating, updating metadata
  - Mapping: moving metadata from one format to another
Why enhance metadata?

- To correct inaccuracies
- To achieve consistency
- To improve “quality”
- To fill gaps
- To provide greater or different functionality
- To foster interoperability

.Metadata accuracy

<DC_record>
<creator>Mitchell, William J.</creator>
<creator>Stevenson, Daniel C.</creator>
<creator>Schoonover, Regina</creator>
<title>Urbanowski, Frank</title>
<subject>City of Bits: Space, Place, and the Infobahn</subject>
<subject>Electronically mediated environments</subject>
<subject>Cyberspace</subject>
<type>Urbanism</type>
<format>Text</format>
<date>1995</date>
<language>http://press.opt.edu/CityOfBits.html</language>
</DC_record>
**Metadata consistency**

- DC records with a `<dc:date>` element
- Most formatted in full W3C-DTF format (e.g., `<dc:date>YYYY-MM-DD</dc:date>`),
- except for:
  - `<dc:date>2000-08</dc:date>`
  - `<dc:date>1996</dc:date>`
  - `<dc:date>July 5, 2001</dc:date>`
  - `<dc:date>2000 Revision</dc:date>`
  - `<dc:date>July 19, 1996</dc:date>`
  - `<dc:date>2001.06.04</dc:date>`

---

**“Quality” metadata**

“Objects, metadata, and collections must now be viewed not only within the context of the projects that created them but as building blocks that others can reuse, repackage, and build services upon.”

Indicators of metadata quality

- Appropriate to the collection, its users, and the use of the objects in it
- Supports interoperability
- Uses standard controlled vocabularies
- States conditions and terms of use
- Possesses the qualities of good objects
- Supports long-term management

Approaches to interoperability

- Convert to a single metadata scheme
- Allow diverse metadata schemes and map to a common scheme for particular purposes
  - For example: access, or sharing metadata
- Use a hybrid approach that involves some uniformity and some mapping
Tools for interoperability

- Metadata standards
- Application profiles
- Community developed best practices
- Community accepted metadata maps (crosswalks)

Metadata mapping

- A formal, repeatable conversion of metadata
  - A potentially ongoing or regularly repeated conversion process
  - Assumes consistent incoming metadata
- Requires a specification (called a “map” or “crosswalk”) that describes how to convert one metadata scheme format to another
Why map metadata?

- To accommodate a change or upgrade in an existing system
- To “ingest” metadata into another system, but maintain original metadata format
- To share metadata with a wider community, improving interoperability
  - Metadata is diverse—we will never all use the same metadata formats
Metadata mapping caveats

- Requires good knowledge of both source and target metadata formats
- Often not a one-to-one correspondence between elements
- Typically involves some conversion operations
  - Data types and values may differ
  - Structure, hierarchy may differ
  - Element optionality/repeatability may differ

Exercise 7: Metadata mapping

- Creating “shareable” metadata
- Designing a detailed metadata map
- Converting from relatively rich metadata to simple Dublin Core records
Session 8: Metadata workflows

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying content for this session)

Session goals

- Understand the components of workflow design
- Understand the management aspects of metadata workflows (tasks, costs, constraints)
- Examine practical aspects of metadata conversion workflows
- Design a metadata workflow
Workflow fundamentals

- The movement of data through a work process
  - Input $\rightarrow$ Transformations $\rightarrow$ Output
- A work process will typically involve multiple components or individual steps (tasks and subtasks)
  - Each task also has its own data movement:
    - Input $\rightarrow$ Transformations $\rightarrow$ Output

Components of workflow design

- Workflow definition and goals
- Identifying constraints
- Defining the metadata workflow tasks and subtasks
- Designing the workflow
- Maintaining the workflow
- Cost considerations and opportunities
Workflow definition and goals

- Defining the workflow objectives
- Analysis of overall work process input and output
  - Understand the characteristics of the workflow input (e.g., source metadata)
  - Understand the characteristics of the workflow output (e.g., target metadata)
- Specifying the required transformation

Identifying constraints

- Resources
  - Money
  - Staff
- Time
- Environmental constraints
- Knowledge and expertise
Defining the tasks

- Breaking overall goal down into tasks and subtasks, small enough to be implemented
- At that level, determine each task’s...
  - Requirements
    - Specifying task input and output
  - Complexity of transformation (input to output)
  - Dependencies
  - Duration
  - Resource requirements

Designing the workflow

- Given the constraints, how do we put all the pieces of the workflow puzzle together in the most optimal way?
- How should tasks be structured in workflow?
  - Sequencing and scheduling of tasks
- Who or what will perform each task?
- What are the communication needs of the workflow?
Maintaining the workflow

- How will the workflow and its tasks be tracked and evaluated?
  - Who is responsible for the workflow?
  - How will improvements or other changes to the workflow be made?

- Once operational, what are the workflow’s ongoing management requirements?
  - How much human oversight is needed?
  - How much tracking can be automated?

Workflow cost considerations

- Workflow setup
  - What is the current and required level of staff expertise with source and target metadata schemes?
  - What staff skills are required to implement workflow transformations?
  - What can be automated?
    - Are there existing, re-usable tools available?
  - What must be done manually?
    - Any prior experience with this type of processing?
Workflow cost considerations

- Workflow maintenance
  - We need to quantify the type and extent of ongoing support and maintenance the workflow will require
  - Cost differences in maintaining manual vs. automated workflows
  - How much management oversight does the workflow require?

Opportunities and benefits

- Increased knowledge and expertise
- Revenue potential
- Greater use of collections and resources
- Greater visibility of institution
Practical aspects of workflows

- Types of workflows
- Characteristics of source and target metadata, and the impact on workflow design
- When to convert metadata
- How to convert metadata

Types of metadata workflows

- Enhancement and mapping
  Source data → Transformations → Target data
- Other workflows:
  - Augmentation of records
  - Analysis or evaluation
  - Quality control/assurance
Metadata conversion workflows

- Many aspects of the workflow will depend on the characteristics of the source and target metadata
  - Static vs. dynamic source metadata
  - Other source metadata considerations
  - Target metadata

Source metadata

- Static source metadata suggests...
  - A one time transfer of metadata from the creator or supplier
  - The creator or supplier is, or will eventually be, out of the picture

- Dynamic source metadata implies...
  - An ongoing, periodic transfer of the same, updated, or augmented metadata
The impact on workflow of...

- Static source metadata
  - Manual processing is at least feasible
    - No disincentive to apply manual work, except for cost
  - A more extensive and subtle range of data enhancement is possible
  - Workflow may not be directly reusable

- Dynamic source metadata
  - Much depends upon the nature and rate of change of the source metadata
  - There is a disincentive to use manual processing
    - Correcting errors
    - Manual “value-add” features
  - There is an incentive to apply programmable transformations
  - Workflow processes must be re-usable to be cost-effective
Source metadata:
Other considerations

- What or who created or supplied the metadata?
  - Is there a clear and single owner?
  - Multiple suppliers?
- Is the source metadata complete?
- Why was this metadata created?
  - Was it created to meet specific functional needs?

Target metadata

- What purposes are the metadata serving?
- Is this a locally defined element set or larger community standard?
- Is the metadata format supported, and by whom?
  - Is there documentation?
  - Is the format maintained and evolved over time?
When/how to convert metadata

- Will depend on the type of metadata conversion required
- Two broad categories or types of metadata conversion work:
  - Enhancement: cleaning up, adding, expanding, disambiguating, updating metadata
  - Mapping: moving metadata from one format to another

When to convert metadata?

- Once and only once
  - Abandon source metadata in favor of improved set
- Continuously
  - On-the-fly, when requested
  - To feed some downstream processes
- Only when you have to
  - Fixing problems when they are pointed out
How to convert metadata?

- Manually, record-by-record
- In batch, with automated processes
  - Planning, testing, evaluation, more planning...
  - Conversion
  - Final, or ongoing, evaluation
- A hybrid approach, with some manual and some automated processing

Exercise 8: Metadata workflow

Library publishing—designing a workflow for a metadata conversion project
Session 9: Digital library development project exercise

(Many thanks to David Ruddy, Library of Congress, and ALCTS for supplying content for this session)

Exercise 9: Digital library development project— the slide library
Day 3 debriefing

Session 10: Digital initiative business planning

(Many thanks to Oya Rieger and David Ruddy for sharing their content for this session)
Why business planning?

- Present a clear picture of your goals and plan of action to ensure clear communication and to get support
- Address sustainability and long-term viability issues
- Assess your organization’s ability and commitment
- Enable assessment and evaluation of your plan

Components of a business plan

- Business description and objectives
- Service model
- Needs assessment
  - Market analysis and plan
  - Risk assessment and contingency planning
- Management team
- Financial information and forecasts (see next slide)
  - Resources: staffing, facilities, equipment
- Communication plan
  - Assessment and evaluation
Fiscal management

- Cost model elements
  - Per page vs. per image costs
  - Planning, production, management, preservation

- Calculating hourly costs
  - Billable hours: 1200 hrs/year
  - Direct expenses: Salary, benefits, training, equipment, networking, etc.
  - Indirect expenses: Equipment depreciation, service center administration, utilities, etc.

- Development vs. production costs

Digital initiative cost elements

- Selection
- Preparation
- Digitization
- Metadata
- System deployment
- Assessment
- Digital preservation
Needs Assessment and Decision-Making

Object type
Condition
Metadata attributes
Selection criteria
Usage restrictions
Relation to other collections

User requirements
Use type
Frequency of use
Use mode
Support needs

Staff and skills
Systems, hardware, software
Stakeholders
Organizational guidelines
Master plans and strategies

Resource planning strategies
To meet resource needs we are faced with two basic approaches...
Dividing up the pie

More pie
Fixed resources

- Allocating funds differently
- Assessing opportunity costs (trade-offs)
  - What activities must be given up or scaled back to support digital initiatives?
  - How to ensure that the opportunity cost is as low as possible?

Increased resources

- External funding sources (donations, grants)?
- Fees for service or use?
- Collaborative or business relationships across organizational boundaries?
Service model case study

Digital Consulting and Production Services
Cornell University Library

- Support effective and efficient creation of digital collections
- Offer a single-point of entry to services for planning, creation, and management of digital collections
- Implement best practices and standards to create digital collections with long-lasting value
- Develop a business model to sustain services
Business planning tool examples

- Project planning outline
- Budget overview form
- Metadata plan-of-work checklist
- Metadata plan of work
Exercise 10: Business planning questionnaire

Discussion questions

- What are your opportunities and challenges in planning and managing a digital initiative?
- What are your institution’s strengths in developing a digital initiative?
- Which areas need additional capacity and resources?
- What are your training needs?
For further study:

Business Planning for Cultural Heritage Institutions
Liz Bishoff and Nancy Allen
January 2004
http://www.clir.org/pubs/pubs.html

Session 11: Digital preservation planning
Reflection (point)

“Digital preservation combines policies, strategies and actions that ensure access to information in digital formats over time”—ALA (ALCTS PARS) working definition

Reflection (counterpoint)

“There are ‘modest’ ways of managing collections without the challenge and responsibility of following the industry standards that are only affordable by large institutions with digital asset management mandates”—Oya Rieger
Well-managed collections, pt. 1

- Adhere to best practices in creating digital content (use common standards and avoid proprietary formats)
- Document your decisions (how files were created, technical specs, copyright issues)
- Create a registry of your digital collections
- Identify a team and assign responsibilities

Well-managed collections, pt. 2

- Have a unified storage plan (inventory, backups)
- Regularly assess what you are doing and how you can improve
- Use reliable access to support preservation
- Join consortial initiatives to replicate content
Life cycle management

- Life cycle management implies institutional commitment and reinforces the connection between development and preservation.
- Recognizes that every initiative has ongoing phases – they are not one-time efforts.
- Places equal emphasis on planning, development, and maintenance stages.
Digital preservation planning

case study

Digital Consulting and Production Services
Cornell University Library

Digital files of concern for DCAPS
Including the master file for this image

For that image file, metadata is here . . .
... and here...

Images from Cornell's Rare Book and Manuscript Collections

Collection Identifier
Cornell University Library
1995923

Title of collection
Images from Cornell Rare Book and Manuscript Collections

Description of collection
A growing collection of images from the general collections of the Cornell University Library includes images from rare books, as well as the University Archives and other manuscript collections.

Collection Type
Collection/registry

Collection logo URL
http://lib.its.cornell.edu/rarebooks/collection_images.jpg

Subject
Art

Language of collection
English

Contact
Please, Dr. (Contact)
Database, Web (Contact Administrator)

Service providing access
Images from Rare Books and Manuscript Collections

Last modified: 2004-04-03 09:47:20

... and here.
Q: Once we have identified an area of concern, how do we begin to address it?

A: Assess and plan

Reflection

“Backing up digital objects is NOT preservation!”—Carl Grant
Exercise 11: Material and organizational assessment for digital preservation planning

Session 12: Metadata and ensuring access over time (collections)
Preservation metadata in context

A preservation data model (PREMIS)

Intellectual Entities

Objects

Rights

Agents

Events
Another look at Cornell's digitization metadata

<table>
<thead>
<tr>
<th>S#</th>
<th># of</th>
<th>Date</th>
<th>Form</th>
<th>Slides</th>
<th>MDL Slides</th>
<th>file name</th>
<th>info</th>
</tr>
</thead>
<tbody>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5028.tif</td>
<td>Marsfile024772baz65643M2004_5028.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5032.tif</td>
<td>Marsfile024772baz65643M2004_5032.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5033.tif</td>
<td>Marsfile024772baz65643M2004_5033.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5050.tif</td>
<td>Marsfile024772baz65643M2004_5050.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5051.tif</td>
<td>Marsfile024772baz65643M2004_5051.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5054.tif</td>
<td>Marsfile024772baz65643M2004_5054.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5043.tif</td>
<td>Marsfile024772baz65643M2004_5043.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5044.tif</td>
<td>Marsfile024772baz65643M2004_5044.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5065.tif</td>
<td>Marsfile024772baz65643M2004_5065.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_5070.tif</td>
<td>Marsfile024772baz65643M2004_5070.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_7044.tif</td>
<td>Marsfile024772baz65643M2004_7044.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_0101.tif</td>
<td>Marsfile024772baz65643M2004_0101.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_0102.tif</td>
<td>Marsfile024772baz65643M2004_0102.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_0206.tif</td>
<td>Marsfile024772baz65643M2004_0206.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_0207.tif</td>
<td>Marsfile024772baz65643M2004_0207.tif</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>550</td>
<td>03/12/2007</td>
<td>BMIC</td>
<td>RMC</td>
<td>RMC2004_0208.tif</td>
<td>Marsfile024772baz65643M2004_0208.tif</td>
<td></td>
</tr>
</tbody>
</table>

Format-specific technical characteristics


Version 2005-05
May 2005

<table>
<thead>
<tr>
<th>FILE INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
</tr>
<tr>
<td>Version</td>
</tr>
<tr>
<td>Platform</td>
</tr>
<tr>
<td>File Name</td>
</tr>
<tr>
<td>File Size</td>
</tr>
</tbody>
</table>

Proceed to Download
Object-level metadata

<table>
<thead>
<tr>
<th>Title:</th>
<th>Hidden house sketch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date:</td>
<td>1943</td>
</tr>
<tr>
<td>Classification:</td>
<td>Publications</td>
</tr>
<tr>
<td>Rights and Permissions:</td>
<td>Held by Artist (Creator). Publisher or Artist's (Creator's) Representatives</td>
</tr>
<tr>
<td>ID Number:</td>
<td>RMC2005_0260</td>
</tr>
<tr>
<td>Subcollection:</td>
<td>Rare Books</td>
</tr>
<tr>
<td>Collection Title Proper:</td>
<td>Literature: privately printed to please a few friends</td>
</tr>
<tr>
<td>Repository:</td>
<td>Division of Rare and Manuscript Collections, Cornell University Library</td>
</tr>
<tr>
<td>Location:</td>
<td>Rare Books FS3547 09342128 Page 85</td>
</tr>
<tr>
<td>Voyager ID:</td>
<td>2363803</td>
</tr>
</tbody>
</table>

Collection-level metadata

Images from Cornell's Rare Book and Manuscript Collections

Collection Identifier
- Cornell ID: rbcollections

Title of collection
- Images from Cornell's Rare Book and Manuscript Collections

Description of collection
- A growing collection of images from the collections of the Division of Rare and Manuscript Collections, Cornell University Library. Images are drawn from a variety of special collections, as well as the University Archives and other manuscript collections.

Collection type
- Collections image

Collection URL
- http://libweb.library.cornell.edu/connections/failed/ pickle/funcicase/11/49.png

Subject
- Art

Language of collection
- English

Contact
- Library, Rare Book and Manuscript Collections

Service providing access
- Images from Cornell's Rare Book and Manuscript Collections

Last Modified: 2005-03-24 11:54:27 EST
Reflection

“More experience with digital preservation is needed to determine the best ways of representing significant properties in general, and of representing modification of significant properties.”—PREMIS Data Dictionary
Exercise 12: Collection-level metadata in a consortial context

Day 4 debriefing
Session 13: Metadata and ensuring access over time (objects)

A data model relevant to our discussions
Reflection

“Most of the PREMIS elements are designed to be automatically supplied by the preservation repository application. (Of course this does not mean that currently available applications do supply them.)”—Caplan, *Understanding PREMIS*

---

What PREMIS is and is not

- **What PREMIS is:**
  - Common data model for organizing/thinking about preservation metadata
  - Guidance for local implementations
  - Standard for exchanging information packages between repositories

- **What PREMIS is not:**
  - Out-of-the-box solution: need to instantiate as metadata elements in repository system
  - All needed metadata: excludes business rules, format-specific technical metadata, descriptive metadata for access, non-core preservation metadata
  - Life cycle management of objects outside repository
  - Rights management: limited to permissions regarding actions taken within repository
Exercise 13: Object-level metadata in a repository context

Digital preservation citations that may interest you

http://www.citeulike.org/group/4576/tag/digital-preservation
Session 14: Project management—team building and work planning

(Many thanks to Mary Woodley, Library of Congress, and ALCTS for supplying content for this session)

Session goals

- Understand the team building process
- Know how to use the team to plan
- Understand the process of building consensus and working together toward common goals
- Learn the process of developing a work plan
- Identify the components of a work plan
Collaboration and partnerships

Success of projects depends on developing a core team of stakeholders

Stakeholders may be part of the institution, parent institution, or partners in the project

Potential stakeholders or team members

- Digital project director
- Grant writer(s)
- Curators
- Project manager
- Specialist in metadata creation
- Specialist in scanning standards
- Conversion specialist
- Hardware / software developer or procurer
- Web page / interface developer
- Marketing and promoter of project
- Staff responsible for implementation tasks
- Assessment specialist
Staffing

Every project will vary

Digital Gutenberg

Project: team of 9
Discuss impact on organization

- Impact on organization
  - Impact on staffing
  - Impact on space, equipment, software
  - Impact on workflow / routines
- Impact on relations with other institutions, organizations
- Selection process

Brainstorming

- Effective tool for hearing multiple viewpoints, issues, and general ideas
- Leads to the development of more specific ideas and solutions to issues
Brainstorming techniques useful for

- Initiating institutional SWOT analysis
  - Strengths
  - Weaknesses
  - Opportunities
  - Threats

- Scope and nature of projects

- Selection

Environment for brainstorming

- Create a relaxed and non-threatening atmosphere
- Decide if all staff involved or representatives from various departments
- Suggest that if participants are representatives that the representative meets with constituents to collect ideas, issues, viewpoints
Brainstorming strategies

- Select a facilitator (sometimes using an outsider has an advantage – facilitator does not have a vested interest in the results, or influence or direct the discussion)

- Write down all comments
  - No evaluation of ideas
  - Everyone has an opportunity to speak
  - Use flip chart, white board or software to record comments

Brainstorming process

1. Define ideas or problems
   - Rephrase idea to make sure that everyone understands the point; write it down concisely

2. Break down broad issues into smaller issues to be “brainstormed” separately

3. Time limit for each section

4. Select the most important issues
Building a consensus

Review all ideas presented then refine by:

- Look for items that duplicate each other
- Combine related concepts
- Narrow list down
- Work towards a consensus: find common ground

Getting to “yes”*

- Decide issues based on their merit
- Look for options that will lead to mutual gains (win-win)
- Avoid arguing from positions
- Focus on the issues/interests, not the people
- Use objective criteria

Stages getting to agreement

1. Analysis stage
   Gather, organize, consider information from all sides

2. Planning stage
   Evaluate the information, think of options

3. Discussion stage
   Communicate interests and options

Active listening skills

1. Hear the message
2. Interpret the message
3. Evaluate the message
4. Respond to the message
Tips for effective listening

- Take notes (locate key points)
- Reflective listening
- Focus on listening
- Build rapport with speaker
- Show respect

Possible impediments to agreement

Causes
- Competing agendas
- Concern about long-term support
- Partners lack skill sets to equally share responsibilities
- Partners fear cultural material will be damaged or lost if “loaned” to lead institution
Reflection

“You can plan the plan, but you can’t plan the outcome.”—Words of wisdom

Planning process*

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal constraints</td>
<td>Organizational mandates</td>
</tr>
<tr>
<td>SWOT analysis</td>
<td>Strengths, Weaknesses, Opportunities and threats</td>
</tr>
<tr>
<td>Mission</td>
<td>Institutional purpose &amp; values</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>Within mission, set realistic goals and objectives / activities</td>
</tr>
<tr>
<td>Stakeholder analysis</td>
<td>“Entities” who have a stake in the results</td>
</tr>
<tr>
<td>Work plan</td>
<td>General description of implementation</td>
</tr>
<tr>
<td>Operating plan</td>
<td>Specifics of work plan for given period</td>
</tr>
<tr>
<td>Vision for success</td>
<td>How the organization will look when plan is implemented</td>
</tr>
</tbody>
</table>

*Based on Bishoff and Allen (2004)
Components of a work plan

The work plan needs to address the following issues:

- What is the need?
- Who is the target audience?
- How is the digital project the best solution?
- What will be the impact on the institution?

Components of a needs analysis

- Determine types of data needed
- Collect and analyze data
- Describe how the digital project is a solution
Types of data needed

- Who is your target audience?
- How are their needs being meet, or not?
- Where are the gaps in service, in content?
- What audience skill, knowledge, or behavior can be improved?
- Environmental scan of other relevant projects

How to find or discover data

- Use governmental statistics
- Use library statistics
  - Size and scope
  - Use statistics
  - Reference desk statistics
  - Published studies
- Surveys
- Focus groups
Audience and needs gap example

The San Fernando Valley, which makes up fully 80 percent of CSUN's service community, is quite diverse ethnically, linguistically, and socio-economically. On the weekends, about 50% of the Library's service requests are by persons who are not affiliated with CSUN such as high school and elementary school students, local historical groups, and individual members of the local business community. [CSUN's] Special Collections and Archives ...contain extensive collections that document the history of the San Fernando Valley through a mixed media of rare illustrated items, drawings, photographs, brochures, pamphlets, maps, official and unofficial reports and studies, personal letters, oral remembrances and related records.

