Metadata and Digital Libraries

Marty Kurth

UAEU Libraries

October 4-8, 2009

Participant Manual

Metadata and Digital Libraries, Marty Kurth, UAEU Libraries, Oct 4-8, 2009
Workshop Schedule

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

Metadata and Digital Libraries

Marty Kurth

UAEU Libraries October 4-8, 2009

1

Workshop approach

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

2

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.

3

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)

4

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

5

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?

6

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)

7

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

8

Svenonius (2000)

- 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

9

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?

10

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

11

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?

12

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?

13

Digital library 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

14

Digital library system objectives

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

15

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?

16

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?

17

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?

18

Session 2: Building digital collections

19

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

20

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

21

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

22

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

23

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

24

Digital collection selection criteria

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

25

Reflection

"There are only local collections, built with local funding, in support of local needs."—Ross Atkinson

26

Exercise 2: Selection for digitization

27

For further study:

De Stefano, Paula. “Selection for Digital Conversion.” *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. Anne R. Kenney and Oya Y. Rieger, eds. Mountain View, CA: Research Libraries Group, 2000, p. 11-23.

28

And just out:

Ooghe, Bart, and Dries Moreels. “Analysing Selection for Digitisation: Current Practices and Common Incentives.” *D-Lib Magazine* 15:9/10 (Sept/Oct 2009). <http://www.dlib.org/dlib/september09/ooghe/09ooghe.html>

29

Session 3: More about metadata and introduction to Dublin Core

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

30

Reflection (point)

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

31

Reflection (counterpoint)

"Metadata is what we know and data is what we're looking for."—David Weinberger

32

Some typical metadata functions

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

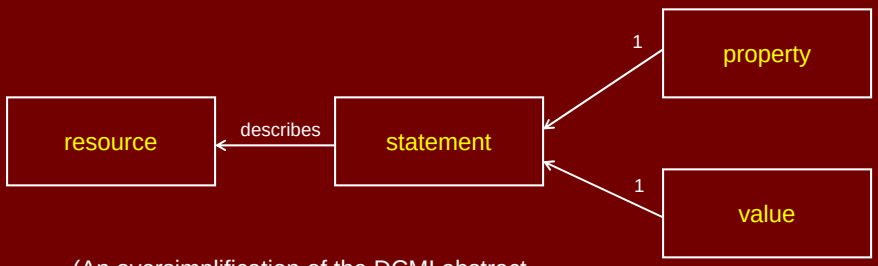
33

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*.

34

Metadata building blocks (in pictures)



(An oversimplification of the DCMI abstract model for resources)

35

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

245 00 $a Amores perros $h [videorecording]

<title>Nueve reinas</title>
<type>MovingImage</type>

36

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

37

Introduction to the Dublin Core

38

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)

39

Dublin Core Metadata Element Set

Creator	Title	Subject
Contributor	Date	Description
Publisher	Type	Format
Coverage	Rights	Relation
Source	Language	Identifier

40

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

41

Resources for which DC is often used

DCMI Type Vocabulary

Collection	Dataset	Event
Image	Interactive Resource	Moving Image
Physical Object	Service	Software
Sound	Still Image	Text

42

Dublin Core principles

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

43

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

44

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>
  - <dcterms:issued>2000-07-11</dcterms:issued>

45

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*

46

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

47

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

48

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

49

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)

50

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

51

Exercise 3: Creating Dublin Core metadata for digital objects

52

Day 1 debriefing

53

Session 4: Understanding functional requirements

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

54

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

55

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

56

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

57

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

58

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

59

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

60

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?

61

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

62

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...

63

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

64

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

65

Requirements and metadata

- Certain functional requirements will depend upon or impact system metadata
- The requirements will inform our decisions about 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

66

Exercise 4: Sample use case

67

Session 5: Metadata and functionality

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

68

Session goals

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

69

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

70

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

71

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

72

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

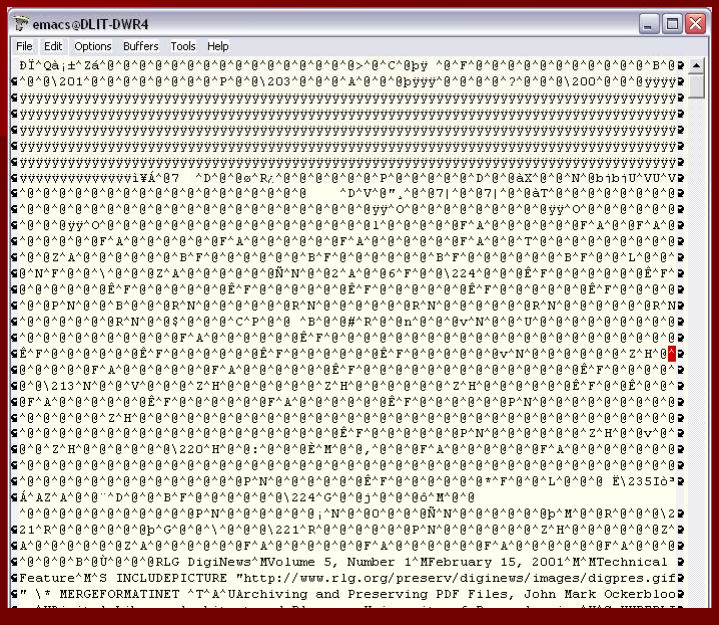
73

```
<?xml version="1.0" encoding="UTF-8" ?>
- <mets:mets xmlns:mets="http://www.loc.gov/METS/" xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:mods="http://www.loc.gov/mods/v3"
  xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
  xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
  http://www.loc.gov/mods/v3 mods-3-0.xsd http://cosimo.stanford.edu/sdr/metsrights/
  http://cosimo.stanford.edu/sdr/metsrights.xsd" OBJID="ark:/13030/kt0h4nc5tp" TYPE="text"
  LABEL="Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability"
  PROFILE="http://sunsite.berkeley.edu/mets/profiles/UCBApplicationContentProfile.xml">
  - <mets:dmdSec ID="DMD1">
    - <mets:mdWrap MDTYPE="MODS">
      - <mets:xmlData>
        - <mods:mods>
          - <mods:titleInfo>
              <mods:title>Proceedings of the Third Berkeley Symposium on Mathematical Statistics
                and Probability (1954/1955)</mods:title>
            </mods:titleInfo>
          - <mods:name type="corporate">
              <mods:namePart>University of California Press</mods:namePart>
            - <mods:role>
                <mods:roleTerm authority="marc">Publisher</mods:roleTerm>
              </mods:role>
            </mods:name>
            <mods:typeOfResource>text</mods:typeOfResource>
            <mods:genre authority="aat">proceedings</mods:genre>
          - <mods:originInfo>
            - <mods:place>
                <mods:placeTerm type="text">Berkeley, California</mods:placeTerm>
              </mods:place>
              <mods:dateCreated>1956</mods:dateCreated>
              <mods:dateCreated encoding="iso8601" point="start">1956</mods:dateCreated>
            </mods:originInfo>
          - <mods:language>
              <mods:languageTerm type="code" authority="iso639-2b">eng</mods:languageTerm>
            </mods:language>
```

74

```
|1|000A|1|1001|
|1|000B|2|1002|
|1|R001|3|1003|
|1|R002|4|1004|
|1|R003|5|1005|
|1|R004|6|1006|
|1|0001|7|1007|
|1|0002|8|1008|
|1|0003|9|1009|
|1|0004|10|1010|
|1|0005|11|1011|
|1|0006|12|1012|
|1|0007|13|1013|
|1|0008|14|1014|
|1|0009|15|1015|
|1|0010|16|1016|
|1|0011|17|1017|
|1|0012|18|1018|
|1|0013|19|1019|
|1|0014|20|1020|
|1|0015|21|1021|
|1|0016|22|1022|
|1|0017|23|1023|
|1|0018|24|1024|
|1|0019|25|1025|
|1|0020|26|1026|
|1|0021|27|1027|
|1|0022|28|1028|
|1|0023|29|1029|
|1|0024|30|1030|
|1|0025|31|1031|
|1|0026|32|1032|
|1|0027|33|1033|
|1|0028|34|1034|
|1|0029|35|1035|
```

75



76

Metadata analysis: Type

- Types of metadata
  - Descriptive
  - Structural
  - Administrative
  - Technical
  - Preservation
  - Access/rights

77

```
_vn 2.2ml
_pd 19970307
_mi M
_t1 CM000257 ANZ7371
_ml 01012081
_ma Goodwin, John Abbot,
_mu
_mt The Pilgrim republic;   an historical review of the colony of New Plymouth, with
sketches of the rise of other New England settlements, the history of Congregationalism,
and the creeds of the period.
_mo
_me
_ci Boston,
_pu Ticknor and Company; [etc., etc.]
_md 1888.
_mc
_mr
_ms New England -- History -- Colonial period, ca. 1600-1775.
_ms Puritans.
_t2 CM000257 ANZ7371 t2
_mn 712
_mp 000A 000B R001 R002 R003 R004 R005 R006 R007 R008 R009 R010 R011 R012 R013
    R014 R015 R016 R017 R018 R019 R020 R021 R022 R023 R024 R025 R026 R027 R028
    R029 R030 R031 R032 R033 R034 R035 R036 R037 R038 R039 R040 R041 R042 R043
    R044 R045 R046 0001 0002 0003 0004 0005 0006 0007 0008 0009 0010 0011 0012
    0013 0014 0015 0016 0017 0018 0019 0020 0021 0022 0023 0024 0025 0026 0027
    0028 0029 0030 0031 0032 0033 0034 0035 0036 0037 0038 0039 0040 0041 0042
    0043 0044 0045 0046 0047 0048 0049 0050 0051 0052 0053 0054 0055 0056 0057
    0058 0059 0060 0061 0062 0063 0064 0065 0066 0067 0068 0069 0070 0071 0072
    0073 0074 0075 0076 0077 0078 0079 0080 0081 0082 0083 0084 0085 0086 0087
    0088 0089 0090 0091 0092 0093 0094 0095 0096 0097 0098 0099 0100 0101 0102
    0103 0104 0105 0106 0107 0108 0109 0110 0111 0112 0113 0114 0115 0116 0117
    0118 0119 0120 0121 0122 0123 0124 0125 0126 0127 0128 0129 0130 0131 0132
    0133 0134 0135 0136 0137 0138 0139 0140 0141 0142 0143 0144 0145 0146 0147
    0148 0149 0150 0151 0152 0153 0154 0155 0156 0157 0158 0159 0160 0161 0162
```