Both the CSUN undergraduate students and the K-12 students seek primary source material about their neighborhood, history of the valley, and history of California missions. It is difficult for them to find reliable information.

Benefits of solution

- Describe the solution
- Detail the benefits
- Describe how the solution will close the gap
- Calculate the cost of the solution
Benefits of solution

The San Fernando ... Digital Library opens accessibility to an unlimited number of client and user groups ... including scholars, teachers, students, local historical societies, and members of the community, material otherwise accessible only by on-site visits. The project will:

- Open holdings to a wider audience
- Heighten interest in the historic development of the Valley
- Provide primary source materials for K-14 classroom use
- Link historical collections throughout the Valley

Why digitize?

- to support collection management and preservation
- to make information and assets more readily available
- to provide material for educational programs and address curriculum needs
- to provide material for curators and researchers (internal and external)
- to eliminate redundant work, and creation of redundant assets (photographs, slides, digital images, etc.)
Presenting your case

“Selling” the project to internal staff, library administrators, campus administrator or governing boards, all may need to hear different content

Explaining the uneasy part without putting people off:
- Labor
- Time
- System support

Explaining what the project is using the right amount of information: products developed

Managing expectations

Selling the project

How does the project help fulfill institutional mission and goals

- Supports community outreach and public relations
- Increases user base
- Increases revenue (through commercial profit but also through donations)
- Creates more efficient workflows
- Helps preserve original materials (less wear and tear)
- Supports educational function of institution
Presenting costs to the administration

- Include a succinct statement of project goals
- Clearly state which (original) collections will be included
- What equipment is needed
- Staffing, how many and what skill sets?
- Hidden costs: “marketing,” benefits for new staff members, grant management costs
- In-kind costs (e.g. staff release time), effect on other projects
- Maintenance, “care and feeding”

Reflection

“There are no short-term cost savings to be realized by digitizing collections”—Lorna Hughes
Cost factors to consider

Every project is unique, costs will vary depending on:

• scope and material of the project
• staff and equipment costs
• database development

Data migration is not “once-in-a-lifetime,” but rather is ongoing.

Cost categories

- Operational
  - Hardware/Software
  - Training
- Organizational
  - Release time
  - Space
- Staffing
Relative Costs

Note: OCR to meet accessibility standards is more costly than indicated here.


Reported cost ranges
In-house and outsourcing: various combinations

- Permanent staff assigned, equipment purchased, software developed locally
- Temporary staff hired, equipment purchased, software developed locally
- Permanent and temporary staff employed, hardware purchased, software “subscription”
- Scanning and metadata creation performed by vendor

Staffing

- Work that can be outsourced:
  - database development
  - Scanning
  - Transcription of audio (e.g., oral histories)
  - Basic tagging (markup) for TEI, or EAD in XML
- In-house labor issues:
  - Release time (“in kind”), duties performed by temporary help?
  - Time supported by project, duties performed by temporary help?
  - New staff hired for project

Labor costs typically represent the largest percentage of costs in a digital project.
Staffing Costs

- Salaries
- Benefits
  - Health
  - Sick Leave
  - Vacation
  - Holidays
- Training
- Attendance at conferences and meetings

Hardware

- Scanners
  - Slide scanners
  - Flatbed scanners
  - Microfilm/Microfiche scanners
- Digital cameras
- Audio/video conversion
- Server for storage/delivery
- Server for streaming audio/video
- Long-term maintenance/replacement
Software

- In-house system deployment:
  - Requires skilled system administrators, programmers
  - How and by whom will the system be updated, enhanced, and maintained?

- Purchase of an off-the-shelf product:
  - Is the vendor reliable, responsive, and likely to stay in business?
  - Are resources available for system enhancements, updates, and ongoing technical support?
  - Is a vendor-hosted option available?

- Documentation of decisions made, code written

Vendor selection

- Visit website whose output you would want to emulate
- Take note of the solutions the project used to create the digital product
- Make a list of desired features and prioritize them
- Decide which features are necessary and which are optional depending on cost
Timeline

- Environmental scan of IT solutions
- Issue RFP
  - Deadline when due
  - Follow up questions
  - Evaluation of responses
- Short-list vendors
  - Site visits
  - Interview current (and past) customers
  - Vendor presentations
- Identify preferred vendor
- Award contract

Request For Proposal (RFP)

- User requirements
- System or technical requirements
- Functional requirements
- Interoperability with other OS / platforms
RFP assessment

- Does the proposed solution meet all the stated requirements?
- To what degree do they meet your ideal solution?
- Contacts and business history
- What support do they provide (e.g., in-house training)?
- Costs/prices clearly delineated
- How well do they communicate with their customer base

Points to remember

- Keep the IT solution in sync with the stated goals of the plan
- Link the “business case” to the goals
- Keep the stakeholders informed of the process
- Remain flexible – it is a dynamic environment
Collaborative Digitization Program

Website provides information about:

- Digital imaging vendors
- Preservation resources
- Software resources
- Technical resources
- Strategic planning documents
- Project manuals and presentations and more

Exercise 14: Team building and work planning
Session 15: Project management—proposal writing and assessment

(Many thanks to Mary Woodley, Library of Congress, and ALCTS for supplying content for this session)

Session goals

- Learn the basics of proposal writing
- Learn how to write an operational or implementation plan
- Understand “outcome-based” evaluation
- Learn why is assessment important
- Learn strategies for deciding
  - who will conduct the evaluation of the project
  - when will it take place
  - what will be the criteria for judging success
Parallels between planning and proposal writing

- Clearly articulated goals and objectives
- Succinct description of the content to be digitized and its relevancy to achieving the goals
- Realistic estimates concerning time, costs, staffing and IT
- Knowledge of the appropriate metadata and scanning standards
- Plan for implementation: workflows
- Defined criteria to measure success

Proposal writing team

- Who are the key players for writing the grant and their responsibilities?
- What is the role of a development officer?
- What is the role of the library director or dean in the process? Of technical services and cataloging staff?
- Whom can you consult with for feedback about the process?
Proposal components

- Title page
- Table of contents
- Summary/abstract
- Introduction
- Statement of need
- Goals and outcomes
- Work plan
- Evaluation/assessment plans
- Budget
- Sustainability
- Marketing

Proposal summary

Concise statement that includes:
- Who you are
- What the project is
- How the project relates to the mission of the organization
- How much funding is required
Proposal introduction

- Describe the organization and its community (adapt to proposal audience)
- What is the significance of the content you plan to digitize
- Does your organization have a track record with similar projects?

Example of library description (abridged)

The University Library is at the heart of the CSU Northridge (CSUN) campus. The building is 235,000 square feet ... The Library is staffed by 23 full and part-time librarians, 51 technical and research specialists, and 250 student assistants. With over 1.2 million volumes, 3 million microforms ... and an extensive historical collection of mixed media, rare books and archives ...
Example of description of the wider community  (abridged)

CSU Northridge (CSUN) is a public University, located in the San Fernando Valley, in the north west section of Los Angeles. As the only major university in this area, CSUN also serves the adjacent incorporated and unincorporated urban and rural areas ... The San Fernando Valley is quite diverse ethnically, linguistically, and socio-economically. (for external audiences)

Statement of need

- What need will be addressed
- Ways in which the project is significant
- Why the need matches funder’s mission (depending on proposal audience)
Audience and needs gap

The San Fernando Valley, which makes up fully 80 percent of CSUN's service community, is quite diverse ethnically, linguistically, and socio-economically. On the weekends, about 50% of the Library's service requests are by persons who are not affiliated with CSUN such as high school and elementary school students, local historical groups, and individual members of the local business community. [CSUN's] Special Collections and Archives …contain extensive collections that document the history of the San Fernando Valley through a mixed media of rare illustrated items, drawings, photographs, brochures, pamphlets, maps, official and unofficial reports and studies, personal letters, oral remembrances and related records.

Both the CSUN undergraduate students and the K-12 students seek primary source material about their neighborhood, history of the valley, and history of California missions. It is difficult for them to find reliable information.

Example of solution to need

The goal of the Digital Library is to provide full, open, and equal access to a wide variety of primary research materials about the socio-economic growth and cultural evolution of the Valley, from its earliest foundation, to its explosive growth post World War 2.
Project goals and objectives

- How does project meet the mission of the institution?
- How does the project provide a solution to the need stated earlier?
- Who is involved?
- Who is being served?
- Is it realistic or overly ambitious?

Example of a goal statement

In the first year, the project will make freely available to the academic community as well as the community at large, 1400 digital objects accompanied by full descriptions. These digital objects will directly support general interest in the fauna of the valley as well as K-12 biology courses. The school district will create 6 curriculum packages based on the digital objects and state curriculum standards.
Project work plan

- What are the quantifiable outcomes?
- What is the work plan to accomplish project?
  - Time frame
  - Space
  - Equipment
  - Staffing
  - Software
  - Metadata
- How do the methods compare to other similar projects?

Digital life cycle

- Activities surrounding the creation and maintenance of digital objects
  - Sequential
  - Parallel
  - Iterative
Digitization issues

- Metadata standards
- Digital standards: imaging and file formats
- Delivery of digitized content
- Rights management
- Preservation
Example of standards statement

The ... Digital Library will conform to [*State* Digital Library Digital Image Format Standards (2001) for documents, photographs, graphic material, oral history transcripts, and related items. The [*State* Digital Library Digital Object Standard: Metadata, Content, and Encoding (2001)](https://example.com) and the guidelines established by the Dublin Core Metadata Initiative will be followed to support retrieval, storage and migration of data resources. *Describing archives: A content standard* (2004) will guide the library cataloging of finding guides and related indexes to archival collections.

Documentation

To ensure consistency in the current project and in the future, the project team must develop a suite of documents:

- for workflow
- for cataloging policies and procedures, data standards, etc.
- for system (e.g. CMS, DAM) usage, data integrity, reports, etc.
Examples of measurable project actions

1. Review \{number\} historical documents for possible inclusion

2. \{number\} documents will be digitized and incorporated into a searchable database that is Internet accessible

Project actions timeline

<table>
<thead>
<tr>
<th>Project Month</th>
<th>Action</th>
<th>Steps Taken</th>
<th>Who is responsible</th>
</tr>
</thead>
<tbody>
<tr>
<td>01-03</td>
<td>Hire Project staff, buy equipment</td>
<td>Interview candidates; training</td>
<td>Project director, manager, catalogers</td>
</tr>
<tr>
<td>02-11</td>
<td>Scanning and metadata creation</td>
<td>Project Technicians will scan items and add data</td>
<td>Project technicians</td>
</tr>
<tr>
<td>12-13</td>
<td>Publicity, Presentations, Post-Grant activities</td>
<td>News Media &amp; Listservs contacted; Official opening, Presentations organized</td>
<td>Development Librarian, Outreach Librarian, Library Director, project staff</td>
</tr>
</tbody>
</table>
 Proposed project budget

a. Salaries and benefits
b. Library materials
c. Operation
d. Equipment
e. Indirect costs

Example of budget summary

<table>
<thead>
<tr>
<th></th>
<th>LSTA (1)</th>
<th>Other funds (2)</th>
<th>In-kind (3)</th>
<th>Total (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Salaries &amp; Benefits</td>
<td>$120,945</td>
<td>$52,275</td>
<td>$173,220</td>
<td></td>
</tr>
<tr>
<td>b. Library Materials</td>
<td>$0</td>
<td>$0</td>
<td>$5,000</td>
<td>$5,000</td>
</tr>
<tr>
<td>c. Operation</td>
<td>$3,760</td>
<td>$0</td>
<td>$6,760</td>
<td>$6,760</td>
</tr>
<tr>
<td>d. Equipment ($5K+)</td>
<td>$0</td>
<td>$0</td>
<td>$7,000</td>
<td>$7,000</td>
</tr>
<tr>
<td>e. Total for Objectives</td>
<td>$124,705</td>
<td>$15,000</td>
<td>$52,275</td>
<td>$191,980</td>
</tr>
<tr>
<td>f. Indirect Cost</td>
<td>$12,471</td>
<td>$12,471</td>
<td>$12,471</td>
<td>$12,471</td>
</tr>
<tr>
<td>g. TOTAL</td>
<td>$137,176</td>
<td>$167,471</td>
<td>$164,746</td>
<td>$204,451</td>
</tr>
</tbody>
</table>
Detailed information requests

- Contact info
- Budget details with narrative support for budget
- Client needs and project goals
  - Collection
  - Partners
  - Benefits
  - Relationship between library service and client group
- Measurable objectives and actions
- Timeline
- Reporting of results
- Marketing and publicity
- Sustainability

Project marketing and publicity
Outcome-based evaluation

- Typically encouraged by granting agencies
- Demonstrates that the goals of the digital project were met
- Includes assessment of operations or management (staffing, workflow efficiency)
- Includes quantitative and qualitative measures
- Is user-centric

Outcome-based evaluations consider:

- Impact and benefits that are the result of the project
- Short-term changes
- Long-term changes
Components of outcome-based evaluations

- Inputs
- Activities
- Outputs
- Outcome indicators (quantifiable outcomes)
- Outcome targets
- Outcome measures

Typical inputs

- Staff
- Money
- Equipment
Typical activities

- Assessment of collection
- Processing of archival and special collections
- Preservation activities
- Digitization and metadata creation

Typical outputs

- Number of images / objects scanned or digitized
- Number of metadata records created
- Number of supporting web pages created
  - Project documentation
  - Curriculum packages created
  - Survey or summary of collection
Typical outcome targets

- Size of collection estimated in grant proposal
- Impact on target audience
- Creation of new audience
- Protection of fragile resources (less handling)
- 24/7 access
- Need gap closed

Typical outcome measures

- Indicators of change
- Connected to the stated goals of project
  - Measured against a benchmark through data collection
    - Quantitative
    - Qualitative
Benchmark

- Represents the starting point
- Determine what you plan to measure at the onset of the project
- Examples:
  - How many students and faculty use the archives and special collections for research?
  - How many assignments on local history are answered by library resources, and which resources are used?

Examples of quantitative measures

- Size of the digital collection
- Number of inquires
- Transaction logs
  - Number of visits to the sites
  - Referring urls
  - IP address of user
  - Date and time of searches
  - Number of searches
  - Types of searches
Qualitative outcomes

Qualitative in terms of accessibility, usability, functionality, user satisfaction and expectations

- Focus groups
- Surveys
- Interviews

Check with institution concerning guidelines for using human subjects

Usability

- Assesses the structure of the digital site
- Assess how the user interacts with site
- Measured by:
  - Ease of navigation
  - Features clearly labelled
  - Logic of presentation
**Functionality**

- Does the software and web site perform as intended?
- Can it deliver the results expected?
- Measured by:
  - Precision and recall of search engine
  - Search options allow:
    - Limits
    - Group
    - Basic and advanced

**Accessibility**

Can the site be used by anyone regardless of disability or impairment?
- Hearing access
- Vision access
- Mobility access
- Cognitive access
Some accessibility issues

- Images without alt tags
- Some tables for layout
- Content presented as graphics without text version
- Video and audio clips without text versions
- Older versions of Adobe
- Links that are not text readable

Exercise 15: Proposal writing and assessment
Day 5 and workshop debriefing
Exercise 1a: Library bibliographic system metadata

Tools:

Two MARC records on pages following

Instructions:

Facilitator leads participants in a discussion of the functional objectives that the MARC records are designed to fulfill. How does the MARC metadata support the objectives of the library system? (For example, to find, identify, select, obtain.) What other system objectives can we detect from the system’s metadata? (Allow 30 minutes.)

Das Boot Columbia Pictures presents a Bavaria Atelier production in...
Nikel; production executive, Lutz Hengst; executive producers, Mark Damon, Edward R. Pressman, John W. Hyde.

500 __ |a Originally produced as motion picture in 1981.
500 __ |a Based on the novel by Lothar-Günther Buchheim.
500 __ |a Dolby stereo on linear tracks, mono compatible.
520 __ |a A World War II drama about a German submarine and its crew on patrol in the North Atlantic and their fight for survival.
500 __ |a "From Producers Sales Organization."
500 __ |a "60139"--Spine of container.
650 _0 |a Feature films.
650 _0 |a War films.
700 1_ |a Rohrbach, Günter.
700 1_ |a Bittins, Michael.
700 1_ |a Petersen, Wolfgang. |4 drt
700 1_ |a Prochnow, Jürgen. |4 prf
700 1_ |a Grönlmeyer, Herbert. |4 prf
700 1_ |a Wennemann, Klaus. |4 prf
700 1_ |a Bengsch, Hubertus. |4 prf
700 1_ |a Semmelrogge, Martin. |4 prf
700 1_ |a Buchheim, Lothar Günther. |t Boot.
710 2_ |a Columbia Pictures.
710 2_ |a Bavaria Atelier GmbH.
710 2_ |a Radiant Film GmbH.
710 2_ |a RCA/Columbia Pictures Home Video (Firm)
740 0_ |a Boat.
905 __ |a 19930306120000.0
948 __ |a c:er
948 2_ |a 20041105 |b m |d sjl8 |e cts
995 __ |a Hivolm |d 20070131
L'Atalante [videorecording] / [c a New Yorker Films release ; GFFA ; une production Jacque-Louis Nounez ; scénario, Jean Guinee ; adaptation et dialogues, Jean Vigo, Albert Riera.]

250  Collectors ed.


300  1 videodisc (87 min.) : [b sd., b&d ; |c 4 3/4 in.

538  a DVD.

546  In French with optional English subtitles.

511  Michel Simon, Dita Parló, Jean Daste, Gilles Margaritis, Maurice Gilles, Rafa Diligent, Louis Lefebvre.

508  Director, Jean Vigo; photography, Boris Kaufman; music, Maurice Jaubert.

500  Originally produced as motion picture in 1934.

520  When Juliette marries Jean, she comes to live with him aboard the barge "L'Atalante", where they travel along the Seine River. The bride becomes bored and leaves the ship for Paris.

500  Special features: Documentary, The making of Jean Vigo's l'Atalante; still gallery; poster gallery; Jean Vigo filmography.

650  0 Marriage [v Drama.

650  0 Ship captain’s spouses [z France [v Drama.

650  0 Love [v Drama.

650  0 Comedy films.

650  0 Feature films.

650  0 Motion pictures, French.

700  1 Vigo, Jean, |d 1905-1934.

700  1 Nounez, Jacques-Louis.
Guinée, Jean.
Simon, Michel, d 1895-1975.
Parlo, Orta.
Daste, Jean, d 1904-
Lefebvre, Louis, d 1871-1947.
Kaufman, Boris.
Jaubert, Maurice.
New Yorker Films.
Gaumont Franco-Film Aubert (Firm)
New Yorker Video (Firm)
pfind Pumpelly
20030626 r ty15 cts
20030729 f sok1 cts
\VIDEO-D\35762-D
9115
Exercise 1b: Digital library system metadata

Tools:

Three digital information system records on pages following

Instructions:

Facilitator leads participants in a discussion of the functional objectives that the XML records are designed to fulfill. How does the XML encoded metadata support the objectives of the digital information system? What system objectives can we detect by examining this system’s metadata? (Allow 30 minutes.)

<film>
  <title>Das Boot - Director's Cut</title>
  <cast>Juergen Prochnow, Arthur Gruenemeyer, Martin May</cast>
  <director>Wolfgang Petersen</director>
  <release>1981</release>
  <country>Germany</country>
  <description>This internationally acclaimed account of a German submarine crew was the first film to examine the ordinary German recruit’s experience in WW II, and remains "a moving testament to the wastefulness of battle" (New York Times). The new director's cut features almost an hour of extra footage and a remixed soundtrack.</description>
  <runtime>210</runtime>
  <sound>Yes</sound>
  <color>Yes</color>
  <rating>R</rating>
  <creationDate>8/27/1999</creationDate>
</film>

<film>
  <title>Jim Dine: A Self Portrait on the Walls</title>
  <director>Richard Stilwell</director>
  <release>19951996</release>
  <country>USA</country>
  <description>Two short documentaries about internationally renowned artist Jim Dine. The first records eight days of intense work and quiet rumination as Dine produces an exhibition of huge, bold charcoal drawings directly on the walls of the Ludwigsburg Kunstverein near Stuttgart, Germany. It is an unusual and transitory exhibition in that the drawings remain on the walls for only six weeks before being painted over. All About Looking depicts Dine teaching drawing at the famed Internationale Sommerakademie fur Bildene Kunst in Salzburg, Austria. The class (and the viewer)
learns that the effort is not geared toward the creation of a finished product; it is the process that is all important -- an understanding that is both liberating and fortifying and designed to enable the student to look and to see. Cosponsored with the History of Art Majors' Society.

<film>
  <title>L'Atalante</title>
  <cast>Jean Daste, Dita Parlo</cast>
  <director>Jean Vigo</director>
  <release>1934</release>
  <country>France</country>
  <description>Jean Vigo's genius emerges from this enchanting story of a marriage tested by life on the Atalante, a river barge. His poetic set pieces, artistic vision and life-affirming spirit animate multiple emotional truths.</description>
  <runtime>89</runtime>
  <sound>Yes</sound>
  <creationDate>7/17/2004</creationDate>
</film>
Exercise 2: Selection for digitization

Tools:

Selection for Digitization Questionnaire

Instructions:

Participants break into groups of about four to five people each. Facilitator introduces *Selection for Digitization Questionnaire*. Working together, group members identify UAEU collections that are possible candidates for digitization. From the collections identified, group members agree on one collection for which they will answer the questions in the questionnaire. As the group discusses and answers each question for the collection selected, group members should note particularly important opportunities and challenges that the collection presents. (Allow 30 minutes.)

For the session wrap-up, all groups report on the collections they identified and share significant opportunities and challenges associated with digitizing the collections. (Allow 30 minutes.)
## Selection for Digitization Questionnaire


<table>
<thead>
<tr>
<th>Category</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal</td>
<td>Is the material restricted because of privacy, content, or donor concerns?</td>
</tr>
<tr>
<td></td>
<td>Is it copyright protected? If so, do you have the right to create and disseminate digital reproductions?</td>
</tr>
<tr>
<td>Material</td>
<td>Does the material lend itself to digitization?</td>
</tr>
<tr>
<td></td>
<td>Can the informational content be adequately captured in digital form?</td>
</tr>
<tr>
<td></td>
<td>Do the physical formats and condition of the material represent major impediments?</td>
</tr>
<tr>
<td></td>
<td>How large and complex in terms of object format is the collection?</td>
</tr>
<tr>
<td>Preservation</td>
<td>Would the material be put at risk in the digitization process?</td>
</tr>
</tbody>
</table>
## Selection for Digitization Questionnaire


<table>
<thead>
<tr>
<th>Preservation: Would digital surrogates reduce use of the originals, thereby offering them protection from handling? Or, is the digital reproduction seen as a means to replace the originals?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization: Is the material in a coherent, logically structured order? Is it paginated or is the arrangement suggested by some other means? Is it complete?</td>
</tr>
<tr>
<td>Use: What kinds, level, and frequency of use are envisioned? Is there a clear understanding of user requirements? Can digitization support these uses?</td>
</tr>
<tr>
<td>Use: Will access to the material be significantly enhanced by digitization?</td>
</tr>
<tr>
<td>Use: Are there issues around security or access that must be taken into account (e.g., access restricted to certain people or use under certain conditions?)</td>
</tr>
<tr>
<td>Collections: Is there added incentive to digitize material based on the availability of complementary digital resources (including resources available from other organizations?)</td>
</tr>
</tbody>
</table>
## Selection for Digitization Questionnaire


<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collections: Is there an opportunity for multi-institutional cooperation?</td>
<td></td>
</tr>
<tr>
<td>Collections: Is there an opportunity for building thematic coherence or &quot;critical mass?&quot;</td>
<td></td>
</tr>
<tr>
<td>Duplication: Has the material already been digitized by another trusted source? If so, do the digital files possess sufficient quality, documentation, and functionality to serve your purposes? What conditions govern access and use of those files?</td>
<td></td>
</tr>
<tr>
<td>Infrastructure: Do your institution already have in place to support managing, delivering, and maintaining digitized materials? What components are still needed?</td>
<td></td>
</tr>
<tr>
<td>Fiscal: Can you determine the total cost of image acquisition (selection, preparation, capture, indexing, and quality control)? Is this cost justified based on real or perceived benefits accruing from digitization?</td>
<td></td>
</tr>
<tr>
<td>Fiscal: Are there funds to support the digitization effort? Is there institutional commitment to the on-going management and preservation of these files?</td>
<td></td>
</tr>
</tbody>
</table>
Exercise 3: Creating Dublin Core metadata for digital objects

Tools:

Digital object printouts (choice of two):

1. The National Whig Song (sheet music), two pages
   URL: http://resolver.library.cornell.edu/misc/5458453/3131/
2. Evening Grosbeak (print), one page
   URL: http://resolver.library.cornell.edu/misc/6280355/00606/

Metadata Template
Dublin Core Metadata Element Set, Version 1.1
DCMI Type Vocabulary

Instructions:

Participants break into groups of about four to five people each. Facilitator introduces digital objects, metadata template, DC element list, and DCMI type vocabulary.

Working together, group members create a brief DC metadata record for the object assigned to their group. Records should include at least these elements: title/subtitle, creator/name, resource type, publication information, physical description, subject, URL. Groups should take notes about the issues they consider and the decisions they make. For example: What controlled vocabularies would you consider? Are all fields necessary? How is your approach shaped by the constraints of the Dublin Core? (Allow 30 minutes.)

For the session wrap-up, groups appoint a spokesperson to present the metadata record they created and the main issues they considered. (Allow 30 minutes.)
THE NATIONAL WHIG SONG.

Written by
WILLIAM HAYDEN, ESQ.

and respectfully dedicated to the
Whigs of the United States.

Published by PARKER & DITSON, 135 Washington St.
THE NATIONAL WHIG SONG.

Written by W. HAYDEN, Esq.

adapted to a Popular Air.

ALLEGRO.

I'll sing you now a new Whig song, made to a good old rhyme,


fine true-hearted gentleman all of the olden time; By

born and blood, by kith and kin, a sound true Whig was he, For his father signed the charter, that

made our country free. Like a fine true-hearted gentleman, All of the olden time.

Entered according to Act of Congress, in the year 1840, by Parker & Ditson, in the Clerk's Office of the District Court of Massachusetts.
<table>
<thead>
<tr>
<th>Data Property (Field)</th>
<th>Data Value (Content)</th>
<th>Controlled Vocabulary Yes or No Specify, if any</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dublin Core Metadata Element Set, Version 1.1

<table>
<thead>
<tr>
<th>Identifier:</th>
<th><a href="http://dublincore.org/documents/2008/01/14/dces/">http://dublincore.org/documents/2008/01/14/dces/</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Supersedes:</td>
<td><a href="http://dublincore.org/documents/2006/12/18/dces/">http://dublincore.org/documents/2006/12/18/dces/</a></td>
</tr>
<tr>
<td>Latest version:</td>
<td><a href="http://dublincore.org/documents/dces/">http://dublincore.org/documents/dces/</a></td>
</tr>
<tr>
<td>Date Issued:</td>
<td>2008-01-14</td>
</tr>
<tr>
<td>Status of document:</td>
<td>This is a DCMI Recommendation.</td>
</tr>
<tr>
<td>Description of document:</td>
<td>This document provides ready reference for the Dublin Core Metadata Element Set, Version 1.1. For more detailed documentation and links to historical versioning information, see the document &quot;DCMI Metadata Terms&quot;.</td>
</tr>
</tbody>
</table>

**Term Name: contributor**

| URI: | http://purl.org/dc/elements/1.1/contributor |
| Label: | Contributor |
| Definition: | An entity responsible for making contributions to the resource. |
| Comment: | Examples of a Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity. |

**Term Name: coverage**

| URI: | http://purl.org/dc/elements/1.1/coverage |
| Label: | Coverage |
| Definition: | The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant. |
| Comment: | Spatial topic and spatial applicability may be a named place or a location specified by its geographic coordinates. Temporal topic may be a named period, date, or date range. A jurisdiction may be a named administrative entity or a geographic place to which the resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names [TGN]. Where appropriate, named places or time periods can be used in preference to numeric identifiers such as sets of coordinates or date ranges. |

**Term Name: creator**

<p>| URI: | <a href="http://purl.org/dc/elements/1.1/creator">http://purl.org/dc/elements/1.1/creator</a> |
| Label: | Creator |
| Definition: | An entity primarily responsible for making the resource. |
| Comment: | Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity. |</p>
<table>
<thead>
<tr>
<th>Term Name: date</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
<tr>
<td>Label: Date</td>
</tr>
<tr>
<td>Definition: A point or period of time associated with an event in the lifecycle of the resource.</td>
</tr>
<tr>
<td>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].</td>
</tr>
<tr>
<td>References: [W3CDTF] <a href="http://www.w3.org/TR/NOTE-datetime">http://www.w3.org/TR/NOTE-datetime</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label: Description</td>
</tr>
<tr>
<td>Definition: An account of the resource.</td>
</tr>
<tr>
<td>Comment: Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: format</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/format">http://purl.org/dc/elements/1.1/format</a></td>
</tr>
<tr>
<td>Label: Format</td>
</tr>
<tr>
<td>Definition: The file format, physical medium, or dimensions of the resource.</td>
</tr>
<tr>
<td>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].</td>
</tr>
<tr>
<td>References: [MIME] <a href="http://www.iana.org/assignments/media-types/">http://www.iana.org/assignments/media-types/</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/identifier">http://purl.org/dc/elements/1.1/identifier</a></td>
</tr>
<tr>
<td>Label: Identifier</td>
</tr>
<tr>
<td>Definition: An unambiguous reference to the resource within a given context.</td>
</tr>
<tr>
<td>Comment: Recommended best practice is to identify the resource by means of a string conforming to a formal identification system.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label: Language</td>
</tr>
<tr>
<td>Definition: A language of the resource.</td>
</tr>
<tr>
<td>Comment: Recommended best practice is to use a controlled vocabulary such as RFC 4646 [RFC4646].</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/publisher">http://purl.org/dc/elements/1.1/publisher</a></td>
</tr>
<tr>
<td>Label: Publisher</td>
</tr>
<tr>
<td>Definition: An entity responsible for making the resource available.</td>
</tr>
<tr>
<td>Comment: Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.</td>
</tr>
<tr>
<td>Term Name: relation</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
<tr>
<td>Label: Relation</td>
</tr>
<tr>
<td>Definition: A related resource.</td>
</tr>
<tr>
<td>Comment: Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: rights</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/rights">http://purl.org/dc/elements/1.1/rights</a></td>
</tr>
<tr>
<td>Label: Rights</td>
</tr>
<tr>
<td>Definition: Information about rights held in and over the resource.</td>
</tr>
<tr>
<td>Comment: Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: source</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/source">http://purl.org/dc/elements/1.1/source</a></td>
</tr>
<tr>
<td>Label: Source</td>
</tr>
<tr>
<td>Definition: A related resource from which the described resource is derived.</td>
</tr>
<tr>
<td>Comment: The described resource may be derived from the related resource in whole or in part. Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/subject">http://purl.org/dc/elements/1.1/subject</a></td>
</tr>
<tr>
<td>Label: Subject</td>
</tr>
<tr>
<td>Definition: The topic of the resource.</td>
</tr>
<tr>
<td>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.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: title</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a></td>
</tr>
<tr>
<td>Label: Title</td>
</tr>
<tr>
<td>Definition: A name given to the resource.</td>
</tr>
<tr>
<td>Comment: Typically, a Title will be a name by which the resource is formally known.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: type</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/type">http://purl.org/dc/elements/1.1/type</a></td>
</tr>
<tr>
<td>Label: Type</td>
</tr>
<tr>
<td>Definition: The nature or genre of the resource.</td>
</tr>
<tr>
<td>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.</td>
</tr>
</tbody>
</table>

**Term Name: Collection**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/dcmitype/Collection">http://purl.org/dc/dcmitype/Collection</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Collection</td>
</tr>
<tr>
<td>Definition</td>
<td>An aggregation of resources.</td>
</tr>
<tr>
<td>Comment</td>
<td>A collection is described as a group; its parts may also be separately described.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>Class</td>
</tr>
<tr>
<td>Member Of</td>
<td><a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version</td>
<td><a href="http://dublincore.org/usage/terms/history/#Collection-003">http://dublincore.org/usage/terms/history/#Collection-003</a></td>
</tr>
</tbody>
</table>

**Term Name: Dataset**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/dcmitype/Dataset">http://purl.org/dc/dcmitype/Dataset</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Dataset</td>
</tr>
<tr>
<td>Definition</td>
<td>Data encoded in a defined structure.</td>
</tr>
<tr>
<td>Comment</td>
<td>Examples include lists, tables, and databases. A dataset may be useful for direct machine processing.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>Class</td>
</tr>
<tr>
<td>Member Of</td>
<td><a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version</td>
<td><a href="http://dublincore.org/usage/terms/history/#Dataset-003">http://dublincore.org/usage/terms/history/#Dataset-003</a></td>
</tr>
</tbody>
</table>
### Term Name: Event

<table>
<thead>
<tr>
<th>URI:</th>
<th><a href="http://purl.org/dc/dcmitype/Event">http://purl.org/dc/dcmitype/Event</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
<td>Event</td>
</tr>
<tr>
<td>Definition:</td>
<td>A non-persistent, time-based occurrence.</td>
</tr>
<tr>
<td>Comment:</td>
<td>Metadata for an event provides descriptive information that is the basis for discovery of the purpose, location, duration, and responsible agents associated with an event. Examples include an exhibition, webcast, conference, workshop, open day, performance, battle, trial, wedding, tea party, conflagration.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>Class</td>
</tr>
<tr>
<td>Member Of:</td>
<td><a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version:</td>
<td><a href="http://dublincore.org/usage/terms/history/#Event-003">http://dublincore.org/usage/terms/history/#Event-003</a></td>
</tr>
</tbody>
</table>

### Term Name: Image

<table>
<thead>
<tr>
<th>URI:</th>
<th><a href="http://purl.org/dc/dcmitype/Image">http://purl.org/dc/dcmitype/Image</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
<td>Image</td>
</tr>
<tr>
<td>Definition:</td>
<td>A visual representation other than text.</td>
</tr>
<tr>
<td>Comment:</td>
<td>Examples include images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation. Note that Image may include both electronic and physical representations.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>Class</td>
</tr>
<tr>
<td>Broader Than:</td>
<td><a href="http://purl.org/dc/dcmitype/StillImage">http://purl.org/dc/dcmitype/StillImage</a></td>
</tr>
<tr>
<td>Broader Than:</td>
<td><a href="http://purl.org/dc/dcmitype/MovingImage">http://purl.org/dc/dcmitype/MovingImage</a></td>
</tr>
<tr>
<td>Member Of:</td>
<td><a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version:</td>
<td><a href="http://dublincore.org/usage/terms/history/#Image-004">http://dublincore.org/usage/terms/history/#Image-004</a></td>
</tr>
</tbody>
</table>

### Term Name: InteractiveResource

<table>
<thead>
<tr>
<th>URI:</th>
<th><a href="http://purl.org/dc/dcmitype/InteractiveResource">http://purl.org/dc/dcmitype/InteractiveResource</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
<td>Interactive Resource</td>
</tr>
<tr>
<td>Definition:</td>
<td>A resource requiring interaction from the user to be understood, executed, or experienced.</td>
</tr>
<tr>
<td>Comment:</td>
<td>Examples include forms on Web pages, applets, multimedia learning objects, chat services, or virtual reality environments.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>Class</td>
</tr>
<tr>
<td>Member Of:</td>
<td><a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version:</td>
<td><a href="http://dublincore.org/usage/terms/history/#InteractiveResource-003">http://dublincore.org/usage/terms/history/#InteractiveResource-003</a></td>
</tr>
<tr>
<td>Term Name: MovingImage</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>URI: <a href="http://purl.org/dc/dcmitype/MovingImage">http://purl.org/dc/dcmitype/MovingImage</a></td>
<td></td>
</tr>
<tr>
<td>Label: Moving Image</td>
<td></td>
</tr>
<tr>
<td>Definition: A series of visual representations imparting an impression of motion when shown in succession.</td>
<td></td>
</tr>
<tr>
<td>Comment: Examples include animations, movies, television programs, videos, zoetropes, or visual output from a simulation. Instances of the type Moving Image must also be describable as instances of the broader type Image.</td>
<td></td>
</tr>
<tr>
<td>Type of Term: Class</td>
<td></td>
</tr>
<tr>
<td>Narrower Than: <a href="http://purl.org/dc/dcmitype/Image">http://purl.org/dc/dcmitype/Image</a></td>
<td></td>
</tr>
<tr>
<td>Member Of: <a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
<td></td>
</tr>
<tr>
<td>Version: <a href="http://dublincore.org/usage/terms/history/#MovingImage-003">http://dublincore.org/usage/terms/history/#MovingImage-003</a></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: PhysicalObject</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/dcmitype/PhysicalObject">http://purl.org/dc/dcmitype/PhysicalObject</a></td>
</tr>
<tr>
<td>Label: Physical Object</td>
</tr>
<tr>
<td>Definition: An inanimate, three-dimensional object or substance.</td>
</tr>
<tr>
<td>Comment: Note that digital representations of, or surrogates for, these objects should use Image, Text or one of the other types.</td>
</tr>
<tr>
<td>Type of Term: Class</td>
</tr>
<tr>
<td>Member Of: <a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version: <a href="http://dublincore.org/usage/terms/history/#PhysicalObject-003">http://dublincore.org/usage/terms/history/#PhysicalObject-003</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/dcmitype/Service">http://purl.org/dc/dcmitype/Service</a></td>
</tr>
<tr>
<td>Label: Service</td>
</tr>
<tr>
<td>Definition: A system that provides one or more functions.</td>
</tr>
<tr>
<td>Comment: Examples include a photocopying service, a banking service, an authentication service, interlibrary loans, a Z39.50 or Web server.</td>
</tr>
<tr>
<td>Type of Term: Class</td>
</tr>
<tr>
<td>Member Of: <a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version: <a href="http://dublincore.org/usage/terms/history/#Service-003">http://dublincore.org/usage/terms/history/#Service-003</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/dcmitype/Software">http://purl.org/dc/dcmitype/Software</a></td>
</tr>
<tr>
<td>Label: Software</td>
</tr>
<tr>
<td>Definition: A computer program in source or compiled form.</td>
</tr>
<tr>
<td>Comment: Examples include a C source file, MS-Windows .exe executable, or Perl script.</td>
</tr>
<tr>
<td>Type of Term: Class</td>
</tr>
<tr>
<td>Member Of: <a href="http://purl.org/dc/terms/DCMIType">http://purl.org/dc/terms/DCMIType</a></td>
</tr>
<tr>
<td>Version: <a href="http://dublincore.org/usage/terms/history/#Software-003">http://dublincore.org/usage/terms/history/#Software-003</a></td>
</tr>
<tr>
<td>Term Name: Sound</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>Member Of:</td>
</tr>
<tr>
<td>Version:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: StillImage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>Narrower Than:</td>
</tr>
<tr>
<td>Member Of:</td>
</tr>
<tr>
<td>Version:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>Member Of:</td>
</tr>
<tr>
<td>Version:</td>
</tr>
</tbody>
</table>
Exercise 4: Sample use case

Tools:

Sample Use Case (pre-print web hosting system for research literature)

Instructions:

Facilitator leads participants in an examination of the sample use case from the perspective of the metadata specialist. In the system development scenario suggested by the use case, the metadata specialist would be responsible for the shape of metadata flowing into this system and would have to ensure that this metadata can fulfill the objectives of the system. Which steps in the Flow of Events would depend on or impact system metadata decisions? These would be the steps that the metadata specialist would need to focus on and perhaps raise questions about. (Allow 45 minutes.)
8.3.2 Use Case for advanced search

**Description**: Access to the advanced search page will be available from all pages within the article pre-print system. The system’s advanced search searches the entire database of pre-print records and all associated full-text.

**Priority**: Critical. Advanced search functionality is considered essential to the system's usability.

**Preconditions**:
- The system has content in it (metadata records, one per pre-print, and associated full-text).
- The system and all search tools are operational.
- The USER has selected the "advanced search" option from some page.

**Flow of Events**:

1. The system displays an Advanced Search form, with five query boxes and all other options displayed.
2. The USER enters search terms in one or more query boxes.
3. For each query box, the USER may accept or alter the default search fields associated with that box. The possible search fields on each query box are:
   - author [default in query box 1]
   - title [default in query box 2]
   - abstract [default in query box 3]
   - subject terms [default in query box 4]
   - full-text [default in query box 5]
   - all fields
4. The USER may accept or alter the boolean operator radio buttons between each query box. The possible selections are:
   - and [default, all buttons]
   - or
   - not
5. The USER may accept or alter date (pre-print submission to system) restrictions on searched content by selecting months and years from pull-down menus in "search from" and "search to" boxes. The month pull-down has all twelve months, in three-letter format. The year pull-down has year values from “2000” to the current year. Default values, which place no date restriction on searched content, are:
   - Search from: Jan | 2000 [month and year of earliest pre-prints in system]
   - Search to: MMM | YYYY [current month and year]
6. The USER may accept or alter search results sorting criteria. Possible search result sort orders are:
   - by date [default]
   - by author
   - by title
7. The USER may accept or alter the “search results per page" selection. The options are:
   • 25 per page [default]
   • 50 per page
   • 100 per page
   • 200 per page
8. The USER submits search to system by clicking a “Search” button.
9. The system executes search.
10. The system displays appropriately formatted search results meeting the USER’s search criteria.
    Each search result includes:
    • author names, individually linked to a single-author search
    • title of pre-print, linked to pre-print record
    • extent of pre-print, in pages
    • subject terms

Alternative Events:
• The USER clicks “Reset” button.
  The Advanced Search form is refreshed, with empty search query boxes and setting returned to default values.
• The USER selects another system page from navigational options.
  No search request is sent. The USER is taken to the selected page.
• No content records match search criteria entered by USER.
  The system displays the advanced search form to the USER, with USER's search criteria displayed. A message says explains that no records match criteria and that criteria should be altered.
• A system error occurs.
  The system displays an error page (standard error message), with a button to the advanced search form.
Exercise 5: Metadata analysis

Tools:

Six metadata examples (MARC record, Simple DC record, EAD document, HTML source code for an archival guide, paper archival guide, paper accession record); Metadata Analysis Template; Metadata Analysis Summary (on following pages)

Instructions:

Participants break into small groups and facilitator introduces the metadata examples and associated tools.

Group instructions: The six examples are of metadata having to do with archival collections. Your job is to work together to assess the examples, filling in the Metadata Analysis Template for each. For vocabulary to use in the template, refer to the Metadata Analysis Summary, which draws on the session slides. Look at each example of metadata independently, ignoring the fact that they all describe the same set of archival papers. Assess each one as if it is all you have. Note that the two final examples are not in electronic form—they are only on paper.

You should try to move through the template fairly quickly—don’t spend too much time on each example, there are no trick questions! For “Content Values,” indicate any content data standards or best practices that you know are relevant. You don’t need to evaluate the degree of adherence to a standard. For “Structure,” give a general structural characterization of the metadata. “Intended Use” may invite some consideration. (Allow 30 minutes.)

As a session wrap-up, participants as a whole will discuss the groups’ results. (Allow 30 minutes.)
Vietnam War: statistical analysis and evaluation projects,

000 02721mpc a2200301 a 450
001 2088649
005 20021213151108.0
008 890420i19681972nyu eng d
035 __ |a (CStRLIN)NYCV89A51
035 __ |a (NIC)notisAKS6960
040 __ |a NIC |c NIC |e appm |d NIC |d NIC
100 1_ |a Prince, William G.
245 00 |a Vietnam War: statistical analysis and evaluation projects, |f 1968-1972.
300 __ |a .7 cubic ft.

545 __ |a Documentation was compiled by Prince as part of an Analysis of Vietnamization project, conducted by the Dept. of Applied Science and Technology, Bendix Aerospace Systems Division, sponsored by the Defense Advanced Research Projects Agency, and completed in 1973.

520 __ |a Documentation for various automated systems designed to provide data about the war in Vietnam. Systems include Project Corona Harvest to evaluate the effectiveness of airpower in Southeast Asia; Hamlet Evaluation System (HES), a reporting system designed to gather data on the progress of the rural pacification effort; SEAPRS (Southeast Asia Province file), designed to facilitate analysis of friendly and enemy military and pacification activity at the province level; PAAS (Pacification Attitude Analysis System), an automated system to provide a means of processing and reporting the results of surveys to determine the attitudes of the Vietnamese people toward pacification, the war, and political, social, and economic development; SEER (System for Evaluating the Effectiveness of RVNAF), designed to provide quantified evaluations of Vietnamese armed forces unit combat effectiveness in performance of assigned missions; AIRSUM (Air Summary Data Base), an historical record of all offensive air activity in Southeast Asia from 1965 to
1972; and Project Corona Harvest, an Air Force project designed to evaluate the effectiveness of air power in Southeast Asia from 1954.