78

```
- <mets:structMap TYPE="mixed">
  - <mets:div LABEL="Proceedings of the Third Berkeley Symposium on Mathematical Statistics and
    Probability" TYPE="text" DMDID="DMD1">
    - <mets:div LABEL="Volume 1: Statistics" TYPE="volume" DMDID="DMD2">
      - <mets:div LABEL="Frontmatter" TYPE="frontmatter">
          <mets:fptr FILEID="FID1" />
        </mets:div>
      - <mets:div LABEL="Estimation of Least Squares and by Maximum Likelihood" TYPE="article"
          DMDID="DMD3">
          <mets:fptr FILEID="FID5" />
        </mets:div>
      - <mets:div LABEL="On a Use of the Mann-Whitney Statistic" TYPE="article" DMDID="DMD4">
          <mets:fptr FILEID="FID6" />
        </mets:div>
      - <mets:div LABEL="Index" TYPE="index">
          <mets:fptr FILEID="FID3" />
        </mets:div>
      </mets:div>
    - <mets:div LABEL="Volume 2: Contributions to probabilty theory, part 1" TYPE="volume"
        DMDID="DMD5">
      - <mets:div LABEL="Frontmatter" TYPE="frontmatter">
          <mets:fptr FILEID="FID2" />
        </mets:div>
      - <mets:div LABEL="On a Class of Probability Spaces" TYPE="article" DMDID="DMD6">
          <mets:fptr FILEID="FID7" />
        </mets:div>
      - <mets:div LABEL="Stationarity, Boundedness, Almost Periodicity of Random-valued
          Functions" TYPE="article" DMDID="DMD7">
          <mets:fptr FILEID="FID8" />
        </mets:div>
      - <mets:div LABEL="Index" TYPE="index">
          <mets:fptr FILEID="FID4" />
        </mets:div>
      </mets:div>
      <mets:div LABEL="Volume 2: Contributions to probability theory, part 2" TYPE="volume"
        DMDID="DMD8" />
    </mets:div>
```

79

```
    <DeviceSource>digital still scanner</DeviceSource>
  - <ScanningSystemCapture>
    - <ScanningSystemHardware>
        <ScannerManufacturer>Scitex</ScannerManufacturer>
      - <ScannerModel>
          <ScannerModelName>Leaf Volare</ScannerModelName>
          <ScannerModelNumber>100SX</ScannerModelNumber>
          <ScannerModelSerialNo>1122334455</ScannerModelSerialNo>
        </ScannerModel>
      </ScanningSystemHardware>
    - <ScanningSystemSoftware>
        <ScanningSoftware>Leaf</ScanningSoftware>
        <ScanningSoftwareVersionNo>4.0</ScanningSoftwareVersionNo>
      </ScanningSystemSoftware>
    - <ScannerCaptureSettings>
        <PixelSize>3</PixelSize>
      - <PhysScanResolution>
          <XphysScanResolution>0.307</XphysScanResolution>
          <YphysScanResolution>0.435</YphysScanResolution>
        </PhysScanResolution>
      </ScannerCaptureSettings>
    </ScanningSystemCapture>
  - <DigitalCameraCapture>
      <DigitalCameraManufacturer>PhaseOne</DigitalCameraManufacturer>
      <DigitalCameraModel>H_20</DigitalCameraModel>
    </DigitalCameraCapture>
  - <CameraCaptureSettings>
      <FNumber>5</FNumber>
      <ExposureTime>.5</ExposureTime>
      <Brightness>2.2</Brightness>
      <ExposureBias>3.5</ExposureBias>
      <SubjectDistance min="4.9" max="5.3">5</SubjectDistance>
      <MeteringMode>Average</MeteringMode>
      <SceneIlluminant>D75 Illuminant</SceneIlluminant>
      <ColorTemp>20</ColorTemp>
      <FocalLength>3</FocalLength>
      <Flash>No</Flash>
      <FlashEnergy>7</FlashEnergy>
```

80

Metadata analysis: Semantics

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

81

```
<?xml version="1.0" encoding="UTF-8" ?>
- <mods xmlns:xlink="http://www.w3.org/TR/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
  instance" xmlns="http://www.loc.gov/mods/" xsi:schemaLocation="http://www.loc.gov/mods/
  http://www.loc.gov/standards/mods/mods.xsd">
 - <titleInfo>
    <title>Arithmetic /</title>
   </titleInfo>
 - <name type="personal">
    <namePart>Sandburg, Carl,</namePart>
    <namePart type="date">1878-1967.</namePart>
    <role>creator</role>
   </name>
 - <name type="personal">
    <namePart>Rand, Ted,</namePart>
    <role>ill.</role>
   </name>
   <typeOfResource>text</typeOfResource>
 - <publicationInfo>
    <placeCode authority="marc">cau</placeCode>
    <place>San Diego :</place>
    <publisher>Harcourt Brace Jovanovich,</publisher>
    <dateIssued>c1993.</dateIssued>
    <dateIssued encoding="marc">1993</dateIssued>
    <edition>1st ed.</edition>
    <issuance>monographic</issuance>
   </publicationInfo>
   <language authority="iso639-2b">eng</language>
 - <physicalDescription>
    <extent>1 v. (unpaged) : ill. (some col.) ; 26 cm.</extent>
   </physicalDescription>
   <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>
 - <targetAudience>
    <listValue>juvenile</listValue>
   </targetAudience>
   <note type="statement of responsiblity">Carl Sandburg ; illustrated as an anamorphic adventure by Ted
    Rand.</note>
```

82

```
<?xml version="1.0" ?>
- <dc xmlns="http://purl.org/dc/elements/1.1/">
    <title>Arithmetic /</title>
    <creator>Sandburg, Carl, 1878-1967.</creator>
    <creator>Rand, Ted, ill.</creator>
    <type>text</type>
    <publisher>San Diego :Harcourt Brace Jovanovich,</publisher>
    <date>c1993.</date>
    <language>eng</language>
    <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>
    <description>One Mylar sheet included in pocket.</description>
    <subject>Arithmetic</subject>
    <subject>Children's poetry, American.</subject>
    <subject>Arithmetic</subject>
    <subject>American poetry.</subject>
    <subject>Visual perception.</subject>
  </dc>
```

83

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html
    PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
     "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" lang="en-US" xml:lang="en-US">
<head>
<title>Arithmetic</title>
<meta name="ROBOTS" content="INDEX, FOLLOW" />
<meta name="DC:title" content="Arithmetic" />
<meta name="DC:creator" content="Sandburg, Carl, 1878-1967" />
<meta name="DC:creator" content="Rand, Ted, ill." />
<meta name="DC:publisher" content="San Diego :Harcourt Brace Jovanovich" />
<meta name="DC:date" content="c1993" />
<meta name="DC:language" content="eng" />
<meta name="DC:description" content="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." />
<meta name="DC:description" content="One Mylar sheet included in pocket." />
<meta name="DC:subject" content="Arithmetic" />
<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" />

<script type="text/javascript">//<![CDATA[
        function preload_image(args)
        {
        var myimages=new Array();
        var i = 1;
        args++;
        while (i != args)
        {
        myimages[i] = new Image();
        myimages[i].src = preload_image.arguments[i];
        i++;
        }
        }
```

84

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?

85

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

86

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…

87

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

88

A typology of data standards, pt. 1

Type of Data Standard	Examples
Data structure standards (metadata element sets, schemas). These are "categories" or "containers" of data that make up a record or other information object.	*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*

(From: Gilliland, Anne J. "Setting the Stage." *Introduction to Metadata*, online ed., version 3.0. Murtha Baca, ed. Los Angeles: Getty Research Institute, 2008. <http://www.getty.edu/research/conducting_research/standards/intrometadata/setting.html>

89

A typology of data standards, pt. 2

Type of Data Standard	Examples
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 .	*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)*

(From: Gilliland, Anne J. "Setting the Stage." *Introduction to Metadata*, online ed., version 3.0. Murtha Baca, ed. Los Angeles: Getty Research Institute, 2008. <http://www.getty.edu/research/conducting_research/standards/intrometadata/setting.html>

90

A typology of data standards, pt. 3

Type of Data Standard	Examples
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.	*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*

(From: Gilliland, Anne J. "Setting the Stage." *Introduction to Metadata*, online ed., version 3.0. Murtha Baca, ed. Los Angeles: Getty Research Institute, 2008. <http://www.getty.edu/research/conducting_research/standards/intrometadata/setting.html>

91

Exercise 5: Metadata analysis

92

Session 6: Four characteristics of metadata practice

(From: Kurth, Martin. "Found in Translation: Four Characteristics of Metadata Practice." In *Metadata and Digital Collections: A Festschrift in Honor of Tom Turner*, edited by Elaine Westbrooks and Keith Jenkins. Ithaca, NY: Cornell University Library, 2009.)

93

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

94

To cover

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

95

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.

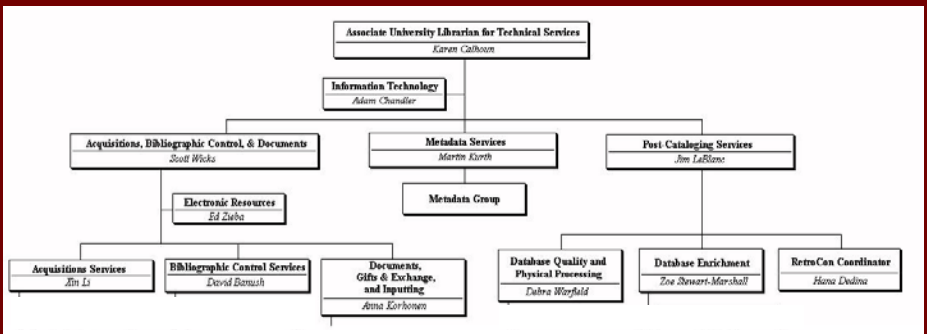
96

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.

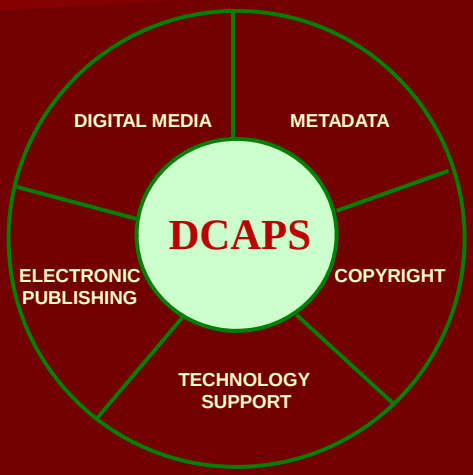
97

Cornell Library Technical Services formed Metadata Services by reallocation in 2002



98

The DCAPS service model



DCAPS: Digital Consulting & Production Services

99

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.

100

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.

101

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.

102

1. Metadata practice approaches metadata in aggregates

103

Practitioners work in the context of

- Projects
- Collections
- Services
- Communities of practice

104

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."

105

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

106

Which means that . . .

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

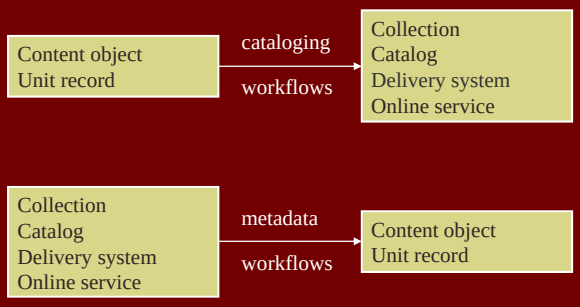
107

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

108

The two approaches contrasted



109

Why is this important?

Cataloging skills still apply in a metadata environment because cataloging and metadata workflows are mirror images of one another.

110

2. Metadata practice comprises interpersonal, informational, and operational layers

111

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

112

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

113

The metadata practitioner's organizational responsibilities

(from Mintzberg by way of Choo)

114

Is it not true that everyone on a team fulfills these roles?

What is unique about metadata work?

115

3. Metadata practice specializes in cross-community translation

116

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

117

In libraries we know how this works

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

. . . acronyms!

118

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

119

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

120

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

121

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-)

122

4. Metadata practice's semantic and syntactic translations support interoperability

123

Metadata is modular (in creation and use)

- ISBN, AACR, LCCN, LCSH
- descriptive, technical, preservation, rights

124

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

125

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>

126

What do mapping and transformation have to do with metadata aggregates, layers of metadata responsibilities, and cross-community translations?

127

NISO *Framework of Guidance for Building Good Digital Collections*

Digital objects, metadata, and collections are building blocks for reuse and integration

128

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.

129

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.

130

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

131

Exercise 6: Metadata analysis scenarios

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

132

Day 2 debriefing

133

Session 7: Metadata conversion

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

134

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

135

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

136

Why enhance metadata?

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

137

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>text/html</date>
  <identifier>1995</identifier>
  <language>http://press.opt.edu/CityOfBits.html</language>
</DC_record>
```

138

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>

139

"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."