555 0_ |a Folder list.

544 __ |3 Additional pamphlets and reports by William G. Prince are |a housed in the Echols Collection, Kroch Library, Cornell University.


650 _0 |a Vietnamese Conflict, 1961-1975.

650 _0 |a Combat |x Statistics |x Information sources.

650 _0 |a Internal security |z Vietnam.

650 _0 |a Insurgency |z Vietnam.

650 _0 |a Military art and science |x Data processing.

650 _0 |a Military art and science |x Automation.

650 _0 |a Military assistance, American |z Southeast Asia |x Computer programs.

856 40 |3 Finding aid |u http://resolver.library.cornell.edu/cgi-bin/EADresolver?id=RMM04406

905 __ |a 19920617120000.0
<?xml version="1.0" encoding="UTF-8" ?>
<dc xmlns="http://purl.org/dc/elements/1.1/">
  <creator>Prince, William G.</creator>
  <subject>Prince, William G.</subject>
  <subject>Vietnamese Conflict, 1961-1975.</subject>
  <subject>Combat--Statistics--Information sources.</subject>
  <subject>Internal security--Vietnam.</subject>
  <subject>Insurgency--Vietnam.</subject>
  <subject>Military art and science--Data processing.</subject>
  <subject>Military art and science--Automation.</subject>
  <subject>Military assistance, American--Southeast Asia--Computer programs.</subject>
  <description>Documentation for various automated systems designed to provide data about the war in Vietnam.</description>
  <publisher>Division of Rare and Manuscript Collections, Cornell University Library</publisher>
  <date>January 2002</date>
  <coverage>1968-1972</coverage>
  <type>collection</type>
  <format>Manuals, reports, questionnaires, correspondence, and other documents</format>
  <identifier>4406</identifier>
  <identifier>http://resolver.library.cornell.edu/cgi-bin/EADresolver?id=RMM04406</identifier>
</dc>
<xml version="1.0" encoding="UTF-8" >
<ead>
 <eadheader langencoding="ISO 639-2">
   <eadid type="SGML catalog">PUBLIC "-/Cornell University::Cornell University Library::Division of Rare and Manuscript Collections//TEXT(US::NIC::RMM04406::Vietnam War: Statistical Analysis and Evaluation Projects)//EN" "RMM04406.xml"</eadid>
 <filedesc>
   <titlestmt>
     <author>Compiled by E. Engst</author>
   </titlestmt>
   <publicationstmt>
     <publisher>Division of Rare and Manuscript Collections, Cornell University Library</publisher>
     <date>January 2002</date>
   </publicationstmt>
 </filedesc>
 <profiledesc>
   <creation>
   </creation>
   <language>ENG</language>
 </profiledesc>
 <eadheader>
   <frontmatter>
     <titlepage>
     </titlepage>
     <publisher>
       Division of Rare and Manuscript Collections
     </publisher>
     <list type="deflist">
       <defitem>
         <label>Contact Information:</label>
         <item>&RMCaddress;</item>
       </defitem>
       <defitem>
         <label>Compiled by:</label>
         <item>E. Engst</item>
       </defitem>
     </list>
   </frontmatter>
 </eadheader>
</ead>
Documentation for various automated systems designed to provide data about the war in Vietnam. Systems include Project Corona Harvest to evaluate the effectiveness of airpower in Southeast Asia; Hamlet Evaluation System (HES), a reporting system designed to gather data on the progress of the rural pacification effort; SEAPRS (Southeast Asia Province file), designed to facilitate analysis of friendly and enemy military and pacification activity at the province level; PAAS (Pacification Attitude Analysis System), an automated system to provide a means of processing and reporting the results of surveys to determine the attitudes of the Vietnamese people toward pacification, the war, and political, social, and economic development; SEER (System for Evaluating the Effectiveness of RVNAF), designed to provide quantified evaluations of Vietnamese armed forces unit combat effectiveness in performance of assigned missions; AIRSUM (Air Summary Data Base), an historical record of all offensive air activity in Southeast Asia from 1965 to 1972; and Project Corona Harvest, an Air Force project designed to evaluate the effectiveness of airpower in Southeast Asia from 1954.

Documentation was compiled by Prince as part of an Analysis of Vietnamization project, conducted by the Dept. of Applied Science and

</scopecontent>
</controlaccess>
</head>
</controlaccess>
</head>
</Subjects>
</controlaccess>
</admininfo>
<head>INFORMATION FOR USERS</head>
<prefercite>
<head>Cite As:</head>
<p>Vietnam War: Statistical Analysis and Evaluation Projects, #4406. Division of Rare and Manuscript Collections, Cornell University Library.</p>
</prefercite>
</admininfo>
<add type="relatedmaterial">
<head>RELATED MATERIAL</head>
<p>Additional pamphlets and reports by William G. Prince are housed in the Echols Collection, Olin Library, Cornell University.</p>
</add>
</relatedmaterial>
</head>
<series>
<label>SERIES I. Project Corona Harvest</label>
<item>Box 1</item>
</series>
Series II. HES 70

Series III. SEAPRS

Series IV. PAAS

Series V. SEER

Series VI. AIRSUM

CONTAINER LIST

Series I. Project Corona Harvest

System to evaluate the effectiveness of airpower in Southeast Asia, encompassing all airpower employed from 1954 to the end of the conflict (Air Force project).

Operating instructions,

1968

Series II. HES 70

Hamlet Evaluation System--reporting system designed to gather data on the progress of the rural pacification effort.

Data gathering instrument--formatted, multiple choice questionnaire:
Respondent--U.S. advisors in the field

Sample size--Every inhabited hamlet and village in south Vietnam

Frequency of reporting--Every hamlet and village reported each month

Comparison of HES 70 and PAAS

Memo, 1970

Appendix A--Question codes, question responses

Hamlet level HES statistics and plots by NMCSSC for village program analysis (1)

VSSG IDX (2)

Security (3)
Collection Number: 4406
Division of Rare and Manuscript Collections
Cornell University Library

Contact Information:
Division of Rare and Manuscript Collections
2B Carl A. Kroch Library
Cornell University
Ithaca, NY 14853
(607) 255-3530
Fax: (607) 255-9524

Compiled by: E. Engst
October 1989

EAD encoding: Martin Heggestad, January 2002

© 2002 Division of Rare and Manuscript Collections, Cornell University Library

Title: Vietnam War: statistical analysis and evaluation projects, 1968-1972
Collection No.: 4406
Creator: William G. Prince
Quantity: .7 cubic ft.
Forms of Material: Manuals, reports, questionnaires, correspondence, and other documents.
Repository: Division of Rare and Manuscript Collections, Cornell University Library

Abstract: Documentation for various automated systems designed to provide data about the war in Vietnam. Systems include Project Corona Harvest to evaluate the effectiveness of airpower in Southeast Asia; Hamlet Evaluation System (HES), a reporting system designed to gather data on the progress of the rural pacification effort; SEAPRS (Southeast
Asia Province file), designed to facilitate analysis of friendly and enemy military and pacification activity at the province level; PAAS (Pacification Analysis System), an automated system to provide a means of processing and reporting the results of surveys to determine the attitudes of the Vietnamese people toward pacification, the war, and political, social, and economic development; SEER (System for Evaluating the Effectiveness of RVNAF), designed to provide quantified evaluations of Vietnamese armed forces unit combat effectiveness in performance of assigned missions; AIRSUM (Air Summary Data Base), an historical record of all offensive air activity in Southeast Asia from 1965 to 1972; and Project Corona Harvest, an Air Force project designed to evaluate the effectiveness of air power in Southeast Asia from 1954.

Documentation was compiled by Prince as part of an Analysis of Vietnamization project, conducted by the Dept. of Applied Science and Technology, Bendix Aerospace Systems Division, sponsored by the Defense Advanced Research Projects Agency, and completed in 1973.

Subjects:
- Combat -- Statistics -- Information sources.
- Internal security -- Vietnam.
- Insurgency -- Vietnam.
- Military art and science -- Data processing.
- Military art and science -- Automation.
- Military assistance, American -- Southeast Asia -- Computer programs.

INFORMATION FOR USERS

Cite As: Vietnam War: Statistical Analysis and Evaluation Projects, #4406. Division of Rare and Manuscript Collections, Cornell University Library.

Additional pamphlets and reports by William G. Prince are housed in the Echols Collection, Kroch Library, Cornell University.

SERIES LIST

- Series I: Project Corona Harvest
  - Box 1
- Series II: HES 70
  - Box 1
- Series III: SEAPRS
  - Box 1
- Series IV: PAAS
  - Box 1
- Series V: SEER
  - Box 1
- Series VI: AIRSUM
  - Boxes 1 and 2
### CONTAINER LIST

<table>
<thead>
<tr>
<th>Description</th>
<th>Container</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series I. Project Corona Harvest</td>
<td></td>
</tr>
<tr>
<td>System to evaluate the effectiveness of airpower in Southeast Asia,</td>
<td></td>
</tr>
<tr>
<td>encompassing all airpower employed from 1954 to the end of the conflict</td>
<td></td>
</tr>
<tr>
<td>(Air Force project).</td>
<td></td>
</tr>
<tr>
<td>Operating instructions, 1968</td>
<td>Box 1</td>
</tr>
<tr>
<td>Comparison of HES 70 and PAAS</td>
<td>Box 1</td>
</tr>
<tr>
<td>Memo, 1970</td>
<td>Box 1</td>
</tr>
<tr>
<td>Appendix A--Question codes, question responses</td>
<td>Box 1</td>
</tr>
<tr>
<td>Hamlet level HES statistics and plots by NMCSSC for village program analysis</td>
<td>Box 1</td>
</tr>
</tbody>
</table>

Collection Number: 4406

Division of Rare and Manuscript Collections
Cornell University Library

Contact Information:
Division of Rare and Manuscript Collections
2B Carl A. Kroch Library
Cornell University
Ithaca, NY 14853
(607) 255-3530
Fax: (607) 255-9524
rareref@cornell.edu
http://rmc.library.cornell.edu

Compiled by: E. Engst
Date completed: October 1989

© 2002 Division of Rare and Manuscript Collections, Cornell University Library

DESCRIPTIVE SUMMARY

Title:
Vietnam War: statistical analysis and evaluation projects, 1968-1972

Collection Number:
4406

Creator:
William G. Prince

Quantity:
.7 cubic ft.

Forms of Material:
Manuals, reports, questionnaires, correspondence, and other documents.

Repository:
Division of Rare and Manuscript Collections, Cornell University Library

Abstract:
Documentation for various automated systems designed to provide data about the war in Vietnam.

COLLECTION DESCRIPTION

Documentation for various automated systems designed to provide data about the war in Vietnam. Systems include Project Corona Harvest to evaluate the effectiveness of airpower in Southeast Asia;
Hamlet Evaluation System (HES), a reporting system designed to gather data on the progress of the rural pacification effort; SEAPRS (Southeast Asia Province file), designed to facilitate analysis of friendly and enemy military and pacification activity at the province level; PAAS (Pacification Attitude Analysis System), an automated system to provide a means of processing and reporting the results of surveys to determine the attitudes of the Vietnamese people toward pacification, the war, and political, social, and economic development; SEER (System for Evaluating the Effectiveness of RVNAF), designed to provide quantified evaluations of Vietnamese armed forces unit combat effectiveness in performance of assigned missions; AIRSUM (Air Summary Data Base), an historical record of all offensive air activity in Southeast Asia from 1965 to 1972; and Project Corona Harvest, an Air Force project designed to evaluate the effectiveness of air power in Southeast Asia from 1954.

Documentation was compiled by Prince as part of an Analysis of Vietnamization project, conducted by the Dept. of Applied Science and Technology, Bendix Aerospace Systems Division, sponsored by the Defense Advanced Research Projects Agency, and completed in 1973.

SUBJECTS

Names:
Prince, William G.

Subjects:
Combat--Statistics--Information sources.
Internal security--Vietnam.
Insurgency--Vietnam.
Military art and science--Data processing.
Military art and science--Automation.
Military assistance, American--Southeast Asia--Computer programs.

INFORMATION FOR USERS

Cite As:
Vietnam War: Statistical Analysis and Evaluation Projects, #4406. Division of Rare and Manuscript Collections, Cornell University Library.

RELATED MATERIAL

Additional pamphlets and reports by William G. Prince are housed in the Echols Collection, Kroch Library, Cornell University.

SERIES LIST

Series I. Project Corona Harvest
Series II. HES 70
Series III. SEAPRS
Series IV. PAAS
CONTAINER LIST

Series I. Project Corona Harvest
System to evaluate the effectiveness of airpower in Southeast Asia, encompassing all airpower employed from 1954 to the end of the conflict (Air Force project).

Operating instructions, 1968 Box 1 Folder 1

Series II. HES 70
Hamlet Evaluation System--reporting system designed to gather data on the progress of the rural pacification effort.
Data gathering instrument--formatted, multiple choice questionnaire:

1. Respondent--U.S. advisors in the field
2. Sample size--Every inhabited hamlet and village in south Vietnam
3. Frequency of reporting--Every hamlet and village reported each month

Comparison of HES 70 and PAAS Box 1 Folder 2
Memo, 1970 Box 1 Folder 3
Appendix A--Question codes, question responses Box 1 Folder 4
Hamlet level HES statistics and plots by NMCSSC for village program analysis (1) Box 1 Folder 5
VSSG IDX (2) Box 1 Folder 6
Security (3) Box 1 Folder 7
Econ. str. (4) Box 1 Folder 8
Econ. stm. (5) Box 1 Folder 9
[unlabeled] (6) Box 1 Folder 10
GVN pol. inf. (8) Box 1 Folder 11
VC pol. inf. Box 1 Folder 12
Prog. effort Box 1 Folder 13
Misc. ques. Box 1 Folder 14

Series III. SEAPRS
Southeast Asia Province file--designed to facilitate analysis of friendly and enemy military and pacification activity at the province level. Summarized data is organized by month for Province Corps, Viet Cong Military Regions, Division Tactical Areas, Special Tactical Zones, and Countrywide. Some fields use data from or indicators developed by HES.

Southeast Asia Province (also includes Hamlet Evaluations System Handbook), 1969 Box 1 Folder 16

Series IV. PAAS
Pacification Attitude Analysis System. Automated system to provide the Military Assistance Command Civil Operations and Rural Development Support, Pacification Study Group a means of processing and reporting the results of surveys
ACCESSION SHEET

Tentative Title or Brief Statement of Content:

Vietnam Statistical Analysis Projects

Final Title (if different from above):

Vietnam War: Statistical Analysis and Evaluation Projects

Name and Address of Donor, Office of Origin, or Other Source:

John Wagner, Bendix Document Storage, 415 Logan Ave., Bld. 10
San Diego, CA

Approximate Inclusive Dates:

1968-1972

Accession Date:

July 29, 1989

Approximate Quantity Upon Arrival:

2 16” boxes

Physical Condition Upon Arrival:

__ Good  ___ Fair  ___ Poor  ___ Other (explain)

Related MSS Collections or Archival Holdings:

Other reports by Prince in Echols

Processing Dates, Personnel, Activities:

Fred Flintstone, Barney Rubble
Arranged, foldered, listed

Number and Size of Boxes and Linear Footage After Processing:

.7 cubic feet

Additional Comments:

See attached damaged and discarded list
## Exercise 5—Metadata Analysis Template

<table>
<thead>
<tr>
<th>File Format</th>
<th>Type of Metadata</th>
<th>Metadata Scheme</th>
<th>Content Values</th>
<th>Intended Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARC Record (pp. 13-14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dublin Core Record (p. 15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAD Document (pp. 16-20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HTML source code (pp. 21-23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper Archival Guide (pp. 24-26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper Accession Record (p. 27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Metadata Analysis Summary

File or data exchange format

Examples:
- SGML / HTML; XML / XHTML; MARC; Plain-text file, perhaps “delimited”; Binary (not plain-text) formats, either open or proprietary.

Type of metadata

Examples:
- Descriptive; Structural; Administrative; Technical; Preservation; Access/Rights.

Considerations:
- What is the informational content of the metadata concerned with?

Semantics (metadata scheme, element set)

Examples:
- MARC21; Dublin Core (DC); EAD; MODS; VRA Core; METS; etc.

Content values

Examples:
- Of content standards or best practices: AACR2/RDA; EAD Best Practice (RLG); CCO; etc.
- Of published and shared vocabularies: LCSH; AAT; TGM; etc.
- Of application profiles: DCMI Libraries AP; DCMI Education AP; DCMI Government AP; etc.

Considerations:
- What is the degree of conformance to any employed standards, practices, or vocabularies?

Structure

Examples:
- Simple unstructured; Simple structured; Richly structured

Considerations:
- Is the record structure flat or hierarchical (nested)?
- How complex are the relationships among data elements?
- Is element qualification allowed?
- What degree of ambiguity exists within the metadata?

Intended Use

Considerations:
- Why was this metadata created? What functional requirements did this metadata support?
- How was it used by its creators?
- What can its intended use tell us about its consistency, reliability, or interoperability?

Status

Examples:
- Static: metadata that is no longer updated, augmented, or maintained. It may be inherited from some source that will no longer contribute to it. It is not likely to change (unless repurposed).
- Dynamic: metadata that is “living,” in the sense that it is maintained by someone, updated when needed, regularly supplemented. Dynamic metadata may change over time.
Exercise 6: Metadata analysis scenarios

Tools:

Four metadata analysis scenarios, on following pages; six metadata examples from Exercise 5 (MARC record, Simple DC record, EAD document, HTML source code for an archival guide, paper archival guide, paper accession record)

Instructions:

Participants break into small groups. Facilitator introduces the metadata analysis scenarios and assigns each scenario to a group.

Group instructions:

Working together, group members analyze the key features of the scenario from the metadata practitioner’s perspective. Answer the question proposed in the scenario in light of the scenario’s constraints and requirements. (Refer to the metadata examples from Exercise 5 when needed to familiarize yourselves with the metadata types to which the scenarios refer.) (Allow 30 minutes.)

As a session wrap-up, groups identify a spokesperson that describes the group’s scenario and reports the group’s results. Participants as a whole discuss how scenario constraints and requirements affected metadata decisions. (Allow 30 minutes.)
Metadata Analysis Scenario A

Your University belongs to a broad consortium of cultural heritage institutions that include universities, museums, and state and local historical societies. The consortium would like to create a centralized discovery system for the consortium's vast range of archival collections. Your University Archives currently creates a MARC collection record and a paper archival guide for each of their archival collections, but it is clear that many of the consortium's smaller institutions have nothing more than an accession record for many of their collections.

As an advisor to the consortium's efforts, what metadata will they need in order to create their centralized discovery system?

Constraints and requirements:

- Not much in the way of resources (funding/staff) to devote to this.
- They would like something as quickly as possible.

Metadata Analysis Scenario B

Your organization belongs to a broad state-wide consortium of cultural heritage institutions that include universities, museums, and state and local historical societies. The state government would like to create a centralized system with in-depth descriptions about all of the consortium's vast range of archival collections. Your organization, like most others in the consortium, have been creating, in MS Word, detailed archival guides for their archival collections and then converting these to HTML for web publication (a simple conversion operation, in MS Word). Other than an accession record, these guides are the only information available about the collections.

As an advisor to the consortium's efforts, what metadata will they need in order to create their system?

Constraints and requirements:

- The state wants archival descriptions at least as detailed as the HTML guides they have now.
- The state wants sophisticated fielded searching capability in their system, such as the ability to limit searches to particular repositories, to collections that contain specific types of materials (such as letters or diaries), or to materials of a certain date range.
- It appears that that state is willing to provide whatever funds are required.
Metadata Analysis Scenario C

Your organization belongs to a broad state-wide consortium of cultural heritage institutions that include universities, museums, and state and local historical societies. The state government would like to create a centralized system with in-depth descriptions about all of the consortium's vast range of archival collections. Your organization, like most others in the consortium, have been creating, in MS Word, detailed archival guides for their archival collections and then converting these to HTML for web publication (a simple conversion operation, in MS Word). Other than an accession record, these guides are the only information available about the collections.

As an advisor to the consortium's efforts, what metadata will they need in order to create their system?

Constraints and requirements:

- The state wants archival descriptions at least as detailed as the HTML guides they have now.
- The state will commit very little funding to this project, so if it is done at all, the project must be carried out at the lowest cost possible.
- The state would like something as quickly as possible.

Metadata Analysis Scenario D

Your University Archives would like to gather management information about their archival collections in an electronic system, so that they can easily retrieve data and generate reports about collection donors, collection values, restricted content, and other management information. The Archives currently creates a MARC collection record and an EAD encoded archival guide for each of their archival collections.

What additional metadata, if any, will the University Archives need to collect to meet their objectives?
Exercise 7: Metadata mapping

Tools:

Four sample article metadata records; Dublin Core element and element refinement descriptions; Metadata Map.

Preliminary instructions:

Participants break into small groups. Facilitator introduces the sample metadata records, DC descriptions, and metadata mapping tool. Participants read the exercise instructions. Facilitator and participants review them and facilitator answers questions.

Group assignment instructions:

Your task is to create a metadata map that takes relatively rich source metadata and converts it to simple Dublin Core. This map will be used by a programmer to create a conversion routine that will automatically translate the source metadata to simple DC.

This is a fairly typical mapping requirement in the library world. In order to create a union catalog of disparate resources, or share metadata from various different sources and systems, we need a common metadata format to map into. Simple DC is often selected for such purposes. Simple DC is also the minimum metadata format for metadata harvesting via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). In other words, if you wish to allow information about the content of your collections to be harvested via OAI, you must at least provide a simple DC record for every resource (you can provide richer metadata).

Source metadata format: on the following pages are four samples of source metadata native to a journal hosting system. Each page represents a single journal issue, and each issue contains child elements that hold metadata about articles. Only one or two articles per issue are listed, as samples of the metadata available in the system.

Target metadata format: simple, or unqualified, Dublin Core. Assume that the appropriate level at which to provide DC records is the article level—one simple DC record per article in the hosting system. Also assume that the simple DC records will be expressed in XML and available for OAI harvesting.

Mapping task: use the Metadata Map template to create your metadata map. This mapping exercise has a fixed target, so begin with the DC elements and ask what source element or elements will be required to populate it. Describe any transformations necessary to get from source to target, or warnings or considerations that a programmer (who will have to implement the map in software code) must know about. The first element is filled in. Feel free to disagree with what has been proposed.
More about Dublin Core: on the following pages, you will find basic descriptions of each of the 15 DC elements possible in a simple, or unqualified, DC record. These are followed by a set of DC element refinements, or qualifiers, which cannot be used in simple DC, but can be useful in determining which elements of the core set to use. This is because DC refinements provide more precision than the core elements, and you may find among them the data definition you are looking for. This then tells you which core element to use—the one which the qualifying term is refining.

Some things to keep in mind about simple Dublin Core:

- All elements are optional
- All elements are repeatable
- Simple DC can contain only the core 15 elements.
- In XML, simple DC elements cannot contain any sub-elements. In other words, no XML or HTML markup is allowed inside of simple DC elements.

Here are some hints about specific DC elements:

Type element: the DC element Type refers to the DCMI Type Vocabulary. That vocabulary includes only the following list of terms. You should choose the most appropriate term.

<table>
<thead>
<tr>
<th>Collection</th>
<th>PhysicalObject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset</td>
<td>Service</td>
</tr>
<tr>
<td>Event</td>
<td>Software</td>
</tr>
<tr>
<td>Image</td>
<td>Sound</td>
</tr>
<tr>
<td>InteractiveResource</td>
<td>StillImage</td>
</tr>
<tr>
<td>MovingImage</td>
<td>Text</td>
</tr>
</tbody>
</table>

Source element: this element is a tricky one. The Dublin Core Libraries Working Group says to use Source “only when the described resource is the result of digitization of non-digital originals. Otherwise, use Relation.” For this exercise, assume that these journals have two separate dissemination streams: one is (still, for now) paper and one is electronic. In other words, the digital version doesn’t result from the digitization of the paper copy—one is not the source of the other.

bibliographicCitation element refinement: important for serial literature is the DC element refinement called “bibliographicCitation.” The DC community has decided that this is the best place to hold typical citation type data (journal name, volume number, year of publication, page range).

General hint: typically, simple DC metadata records are generated in order to increase the discovery of resources by end-users. This is certainly true of most OAI record harvesting. Let’s assume that’s our main goal here. So when faced with any particular mapping decision where several alternatives may be possible, choose to convey information most relevant to the discovery of the resource.
Some risk management problems for firms with internal competition and debt

Consider an optimization of the $\text{Swigler}$ problem, first formulated by Kunst in $\text{Liability Constant Rates}$: a constant liability payment rate $\text{B}$, an average return $\text{R}$, and a risk $\text{N}\times\text{xy}$ proportional to the size of the business unit.
Radiation effect on MHD free-convection flow of a gas at a stretching surface with a uniform free stream

We investigate the problem of free convection heat transfer near an isothermal stretching sheet. This has been done under the simultaneous action of buoyancy, radiation, and transverse magnetic field. The governing equations are solved by the shooting method. The velocity and temperature functions are represented graphically for various values of the flow parameters: radiation parameter $F$, free convection parameter $\text{Gr}$, magnetic parameter $M$, Prandtl number $\text{Pr}$, and the parameter of relative difference between the temperature of the sheet, and the temperature far away from the sheet. The effects of the radiation and magnetic field parameters on the shear stress and heat flux are discussed.
The $C^{1,1}$ regularity of the pluricomplex Brown function

Zbigniew B&őcłaski

32U35

211|215

A note on Pierskorn spheres and the generalized Jones conjecture

Yves Flambaud

57Q45

325|333
## The Dublin Core Metadata Element Set

### Term Name: contributor

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/elements/1.1/contributor">http://purl.org/dc/elements/1.1/contributor</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Contributor</td>
</tr>
<tr>
<td>Definition</td>
<td>An entity responsible for making contributions to the content of the resource.</td>
</tr>
<tr>
<td>Comment</td>
<td>Examples of a Contributor include a person, an organisation, or a service. Typically, the name of a Contributor should be used to indicate the entity.</td>
</tr>
</tbody>
</table>

### Term Name: coverage

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/elements/1.1/coverage">http://purl.org/dc/elements/1.1/coverage</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Coverage</td>
</tr>
<tr>
<td>Definition</td>
<td>The extent or scope of the content of the resource.</td>
</tr>
<tr>
<td>Comment</td>
<td>Coverage will typically include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and that, where appropriate, named places or time periods be used in preference to numeric identifiers such as sets of coordinates or date ranges.</td>
</tr>
</tbody>
</table>

### Term Name: creator

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/elements/1.1/creator">http://purl.org/dc/elements/1.1/creator</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Creator</td>
</tr>
<tr>
<td>Definition</td>
<td>An entity primarily responsible for making the content of the resource.</td>
</tr>
<tr>
<td>Comment</td>
<td>Examples of a Creator include a person, an organisation, or a service. Typically, the name of a Creator should be used to indicate the entity.</td>
</tr>
<tr>
<td>Term Name: date</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td><strong>URI:</strong></td>
<td><a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
<td>Date</td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
<td>A date associated with an event in the life cycle of the resource.</td>
</tr>
<tr>
<td><strong>Comment:</strong></td>
<td>Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and follows the YYYY-MM-DD format.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
</tr>
<tr>
<td><strong>Comment:</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: format</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
</tr>
<tr>
<td><strong>Comment:</strong></td>
</tr>
</tbody>
</table>
## Dublin Core Element Description

### Exercise 7

<table>
<thead>
<tr>
<th>Term Name: identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/identifier">http://purl.org/dc/elements/1.1/identifier</a></td>
</tr>
<tr>
<td>Label: Resource Identifier</td>
</tr>
<tr>
<td>Definition: An unambiguous reference to the resource within a given context.</td>
</tr>
<tr>
<td>Comment: Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Example formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: language</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/language">http://purl.org/dc/elements/1.1/language</a></td>
</tr>
<tr>
<td>Label: Language</td>
</tr>
<tr>
<td>Definition: A language of the intellectual content of the resource.</td>
</tr>
<tr>
<td>Comment: Recommended best practice is to use RFC 3066 [RFC3066], which, in conjunction with ISO 639 [ISO639], defines two- and three-letter primary language tags with optional subtags. Examples include &quot;en&quot; or &quot;eng&quot; for English, &quot;akk&quot; for Akkadian, and &quot;en-GB&quot; for English used in the United Kingdom.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/publisher">http://purl.org/dc/elements/1.1/publisher</a></td>
</tr>
<tr>
<td>Label: Publisher</td>
</tr>
<tr>
<td>Definition: An entity responsible for making the resource available</td>
</tr>
<tr>
<td>Comment: Examples of a Publisher include a person, an organisation, or a service. Typically, the name of a Publisher should be used to indicate the entity.</td>
</tr>
<tr>
<td>Term Name: relation</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: rights</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: source</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>Term Name: subject</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/subject">http://purl.org/dc/elements/1.1/subject</a></td>
</tr>
<tr>
<td>Label: Subject and Keywords</td>
</tr>
<tr>
<td>Definition: The topic of the content of the resource.</td>
</tr>
<tr>
<td>Comment: Typically, a Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: title</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/title">http://purl.org/dc/elements/1.1/title</a></td>
</tr>
<tr>
<td>Label: Title</td>
</tr>
<tr>
<td>Definition: A name given to the resource.</td>
</tr>
<tr>
<td>Comment: Typically, a Title will be a name by which the resource is formally known.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: type</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI: <a href="http://purl.org/dc/elements/1.1/type">http://purl.org/dc/elements/1.1/type</a></td>
</tr>
<tr>
<td>Label: Resource Type</td>
</tr>
<tr>
<td>Definition: The nature or genre of the content of the resource.</td>
</tr>
<tr>
<td>Comment: Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCMITYPE]). To describe the physical or digital manifestation of the resource, use the Format element.</td>
</tr>
</tbody>
</table>
## Dublin Core Element Description Exercise

### Element Refinements

<table>
<thead>
<tr>
<th>Term Name: abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
</tr>
<tr>
<td><strong>Type of Term:</strong></td>
</tr>
<tr>
<td><strong>Refines:</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: accessRights</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
</tr>
<tr>
<td><strong>Comment:</strong></td>
</tr>
<tr>
<td><strong>Type of Term:</strong></td>
</tr>
<tr>
<td><strong>Refines:</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
</tr>
<tr>
<td><strong>Comment:</strong></td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td><strong>Type of Term:</strong></td>
</tr>
<tr>
<td><strong>Refines:</strong></td>
</tr>
</tbody>
</table>

**Term Name: available**

| **URI:** | `http://purl.org/dc/terms/available` |
|**Label:** | Available |
| **Definition:** | Date (often a range) that the resource will become or did become available. |
| **Type of Term:** | `element-refinement` |
| **Refines:** | `http://purl.org/dc/elements/1.1/date` |

**Term Name: bibliographicCitation**

| **URI:** | `http://purl.org/dc/terms/bibliographicCitation` |
|**Label:** | Bibliographic Citation |
| **Definition:** | A bibliographic reference for the resource. |
| **Comment:** | Recommended practice is to include sufficient bibliographic detail to identify the resource as unambiguously as possible, whether or not the citation is in a standard form. |
| **Type of Term:** | `element-refinement` |
| **Refines:** | `http://purl.org/dc/elements/1.1/identifier` |

**Term Name: conformsTo**

<p>| <strong>URI:</strong> | <code>http://purl.org/dc/terms/conformsTo</code> |</p>
<table>
<thead>
<tr>
<th>Label</th>
<th>Conforms To</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>A reference to an established standard to which the resource conforms.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
</tbody>
</table>

**Term Name: created**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/terms/created">http://purl.org/dc/terms/created</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Created</td>
</tr>
<tr>
<td>Definition</td>
<td>Date of creation of the resource.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
</tbody>
</table>

**Term Name: dateAccepted**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/terms/dateAccepted">http://purl.org/dc/terms/dateAccepted</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Date Accepted</td>
</tr>
<tr>
<td>Definition</td>
<td>Date of acceptance of the resource (e.g. of thesis by university department, of article by journal, etc.).</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
</tbody>
</table>

**Term Name: dateCopyrighted**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/terms/dateCopyrighted">http://purl.org/dc/terms/dateCopyrighted</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td></td>
</tr>
<tr>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>Label</td>
<td>Date Copyrighted</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>Definition</td>
<td>Date of a statement of copyright.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
</tbody>
</table>

**Term Name: dateSubmitted**

<table>
<thead>
<tr>
<th>Label</th>
<th>Date Submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>Date of submission of the resource (e.g. thesis, articles, etc.).</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
</tbody>
</table>

**Term Name: educationLevel**

<table>
<thead>
<tr>
<th>Label</th>
<th>Audience Education Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>A general statement describing the education or training context. Alternatively, a more specific statement of the location of the audience in terms of its progression through an education or training context.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/terms/educationLevel">http://purl.org/dc/terms/educationLevel</a></td>
</tr>
</tbody>
</table>

**Term Name: extent**

<p>| Label | <a href="http://purl.org/dc/terms/extent">http://purl.org/dc/terms/extent</a>                       |</p>
<table>
<thead>
<tr>
<th>Label</th>
<th>Extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>The size or duration of the resource.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/format">http://purl.org/dc/elements/1.1/format</a></td>
</tr>
</tbody>
</table>

**Term Name: hasFormat**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/terms/hasFormat">http://purl.org/dc/terms/hasFormat</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Has Format</td>
</tr>
<tr>
<td>Definition</td>
<td>The described resource pre-existed the referenced resource, which is essentially the same intellectual content presented in another format.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Date Issued</td>
<td>2000-07-11</td>
</tr>
</tbody>
</table>

**Term Name: hasPart**

<table>
<thead>
<tr>
<th>URI</th>
<th><a href="http://purl.org/dc/terms/hasPart">http://purl.org/dc/terms/hasPart</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Has Part</td>
</tr>
<tr>
<td>Definition</td>
<td>The described resource includes the referenced resource either physically or logically.</td>
</tr>
<tr>
<td>Type of Term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
</tbody>
</table>

**Term Name: hasVersion**
<table>
<thead>
<tr>
<th>Dublin Core Element Description Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><strong>Label:</strong></td>
</tr>
<tr>
<td><strong>Definition:</strong></td>
</tr>
<tr>
<td><strong>Type of Term:</strong></td>
</tr>
<tr>
<td><strong>Refines:</strong></td>
</tr>
</tbody>
</table>

**Term Name: isFormatOf**

| **URI:** | http://purl.org/dc/terms/isFormatOf |
| **Label:** | Is Format Of |
| **Definition:** | The described resource is the same intellectual content of the referenced resource, but presented in another format. |
| **Type of Term:** | element-refinement |
| **Refines:** | http://purl.org/dc/elements/1.1/relation |

**Term Name: isPartOf**

| **URI:** | http://purl.org/dc/terms/isPartOf |
| **Label:** | Is Part Of |
| **Definition:** | The described resource is a physical or logical part of the referenced resource. |
| **Type of Term:** | element-refinement |
| **Refines:** | http://purl.org/dc/elements/1.1/relation |

**Term Name: isReferencedBy**
<table>
<thead>
<tr>
<th>URI:</th>
<th><a href="http://purl.org/dc/terms/isReferencedBy">http://purl.org/dc/terms/isReferencedBy</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
<td>Is Referenced By</td>
</tr>
<tr>
<td>Definition:</td>
<td>The described resource is referenced, cited, or otherwise pointed to by the referenced resource.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
</tbody>
</table>

**Term Name: isReplacedBy**

<table>
<thead>
<tr>
<th>URI:</th>
<th><a href="http://purl.org/dc/terms/isReplacedBy">http://purl.org/dc/terms/isReplacedBy</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
<td>Is Replaced By</td>
</tr>
<tr>
<td>Definition:</td>
<td>The described resource is supplanted, displaced, or superseded by the referenced resource.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
</tbody>
</table>

**Term Name: isRequiredBy**

<table>
<thead>
<tr>
<th>URI:</th>
<th><a href="http://purl.org/dc/terms/isRequiredBy">http://purl.org/dc/terms/isRequiredBy</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label:</td>
<td>Is Required By</td>
</tr>
<tr>
<td>Definition:</td>
<td>The described resource is required by the referenced resource, either physically or logically.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
<tr>
<td>Term Name: issued</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>URI:</td>
<td><a href="http://purl.org/dc/terms/issued">http://purl.org/dc/terms/issued</a></td>
</tr>
<tr>
<td>Label:</td>
<td>Issued</td>
</tr>
<tr>
<td>Definition:</td>
<td>Date of formal issuance (e.g., publication) of the resource.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/date">http://purl.org/dc/elements/1.1/date</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: isVersionOf</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>Refines:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: license</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>Type of element-refinement:</td>
</tr>
<tr>
<td>Term:</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Term Name: medium</td>
</tr>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>Refines:</td>
</tr>
</tbody>
</table>

| Term Name: modified |
| URI: | http://purl.org/dc/terms/modified |
| Label: | Modified |
| Definition: | Date on which the resource was changed. |
| Type of Term: | element-refinement |
| Refines: | http://purl.org/dc/elements/1.1/date |

<p>| Term Name: references |
| URI: | <a href="http://purl.org/dc/terms/references">http://purl.org/dc/terms/references</a> |
| Label: | References |
| Definition: | The described resource references, cites, or otherwise points to the referenced resource. |</p>
<table>
<thead>
<tr>
<th>Type of Term:</th>
<th>element-refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
<tr>
<td>Term Name:</td>
<td>replaces</td>
</tr>
<tr>
<td>URI:</td>
<td><a href="http://purl.org/dc/terms/replaces">http://purl.org/dc/terms/replaces</a></td>
</tr>
<tr>
<td>Label:</td>
<td>Replaces</td>
</tr>
<tr>
<td>Definition:</td>
<td>The described resource supplants, displaces, or supersedes the referenced resource.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Term:</th>
<th>element-refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
<tr>
<td>Term Name:</td>
<td>requires</td>
</tr>
<tr>
<td>URI:</td>
<td><a href="http://purl.org/dc/terms/requires">http://purl.org/dc/terms/requires</a></td>
</tr>
<tr>
<td>Label:</td>
<td>Requires</td>
</tr>
<tr>
<td>Definition:</td>
<td>The described resource requires the referenced resource to support its function, delivery, or coherence of content.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Term:</th>
<th>element-refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/relation">http://purl.org/dc/elements/1.1/relation</a></td>
</tr>
<tr>
<td>Term Name:</td>
<td>spatial</td>
</tr>
<tr>
<td>URI:</td>
<td><a href="http://purl.org/dc/terms/spatial">http://purl.org/dc/terms/spatial</a></td>
</tr>
<tr>
<td>Label:</td>
<td>Spatial</td>
</tr>
<tr>
<td>Definition:</td>
<td>Spatial characteristics of the intellectual content of the resource.</td>
</tr>
<tr>
<td>Term Name: tableOfContents</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td>URI:</td>
<td><a href="http://purl.org/dc/terms/tableOfContents">http://purl.org/dc/terms/tableOfContents</a></td>
</tr>
<tr>
<td>Label:</td>
<td>Table Of Contents</td>
</tr>
<tr>
<td>Definition:</td>
<td>A list of subunits of the content of the resource.</td>
</tr>
<tr>
<td>Type of Term:</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines:</td>
<td><a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: temporal</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>Refines:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term Name: valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Label:</td>
</tr>
<tr>
<td>Definition:</td>
</tr>
<tr>
<td>Type of Term:</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Refines:</td>
</tr>
</tbody>
</table>
## Metadata Map

<table>
<thead>
<tr>
<th>Source Metadata (native publisher metadata)</th>
<th>Transformation Rules</th>
<th>Target Metadata (simple DC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>journal_issue/record/title</td>
<td>Take source element as is. If multiple title elements in source, take all; each source title gets a separate dc:title element.</td>
<td>title</td>
</tr>
<tr>
<td>Source Metadata (native publisher metadata)</td>
<td>Transformation Rules</td>
<td>Target Metadata (simple DC)</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Exercise 8: Metadata workflow

Tools:

Two metadata workflow scenarios; Scenario A & B Deliverables; five MARC source metadata samples; target metadata template; Characteristics of Workflow’s Source and Target Metadata; Metadata Map.

Preliminary instructions:

Participants break into small groups and facilitator assigns scenarios to groups. Group members read their scenario and the list of deliverables. Facilitator and participants review the deliverables and facilitator answers questions.

After groups complete the deliverables, each group identifies a spokesperson who reports on group results. Participants as a whole discuss the impact of scenario constraints and requirements on metadata decisions.

Allow 40 minutes for group work and 30 minutes for reporting and discussion.
Metadata Workflow—Scenario A—Big Dreams for Library Publishing

The library is considering whether to participate in a publishing operation run by a successful commercial internet company, Doodle. Doodle offers full-text access to online versions of public domain books to participating academic libraries and their communities. The number of titles in the system is now approaching twenty million.

Access to all titles in Doodle’s system is free to all participating libraries. Doodle also lets the general public search and browse its metadata holdings and purchase the full-text of books or parts of books if desired. A portion of the revenue from these book sales are distributed back to the participating libraries. The pay-out distribution is based on the volume of books sold contributed by each library. In other words, if your library contributes books that sell many copies, then your library will receive a corresponding greater amount of the distributed revenue. For some libraries, this has been quite lucrative, allowing for the digital reformatting of hundreds of books a year and the hiring of additional staff to manage and carry out this work.

As part of the library’s obligation, they would need to provide Doodle with electronic files of the scanned books, together with metadata for these titles in a specified format. The titles must be in the library’s permanent collection.

The library has established a digital library team to come up with a proposal for working with Doodle. On this team is a metadata specialist, and he has been asked to come up with a plan for establishing an efficient workflow to generate the metadata files required by Doodle.

The library has assigned a group of selectors to decide on which books to contribute. This group will work with a publishing market consultant from Doodle to establish criteria. At this point, they plan to identify and provide (including metadata) up to 50 titles every quarter. If all goes well, they anticipate doubling this after the first year.

Doodle has shared one important piece of information with the library. Their statistics have clearly demonstrated a direct correlation between the amount of metadata provided and number of book sales. In other words, books with relatively more metadata are not only more likely to be purchased, but to be purchased multiple times. Since the library administration would like to see this venture succeed ($$$), they have encouraged the metadata specialist to take this into consideration. They have also said, recognizing this relationship between metadata richness and sales, that they are willing to find some extra staff time for metadata work during the next two years (this time would come from technical services, and they’ve asked the metadata specialist what he needs). After that, metadata staff associated with this project would need to be funded from project revenues.

The metadata specialist begins his investigations and has thus far learned…

- that any updates to records already shipped to Doodle are to be handled by resubmitting the monograph again. Every monograph has an ID value, and Doodle will completely replace any monograph in its system, if a new submission has the same ID.
that the head of technical services is being very uncooperative. He has agreed to update catalog records to reflect that an electronic version of a monograph is available. But he absolutely refuses to allow additional data to be inserted into MARC records. His reasoning has to do with record consistency across all holdings. The metadata specialist knows the library administration will never go against his wishes.

The deliverables for Exercise 8 are described after Scenario B.
The library is considering whether to participate in a publishing operation, Books-R-Us, coordinated by a university consortium. The books accessible through its system are full-text electronic versions of public domain books, which have been selected and contributed by participating libraries but are actually hosted locally by the contributing library. Books-R-Us merely provides a portal to the entire collection of scanned books by merging book records into a searchable union catalog and allowing it to be searched.

Access to the online versions of these books is free to all participating library members. As of yet, the number of participants is fairly small, but the hope is that in time, this consortium project would allow participating libraries access to much larger book collections than possible on their own.

The consortium collects fees from member libraries. It then uses this money to finance the scanning of books in a centralized location. Once scanned, the files go back to the contributing library. The library’s obligation is to host the resulting electronic files on local servers, and also to provide Books-R-Us with metadata for these titles in a specified format. Books-R-Us creates its union catalog from these records.

The library has asked its metadata specialist to come up with a plan for establishing an efficient workflow to generate the metadata files required by Books-R-Us.

The library has assigned a group of selectors to decide on which books to contribute. This group expects to identify about 200 titles initially. The plan is to convert 25 of these each quarter. After two years, the project will be re-evaluated.

The library has said they have programming staff that they can dedicate to this project, in order to see it successfully setup. Once established, however, there are no plans for ongoing technical support. Further, it does not look like the library will be able to devote other ongoing staff resources to this project.