A Framework of Guidance for Building Good Digital Collections. 2nd ed. NISO, 2004.
http://www.niso.org/framework/framework2.html

140

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

141

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

142

Tools for interoperability

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

143

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

144

DC Element	MARC Fields	Implementation notes
Title	245	
Creator	100, 110, 111, 700, 710, 711	See Appendix 1 below; Contributor element not used.
	720	
Subject	600, 610, 611, 630, 650, 653	
Description	500-599, except 506, 530, 540, 546	
Contributor		See Appendix 1 below; Contributor element not used.
Publisher	260ab	
Date	260$c	
Type	Leader06, Leader07	See Appendix 2 for Leader-Type rules
	655	
Format	856$q	
Identifier	856$u	
Source	786ot	
Language	008/35-37	
	546	
Relation	530, 760-787ot	
Coverage	651	
	752	
Rights	506, 540	

http://www.loc.gov/marc/marc2dc.html

145

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

146

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

147

Exercise 7: Metadata mapping

- Creating “shareable” metadata
- Designing a detailed metadata map
- Converting from relatively rich metadata to simple Dublin Core records

148

Session 8: Metadata workflows

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

149

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

150

Workflow fundamentals

- The movement of data through a work process

  Input → Transformations → Output

- A work process will typically involve multiple components or individual steps (tasks and subtasks)
  - Each task also has its own data movement:
    - Input → Transformations → Output

151

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

152

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

153

Identifying constraints

- Resources
  - Money
  - Staff
- Time
- Environmental constraints
- Knowledge and expertise

154

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

155

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?

156

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?

157

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?

158

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?

159

Opportunities and benefits

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

160

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

161

Types of metadata workflows

- Enhancement and mapping

  Source data → Transformations → Target data
- Other workflows:
  - Augmentation of records
  - Analysis or evaluation
  - Quality control/assurance

162

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

163

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

164

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

165

The impact on workflow of...

- 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

166

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?

167

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?

168

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

169

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

170

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

171

Exercise 8: Metadata workflow

Library publishing—designing a workflow for a metadata conversion project

172

Session 9: Digital library development project exercise

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

173

Exercise 9: Digital library development project—the slide library

174

Day 3 debriefing

175

Session 10: Digital initiative business planning

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

176

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

177

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

178

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

179

Digital initiative cost elements

- Selection
- Preparation
- Digitization
- Metadata
- System deployment
- Assessment
- Digital preservation

180

Needs Assessment and Decision-Making

objects

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

users

User requirements
Use type
Frequency of use
Use mode
Support needs

resources

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

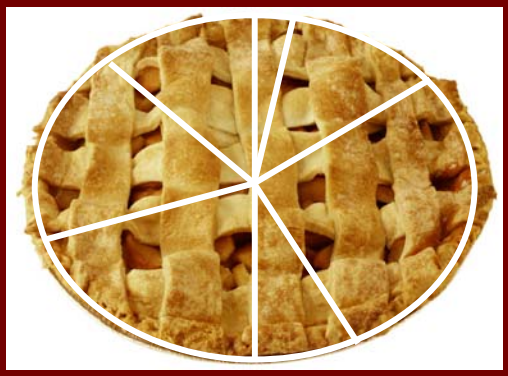
181

Resource planning strategies

To meet resource needs we are faced with two basic approaches . . .

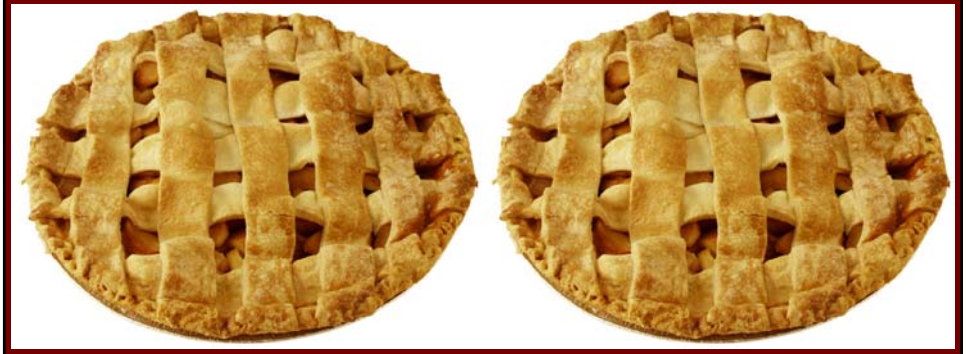
182

Dividing up the pie



183

More pie



184

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?

185

Increased resources

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

186

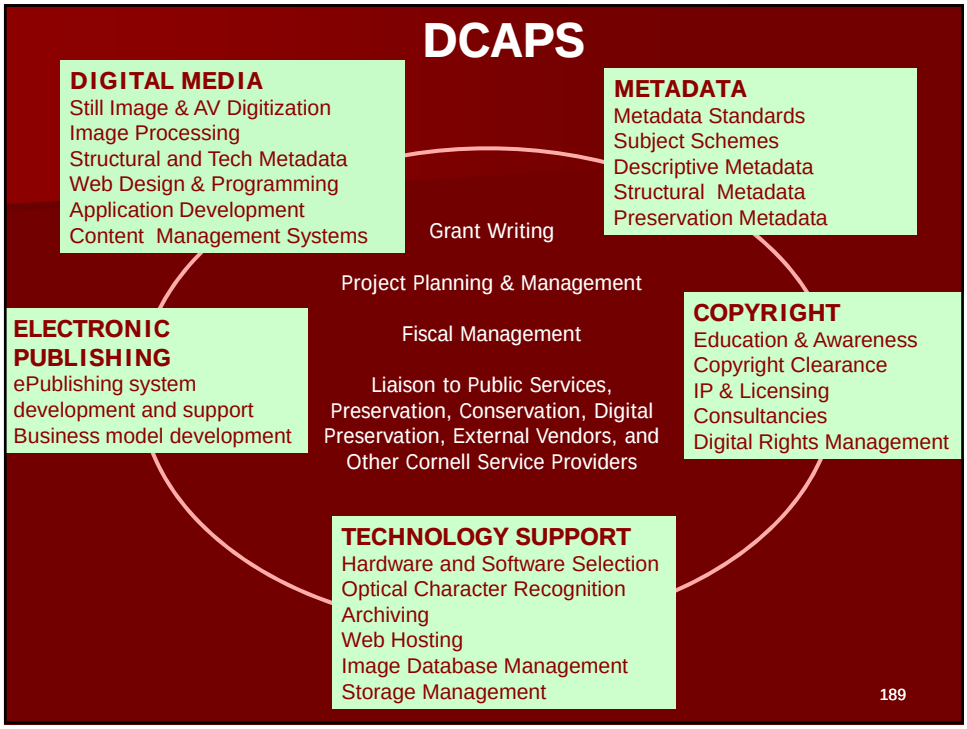
Service model case study

Digital Consulting and Production Services
Cornell University Library

187

- 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

188



Business planning tool examples

- Project planning outline
- Budget overview form
- Metadata plan-of-work checklist
- Metadata plan of work

190

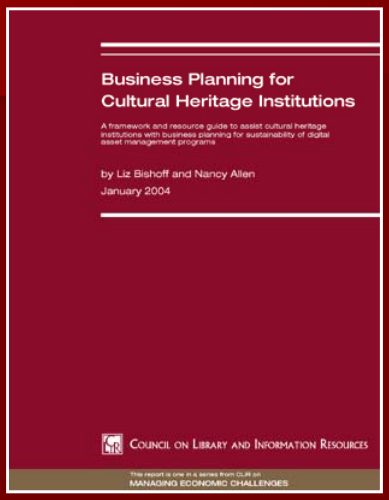
Exercise 10: Business planning questionnaire

191

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?

192



For further study:

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

193

Session 11: Digital preservation planning

194

Reflection (point)

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

195

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

196

Center for Research Libraries
about CRL
membership
collections
preservation
projects & programs
sales
catalog
Catalog Search
Keyword
go
quick links
Collection Search Links
e-Collections
Requesting Materials
Member Services
Member List
FOCUS Newsletter
Area Studies Projects
Global Resources Network
Employment
preservation
Digital Archives
Certification and Assessment of Portico and HathiTrust
Certification of Digital Archives Project
Digital Archive Reports
Long-Lived Digital Collections
Metrics for Assessing and Certifying
Core Requirements
TRAC Checklist & Report
International Coalition on Newspapers (ICON)
Microfilming Projects
Political Communications Web Archive Project
Print Archives
auditing and certification of
Digital Archives
Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist
The Audit Checklist for the Certification of Trusted Digital Repositories, originally developed by RLG-NARA Digital Repository Certification Task Force, has been revised and expanded. The revised version is entitled: Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC).
Introduction
Table of Contents
Using this checklist for audit and certification
Audit and certification criteria
Section A: Organizational Infrastructure
Section B: Digital Object Management
Section C: Technologies, Technical Infrastructure, & Security
Audit Checklist: Criteria for measuring trustworthiness of digital repositories and archives
References
Appendix 1: Glossary
Appendix 2: Understandability and use
Appendix 3: Minimum required documents
Appendix 4: A Perspective on ingest
Appendix 5: Preservation planning and strategies
Appendix 6: Understanding digital repositories and access functionality
Last updated 07/11/2008
search the site
site map
contact us
help
©2002 - 2008 Center for Research Libraries

DRAMBORA interactive
Digital Repository Audit Method Based on Risk Assessment
DCC
Register for DRAMBORA
Username:
Password:
Remember:
Log In
Forgot Password?
Home
Latest News
Get Expert Help
Download Offline Version
Submit Feedback
DRAMBORA Training
About
Objectives
Benefits
The DRAMBORA Team
Collaborations
Dissemination
DRAMBORA Users
Welcome to DRAMBORA Interactive: Log in or Register to Use the Toolkit
Welcome to DRAMBORA Interactive
Please register, log in or read about the toolkit, get an overview of its objectives, or learn of the benefits that it can offer you and your repository.
The Digital Curation Centre (DCC) and DigitalPreservationEurope (DPE) are delighted to announce the release of the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) toolkit. This toolkit is intended to facilitate internal audit by providing repository administrators with a means to assess their capabilities, identify their weaknesses, and recognise their strengths. Digital repositories are still in their infancy and this model is designed to be responsive to the rapidly developing landscape. The development of the toolkit follows a concentrated period of repository pilot audits undertaken by the DCC, conducted at a diverse range of organisations including national libraries, scientific data centres and cultural and heritage data archives.
For more information, why not download the DRAMBORA Flyer in PDF or JPEG formats?
News
DRAMBORA for the ISO/TC 46/SC 11 standard
6th Feb 2009
DRAMBORA Auditor's Training Courses 2009
5th Jan 2009
DRAMBORA Interactive around the world
15th Dec 2008
Download
Please register to download the offline copy of the DRAMBORA Toolkit
Comment
Submit feedback via our feedback form or email us.
Get Help
Get help in the DRAMBORA evaluation process by volunteering to host a DRAMBORA facilitated assessment.
DRAMBORA Consortium Sponsors:
JISC
e-Science
Date Published: 1 February 2008 | Last Modified: 1 February 2008
© Digital Curation Centre and DigitalPreservationEurope, except where otherwise noted | Notices

Well-managed collections, pt. 1

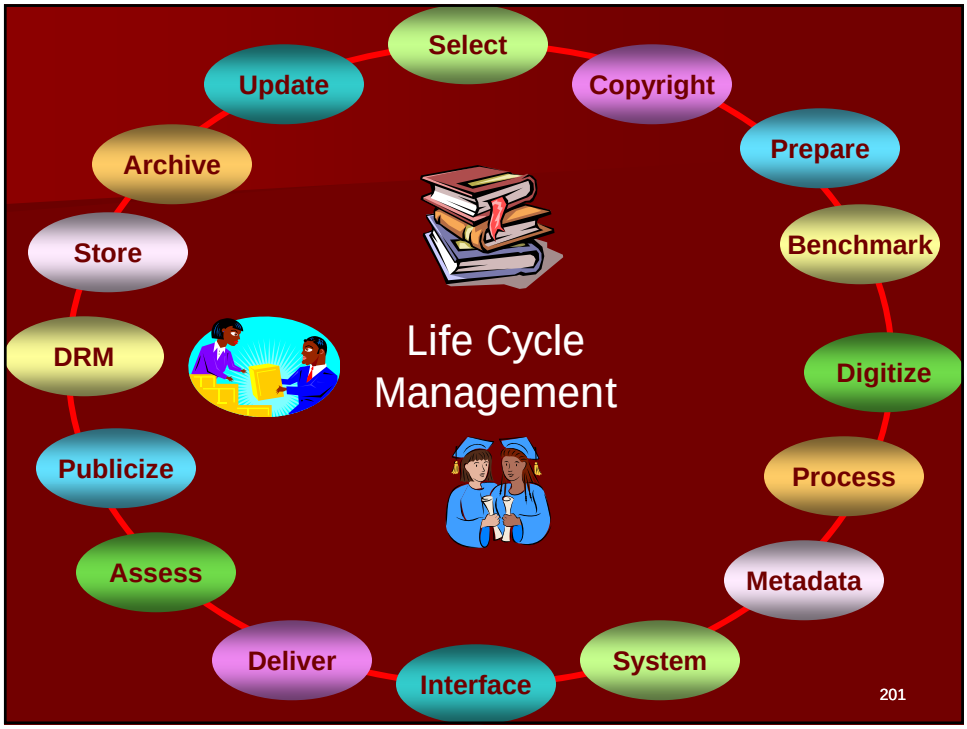
- Adhere to best practices in creating digital content (use common standards and avoid proprietary formats)
- Document your decisions (how files were created, technical specs, copyright issues)
- Create a registry of your digital collections
- Identify a team and assign responsibilities

199

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

200



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

202

Digital preservation planning case study

Digital Consulting and Production Services
Cornell University Library

203

Digital files of concern for DCAPS

204

Including the master file for this image

205

For that image file, metadata is here . . .

	B	C	D	E	F	G
1	DVD #	Box #	Burning Date	Coll.	file name	info
35991	568	8	3/12/2007	RMC	RMC2004_5028.tif	LunaTiffsRMC0072\sRGB\RMC2004_5028.tif
35992	568	8	3/12/2007	RMC	RMC2004_5032.tif	LunaTiffsRMC0072\sRGB\RMC2004_5032.tif
35993	568	8	3/12/2007	RMC	RMC2004_5033.tif	LunaTiffsRMC0072\sRGB\RMC2004_5033.tif
35994	568	8	3/12/2007	RMC	RMC2004_5039.tif	LunaTiffsRMC0072\sRGB\RMC2004_5039.tif
35995	568	8	3/12/2007	RMC	RMC2004_5040.tif	LunaTiffsRMC0072\sRGB\RMC2004_5040.tif
35996	568	8	3/12/2007	RMC	RMC2004_5043.tif	LunaTiffsRMC0072\sRGB\RMC2004_5043.tif
35997	568	8	3/12/2007	RMC	RMC2004_5044.tif	LunaTiffsRMC0072\sRGB\RMC2004_5044.tif
35998	568	8	3/12/2007	RMC	RMC2004_5066.tif	LunaTiffsRMC0072\sRGB\RMC2004_5066.tif
35999	568	8	3/12/2007	RMC	RMC2004_5070.tif	LunaTiffsRMC0072\sRGB\RMC2004_5070.tif
36000	568	8	3/12/2007	RMC	RMC2004_7044.tif	LunaTiffsRMC0072\GrayG22\RMC2004_7044.tif
36001	568	8	3/12/2007	RMC	RMC2005_0181.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0181.tif
36002	568	8	3/12/2007	RMC	RMC2005_0182.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0182.tif
36003	568	8	3/12/2007	RMC	RMC2005_0266.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0266.tif
36004	568	8	3/12/2007	RMC	RMC2005_0267.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0267.tif
36005	**568**	**8**	**3/12/2007**	**RMC**	**RMC2005_0268.tif**	**LunaTiffsRMC0072\adobe1998\RMC2005_0268.tif**
36006	568	8	3/12/2007	RMC	RMC2005_0269.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0269.tif
36007	568	8	3/12/2007	RMC	RMC2005_0270.tif	LunaTiffsRMC0072\sRGB\RMC2005_0270.tif
36008	568	8	3/12/2007	RMC	RMC2005_0374.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0374.tif
36009	568	8	3/12/2007	RMC	RMC2005_0375.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0375.tif
36010	568	8	3/12/2007	RMC	RMC2005_0386.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0386.tif
36011	568	8	3/12/2007	RMC	RMC2005_0387.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0387.tif
36012	568	8	3/12/2007	RMC	RMC2005_0388.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0388.tif
36013	568	8	3/12/2007	RMC	RMC2005_0397.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0397.tif
36014	568	8	3/12/2007	RMC	RMC2005_0402.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0402.tif
36015	568	8	3/12/2007	RMC	RMC2005_0415.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0415.tif
36016	568	8	3/12/2007	RMC	RMC2005_0416.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0416.tif
36017	568	8	3/12/2007	RMC	RMC2005_0417.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0417.tif
36018	568	8	3/12/2007	RMC	RMC2005_0418.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0418.tif
36019	568	8	3/12/2007	RMC	RMC2005_0419.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0419.tif
36020	568	8	3/12/2007	RMC	RMC2005_0460.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0460.tif
36021	568	8	3/12/2007	RMC	RMC2005_0461.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0461.tif

206

. . . and here . . .

Adobe

Adobe® RGB (1998) Color Image Encoding

Version 2005-05
May 2005

FILE INFORMATION

Product	ICC profiles
Version	CS4
Platform	Macintosh
File Name	AdobeICCProfilesCS4Mac_end-user.zip
File Size	6.2MB

Proceed to Download

207

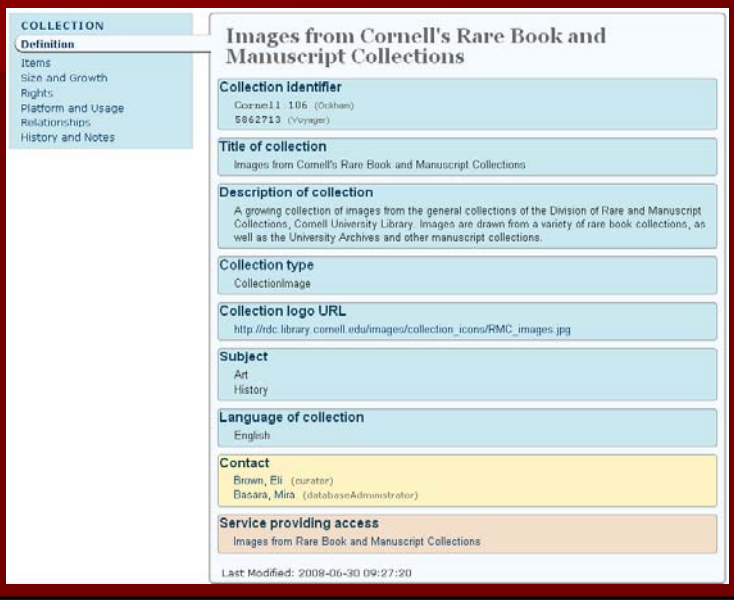
. . . and here . . .

Image ID:	105662
Resolution Size:	2
Format:	JPEG
Media Type:	Image
File Name:	RMC2005_0268.jpg
Width:	271
Height:	384

Title:	Hidden house sketch
Date:	1943
Classification:	Publications
Rights and Permissions:	Held by Artist (Creator), Publisher or Artist's (Creator's) Representatives
ID Number:	RMC2005_0268
Subcollection:	Rare Books
Collection Title Proper:	Illiterature : privately printed to please a few friends
Repository:	Division of Rare and Manuscript Collections, Cornell University Library
Location:	Rare Books PS3547.O9342 I29 Page 65
Voyager ID:	2363803

208

. . . and here . . .



209

. . . and here.