The metadata specialist begins his investigations and has thus far learned…

- that Books-R-Us expects to harvest complete metadata shipments monthly from participating libraries. This is how record updates and additions are handled. Books-R-Us completely rebuilds its entire aggregated metadata collection every month.
- that Books-R-Us insists that the metadata they receive be kept in sync with the library catalog. They do not themselves provide direct access to online books, but rather link back to a contributing library’s catalog record, which in turn will link to the electronic files. In this sense, Books-R-Us only serves as a union catalog of all online books available to participating libraries.
- that Books-R-Us is using a book metadata format more typical of the commercial publishing industry, and that they are encouraging participating libraries to provide them with as much metadata per title as possible.
Exercise 8: Scenario A & B Deliverables

You are the metadata specialist on a larger team that will carry out this project. The workflow you are designing now (for this exercise) is concerned only with the metadata portion of the project. There will likely be places where your workflow will need to “interface” with other activities of the project, perhaps another workflow, but don’t get sidetracked by developing non-metadata aspects of the project into your workflow (such as selection, or scanning, etc.).

1. Workflow Definition and Goals, Input/Output Analysis
   a) In a sentence or two, define the overall metadata workflow objective. Remember, this workflow is only concerned with the metadata portion of the project.
   b) Using the worksheet provided, briefly describe the characteristics of the workflow source metadata (samples of source metadata follow).
   c) Using the worksheet provided, briefly describe the characteristics of the workflow target metadata (samples of target metadata follow).

   Work through 1 (b) and (c) quickly. The characteristics to pay attention to are those that will impact workflow, such as, “status,” especially.

   The source and target metadata examined here are at the overall project level. As you define the tasks below, there may be “transitional” metadata, and thus transitional source and target metadata requiring their own mappings. In other words, it may not be feasible or efficient to convert, in one step, the project’s source metadata to the project’s target metadata.

2. Identifying constraints
   a) List the constraints that you face in setting up and maintaining this metadata workflow. (See slide 154.)

3. Defining the workflow tasks
   a) Start with the overall workflow objective and begin to break it down into smaller and smaller tasks and subtasks. What you should end up with is a list of discrete and manageable tasks, ones that could feasibly be carried out within an actual workflow operation.
   b) For each of the tasks above, answer the following:
      o What are the task’s requirements? Specify what is required to begin this task (what input must the task have? what is required of that input?). Specify what is required of the task output (what requirements must it fulfill?).
      o What is the level of complexity required to transform input to output?
      o What are the task dependencies? What is the task dependent upon in order to successfully transform input to output? (Your understanding of workflow constraints should help here.)
      o What is the projected duration of this task? Is it a one-time task (writing software to do something), or is it a recurring, ongoing activity? How certain is the projected...
duration of the task? Do workflow constraints or task dependencies make duration difficult to predict?
  o What are the resource requirements of this task? What or who needs to be involved in accomplishing this task? If people, how many and what level of expertise and experience is required?

4. Designing the workflow
  a) How should all the tasks defined in step 3 above be sequenced? Which tasks can occur simultaneously and which are dependent on a previous task?
  b) What are the communication needs of the workflow?

5. Maintaining the workflow
  a) Is this workflow a one-time data conversion project, or will it be an ongoing, regular part of library operations?
  b) If ongoing, what type of tracking and oversight is required to ensure the workflow is successfully meeting its objectives?
  c) How much human oversight will the workflow require and what will it involve?
  d) How much automated tracking is possible, and how would that tracking process work?

6. Workflow cost considerations
  a) Make an estimate of how many FTEs over what period of time would be required to setup this metadata workflow. (FTE is “full-time equivalent”—1.0 FTE equals one person working full-time on this project.)
  b) What level of staff expertise and experience will be required to setup this workflow.
  c) Make an estimate of how many FTEs would be required to maintain this metadata workflow going forward (after setup and initial operation), if that is required.
  d) What level of staff expertise and experience will be required to maintain this workflow.
  e) Do these staffing requirements match with workflow constraints? If not, how are you planning to deal with the mismatch?

7. Opportunities and benefits
  a) List all the benefits you can think of that may result from setting up and maintaining this workflow.

8. Metadata workflow conversion maps
  a) It is likely that at least one of the tasks in step 3 above involved some metadata mapping. Choose one of the mapping tasks and, using the metadata map template provided, develop the source-to-target mapping rules.
Report on the manuscripts of Allan George Finch, esq., of...

000 01867cam a2200349 450
001 3645091
005 20060504103643.0
008 750522m19139999enk f000 0 eng c
010 __ |a ac 35001225 //r
035 __ |a (NIC)notisASL2364
035 __ |a (OCoLC)ocm01350566
040 __ |a New York. Public Libr. |c TOL |d SER |d OCL |d OUN |d OCL |d NIC
050 0_ |a DA25.M2 |b F4
082 __ |a 942.06

110 1_ |a Great Britain. |b Royal Commission on Historical Manuscripts.

245 00 |a Report on the manuscripts of Allan George Finch, esq., of Burley-on-the-Hill, Rutland...

260 __ |a London, |b Published by H.M. Stationery Off., |c 1913-

300 __ |a Vols. 1-2 issued in the Parliamentary series as Cd. 6508, 8383; v. 3 issued as no. 71 of the commission’s Publications.

500 __ |a At head of title: Historical manuscripts commission.

500 __ |a Title varies slightly.


500 __ |a Vols. 1-2 deal with 16th and early 17th century letters of the Finch family; the correspondence of Heneage, earl of Winchilsea, during his embassy to Constantinople, 1660-1668; letters and papers of his cousin, Sir John Finch, who followed him as ambassador to Turkey; letters and papers of Heneage, earl of Nottingham, and his family; and the correspondence of Daniel, earl of Winchilsea and Nottingham, secretary of state from 1688-1693. cf. v. 1, p. [v]
651 _0 |a Great Britain |x History |y 1485- |v Sources.
651 _0 |a Turkey |x History |y 1453-1683 |v Sources.
651 _0 |a Great Britain |x Foreign relations |z Turkey.
651 _0 |a Turkey |x Foreign relations |z Great Britain.
700 1_ |a Finch, Allan George, |d 1863-1914.
905 __ |a 19991204120000.0
950 __ |l OLIO1 |x 175 |a DA25.M2 |b F49 |d \\+
955 __ |l OLIO1 |a DA25.M2 |b F49 |c 1:v.1-4
999 __ |l OLIO1 |a DA25.M2 |b F49 |d \\+ |c 1 |v v.1-4
Projektive geometrie der ebene, unter benutzung der punktrechnung...
The war of the rebellion: a compilation of the official records of the Union and Confederate armies / Pub. under the direction of the secretary of war ...
Robert N. Scott compiled and edited v. 1-18, 1880-87, and also collected the greater part of the material for v. 19-36, 1887-91. After his death in 1887 the work was continued by Henry M. Lazelle, 1887-89, and by a board of publication, 1889-99, consisting of George B. Davis, 1889-97, Leslie J. Perry, 1889-99, Joseph W. Kirkley, 1889-99, and Fred C. Ainsworth, 1898-99; from 1899-1901 edited by Fred C. Ainsworth and Joseph W. Kirkley.

505 0_ |a ser. I. v. 1-53 [serial no. 1-111] Formal reports, both Union and Confederate, of the first seizures of United States property in the southern states, and of all military operations in the field, with the correspondence, orders and returns relating specially thereto. 1880-98. 111 v.--ser. II. v. 1-8 [serial no. 114-121] Correspondence, orders, reports and returns, Union and Confederate, relating to prisoners of war ... and to state or political prisoners. 1894 [i. e. 1898]-1899. 8 v.--ser. III. v. 1-5 [serial no. 122-126] Correspondence, orders, reports and return of the Union authorities (embracing their correspondence with the Confederate officials) not relating specially to the subjects of the first and second series. It embraces the reports of the secretary of war, of the general-in-chief and of the chiefs of the several staff corps and departments ... 1899-1900. 5 v.--ser. IV. v. 1-3 [serial no. 127-129] Correspondence, orders, reports and returns of the Confederate authorites, similar to that indicated for the Union officials, as of the third series, but includeing the correspondence between the Union and Confederate authorities, given in that series. 1900. 3 v.--[serial no. 130] General index and additions and corrections. Mr. John S. Moodey, indexer. Preface [by Elihu Root, secretary of war] Explanations. Synopsis of the contents of volumes. Special index for the principal armies, army corps, military divisions

505 8_ |a and departments. General index. Additions and corrections ... 1901.

651 _0 |a United States |x History |y Civil War, 1861-1865 |x Maps.

651 _0 |a United States |x History |y Civil War, 1861-1865 |x Sources.

651 _0 |a United States |x History |y Civil War, 1861-1865 |x Regimental histories.
610 20 |a Confederate States of America |x History |x Sources.

710 1_ |a United States. |b Record and Pension Office.

710 1_ |a United States. |b War Records Office.

700 1_ |a Moodey, John S. |q (John Sheldon), |d b. 1842.

710 1_ |a Cowles, Calvin D. |q (Calvin Duvall), |d b. 1849.

700 1_ |a Ainsworth, Fred C. |q (Fred Crayton), |d 1852-1934.

700 1_ |a Scott, Robert N. |q (Robert Nicholson), |d 1838-1887.

700 1_ |a Davis, George B. |q (George Breckenridge), |d 1847-1914.

700 1_ |a Perry, Leslie J.

700 1_ |a Kirkley, Joseph W. |q (Joseph William), |d 1841-1912.

740 0_ |a Official records of the Union and Confederate armies.

773 0_ |7 nnbc |t Burt Green Wilder papers. |w (CStRLIN) NYCV86-A116.

948 1_ |a 20031211 |b c |d lbb4 |e rmc |f ? |h ?

948 2_ |a 20040811 |b m |d jm17 |e cts

948 2_ |a 20061002 |b m |d bmt1 |e cts
History of the city of New York, from its earliest settlement to the present time.

By Mary L. Booth. Illus. with over one hundred engravings.


xix, <21>-846 p. incl. illus., plates, ports. front. 24 cm.
Celestine, being the diary of a chambermaid. By Octave Mirbeau. Translated...
<?xml version="1.0"?>
<monograph id="" lang="">
  <monograph_data>
    <identifiers>
      <identifier type="doi"/>
    </identifiers>
    <author>
      <name>
        <given_name>
        <surname>
      </name>
      <affiliation>
        <prof_title>
        <organization>
      </affiliation>
      <bio>
        <p/>
      </bio>
      <image type="" id=""/>
    </author>
    <contributor>
      <role/>
      <name>
        <given_name>
        <surname>
      </name>
      <affiliation>
        <prof_title>
        <organization>
      </affiliation>
      <bio>
        <p/>
      </bio>
      <image type="" id=""/>
    </contributor>
    <title lang=""></title>
    <edition/>
    <publisher/>
    <publisher_location>
    <publication_date iso8601-date=""></publication_date>
    <abstract lang="">
      <p/>
    </abstract>
    <subjects>
      <subject scheme=""/>
    </subjects>
  </monograph_data>
</monograph>
All elements in the target format are optional, except:
  monograph
  monograph_data
  title [required when parent element is used]
  section [at least one required when parent element is used]

Repeatable elements include:
  identifier
  author
  contributor
  affiliation
  abstract
  p (paragraph)
  subject
  section
# Characteristics of Workflow’s Source and Target Metadata

## Characteristics of Source Metadata

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>File format</td>
<td></td>
</tr>
<tr>
<td>Type of metadata</td>
<td></td>
</tr>
<tr>
<td>Metadata scheme</td>
<td></td>
</tr>
<tr>
<td>Scheme Support and Documentation</td>
<td></td>
</tr>
<tr>
<td>Content values</td>
<td></td>
</tr>
<tr>
<td>Structure</td>
<td></td>
</tr>
<tr>
<td>Intended use</td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td></td>
</tr>
</tbody>
</table>

## Characteristics of Target Metadata

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>File format</td>
<td></td>
</tr>
<tr>
<td>Type of metadata</td>
<td></td>
</tr>
<tr>
<td>Metadata scheme</td>
<td></td>
</tr>
<tr>
<td>Scheme Support and Documentation</td>
<td></td>
</tr>
<tr>
<td>Content values</td>
<td></td>
</tr>
<tr>
<td>Structure</td>
<td></td>
</tr>
<tr>
<td>Intended use</td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td></td>
</tr>
</tbody>
</table>
## Metadata Map

<table>
<thead>
<tr>
<th>Source</th>
<th>Transformation</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Metadata Map

<table>
<thead>
<tr>
<th>Source</th>
<th>Transformation</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Exercise 9: Digital library development project—the slide library

Tools:

The Slide Collection Digital Library Project—Background; four Slide Collection Digital Library Project scenarios; The Slide Collection Digital Library Project—Deliverables; Filemaker Pro native database records; sample VRA Core record; Characteristics of Source Metadata; Characteristics of Target Metadata; Metadata Map.

Preliminary instructions:

Participants break into small groups and facilitator assigns scenarios to groups. Group members read the project background, their scenario, and the list of deliverables. Facilitator and participants review the project and deliverables and facilitator answers questions.

After groups complete the deliverables, each group identifies a spokesperson who reports on group results. Participants as a whole discuss the impact of scenario constraints and requirements on metadata decisions.

Allow 45 minutes for group work and 45 minutes for reporting and discussion.
Project Background—All Scenarios

The Slide Collection Digital Library Project—Background

University College has about 12,000 slides in its Art & Architecture Library slide collection, and the A&A Librarian, Mark Michelson, would like to make these accessible via the web. Most of the A&A faculty use slides in their classes and many are mentioning some system that they’ve heard of that allows faculty to select slides (while at home!) from a database and then show them in class, somehow. Mark isn’t too sure how this would work, but he is sure there must be some software system that does all this.

One professor of Architecture in particular, Bram Bristle, is quite vocal about this and he has written to Mark and the Library Director. In an effort to appease him, the Library Director has asked the Digital Library Group (DLG) to meet with the A&A Librarian and see if there’s something that can be done. Sarah Scanner, the head of DLG, meets with Mark and Bram and several other A&A faculty one afternoon. Here’s what she learns:

- Currently, the Art Library has a simple Filemaker Pro database of all the slides in the collection. This was created several years ago, from a card file that was begun in the 1950s. All new slides added since the creation of the database are entered directly into the Filemaker Pro database. The database grows by about 100 entries per year.

- Each database record has about 20 elements. Records are not keyed to each other, although many, especially of the same building or place, share the same terms. The A&A Librarian seems very knowledgeable about each of the record elements, regarding what information they contain and how that information should be recorded. Mark admits, however, that there are unfortunate problems with some of the data, which he attributed to the lack of data standards in the early days of the slide collection catalog (before his time!), or to students, who occasionally enter data for him. One point that Mark kept insisting on was the accurate use, throughout the entire set of records, of the classification code. The code seemed impenetrable to Sarah, but Mark said that although the code was arcane, it was packed with information and uniformly applied throughout, since he personally applied it. When Sarah asked Mark if anyone else used the code and how, Mark said no, it was used to classify every image.

- What Sarah hears that the faculty want most are these functions:
  - The ability to access the image database over the web from their offices or homes.
  - The ability to search on the title of a work and bring up all images of that work.
  - The ability to search by period or date and bring up all works of that period or date.
  - The ability to search by type of work, like “painting,” or “temple,” and bring up all images pertaining to that type.
  - The ability to see small thumbnails in search results and browse mode.
  - The ability to select images and view all the metadata associated with that image.
  - The ability to select and save images to some sort of work list, so that they could be easily recovered during a classroom presentation.
Sarah assembles a team to carry out this project. The Library Director agrees that Kat Krammer, the library’s metadata specialist, can be on the team. Sarah also enlists Paul Plotter, a programmer.

On the following pages are four scenarios for how this project continues. Read the one that is assigned to your group.

Then read the list of deliverables assigned to the metadata specialist.
The head of the Digital Library Group, Sarah Scanner, does not have an existing system that will meet the needs of this project. From peers at other universities, however, she’s aware of ArtBox, a digital library system that is increasingly seeking the art library market. She thinks she can convince the Library Director to license ArtBox.

The current version of ArtBox only accepts simple Dublin Core (DC) records. The vendor apparently believes this is a drawback and has promised that the software will accept VRA Core (Visual Resources Association Core Categories) records within the next two years.

Sarah asks Kat to evaluate the Filemaker Pro metadata as well as DC and VRA Core, about which Sarah knows very little other than that these seem to be accepted standards. Kat spends some time looking at the native database records and at VRA Core. As an exercise, she creates a potential VRA Core record from a typical record in the native A&A database (see sample VRA Core record). She also begins to understand VRA Core’s distinction between “work” and “image,” and thinks this distinction may be useful to the project.

Sarah asks her programmer, Paul Plotter, to make an initial evaluation of ArtBox features, to see if the system has the functionality desired by the A&A faculty. Paul reports back that ArtBox supports web access worldwide (with proper authentication); the ability to associate thumbnails with every image, which are then visible in search results and browsing; the ability for authenticated users to create profiles and save image lists there for later use; and the ability to see all metadata (currently only simple Dublin Core) associated with an image. Paul says that ArtBox can be easily configured to search on any metadata field in the Dublin Core records and return all matching records. As to whether a search on a work will return all images of that work, Paul says “sure, no problem, as long as all the images have the same title.” He says the same thing about searching on date and type of work.

Prompted by a question from Kat about metadata workflow and data ingest, the team takes a look at the ArtBox cataloging client. This client provides a way to create and modify individual database entries directly into ArtBox. Since ArtBox currently uses simple Dublin Core records, that’s what the cataloging client guides the data imputer to create. Sarah and Kat show this client to the A&A Librarian, and he is troubled by it. The data fields are not nearly as sophisticated as those in his database, he feels, and he doesn’t like the idea of using it. Besides this method of data ingest, ArtBox can also accept any number of simple DC records via a batch process.
The head of the Digital Library Group, Sarah Scanner, does not have an existing system that will meet the needs of this project. From peers at other universities, however, she’s aware of ArtBox, a digital library system that is increasingly seeking the art library market. She thinks she can convince the Library Director to license ArtBox.

The current version of ArtBox only accepts simple Dublin Core (DC) records. The vendor apparently believes this is a drawback and has promised that the software will accept VRA Core (Visual Resources Association Core Categories) records within the next two years.

Sarah asks Kat to evaluate the Filemaker Pro metadata as well as DC and VRA Core, about which Sarah knows very little other than that these seem to be accepted standards. Kat spends some time looking at the native database records and at VRA Core. As an exercise, she creates a potential VRA Core record from a typical record in the native A&A database (see sample VRA Core record). She also begins to understand VRA Core’s distinction between “work” and “image,” and thinks this distinction may be useful to the project.

Sarah asks her programmer, Paul Plotter, to make an initial evaluation of ArtBox features, to see if the system has the functionality desired by the A&A faculty. Paul reports back that ArtBox supports web access worldwide (with proper authentication); the ability to associate thumbnails with every image, which are then visible in search results and browsing; the ability for authenticated users to create profiles and save image lists there for later use; and the ability to see all metadata (currently only simple Dublin Core) associated with an image. Paul says that ArtBox can be easily configured to search on any metadata field in the Dublin Core records and return all matching records. As to whether a search on a work will return all images of that work, Paul says “sure, no problem, as long as all the images have the same title.” He says the same thing about searching on date and type of work.

Prompted by a question from Kat about metadata workflow and data ingest, the team discovers that, remarkably, ArtBox has no cataloging client. They question the vendor on this and are told that work on a client is underway. But after more questioning, they discover that development of the cataloging client is tied to the move to VRA Core. In other words, the client won’t be available before the VRA Core compliant release of ArtBox. The software does have, of course, the ability to ingest any number of DC records via a batch process.
The head of the Digital Library Group, Sarah Scanner, does not have an existing system that will meet the needs of this project. Further, the Library Director tells her that they do not have much money to devote to the project, especially on an ongoing basis. On the other hand, the Director feels strongly that they must make an effort to satisfy Bram Bristle.

After talking to her programmer, Paul Plotter, Sarah decides that it is feasible to create a web interface into the existing Filemaker Pro database. This would allow the Filemaker Pro database to continue to be used. Paul is convinced that he can add all the desired functionality through such a system, such as web access worldwide (with proper authentication); the ability to associate thumbnails with every image, which are then visible in search results and browsing; the ability for authenticated users to create profiles and save image lists there for later use; and the ability to see all metadata associated with an image. As far as searching, Paul says he can search on any of the data in the database and deliver search results. As to whether a search on a work will return all images of that work, Paul says “sure, no problem, as long as all the images have the same title.” He says the same thing about searching on date and type of work.

In the midst of their initial explorations, Bram goes to the Library Director with a new idea he has just heard about. It involves participating in a larger network of university art and architecture department slide collections. After the meeting, the Library Director writes an email to Sarah telling her about this and asking her to investigate what would be involved. Sarah finds the project’s web site and discovers what the requirements for participation are: participating institutions need to make metadata records for their collections available for harvesting via OAI (Open Archives Initiative) in both the simple Dublin Core and VRA Core (Visual Resources Association Core Categories) standards.

Sarah asks Kat to evaluate the Filemaker Pro metadata as well as DC and VRA Core, about which Sarah knows very little other than that these seem to be accepted standards. Kat spends some time looking at the native database records and at VRA Core. As an exercise, she creates a potential VRA Core record from a typical record in the native A&A database (see sample VRA Core record). She also begins to understand VRA Core’s distinction between “work” and “image,” and thinks this distinction may be useful to the project.

Prompted by a question from Kat about metadata workflow and data input, the team discusses this topic. Paul is convinced that in a month or less he can move the Filemaker Pro data to another database, one that can store and ingest VRA Core records. Sarah decides that she will clear Paul’s schedule so that he can concentrate exclusively on this over the next three months. Paul also plans to build an OAI component to export the DC and VRA Core records, but that should be easy, he says.
The head of the Digital Library Group, Sarah Scanner, does not have an existing system that will meet the needs of this project. Further, the Library Director tells her that they do not have much money to devote to the project, especially on an ongoing basis. On the other hand, the Director feels strongly that they must make an effort to satisfy Bram Bristle.

After talking to her programmer, Paul Plotter, Sarah decides that it is feasible to create a web interface into the existing Filemaker Pro database. This would allow the Filemaker Pro database to continue to be used. Paul is convinced that he can add all the desired functionality through such a system, such as web access worldwide (with proper authentication); the ability to associate thumbnails with every image, which are then visible in search results and browsing; the ability for authenticated users to create profiles and save image lists there for later use; and the ability to see all metadata associated with an image. As far as searching, Paul says he can search on any of the data in the database and deliver search results. As to whether a search on a work will return all images of that work, Paul says “sure, no problem, as long as all the images have the same title.” He says the same thing about searching on date and type of work.

In the midst of their initial explorations, Bram goes to the Library Director with a new idea he has just heard about. It involves participating in a larger network of university art and architecture department slide collections. After the meeting, the Library Director writes an email to Sarah telling her about this and asking her to investigate what would be involved. Sarah finds the project’s web site and discovers what the requirements for participation are: participating institutions need to make metadata records for their collections available for harvesting via OAI (Open Archives Initiative) in both the simple Dublin Core and VRA Core (Visual Resources Association Core Categories) standards.

Sarah asks Kat to evaluate the Filemaker Pro metadata as well as DC and VRA Core, about which Sarah knows very little other than that these seem to be accepted standards. Kat spends some time looking at the native database records and at VRA Core. As an exercise, she creates a potential VRA Core record from a typical record in the native A&A database (see sample VRA Core record). She also begins to understand VRA Core’s distinction between “work” and “image,” and thinks this distinction may be useful to the project.

Prompted by a question from Kat about metadata workflow and data input, the team discusses this topic. Paul thinks that he can move the Filemaker Pro data to another database, one that can store VRA Core records. But Sarah realizes that she won’t be able to devote Paul to this work for another two years, given all the other projects needing attention. So for now, it seems they must accept the Filemaker Pro database for data entry. Sarah does see that she will need to allow Paul to build an OAI component to export the DC and VRA Core records.
The Slide Collection Digital Library Project—Deliverables

You are the metadata specialist (Kat Krammer) on this project team, and you are being asked to contribute your expertise and understanding of metadata and metadata processing work to help the project leader develop appropriate and feasible plans for accomplishing the overall project objectives. Remember to keep your focus on aspects of the project related to metadata (and there are plenty of them), as opposed to overall project management. Specifically, the metadata specialist is asked to take responsibility for the following deliverables:

1. System functional requirements and metadata
   
a) Work through the desired functional requirements described in the Project Background and list those requirements that will depend on descriptive metadata. For each…
   
i) List the metadata element or elements involved in fulfilling this functional requirement.
   
ii) Describe what demands the desired functionality will make on these metadata elements.
   
iii) Are there different strategies for meeting this functional requirement, especially in terms of metadata?

b) List any decisions you are aware of that the project team must make regarding functionality and metadata. You can add to this list as you work through the rest of the exercise.

2. Metadata conversion/mapping
   
a) Make a list of all the metadata conversion processes that will be required for this project. For each conversion process on the list…
   
i) Is this a one-time conversion of metadata, or an ongoing, recurring conversion? If an ongoing conversion, is it for a defined or indefinite duration?
   
ii) Using the source metadata template, briefly describe the characteristics of the source metadata.
   
iii) Using the target metadata template, briefly describe the characteristics of the target metadata.

b) Make a list of all the metadata maps required for the metadata work on this project. For each, use the metadata map templates to describe the transformation rules necessary for the mapping.

3. Metadata workflow design
   
a) Make a list all the metadata conversion workflows, both immediate and future, that this project will require. Does this match 2 (a) above? For each workflow…
   
i) Give a very brief description of the main objective (transformation) of this workflow.
ii) Identify the constraints that will impact carrying out this workflow.

iii) Begin to break down the main objective of this workflow until you have a list of manageable tasks and subtasks that can feasibly be implemented.

iv) Define the sequencing of the tasks listed in (iii) above. Which can be scheduled simultaneously and which require the completion of some prior task?

v) For any ongoing, regularly recurring conversion workflows, describe what tracking and oversight is required to maintain it. What sorts of automated tracking is feasible? How much and what type of human oversight is needed?

vi) What level of staff expertise and experience will be required to setup and/or maintain this workflow? Are these staffing needs in line with the constraints identified in (ii) above?
Filemaker Pro native database records (samples)  

Exercise 9

ACCESSION_NO: 91000031  
IMAGE_FILE_ID: MDD_02250  
CLASSIFICATION: B-Q5 Mdu 3.4 Meenk 5-2  
PERIOD: Nayak Rulers  
CURRENT LOCATION: Madurai  
REGION: Tamil Nadu  
COUNTRY: India  
TITLE_1: Meenakshi Sundaresvara Temple  
TITLE_2: temple; gate; reservoir  
DATE: ca. 1500-1700  
VIEW: Ext.: East towers and Golden Lily tank from Southwest  
KEYWORDS: reservoirs; gopura; columns  
NOTES: Tank is mentioned in the legend of the siting of Madurai. Dates: 16th to 17th C. Built by Nayak Rulers.  
SOURCE_DONOR: Francis Max Collection  
FILE_TYPE: Color: K5073, 1981  
CITATION_1: Thiagarajan, K.; Meenakshi Temple, Masurai; Madurai: Meenakshi Sundareswarar Temple Renovation Committee; 1965; Fine Arts; NA6008.M28 T42  
CITATION_2: Grove Dictionary of Art  
CITATION_3:

ACCESSION_NO: 91000090  
IMAGE_FILE_ID: MDD_02251  
CLASSIFICATION: B-Q5 Mdu 3.4 Meenk 5-3  
PERIOD: Nayak Rulers  
CURRENT LOCATION: Madurai  
REGION: Tamil Nadu  
COUNTRY: India  
TITLE_1: Meenakshi Sundaresvara Temple  
TITLE_2: temple; gate; reservoir  
DATE: ca. 1500-1700  
VIEW: Ext.: West gopuram  
KEYWORDS: gopura; streets; people  
NOTES: The temple comprises two east-facing shrines dedicated to the goddess Meenakshi and to Sundareshvara. The gopuras at madurai are known for their sweeping concave profiles and profusion of images. Dates: 16th to 17th C. Built by the Nayak Rulers.  
SOURCE_DONOR: Francis Max Collection  
FILE_TYPE: Color: K, 1984  
CITATION_1: Thiagarajan, K.; Meenakshi Temple, Masurai; Madurai: Meenakshi Sundareswarar Temple Renovation Committee; 1965; Fine Arts; NA6008.M28 T42  
CITATION_2: Grove Dictionary of Art  
CITATION_3:

ACCESSION_NO: 91000139  
IMAGE_FILE_ID: MDD_02252  
CLASSIFICATION: B-Q5 Mdu 3.4 Meenk 5-4  
PERIOD: Nayak Rulers  
CURRENT LOCATION: Madurai  
REGION: Tamil Nadu  
COUNTRY: India  
TITLE_1: Sundaresvara Meenakshi Temple  
TITLE_2: temple  
DATE: 16th-18th century
VIEW: Ext.: one end of the Golden Lily tank
KEYWORDS: reservoirs; gopura; columns
NOTES: The temple comprises two east-facing shrines dedicated to the goddess Meenakshi and to Sundareshvara. The gopuras at Madurai are known for their sweeping concave profiles and profusion of images. Dates: 16th to 17th C. Built by the Nayak Rulers.
SOURCE DONOR: Francis Max Collection
FILM TYPE: Color: K5073, 1981
CITATION 1: Balaram Iyer, T. G. S.; History & Description of Sri Meenakshi Temple; Madurai: Sri Karthikeiya Publication; 1976.
CITATION 2: Grove Dictionary of Art
CITATION 3:

ACCESSION NO: 91000141
IMAGE_FILE_ID: MCC_0367
CLASSIFICATION: B-Q5 Mdu 3.4 Meenk 5-1
PERIOD: Nayak Rulers
CURRENT LOCATION: Madurai
REGION: Tamil Nadu
COUNTRY: India
TITLE 1: Meenakshi Temple
TITLE 2:
WORK TYPE: temple
DATE: ca. 17th century
VIEW: Ext.: a prakara with devotees resting on the floor
KEYWORDS:
NOTES:
SOURCE DONOR: Miles Blimmer
FILM TYPE: Color: K5034*
CITATION 1:
CITATION 2:
CITATION 3:

-----

ACCESSION NO: 91000303
IMAGE_FILE_ID: DVB_08723
CLASSIFICATION: B-Q5 Pal 3.4 Shat/AdiB/AdiB 5-1
PERIOD:
CURRENT LOCATION: Palitana
REGION: Gujarat
COUNTRY: India
TITLE 1: Adisvara Bhagavan Temple
TITLE 2:
WORK TYPE: temple
DATE: ca. 925-975
VIEW: Ext.: from Southwest, Adisvara Bhagavan w/corner of Samet Sikhar on right
KEYWORDS: sanctums; shrines
NOTES: Located at end of S summit, shrine is basically of chaumukh type. Dates: mid 10th C., before 961. Built by Javada Sah.
SOURCE DONOR: Francis Max Collection
FILM TYPE: Color: K5032, 1978
CITATION 1: Burgess, James; The Temples of Satrunjaya; Calcutta: Jain Bhawan; 1977
CITATION 2:
CITATION 3:
ACCESSION_NO: 91000683
IMAGE_FILE_ID: DVB_08365
CLASSIFICATION: B-Q5 Sri 6.3 Dal 4
PERIOD:
CURRENT LOCATION: Srinagar
REGION: Jammu and Kashmir
COUNTRY: India
TITLE_1: Dal Lake
TITLE_2:
WORK_TYPE: lake; market
DATE:
VIEW: Close up of vegetable vendor weighing vegetables
KEYWORDS: shikara; markets; lakes
NOTES: The 6.4 km long and 4 km wide natural lake is divided into four man-
made causeways: Gangribal, Lokut Dal, Bod Dal, and Nagin. Small islands on
the lake are willow-covered. The Mihrbari people have traditionally lived
around and on the lake in boats.
SOURCE_DONOR:
FILM_TYPE: Color: E, 1985
CITATION_1:
CITATION_2: Grove Dictionary of Art
CITATION_3: http://srinagar.nic.in/

ACCESSION_NO: 91000646
IMAGE_FILE_ID: DVB_02736
CLASSIFICATION: B-Q5 Sri 1.1 Dal 7
PERIOD:
CURRENT LOCATION: Srinagar
REGION: Jammu and Kashmir
COUNTRY: India
TITLE_1: Dal Lake
TITLE_2:
WORK_TYPE: lake
DATE:
VIEW: Travelling on a shikara through Dal Lake
KEYWORDS: shikara; lakes; flora
NOTES: The 6.4 km long and 4 km wide natural lake is divided into four man-
made causeways: Gangribal, Lokut Dal, Bod Dal, and Nagin. Small islands on
the lake are willow-covered. The Mihrbari people have traditionally lived
around and on the lake in boats.
SOURCE_DONOR:
FILM_TYPE: Color: K, 1985
CITATION_1:
CITATION_2: Grove Dictionary of Art
CITATION_3: http://srinagar.nic.in/

ACCESSION_NO: 91000647
IMAGE_FILE_ID: DVB_00436
CLASSIFICATION: B-Q5 Sri 1.1 Dal 8
PERIOD:
CURRENT LOCATION: Srinagar
REGION: Jammu and Kashmir
COUNTRY: India
TITLE_1: Dal Lake
TITLE_2:
WORK_TYPE: lake
DATE:
VIEW: Tourist shikaras on Dal Lake
KEYWORDS: shikara; houseboats; lakes
NOTES: The 6.4 km long and 4 km wide natural lake is divided into four man-made causeways: Gangribal, Lokut Dal, Bod Dal, and Nagin. Small islands on the lake are willow-covered. The Mihrbari people have traditionally lived around and on the lake in boats.
SOURCE_DONOR: Miles Blimmer
FILM_TYPE: Color: E5074, 1985
CITATION_1: Grove Dictionary of Art
CITATION_2: http://srinagar.nic.in/

ACCESSION_NO: 91000622
IMAGE_FILE_ID: DVB_07564
CLASSIFICATION: B-Q5 Sri 1.1 Dal 4
PERIOD:
CURRENT LOCATION: Srinagar
REGION: Jammu and Kashmir
COUNTRY: India
TITLE_1: Dal Lake
TITLE_2: Weed collecting on Dal Lake
WORK_TYPE: lake; documentary photograph
DATE:
VIEW: Weed collecting on Dal Lake
KEYWORDS: shikara; weeds; gardens; people
NOTES: Locals tend to floating vegetable beds that are shielded with weeds. Natural lake is 6.4 km long and 4 km wide, and is divided into four man-made causeways: Gangribal, Lokut Dal, Bod Dal, and Nagin.
SOURCE_DONOR: Miles Blimmer
FILM_TYPE: Color: K5034, 1985*
CITATION_1: Grove Dictionary of Art
CITATION_2: http://srinagar.nic.in/

ACCESSION_NO: 91000204
IMAGE_FILE_ID: missing
CLASSIFICATION: B-Q5 Had 3.4 Stu 1-1
PERIOD:
CURRENT LOCATION: Hadda
REGION:
COUNTRY: India
TITLE_1: Stupa
TITLE_2: Plan: Tepe shutur
WORK_TYPE: shrine
DATE:
VIEW: Plan: Tepe shutur
KEYWORDS: stupas
NOTES:
SOURCE_DONOR: Miles Blimmer
FILM_TYPE: Color: E5017
CITATION_1: 
CITATION_2: 
CITATION_3: 

ACCESSION_NO: 91000691
IMAGE_FILE_ID: DVB_00326
CLASSIFICATION: B-Q5 Kas 2.171 Hor 2
PERIOD:
CURRENT LOCATION: Kashmir
COUNTRY: India
TITLE_1: Packed Horses on Road in Kashmir in 1985
TITLE_2: transportation
DATE:
VIEW:
KEYWORDS: transportation; horses
NOTES:
SOURCE_DONOR:
FILM_TYPE: Color: E5034, 1985
CITATION_1:
CITATION_2:
CITATION_3:

ACCESSION_NO: 91000110
IMAGE_FILE_ID: missing
CLASSIFICATION: B-Q5 Kum 3.4 SSrg 9a-1
PERIOD:
CURRENT LOCATION: Kumbakonam
REGION: Tamil Nadu
COUNTRY: India
TITLE_1: Sri Sarangapani Temple
TITLE_2: temple
DATE: ca. 1100-1175
VIEW: Ext.det.: front gopuram
KEYWORDS: Vaishnavite temple
NOTES: Among the most important temples dedicated to Vishnu. The gopura is 11 stories and 44 meters high. Dates: early to mid 12th C., 1121 onward. Built by a Chola Ruler (possibly Vikrama Chola).
SOURCE_DONOR:
FILM_TYPE: Color: K5032, 1978
CITATION_1: Meena, V.; Temples of South India; Kanyakumari: Hari Kumari Arts; 1976; Fine Arts; NA6007.S6 M51
CITATION_2:
CITATION_3:
Sacred tank lies in the NE sector of the fourth prakara. Dates: mid 11th to 14th C. Extensive building by Kulottunga Chola I and his son Vikrana Chola.
Characteristics of Source Metadata

File format:
Type of metadata:
Metadata scheme:
Scheme Support and Documentation:
Content values:
Structure:
Intended use:
Status:
Characteristics of Target Metadata

File format:
Type of metadata:
Metadata scheme:
Scheme Support and Documentation:
Content values:
Structure:
Intended use:
Status:
<table>
<thead>
<tr>
<th>Source</th>
<th>Transformation</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Transformation</td>
<td>Target</td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Metadata Map

<table>
<thead>
<tr>
<th>Source</th>
<th>Transformation</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Metadata Map

<table>
<thead>
<tr>
<th>Source</th>
<th>Transformation</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Exercise 10: Business planning questionnaire

Tools: Cornell business planning tool examples (*Digital Imaging Project Planning Outline*; DCAPS budget overview form; *Metadata Plan of Work Checklist; Metadata Plan of Work*)

Instructions:

Participants break into groups of four or five people each. Working individually, participants take up to 10 minutes to complete the Business Planning Questionnaire below, referring to Cornell business planning tool examples as needed. Once everyone in the small group is done, each person takes about 5 minutes to report to the other group members his/her responses. Listeners should ask questions or offer suggestions to help expand the reporter’s responses. As well, listeners should use this information to augment their own lists where appropriate. Try to track responses that are common across reports.

For the session wrap-up, all participants together discuss common responses across reports.

#### Business Planning Questionnaire

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>List opportunities for your organization in planning and managing a digital initiative.</td>
<td></td>
</tr>
<tr>
<td>List challenges for your organization in planning and managing a digital initiative.</td>
<td></td>
</tr>
<tr>
<td>List your institution's strengths in developing a digital initiative.</td>
<td></td>
</tr>
<tr>
<td>List areas where your institution will need additional capacity and resources to plan and manage a digital initiative.</td>
<td></td>
</tr>
<tr>
<td>List areas where your institution will need additional staff training to plan and manage a digital initiative.</td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>Tasks</td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Overall Project Management</td>
<td>Define requirements, develop a work plan and a time line, assess needs, identify resource requirements, monitor progress and product quality, evaluate, facilitate communication, prepare project reports, financial management</td>
</tr>
<tr>
<td>Selection</td>
<td>Select materials, identify incomplete items or pages in poor condition</td>
</tr>
<tr>
<td>Copyright</td>
<td>Determine legal restrictions</td>
</tr>
<tr>
<td>Physical Preparation</td>
<td>Retrieve documents, record physical conditions, check out if necessary, repair pages, disbind if necessary</td>
</tr>
<tr>
<td>Intellectual Preparation</td>
<td>Create documentation for indexing, tagging, and other information such as special treatment of illustrations and photographs</td>
</tr>
<tr>
<td>User Requirements</td>
<td>Determine user requirements for legibility, navigation, color fidelity, file formats, etc.</td>
</tr>
<tr>
<td>Image Benchmarking</td>
<td>Determine imaging requirements including scanning equipment, resolution, bit depth, image enhancement and correction, compression, file format, file headers</td>
</tr>
<tr>
<td>Metadata Benchmarking</td>
<td>Define metadata needs such as descriptive metadata, file naming and structuring, DC, controlled vocabulary, Voyager record, finding aids, technical metadata, and other resource discovery tools such as Meta Tags</td>
</tr>
<tr>
<td>Image Processing</td>
<td>OCR (determine accuracy rate, software used, etc.), SGML or XML &amp; necessary DTD</td>
</tr>
<tr>
<td>Image Quality Control</td>
<td>Determine methodology (percentage of QC and method – on screen vs. on paper, QC tools)</td>
</tr>
<tr>
<td>Metadata Quality Control</td>
<td>Set accuracy levels for consistency of file naming, file headers, and other manually created metadata</td>
</tr>
<tr>
<td>Derivative Creation</td>
<td>Decide on access file formats, on-the-fly vs. static derivative creation, need for facsimile creation</td>
</tr>
<tr>
<td>Storage</td>
<td>Identify needs and purchase storage hardware and software Plan backups, file loading, and data integrity</td>
</tr>
<tr>
<td>Image Database</td>
<td>Identify access requirements and select image management software</td>
</tr>
<tr>
<td>Digital Preservation</td>
<td>Decide on preservation metadata, backups or mirror sites, any other policy decisions</td>
</tr>
<tr>
<td>Web Delivery</td>
<td>Website interface and graphical design, scripting, write project narrative, prepare user guides</td>
</tr>
<tr>
<td>Project Website</td>
<td>Information about the project including time line, quarterly reports, and other related documents</td>
</tr>
<tr>
<td>Publicity</td>
<td>Inform CUL staff, scholars and researchers, other cultural institutions and scholarly communities of the imaging project</td>
</tr>
<tr>
<td>User Support</td>
<td>Work with PS staff in identifying user support needs, such as training sessions and handouts</td>
</tr>
<tr>
<td>Digital Rights Management</td>
<td>Statement about getting permission to use or reproduce images</td>
</tr>
<tr>
<td>Assessment</td>
<td>Determine criteria for outcome assessments, design and implement assessment, conduct a user study</td>
</tr>
<tr>
<td>Financial Management</td>
<td>Financial reporting and CUL Accounting liaison for managing grant funds</td>
</tr>
</tbody>
</table>

Last Modified Oct 2004, oyr

MDL Exercises 112
<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Pages</th>
<th>Cost/Unit</th>
<th>Subtotals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PREPARATION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User &amp; Content Requirements</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Collation, Tagging</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Conservation and Repairs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DIGITIZATION, STRUCTURING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digitization (Specify)</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Image Processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structuring &amp; Tagging</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>POST PROCESSING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCR Preparation</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Derivative Creation (e.g., PDF)</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td><strong>QUALITY CONTROL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Image QC</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td><strong>METADATA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td><strong>DIGITAL CONTENT DELIVERY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Preparation &amp; Ingest</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Assessment and Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WEB DEVELOPMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Website Design</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Usability Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>STORAGE/YEAR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage and Maintenance</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Preservation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RIGHTS MANAGEMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copyright &amp; IPR Consulting</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td><strong>PROJECT MANAGEMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Management</td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL ESTIMATE</strong></td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
</tbody>
</table>
Metadata Plan of Work Checklist

Follow these steps before preparing a Metadata Plan of Work.

**Determine functional requirements (including specifying which requirements mandatory, which optional)**

1. Interview clients and other key stakeholders
2. Do preliminary analysis of data (use cases, etc.)
3. Gather existing documents: budget (e.g., grant proposal), system documentation, time constraints
4. Identify appropriate methods to meet project goals
5. Get feedback from stakeholders re alternative solutions until agreement is reached

**Determine deliverables**

- Detail solution (including development components & costs)
- Use detailed solution info to create deliverables list

**Determine available staff, time, resources to create timeline**

- Assign staff
- Talk with assigned staff about time requirements
- Develop budget (if not already set)
- Develop timeline
- Establish key deadlines for segments
Metadata Plan of Work

Date:
Project Name:
Technical Lead:
Client Contact:

Project overview and critical elements:
[Describe the project briefly and give key facts.]

Functional requirements this plan addresses:

Deliverables:
[Provide as detailed a list as is reasonable, including verification plan.]

Timeline:
[Include key deadlines, responsible staff, and external dependencies.]
Exercise 11: Material and organizational assessment for digital preservation planning

Tools:

*Preservation Planning Launch Questionnaire—Needs Assessment and Analysis*
*Assessing an Organization’s Commitment to Digital Preservation*

Instructions:

Participants break into groups of about four or five people each. Facilitator introduces *Preservation Planning Questionnaire—Needs Assessment and Analysis*.

Working individually, participants take up to 10 minutes to complete the questionnaire for existing digital objects they are aware of that are candidates for preservation. Once everyone in the small group is done, each person takes about 5 minutes to report to the other small group members about the questionnaire responses. Reporters pay particular attention to: object formats and relationships, known metadata gaps, rights or access issues, and nagging vulnerabilities. Listeners ask questions to clarify anything they don’t understand about the objects their colleagues have identified. (Allow 30 minutes.)

Once all participants have reported and each group has a common understanding of all group members’ reports, facilitator introduces *Assessing an Organization’s Commitment to Digital Preservation*.