000 01025cam a22003011i 450
001 2363803
005 20070214082232.0
008 931215m19371943nyua 000 0 eng
035 __ **|a** (CStRLIN)NYCX93B114260
035 __ **|a** (NIC)notisALX7136
035 __ **|a** (OCoLC)64077320
035 __ **|a** 2363803
040 __ **|a** RPB **|c** RPB **|d** CStRLIN **|d** NIC
100 1_ **|a** Young, George.
245 10 **|a** Illiterature : **|b** vacation verses, privately printed to please a few friends / **|c** George and Helen Young.
260 __ **|a** Ithaca, N.Y. : **|b** [Cayuga Press] **|c** 1937-43.
300 __ **|a** 2 v. : **|b** ill. ; **|c** 20 cm.
500 __ **|a** Vol. 1 has subtitle: Vacation verses.
700 1_ **|a** Young, Helen.
905 __ **|a** 19940727120000.0
948 __ **|a** RET
950 __ **|l** OLIN **|a** PS3547.O9342 **|b** I29 **|v** v.1-2 **|i** 12/15/93 N
950 __ **|l** RARE **|a** PS3547.O9342 **|b** I29 **|i** 12/15/93 N
955 __ **|l** OLIN **|c** 1 **|i** 12/15/93 C
955 __ **|l** RARE **|c** 1:v.1 **|i** 12/15/93 C
955 __ **|l** RARE **|c** 1:v.2 **|i** 12/15/93 C
998 __ **|a** 12/15/93 **|t** c **|s** 9125 **|n** NIC **|w** RIBG5225736B **|d** 12/15/93 **|c** RET **|b** ANN **|i** 931215 **|l** NYCX
995 __ **|a** Hivolm **|d** 20070214

210

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

A: Assess and plan

211

Reflection

"Backing up digital objects is NOT preservation!"—Carl Grant

212

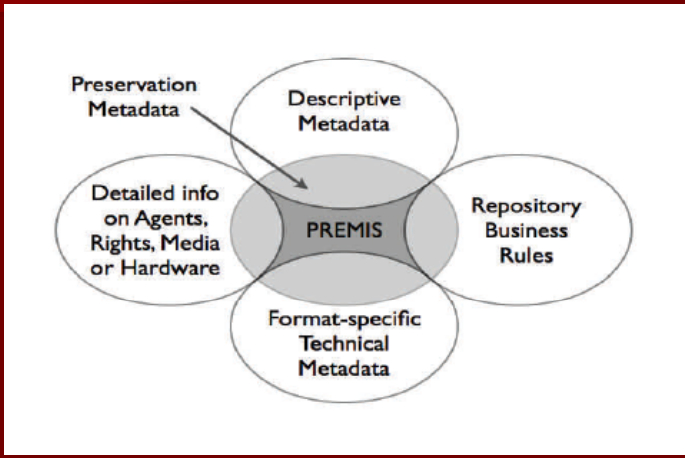
Exercise 11: Material and organizational assessment for digital preservation planning

213

Session 12: Metadata and ensuring access over time (collections)

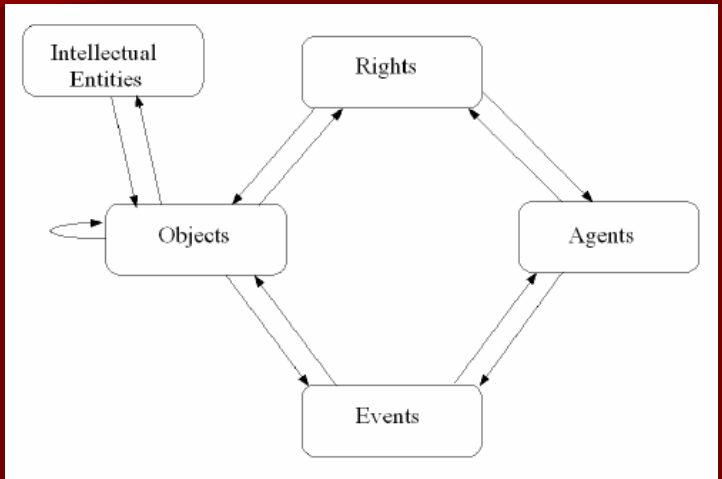
214

Preservation metadata in context



215

A preservation data model (PREMIS)



216

Another look at Cornell's digitization metadata

	B	C	D	E	F	G
1	DVD #	Box #	Burning Date	Coll.	file name	info
35991	568	8	3/12/2007	RMC	RMC2004_5028.tif	LunaTiffsRMC0072\sRGB\RMC2004_5028.tif
35992	568	8	3/12/2007	RMC	RMC2004_5032.tif	LunaTiffsRMC0072\sRGB\RMC2004_5032.tif
35993	568	8	3/12/2007	RMC	RMC2004_5033.tif	LunaTiffsRMC0072\sRGB\RMC2004_5033.tif
35994	568	8	3/12/2007	RMC	RMC2004_5039.tif	LunaTiffsRMC0072\sRGB\RMC2004_5039.tif
35995	568	8	3/12/2007	RMC	RMC2004_5040.tif	LunaTiffsRMC0072\sRGB\RMC2004_5040.tif
35996	568	8	3/12/2007	RMC	RMC2004_5043.tif	LunaTiffsRMC0072\sRGB\RMC2004_5043.tif
35997	568	8	3/12/2007	RMC	RMC2004_5044.tif	LunaTiffsRMC0072\sRGB\RMC2004_5044.tif
35998	568	8	3/12/2007	RMC	RMC2004_5066.tif	LunaTiffsRMC0072\sRGB\RMC2004_5066.tif
35999	568	8	3/12/2007	RMC	RMC2004_5070.tif	LunaTiffsRMC0072\sRGB\RMC2004_5070.tif
36000	568	8	3/12/2007	RMC	RMC2004_7044.tif	LunaTiffsRMC0072\GrayG22\RMC2004_7044.tif
36001	568	8	3/12/2007	RMC	RMC2005_0181.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0181.tif
36002	568	8	3/12/2007	RMC	RMC2005_0182.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0182.tif
36003	568	8	3/12/2007	RMC	RMC2005_0266.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0266.tif
36004	568	8	3/12/2007	RMC	RMC2005_0267.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0267.tif
36005	**568**	**8**	**3/12/2007**	**RMC**	**RMC2005_0268.tif**	**LunaTiffsRMC0072\adobe1998\RMC2005_0268.tif**
36006	568	8	3/12/2007	RMC	RMC2005_0269.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0269.tif
36007	568	8	3/12/2007	RMC	RMC2005_0270.tif	LunaTiffsRMC0072\sRGB\RMC2005_0270.tif
36008	568	8	3/12/2007	RMC	RMC2005_0374.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0374.tif
36009	568	8	3/12/2007	RMC	RMC2005_0375.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0375.tif
36010	568	8	3/12/2007	RMC	RMC2005_0386.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0386.tif
36011	568	8	3/12/2007	RMC	RMC2005_0387.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0387.tif
36012	568	8	3/12/2007	RMC	RMC2005_0388.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0388.tif
36013	568	8	3/12/2007	RMC	RMC2005_0397.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0397.tif
36014	568	8	3/12/2007	RMC	RMC2005_0402.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0402.tif
36015	568	8	3/12/2007	RMC	RMC2005_0415.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0415.tif
36016	568	8	3/12/2007	RMC	RMC2005_0416.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0416.tif
36017	568	8	3/12/2007	RMC	RMC2005_0417.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0417.tif
36018	568	8	3/12/2007	RMC	RMC2005_0418.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0418.tif
36019	568	8	3/12/2007	RMC	RMC2005_0419.tif	LunaTiffsRMC0072\GrayG22\RMC2005_0419.tif
36020	568	8	3/12/2007	RMC	RMC2005_0460.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0460.tif
36021	568	8	3/12/2007	RMC	RMC2005_0461.tif	LunaTiffsRMC0072\adobe1998\RMC2005_0461.tif

217

Format-specific technical characteristics

Adobe

Adobe® RGB (1998)
Color Image Encoding

Version 2005-05
May 2005

FILE INFORMATION

Product	ICC profiles
Version	CS4
Platform	Macintosh
File Name	AdobeICCProfilesCS4Mac_end-user.zip
File Size	6.2MB

Proceed to Download

218

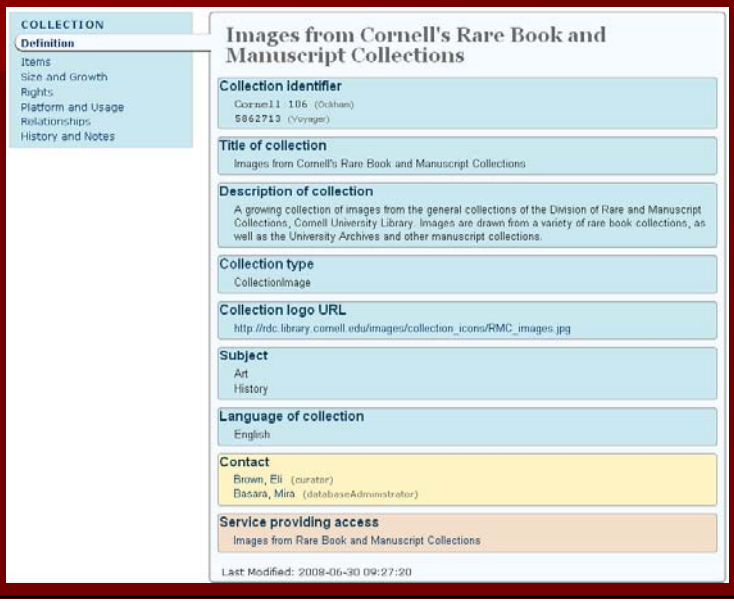
Object-level metadata

Image ID:	105662
Resolution Size:	2
Format:	JPEG
Media Type:	Image
File Name:	RMC2005_0268.jpg
Width:	271
Height:	384

Title:	Hidden house sketch
Date:	1943
Classification:	Publications
Rights and Permissions:	Held by Artist (Creator), Publisher or Artist's (Creator's) Representatives
ID Number:	RMC2005_0268
Subcollection:	Rare Books
Collection Title Proper:	Illiterature : privately printed to please a few friends
Repository:	Division of Rare and Manuscript Collections, Cornell University Library
Location:	Rare Books PS3547.O9342 I29 Page 65
Voyager ID:	2363803

219

Collection-level metadata



220

Metadata for a related entity

000 01025cam a22003011i 450
001 2363803
005 20070214082232.0
008 931215m19371943nyua 000 0 eng
035 __ **|a** (CStRLIN)NYCX93B114260
035 __ **|a** (NIC)notisALX7136
035 __ **|a** (OCoLC)64077320
035 __ **|a** 2363803
040 __ **|a** RPB **|c** RPB **|d** CStRLIN **|d** NIC
100 1_ **|a** Young, George.
245 10 **|a** Illiterature : **|b** vacation verses, privately printed to please a few friends / **|c** George and Helen Young.
260 __ **|a** Ithaca, N.Y. : **|b** [Cayuga Press] **|c** 1937-43.
300 __ **|a** 2 v. : **|b** ill. ; **|c** 20 cm.
500 __ **|a** Vol. 1 has subtitle: Vacation verses.
700 1_ **|a** Young, Helen.
905 __ **|a** 19940727120000.0
948 __ **|a** RET
950 __ **|l** OLIN **|a** PS3547.O9342 **|b** I29 **|v** v.1-2 **|i** 12/15/93 N
950 __ **|l** RARE **|a** PS3547.O9342 **|b** I29 **|i** 12/15/93 N
955 __ **|l** OLIN **|c** 1 **|i** 12/15/93 C
955 __ **|l** RARE **|c** 1:v.1 **|i** 12/15/93 C
955 __ **|l** RARE **|c** 1:v.2 **|i** 12/15/93 C
998 __ **|a** 12/15/93 **|t** c **|s** 9125 **|n** NIC **|w** RIBG5225736B **|d** 12/15/93 **|c** RET **|b** ANN **|i** 931215 **|l** NYCX
995 __ **|a** Hivolm **|d** 20070214