Staying in the small groups, group members take turns talking about where they think the organization falls on the organizational, fiscal, and technological continuums, and why. What are your areas of greatest concern? What are some concrete steps that your organization could take to move itself along the continuums of concern? Which individuals or groups within the organization would need to be on board in order to take those steps? Which steps would have the greatest impact? (Allow 30 minutes.)
<table>
<thead>
<tr>
<th>Needs Assessment and Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What digital objects already exist that you want to preserve?</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>In which formats are the objects?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>How many objects are there and how big are the files?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Are some objects related to others in important ways? (versions, derivatives, component parts)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What is the current growth rate of the body of objects you want to preserve?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What new collections and object types are you planning to acquire?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What metadata do you have and what metadata do you want that you don’t have?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Where does existing metadata reside?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What copyright and access restrictions apply to the objects?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>In which essential ways do you want to be able to search for the preserved objects?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What are the points of greatest vulnerability regarding existing objects and metadata?</td>
</tr>
</tbody>
</table>
## Assessing an Organization's Commitment to Digital Preservation

<table>
<thead>
<tr>
<th></th>
<th>Acknowledge</th>
<th>Act</th>
<th>Consolidate</th>
<th>Institutionalize</th>
<th>Externalize</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizational</strong></td>
<td>No, implicit, or high-level policy</td>
<td>General policy; some commitment shown</td>
<td>Basic policies defining essentials</td>
<td>Sustained oversight; comprehensive policies</td>
<td>Planning assumes collaboration; policy integration with collective</td>
</tr>
<tr>
<td><strong>Fiscal</strong></td>
<td>Low, ad hoc commitment</td>
<td>Project-based funding</td>
<td>Some support beyond projects</td>
<td>Sustainable funding for core functions</td>
<td>Sustainable, possibly distributed fiscal management</td>
</tr>
<tr>
<td><strong>Technological</strong></td>
<td>No or disconnected components</td>
<td>Project-based and reactive components</td>
<td>Assess current components and identify those desired</td>
<td>Planning anticipates needs; planned implementation underway</td>
<td>Distributed, integrated components; cooperative services</td>
</tr>
</tbody>
</table>

Exercise 12: Collection-level metadata in a consortial context

Tools:

*MetaArchive Collection-Level Metadata Worksheet*

Instructions:

Participants break into groups of about four or five people each that are different from the groups in the previous session. Facilitator introduces *MetaArchive Collection-Level Metadata Worksheet*.

Exercise 12 assumes that the organization is part of a consortium that will cooperate in an effort to ensure ongoing access to their organization’s digital content.

Group members together work through the elements in the worksheet. For each element discuss whether the element should be mandatory (M) or optional (O) for the consortium and why. Note any elements for which the organization already records the data that would be included in the element. For those elements, where is the data recorded?

When considering the Risk Factors and Risk Rank elements: Each group member briefly describes the objects they reported on in Exercise 11 and describes the risk factors associated with them. Group members together assign Risk Ranks for the objects reported by the members. (Allow 30 minutes.)

For the session wrap-up, each group identifies a spokesperson who reports on the objects and associated risk factors and risk ranks that the group discussed. (Allow 30 minutes.)
<table>
<thead>
<tr>
<th>Metadata Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Access Rights</strong> [dcterms:accessRights]</td>
<td>A statement of any access restrictions placed on a collection, including allowed users, charges, etc.</td>
</tr>
<tr>
<td><strong>Accrual Periodicity</strong> [cld:accrualPeriodicity]</td>
<td>Frequency with which items are expected to be added to a collection.</td>
</tr>
<tr>
<td><strong>Accrual Policy</strong> [cld:accrualPolicy]</td>
<td>Approach adopted to add items to the collection or a statement about anticipated growth of the collection, including quantity and frequency.</td>
</tr>
<tr>
<td><strong>Bytes</strong> [dcterms:extent]</td>
<td>The total calculated file size, expressed in computer bytes, of the digital collection being described.</td>
</tr>
<tr>
<td><strong>Cataloged Status</strong> [ma:catalogedstatus]</td>
<td>Indication of level of cataloging of the collection beyond the collection level.</td>
</tr>
<tr>
<td><strong>Custodial History</strong> [dcterms:provenance]</td>
<td>A statement of any changes in ownership and custody of the collection since its creation in digital format that are significant for its authenticity, integrity and interpretation. Other provenancial information such as the custodial history of a physical source for the digital collection may also be included.</td>
</tr>
<tr>
<td><strong>Description</strong> [dc:description]</td>
<td>A summary of the content of the digital collection.</td>
</tr>
</tbody>
</table>
### MetaArchive Collection-Level Metadata Worksheet

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Format Characteristics [dc:format]</strong></td>
<td>The digital manifestation(s) represented in the digital collection being described.</td>
</tr>
<tr>
<td><strong>Institution Collection Identifier [dc:identifier]</strong></td>
<td>An institution assigned identifier for the digital collection, according to local conventions.</td>
</tr>
<tr>
<td><strong>Is Available Via [cld_gen:isAvailableAt]</strong></td>
<td>The service that provides access to the digital collection, such as an online searchable database.</td>
</tr>
<tr>
<td><strong>Manifestation [mods:physicalDescription]</strong></td>
<td>The role (e.g., access, preservation, replacement) of individual files within the described digital collection as part of a digital preservation strategy.</td>
</tr>
<tr>
<td><strong>OAI Provider [ma:oaiprovider]</strong></td>
<td>This is the URL that specifies the location of the OAI data provider for the collection, if one exists.</td>
</tr>
<tr>
<td><strong>Publisher [dc:publisher]</strong></td>
<td>The institution primarily responsible for making available the digital collection described.</td>
</tr>
<tr>
<td><strong>Recommended Harvest Procedure [ma:harvestproc]</strong></td>
<td>Harvest procedure to be used by MetaArchive preservation network, either LOCKSS web crawl or LOCKSS OAI harvest.</td>
</tr>
<tr>
<td><strong>Rights [dc:rights]</strong></td>
<td>A statement of (or link to) any rights (copyrights, etc.) held in/over the collection.</td>
</tr>
</tbody>
</table>
### Risk Factors [ma:riskfactors]
Describes the risk factors that put this collection at risk and justify a particular risk rank.

M/O:

### Risk Rank [ma:riskrank]
A numeric ranking of the degree to which the collection is at risk. Examples and criteria:  
5: Extreme risk, no one is responsible for preservation, no other copies of the digital content are preserved beyond the available copy under consideration, no regular backups or data migration.  
4: Significant risk, responsibility under discussion, curators fretting about who will take responsibility for preservation. 3: High risk, only one backup copy of digital masters on CD-ROM, no regular backups or data migration. 2: Moderate risk, some danger that collection backups might be lost in future. 1: Low risk, copies are backed up regularly with a long term maintenance plan in some other trusted digital archive.

M/O:

### Subject [dc:subject]
Terms that describe the topic(s) of the content of the resource.

M/O:

### Temporal Coverage [cld:temporal]
The temporal coverage of the intellectual content of the items in the collection.

M/O:

### Title [dc:title]
A name given to the digital collection.

M/O:

### Type [dc:type]
The nature or genre of the content of the resource; the name of the object(s) or resource(s) represented by the digital item(s) in the collection. Here is a suggested list of collection types: Computer Animations, Complex or Learning Objects, Databases, Datasets, Events, Interactive Resources, Moving Images, Physical Object, Services, Software, Sound, Still Images, Text.

M/O:
Exercise 13: Object-level metadata in a repository context

Tools:

PREMIS Object Entity Metadata Worksheet

Instructions:

With participants in small groups, facilitator introduces PREMIS Object Entity Metadata Worksheet, pointing out the elements that are nested in container elements.

Exercise 13 assumes each small group will participate in an effort to ensure ongoing access to their organization’s digital content.

Group members together work through the elements in the worksheet. For each element decide whether the element should be mandatory (M) or optional (O). Note whether the organization would need to supply the data for the element or whether group members would expect the repository to supply the data. For elements where the organization would need to supply data, note instances in which the organization already records the data and in which it does not. For those instances in which the organization already records the data, where is it recorded? For instances in which the organization does not already gather the data, rank the impact of gathering the data on this subjective scale: low, medium, high.

When considering the significantPropertiesType and significantPropertiesValue elements: For the objects the group members reported on in Exercises 11 and 12, work together to propose significant property types and significant property values for them.

Finally, prepare for the session wrap-up by considering this situation: Choose one type of object from your group. Assume that for this object type its significant properties will be lost because the required software/hardware configuration will no longer be maintained by the organization. Propose some strategies for mitigating this loss. (Allow 40 minutes.)

For the session wrap-up, groups report on their significant-property loss scenario and all participants together discuss when such situations might arise and how organizations might anticipate them and manage the consequences. (Allow 30 minutes.)
### PREMIS Object Entity Metadata Worksheet

[Content is a subset of PREMIS Data Dictionary for Preservation Metadata version 2.0, <http://www.loc.gov/premis/v2/premis-2-0.pdf>]

<table>
<thead>
<tr>
<th>1.1 objectIdentifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>A designation used to uniquely identify the object within the preservation repository system in which it is stored.</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.2 objectCategory</th>
</tr>
</thead>
<tbody>
<tr>
<td>The category of object to which the metadata applies. Suggested values: representation, file, bitstream.</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.3.1 preservationLevelValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>A value indicating the set of preservation functions expected to be applied to the object. Examples: bit-level, full, 0, 1, 2, fully supported with future migrations.</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.3.4 preservationLevelDateAssigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>The date, or date and time, when a particular preservationLevelValue was assigned to the object.</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.4 significantProperties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of a particular object subjectively determined to be important to maintain through preservation actions. Significant properties may pertain to all objects of a certain class; in other cases, e.g., for media art, the significant properties may be unique to each individual object.</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.4.1 significantPropertiesType</th>
</tr>
</thead>
<tbody>
<tr>
<td>The aspect, facet, or attribute of an object about which significant properties are being described. Examples: content, structure, behavior, hyperlinks, image count, page count, color space</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.4.2 significantPropertiesValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description of the characteristics of a particular object subjectively determined to be important to maintain through preservation actions. Examples: content only, editable, Adobe RGB 1998.</td>
</tr>
<tr>
<td>M/O:</td>
</tr>
</tbody>
</table>
# PREMIS Object Entity Metadata Worksheet

## 1.5.2 fixity
Information used to verify whether an object has been altered in an undocumented or unauthorized way.

M/O:

## 1.5.2.1 messageDigestAlgorithm
The algorithm used to construct the message digest for the digital object in a fixity check.

M/O:

## 1.5.2.2 messageDigest
The output of the message digest algorithm in a fixity check.

M/O:

## 1.5.3 size
The size in bytes of the file or bitstream stored in the repository.

M/O:

## 1.5.4.1.1 formatName
A designation of the format of the file or bitstream.

M/O:

## 1.5.4.1.2 formatVersion
The version of the format named in formatName.

M/O:

## 1.5.4.2.1 formatRegistryName
A designation identifying a format registry that gives further information about the format.

M/O:

## 1.5.4.2.2 formatRegistryKey
The unique key used to reference an entry for this format in a format registry.

M/O:
## PREMIS Object Entity Metadata Worksheet

<table>
<thead>
<tr>
<th>1.5.5.1 creatingApplicationName</th>
<th>A designation for the name of the software program that created the object.</th>
<th>M/O:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5.5.2 creatingApplicationVersion</td>
<td>The version of the software program that created the object.</td>
<td>M/O:</td>
</tr>
<tr>
<td>1.5.5.3 dateCreatedByApplication</td>
<td>The actual or approximate date and time the object was created.</td>
<td>M/O:</td>
</tr>
<tr>
<td>1.5.6 inhibitors</td>
<td>Features (e.g., encryption) of the object intended to inhibit access, use, or migration.</td>
<td>M/O:</td>
</tr>
<tr>
<td>1.5.6.1 inhibitorType</td>
<td>The inhibitor method employed. Examples: DES, PGP, Blowfish, Password protection.</td>
<td>M/O:</td>
</tr>
<tr>
<td>1.5.6.3 inhibitorKey</td>
<td>The decryption key or password.</td>
<td>M/O:</td>
</tr>
<tr>
<td>1.6 originalName</td>
<td>The name of the object as submitted to or harvested by the repository, before any renaming by the repository.</td>
<td>M/O:</td>
</tr>
<tr>
<td>1.8.5.1 swName</td>
<td>Manufacturer and title of the software application required to render or use the object.</td>
<td>M/O:</td>
</tr>
</tbody>
</table>
### PREMIS Object Entity Metadata Worksheet

<table>
<thead>
<tr>
<th><strong>1.8.5.2 swVersion</strong></th>
<th>The version or versions of the software referenced in swName.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.8.5.3 swType</strong></th>
<th>Class or category of software. Examples: renderer, ancillary, operatingSystem, driver.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.8.6.1 hwName</strong></th>
<th>Manufacturer, model, and version (if applicable) of the hardware components needed by the software referenced in swName or the human user of the referenced software.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.8.6.2 hwType</strong></th>
<th>Class or category of the hardware. Suggested values: processor, memory, input/output device, storage device.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.10 relationship</strong></th>
<th>Information about a relationship between this object and one or more other objects.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.10.1 relationshipType</strong></th>
<th>A high-level categorization of the nature of the relationship. Suggested: structural, derivation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.10.2 relationshipSubType</strong></th>
<th>A specific characterization of the nature of the relationship documented in relationshipType. Suggested values: has sibling, has part, is part of, has source, is source of, has root.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1.10.3.2 relatedObjectIdentifierValue</strong></th>
<th>The value of the identifier of the related object.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/O:</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>1.10.3.3 relatedObjectSequence</strong></td>
<td>The order of the related object relative to other objects with the same type of relationship.</td>
</tr>
<tr>
<td>M/O:</td>
<td></td>
</tr>
<tr>
<td><strong>1.10.4 relatedEventIdentification</strong></td>
<td>The identifier and contextual sequence of an event (e.g., migration) associated with the relationship.</td>
</tr>
<tr>
<td>M/O:</td>
<td></td>
</tr>
<tr>
<td><strong>1.10.4.2 relatedEventIdentifierValue</strong></td>
<td>The eventIdentifierValue of the related event.</td>
</tr>
<tr>
<td>M/O:</td>
<td></td>
</tr>
<tr>
<td><strong>1.12 linkingIntellectualEntityIdentifier</strong></td>
<td>An identifier for an intellectual entity associated with the object. (This may be a link to descriptive metadata that describes the entity. This link will likely be to an identifier of an object that is at a higher conceptual level than the object for which the metadata is provided.)</td>
</tr>
<tr>
<td>M/O:</td>
<td></td>
</tr>
<tr>
<td><strong>1.13 linkingRightsStatementIdentifier</strong></td>
<td>An identifier for a rights statement associated with the object.</td>
</tr>
<tr>
<td>M/O:</td>
<td></td>
</tr>
</tbody>
</table>
Exercise 14: Team building and work planning

Tools:

Three digital project scenarios (on following pages)

Instructions:

Participants break into groups of about four or five people each. Facilitator introduces the three digital project scenarios and assigns a scenario to each group.

Each group is a planning team for its digital project. Group members discuss and address: What are strengths, weaknesses, opportunities, and threats associated with the project? Who are possible implementation team members for the project? Who are key stakeholders? What is the target audience for the project? What need does the project address? Write a brief “benefits of solution” statement for the project. What are the primary cost categories involved? (Categories only, no amounts.) Would the organization(s) involved be able to support an in-house system deployment? (Allow 30 minutes.)

For the session wrap-up, each group identifies a spokesperson that briefly describes the project and reports on the group’s results. (Allow 30 minutes.)
Digital Project Scenario 1
Collaborative Oral History: Pioneers of the Great Plateau

The State Historical Society, a member of the Consortia of the Great Plateau, plans to write a grant for $150,000 to fund the creation of a digital archive of the oral histories made of the original pioneers who settled in the Great Plateau at the turn of the century. The Consortia estimates that they have a total of 500 oral histories of the founders of the Plateau, with another 5,000 documents supporting the oral histories (letters, maps, photographs, newspaper clippings from 1900-1933, brochures, postcards). The Consortia consists of the State Historical Society and five historical societies that support the local communities of the Great Plateau. The State Historical Society has 6 paid employees: a director, 2 curators, 2 museum registrars/technicians, and 1 secretary. The Historical Society of Mountain Pass has a director/curator, a secretary, and two assistants who provide guided tours for schools and the general public. The other four local historical societies are managed by volunteers who provide access to their materials on a limited basis (weekends and some holidays). The collections of the volunteer historical societies represent about 35% of the total estimated material. The State Historical Society has 40% and the Historical Society of Mountain Pass has the remaining 25% of the material. The original oral histories are a mix of media types and states of preservation. In some cases, only the written transcript remains accessible.

Digital project scenario 2
Charles Dickens collection

University of Midwest is a private mid-sized university who received a major archival gift five years ago from the endowed chair of the English Department. The archives include her extensive collection of works by Charles Dickens and works based on the original works. The collection includes first editions of the books, some newspaper clippings in albums of the serialization as the novels first appeared, and subsequent materials adapted by others based on the works of Charles Dickens: scores to songs, musical scores, children’s editions, prints (illustrations), as well as modern books on tape, film versions of the novels (reel-to-reel, videos and DVDs), three dimensional objects (glass figurines, music boxes…). The collection also includes scholarly and popular works about Charles Dickens and his works. The University plans to digitize much of the work for which there a few originals. The intention is to develop curriculum packages that would support K-12 education as well the University courses on nineteenth-century literature. The Special Collections curator has a reading room supervisor and one staff member to help process material. The University has a systems librarian and a web developer on staff. Technical Services has 3 professional catalogers and 15 paraprofessionals in cataloging, acquisitions and processing. They are excited about the project but are worried about the impact on the workflow. They do not have a backlog and hope to maintain that trend.
Digital project scenario 3
Digitization of local newspapers

The University of the Northeast is a large library with a significant microfilm collection of local neighborhood newspapers that are a rich primary resource for history, political science, urban studies, geography, economics, and population studies. The University has become aware that two other colleges in the state have similar collections that complement their holdings. In all they have 5,000 reels of microfilm. The collection dates from the period of the American Revolution up to 1965 at which point most of the regional newspapers were absorbed by large newspaper conglomerates or just faded away. The intention is to provide access to this primary resource for curriculum needs of K-12 as well as higher education and researchers. The University of the Northeast has 1 systems librarian and 3 programmers on staff; however, they are also responsible for all the open labs on campus. Their Technical Services department is in the process of reorganizing after migrating to a new ILS. The other 2 colleges only have 1 systems librarian each. Their Technical Services departments only have 2 professional librarians and 12 paraprofessionals in cataloging and acquisitions. Librarians and staff at U.N. feel that they can absorb the project with the new ILS making their current work more efficient.
Exercise 15: Proposal writing and assessment

Tools:

Three digital project scenarios (on following pages)

Instructions:

Participants stay in the same groups as Exercise 14 and work with the same digital project scenarios.

Group members discuss and address: Specify an overall time frame and quantifiable outcomes for the project. In general terms, list space, equipment, staffing, system, and metadata needs. What additional information would you need to gather to answer key resource questions. Draft an action timeline for the project that covers key steps. What are your sustainability concerns? Briefly draft an evaluation plan. Which outcomes would you measure? How do you propose measuring the impact of the project on its target audience? (Allow 30 minutes.)

For the session wrap-up, each group identifies a spokesperson (not the same spokesperson as Exercise 14) that reports on the group’s results. (Allow 30 minutes.)
Digital Project Scenario 1
Collaborative Oral History: Pioneers of the Great Plateau

The State Historical Society, a member of the Consortia of the Great Plateau, plans to write a grant for $150,000 to fund the creation of a digital archive of the oral histories made of the original pioneers who settled in the Great Plateau at the turn of the century. The Consortia estimates that they have a total of 500 oral histories of the founders of the Plateau, with another 5,000 documents supporting the oral histories (letters, maps, photographs, newspaper clippings from 1900-1933, brochures, postcards). The Consortia consists of the State Historical Society and five historical societies that support the local communities of the Great Plateau. The State Historical Society has 6 paid employees: a director, 2 curators, 2 museum registrars/technicians, and 1 secretary. The Historical Society of Mountain Pass has a director/curator, a secretary, and two assistants who provide guided tours for schools and the general public. The other four local historical societies are managed by volunteers who provide access to their materials on a limited basis (weekends and some holidays). The collections of the volunteer historical societies represent about 35% of the total estimated material. The State Historical Society has 40% and the Historical Society of Mountain Pass has the remaining 25% of the material. The original oral histories are a mix of media types and states of preservation. In some cases, only the written transcript remains accessible.

Digital project scenario 2
Charles Dickens collection

University of Midwest is a private mid-sized university who received a major archival gift five years ago from the endowed chair of the English Department. The archives include her extensive collection of works by Charles Dickens and works based on the original works. The collection includes first editions of the books, some newspaper clippings in albums of the serialization as the novels first appeared, and subsequent materials adapted by others based on the works of Charles Dickens: scores to songs, musical scores, children’s editions, prints (illustrations), as well as modern books on tape, film versions of the novels (reel-to-reel, videos and DVDs), three dimensional objects (glass figurines, music boxes…). The collection also includes scholarly and popular works about Charles Dickens and his works. The University plans to digitize much of the work for which there a few originals. The intention is to develop curriculum packages that would support K-12 education as well the University courses on nineteenth-century literature. The Special Collections curator has a reading room supervisor and one staff member to help process material. The University has a systems librarian and a web developer on staff. Technical Services has 3 professional catalogers and 15 paraprofessionals in cataloging, acquisitions and processing. They are excited about the project but are worried about the impact on the workflow. They do not have a backlog and hope to maintain that trend.
Digital project scenario 3
Digitization of local newspapers

The University of the Northeast is a large library with a significant microfilm collection of local neighborhood newspapers that are a rich primary resource for history, political science, urban studies, geography, economics, and population studies. The University has become aware that two other colleges in the state have similar collections that complement their holdings. In all they have 5,000 reels of microfilm. The collection dates from the period of the American Revolution up to 1965 at which point most of the regional newspapers were absorbed by large newspaper conglomerates or just faded away. The intention is to provide access to this primary resource for curriculum needs of K-12 as well as higher education and researchers. The University of the Northeast has 1 systems librarian and 3 programmers on staff; however, they are also responsible for all the open labs on campus. Their Technical Services department is in the process of reorganizing after migrating to a new ILS. The other 2 colleges only have 1 systems librarian each. Their Technical Services departments only have 2 professional librarians and 12 paraprofessionals in cataloging and acquisitions. Librarians and staff at U.N. feel that they can absorb the project with the new ILS making their current work more efficient.