221

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*

222

Exercise 12: Collection-level metadata in a consortial context

223

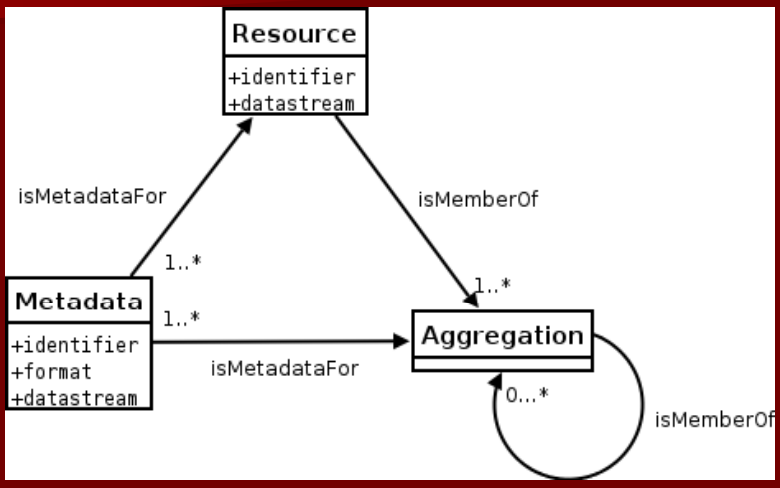
Day 4 debriefing

224

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

225

A data model relevant to our discussions



226

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*

227

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

228

Exercise 13: Object-level metadata in a repository context

229

Digital preservation citations that may interest you

http://www.citeulike.org/group/4576/tag/digital-preservation

230

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)

231

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

232

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

233

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

234

Staffing

Every project will vary

Project Staff

Project Director	Tony Gardner, *Curator of Special Collections and Archivist*
Project Manager	Robert Marshall, *Head Archivist*
Metadata Librarian	Mary Woodley, *Librarian*
Project Archivist	Charla Bench
Project Technician	Diane Dreiss
Metadata Specialists	Rachael Gordon, *Librarian* Adina Lerner, *Librarian* Sally Mendelsohn, *Admin. Support Asst.* Lori Saavedra, *Librarian* Jina Wakasono, *Librarian* Annie Zeidman, *Librarian*

Library Administration

Dean	Susan C. Curzon
Associate Dean	Susan Parker
Project Publicity	Cindy Vermilyeth, *Director of Development* Jocelyn Dunham, *Development Assistant*
Systems Administrator	Eric Willis
Web Designer	Hui-Ling Tang, *Library Assistant*

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

237

Brainstorming

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

238

Brainstorming techniques useful for

- Initiating institutional SWOT analysis
  - Strengths
  - Weaknesses
  - Opportunities
  - Threats
- Scope and nature of projects
- Selection

239

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

240

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

241

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

242

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

243

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

*From: Fisher, Roger, and William Fry. *Getting to Yes: Negotiating Agreement Without Giving In*. 2nd ed. New York: Penguin, 1991.

244

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

245

Active listening skills

1. Hear the message
2. Interpret the message
3. Evaluate the message
4. Respond to the message

246

Tips for effective listening

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

247

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

248

Reflection

"You can plan the plan, but you can't plan the outcome."—Words of wisdom

249

Planning process*

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

*Based on Bishoff and Allen (2004)

250

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?

251

Components of a needs analysis

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

252

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

253

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

254

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.

255

Benefits of solution

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

256

Benefits of solution

example

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

257

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.)

258

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

259

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

260

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"

261

Reflection

"There are no short-term cost savings to be realized by digitizing collections"—Lorna Hughes

262

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.

263

Cost categories

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

264

Relative Costs

Table 6: Comparative Cost Ratings Based on Overall Average Projections

Digitization Category	Digitizing	Metadata Creation	Other	Overall Costs
Mixed Collections	Higher	Slightly Higher	Lower	Slightly Higher
Single Items	Lower	Slightly Higher	Higher	Higher
Photographs	Slightly Higher	Lower	Average	Lower
Books/Pamphlets	Lower	Lower	Lower	Lower
Re-keyed Text	Lower	Lower	Lower	Lower
OCR	Very Low	Very Low	Very Low	Very Low

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

(From: Puglia, Steven. "Costs of Digital Imaging Projects" *RLG DigiNews* 3:5 (1999) <http://www.rlg.org/preserv/diginews/diginews3-5.html >)

Reported cost ranges

Table 5: Reported Cost Ranges for Various Digitization Processes

Digitization Category	Digitizing	Metadata Creation	Other	Overall Costs
Overall Projections	$0.25-$19.80	$0.75-$34.65	$0.45-$50.20	$1.85-$96.45
Adjusted Projections	$0.25-$16.65	$0.75-$17.25	$0.45-$28.15	$1.85-$42.45
Mixed Collections	$3.45-$16.50	$2.85-$17.25	$4.50-$21.55	$3.25-$40.50
Single Items	$1.90-$8.00	$5.75-$12.85	$7.60-$28.15	$23.10-$35.80
Photographs	$2.30-$16.65	$4.85-$6.45	$3.35-$24.65	$5.20-$42.45
Books/Pamphlets	$2.10-$6.10	$1.50-$11.10	$1.35-$6.90	$4.60-$14.40
Re-keyed Text	$2.55-$5.00	$2.35-$5.70	Limited Data	Limited Data
OCR	$0.25-$3.60	$0.75-$2.40	$0.40-$2.10	$1.85-$7.65

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

267

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

268

Staffing Costs

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

269

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

270

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

271

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

272

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

273

Request For Proposal (RFP)

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

274

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

275

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

276

Collaborative Digitization Program

http://www.bcr.org/cdp/

Website provides information about:

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

277

Exercise 14: Team building and work planning

278

Session 15: Project management—proposal writing and assessment

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

279

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

280

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

281

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?

282

Proposal components

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

283

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

284

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?

285

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 of collection of mixed media, rare books and archives ...

286

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)

287

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)

288

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.

289

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.

290

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?

291

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.

292

Project work plan

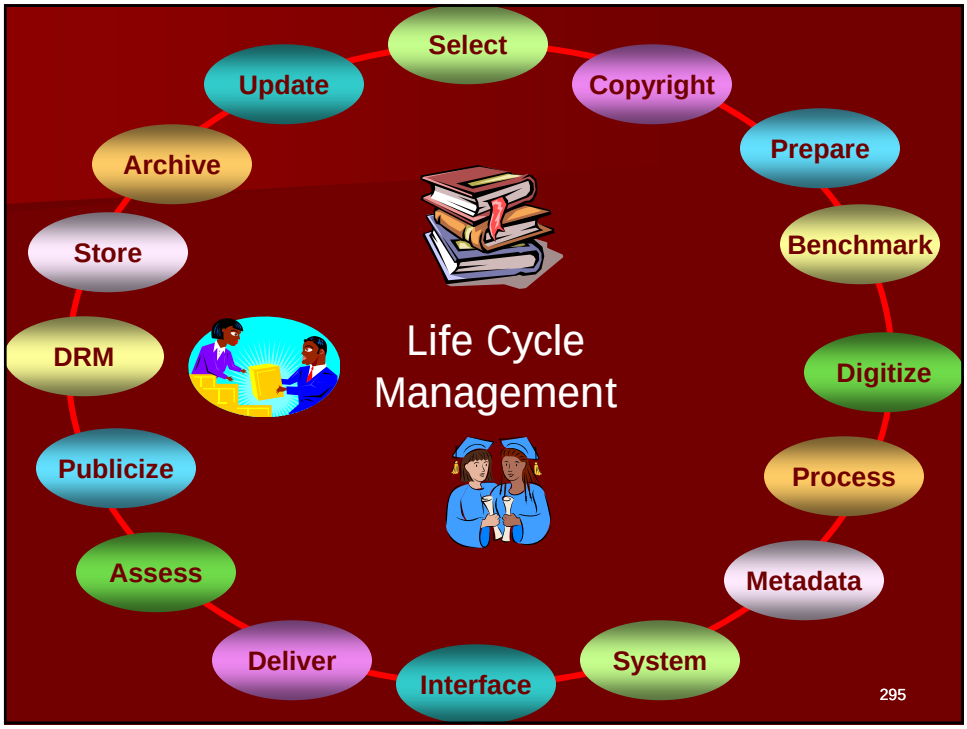
- What are the quantifiable outcomes?
- What is the work plan to accomplish project?
  - Time frame
  - Space
  - Equipment
  - Staffing
  - Software
  - Metadata
- How do the methods compare to other similar projects?

293

Digital life cycle

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

294



Digitization issues

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

296

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) 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.

297

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.

298

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

299

Project actions timeline

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

300

Proposed project budget

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

301

Example of budget summary

10. Budget Summary	LSTA (1)	Other funds (2)	In-kind (3)	Total (4)
a. Salaries & Benefits	$120,945	----------	$52,275	$173,220
b. Library Materials	0	----------	0	$ 5,000
c. Operation	$ 3,760	----------	0	$ 6,760
d. Equipment ($5K+)	0	----------	0	$ 7,000
e. Total for Objectives	$124,705	$15,000	$52,275	$191,980
f. Indirect Cost	$ 12,471			$ 12,471
g. TOTAL	$137,176			$204,451

302

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

303

Project marketing and publicity

304

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

305

Outcome-based evaluations consider:

- Impact and benefits that are the result of the project
- Short-term changes
- Long-term changes

306

Components of outcome-based evaluations

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

307

Typical inputs

- Staff
- Money
- Equipment

308

Typical activities

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

309

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

310

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

311

Typical outcome measures

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

312

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?

313

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

314

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

315

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

316

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

317

Accessibility

Can the site be used by anyone regardless of disability or impairment?

- Hearing access
- Vision access
- Mobility access
- Cognitive access

318

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

319

Exercise 15: Proposal writing and assessment

320

Day 5 and workshop debriefing

321

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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...

000 02463mgm a2200553 a 450
001 1968579
005 20070131091541.0
007 vf|cbahos
008 861120p19861981gw 150 e vlger d
020 __ |**a** 080010529X
035 __ |**a** (NIC)notisAKD5959
035 __ |**a** (OCoLC)23881014
035 __ |**a** 1968579
037 __ |**a** 60139 |**b** RCA/Columbia Pictures Home Video
040 __ |**a** NIC |**c** NIC
041 __ |**a** ger |**b** eng
245 04 |**a** Das Boot |**h** [videorecording] = The Boat / |**c** Columbia Pictures presents a Bavaria Atelier production in association with Radiant Films a film by Wolfgang Petersen ; produced by Günter Rohrbach ; co-producer, Michael Bittins ; written and directed by Wolfgang Petersen.
260 __ |**a** Burbank, CA : |**b** RCA/Columbia Pictures Home Video, |**c** [1986]
300 __ |**a** 1 videocassette (150 min.) : |**b** sd., col. ; |**c** 1/2 in.
538 __ |**a** VHS Hi-fi.
500 __ |**a** German dialogue, English subtitles.
511 1_ |**a** Jürgen Prochnow, Herbert Grönemeyer, Klaus Wennemann, Hubertus Bengsch, Martin Semmelrogge.
508 __ |**a** Director of photography, Jost Vacano; miniature photography, Ernst Wild; art director, Gotz Weidner; music, Klaus Doldinger; naval advisors, Achim Krug, Karl Böhm; film editor, Hannes

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** Grönemeyer, 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 a New Yorker Films release ; GFFA ; une production Jacque-Louis...

000 02313cgm a2200577 a 450
001 4699740
005 20060804141404.0
007 vd bvaizu
008 030416p20031934fr 087 g vlfre d
020 __ |**a** 1567301975
035 __ |**a** (MiU)notisBBV8582
035 __ |**a** (OCoLC)52077823
035 __ |**a** (III)iiio4197220x
035 __ |**a** (CStRLIN)MIUGBBV8582-F
035 __ |**a** 4699740
040 __ |**a** TXA |**c** TXA |**d** EYM
041 1_ |**a** fre |**b** eng
245 02 |**a** L'Atalante |**h** [videorecording] / |**c** a New Yorker Films release ; GFFA ; une production Jacque-Louis Nounez ; scénario, Jean Guinee ; adaptation et dialogues, Jean Vigo, Albert Riera.
250 __ |**a** Collectors ed.
260 __ |**a** [New York] : |**b** New Yorker Video, |**c** 2003.
300 __ |**a** 1 videodisc (87 min.) : |**b** sd., b&w ; |**c** 4 3/4 in.
538 __ |**a** DVD.
546 __ |**a** In French with optional English subtitles.
511 1_ |**a** Michel Simon, Dita Parló, Jean Daste, Gilles Margaritis, Maurice Gilles, Rafa Diligent, Louis Lefebvre.
508 __ |**a** Director, Jean Vigo; photography, Boris Kaufman; music, Maurice Jaubert.
500 __ |**a** Originally produced as motion picture in 1934.
520 __ |**a** 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 __ |**a** Special features: Documentary, The making of Jean Vigo's l'Atalante; still gallery; poster gallery; Jean Vigo filmography.
650 _0 |**a** Marriage |**v** Drama.
650 _0 |**a** Ship captain's spouses |**z** France |**v** Drama.
650 _0 |**a** Love |**v** Drama.
650 _0 |**a** Comedy films.
650 _0 |**a** Feature films.
650 _0 |**a** Motion pictures, French.
700 1_ |**a** Vigo, Jean, |**d** 1905-1934.
700 1_ |**a** Nounez, Jacques-Louis.

700 1_ |**a** Guinée, Jean.
700 1_ |**a** Simon, Michel, |**d** 1895-1975.
700 1_ |**a** Parlo, Orta.
700 1_ |**a** Dasté, Jean, |**d** 1904-
700 1_ |**a** Lefebvre, Louis, |**d** 1871-1947.
700 1_ |**a** Kaufman, Boris.
700 1_ |**a** Jaubert, Maurice.
710 2_ |**a** New Yorker Films.
710 2_ |**a** Gaumont Franco-Film Aubert (Firm)
710 2_ |**a** New Yorker Video (Firm)
902 __ |**a** pfnd |**b** Pumpelly
948 0_ |**a** 20030626 |**b** r |**d** ty15 |**e** cts
948 1_ |**a** 20030729 |**b** f |**d** sok1 |**e** cts
950 __ |**b** \VIDEO-D\35762-D
998 __ |**s** 9115

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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>
    <cast />
    <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.</description>
    <runtime>57</runtime>
    <sound>Yes</sound>
    <color>Yes</color>
    <creationDate>3/10/2006</creationDate>
</film>


<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>
```

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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

[Adapted from: *Moving Theory into Practice: Digital Imaging Tutorial*, Cornell University Library Research Dept., 2000-2003 <http://www.library.cornell.edu/preservation/tutorial/index.html>]

Legal: Is the material restricted because of privacy, content, or donor concerns?
Legal: Is it copyright protected? If so, do you have the right to create and disseminate digital reproductions?
Material: Does the material lend itself to digitization?
Material: Can the informational content be adequately captured in digital form?
Material: Do the physical formats and condition of the material represent major impediments?
Material: How large and complex in terms of object format is the collection?
Preservation: Would the material be put at risk in the digitization process?

Selection for Digitization Questionnaire

[Adapted from: *Moving Theory into Practice: Digital Imaging Tutorial*, Cornell University Library Research Dept., 2000-2003 <http://www.library.cornell.edu/preservation/tutorial/index.html>]

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?
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?
Organization: Is there adequate descriptive, navigational, or structural information about the material, such as bibliographic records or a detailed finding aid?
Use: What kinds, level, and frequency of use are envisioned? Is there a clear understanding of user requirements? Can digitization support these uses?
Use: Will access to the material be significantly enhanced by digitization?
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?)
Collections: Is there added incentive to digitize material based on the availability of complementary digital resources (including resources available from other organizations?)

Selection for Digitization Questionnaire

[Adapted from: *Moving Theory into Practice: Digital Imaging Tutorial*, Cornell University Library Research Dept., 2000-2003 <http://www.library.cornell.edu/preservation/tutorial/index.html>]

Collections: Is there an opportunity for multi-institutional cooperation?
Collections: Is there an opportunity for building thematic coherence or "critical mass?"
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?
Infrastructure: What components does your institution already have in place to support managing, delivering, and maintaining digitized materials? What components are still needed?
Infrastructure: Do your principal users have adequate computing and connectivity to make effective use of these materials?
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?
Fiscal: Are there funds to support the digitization effort? Is there institutional commitment to the on-going management and preservation of these files?

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.

W. H. Harrison.

Written by

William Hayden, Esq.

and respectfully dedicated to the

Price 25 cts. nett.

Whigs of the United States.

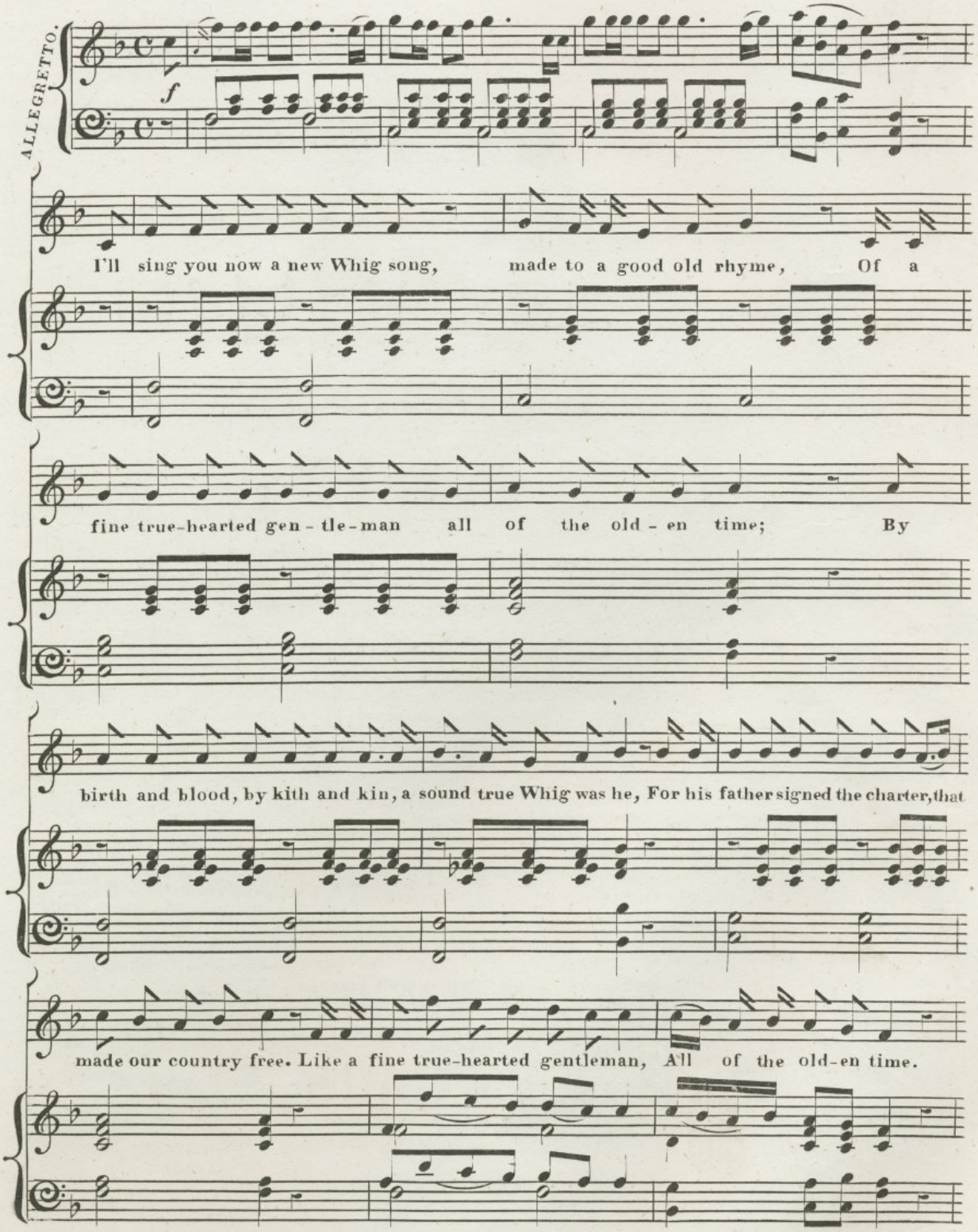
BOSTON.

Published by PARKER & DITSON, *135 Washington St.*

Entered according to Act of Congress in the year 1840 by Parker & Ditson, in the Clerks office of the District Court of Massachusetts.

THE NATIONAL WHIG SONG.

Written by W. HAYDEN, Esq. adapted to a Popular Air.



Entered according to Act of Congress, in the year 1840 by Parker & Ditson, in the Clerk's Office of the District Court of Massachusetts.

Evening Grosbeak.

1. Male. 2. Female. 3 Young Male.

Drawn from Nature by J. J. Audubon. F.R.S.F.L.S.

Lith.d Printed & Col.d by J. T. Bowen. Philad.

Metadata Template

Data Property (Field)	Data Value (Content)	Controlled Vocabulary Yes or No Specify, if any

Dublin Core Metadata Element Set, Version 1.1

Identifier:	http://dublincore.org/documents/2008/01/14/dces/
Supersedes:	http://dublincore.org/documents/2006/12/18/dces/
Latest version:	http://dublincore.org/documents/dces/
Date Issued:	2008-01-14
Status of document:	This is a DCMI Recommendation.
Description of document:	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 "DCMI Metadata Terms".
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.
References:	[TGN] http://www.getty.edu/research/tools/vocabulary/tgn/index.html
Term Name: creator	
URI:	http://purl.org/dc/elements/1.1/creator
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.

Term Name: date	
URI:	http://purl.org/dc/elements/1.1/date
Label:	Date
Definition:	A point or period of time associated with an event in the lifecycle of the resource.
Comment:	Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF].
References:	[W3CDTF] http://www.w3.org/TR/NOTE-datetime
Term Name: description	
URI:	http://purl.org/dc/elements/1.1/description
Label:	Description
Definition:	An account of the resource.
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.
Term Name: format	
URI:	http://purl.org/dc/elements/1.1/format
Label:	Format
Definition:	The file format, physical medium, or dimensions of the resource.
Comment:	Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types [MIME].
References:	[MIME] http://www.iana.org/assignments/media-types/
Term Name: identifier	
URI:	http://purl.org/dc/elements/1.1/identifier
Label:	Identifier
Definition:	An unambiguous reference to the resource within a given context.
Comment:	Recommended best practice is to identify the resource by means of a string conforming to a formal identification system.
Term Name: language	
URI:	http://purl.org/dc/elements/1.1/language
Label:	Language
Definition:	A language of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as RFC 4646 [RFC4646].
References:	[RFC4646] http://www.ietf.org/rfc/rfc4646.txt
Term Name: publisher	
URI:	http://purl.org/dc/elements/1.1/publisher
Label:	Publisher
Definition:	An entity responsible for making the resource available.
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.

Term Name: relation	
URI:	http://purl.org/dc/elements/1.1/relation
Label:	Relation
Definition:	A related resource.
Comment:	Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.
Term Name: rights	
URI:	http://purl.org/dc/elements/1.1/rights
Label:	Rights
Definition:	Information about rights held in and over the resource.
Comment:	Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.
Term Name: source	
URI:	http://purl.org/dc/elements/1.1/source
Label:	Source
Definition:	A related resource from which the described resource is derived.
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.
Term Name: subject	
URI:	http://purl.org/dc/elements/1.1/subject
Label:	Subject
Definition:	The topic of the resource.
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.
Term Name: title	
URI:	http://purl.org/dc/elements/1.1/title
Label:	Title
Definition:	A name given to the resource.
Comment:	Typically, a Title will be a name by which the resource is formally known.
Term Name: type	
URI:	http://purl.org/dc/elements/1.1/type
Label:	Type
Definition:	The nature or genre of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary [DCMITYPE]. To describe the file format, physical medium, or dimensions of the resource, use the Format element.
References:	[DCMITYPE] http://dublincore.org/documents/dcmi-type-vocabulary/

Errata 2008-06-09. Updated URIs for ANSI/NISO Z39.85 and ISO 15836.

DCMI Type Vocabulary

Title:	DCMI Type Vocabulary
Creator:	DCMI Usage Board
Identifier:	http://dublincore.org/documents/2008/01/14/dcmi-type-vocabulary/
Date Issued:	2008-01-14
Latest Version:	http://dublincore.org/documents/dcmi-type-vocabulary/
Replaces:	http://dublincore.org/documents/2006/08/28/dcmi-type-vocabulary/
Replaced By:	Not applicable
Translations:	http://dublincore.org/resources/translations/
Document Status:	This is a DCMI Recommendation.
Description:	The DCMI Type Vocabulary provides a general, cross-domain list of approved terms that may be used as values for the Resource Type element to identify the genre of a resource. The terms documented here are also included in the more comprehensive document "DCMI Metadata Terms" at http://dublincore.org/documents/dcmi-terms/.
Term Name: Collection	
URI:	http://purl.org/dc/dcmitype/Collection
Label:	Collection
Definition:	An aggregation of resources.
Comment:	A collection is described as a group; its parts may also be separately described.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Collection-003
Term Name: Dataset	
URI:	http://purl.org/dc/dcmitype/Dataset
Label:	Dataset
Definition:	Data encoded in a defined structure.
Comment:	Examples include lists, tables, and databases. A dataset may be useful for direct machine processing.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Dataset-003

Term Name: Event	
URI:	http://purl.org/dc/dcmitype/Event
Label:	Event
Definition:	A non-persistent, time-based occurrence.
Comment:	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.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Event-003
Term Name: Image	
URI:	http://purl.org/dc/dcmitype/Image
Label:	Image
Definition:	A visual representation other than text.
Comment:	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.
Type of Term:	Class
Broader Than:	http://purl.org/dc/dcmitype/StillImage
Broader Than:	http://purl.org/dc/dcmitype/MovingImage
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Image-004
Term Name: InteractiveResource	
URI:	http://purl.org/dc/dcmitype/InteractiveResource
Label:	Interactive Resource
Definition:	A resource requiring interaction from the user to be understood, executed, or experienced.
Comment:	Examples include forms on Web pages, applets, multimedia learning objects, chat services, or virtual reality environments.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#InteractiveResource-003

Term Name: MovingImage	
URI:	http://purl.org/dc/dcmitype/MovingImage
Label:	Moving Image
Definition:	A series of visual representations imparting an impression of motion when shown in succession.
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.
Type of Term:	Class
Narrower Than:	http://purl.org/dc/dcmitype/Image
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#MovingImage-003
Term Name: PhysicalObject	
URI:	http://purl.org/dc/dcmitype/PhysicalObject
Label:	Physical Object
Definition:	An inanimate, three-dimensional object or substance.
Comment:	Note that digital representations of, or surrogates for, these objects should use Image, Text or one of the other types.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#PhysicalObject-003
Term Name: Service	
URI:	http://purl.org/dc/dcmitype/Service
Label:	Service
Definition:	A system that provides one or more functions.
Comment:	Examples include a photocopying service, a banking service, an authentication service, interlibrary loans, a Z39.50 or Web server.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Service-003
Term Name: Software	
URI:	http://purl.org/dc/dcmitype/Software
Label:	Software
Definition:	A computer program in source or compiled form.
Comment:	Examples include a C source file, MS-Windows .exe executable, or Perl script.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Software-003

Term Name: Sound	
URI:	http://purl.org/dc/dcmitype/Sound
Label:	Sound
Definition:	A resource primarily intended to be heard.
Comment:	Examples include a music playback file format, an audio compact disc, and recorded speech or sounds.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Sound-003
Term Name: StillImage	
URI:	http://purl.org/dc/dcmitype/StillImage
Label:	Still Image
Definition:	A static visual representation.
Comment:	Examples include paintings, drawings, graphic designs, plans and maps. Recommended best practice is to assign the type Text to images of textual materials. Instances of the type Still Image must also be describable as instances of the broader type Image.
Type of Term:	Class
Narrower Than:	http://purl.org/dc/dcmitype/Image
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#StillImage-003
Term Name: Text	
URI:	http://purl.org/dc/dcmitype/Text
Label:	Text
Definition:	A resource consisting primarily of words for reading.
Comment:	Examples include books, letters, dissertations, poems, newspapers, articles, archives of mailing lists. Note that facsimiles or images of texts are still of the genre Text.
Type of Term:	Class
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Text-003

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.)

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

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.

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

Exercise 5: Metadata analysis

Tools:

Six metadata examples (MARC record, Simple DC record, EAD document, HMTL 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.

524 __ |a Prince, William G. Vietnam War: Statistical Analysis and Evaluation Systems, #4406. Division of Rare and Manuscript Collections, Cornell University Library.

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/">
        <title>Guide to the Vietnam War: Statistical Analysis and Evaluation
        Projects, 1968-1972</title>
        <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>
             <titleproper>Guide to the Vietnam War: Statistical Analysis and Evaluation
             Projects, 1968-1972</titleproper>
             <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>
             Martin Heggestad, January 2002. File last modified:
             <date>January 2002</date>
          </creation>
          <langusage>
             <language>ENG</language>
          </langusage>
       </profiledesc>
    </eadheader>
    <frontmatter>
      <titlepage>
          <titleproper>
             Guide to the Vietnam War: Statistical Analysis and Evaluation Projects,
             <lb />
             1968-1972
          </titleproper>
          <num>Collection Number: 4406</num>
          <publisher>
             Division of Rare and Manuscript Collections
             <lb />
             Cornell University Library
          </publisher>
          <list type="deflist">
             <defitem>
                <label>Contact Information:</label>
                <item>&RMCaddress;</item>
             </defitem>
             <defitem>
                <label>Compiled by:</label>
                <item>E. Engst</item>
             </defitem>
```

```
            <defitem>
              <label>Date completed:</label>
              <item>October 1989</item>
            </defitem>
            <defitem>
              <label>EAD encoding:</label>
              <item>Martin Heggestad, January 2002</item>
            </defitem>
          </list>
          <date>© 2002 Division of Rare and Manuscript Collections, Cornell University
            Library</date>
        </titlepage>
      </frontmatter>
    <archdesc level="collection" langmaterial="ENG">
      <did>
        <head>DESCRIPTIVE SUMMARY</head>
        <unittitle label="Title:">
          Vietnam War: statistical analysis and evaluation projects,
          <unitdate type="inclusive">1968-1972</unitdate>
        </unittitle>
        <unitid label="Collection Number:">4406</unitid>
        <origination label="Creator:">
          <persname normal="Prince, William G.">William G. Prince</persname>
        </origination>
        <physdesc label="Quantity:">.7 cubic ft.</physdesc>
        <physdesc label="Forms of Material:">Manuals, reports, questionnaires,
          correspondence, and other documents.</physdesc>
        <repository label="Repository:">Division of Rare and Manuscript Collections,
          Cornell University Library</repository>
        <abstract label="Abstract:">Documentation for various automated systems
          designed to provide data about the war in Vietnam.</abstract>
      </did>
      <scopecontent>
        <head>COLLECTION DESCRIPTION</head>
        <p>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.</p>
        <p>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.</p>
</scopecontent>
<controlaccess>
  <head>SUBJECTS</head>
  <controlaccess>
     <head>Names:</head>
     <persname encodinganalog="MARC 100">Prince, William G.</persname>
  </controlaccess>
<controlaccess>
  <head>Subjects:</head>
  <subject encodinganalog="MARC 650">Vietnamese Conflict, 1961-1975.</subject>
  <subject encodinganalog="MARC 650">Combat--Statistics--Information
     sources.</subject>
  <subject encodinganalog="MARC 650">Internal security--Vietnam.</subject>
  <subject encodinganalog="MARC 650">Insurgency--Vietnam.</subject>
  <subject encodinganalog="MARC 650">Military art and science--Data
     processing.</subject>
  <subject encodinganalog="MARC 650">Military art and science--
     Automation.</subject>
  <subject encodinganalog="MARC 650">Military assistance, American--Southeast
     Asia--Computer programs.</subject>
 </controlaccess>
</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">
  <relatedmaterial>
     <head>RELATED MATERIAL</head>
     <p>
        Additional pamphlets and reports by William G. Prince are housed in the
        <extref
        href="http://www.library.cornell.edu/Asia/ECHOLS/index.htm">Echols
        Collection</extref>
        , Olin Library, Cornell University.
     </p>
  </relatedmaterial>
</add>
<organization>
  <head>SERIES LIST</head>
     <list type="deflist">
     <defitem>
        <label>Series I. Project Corona Harvest</label>
        <item>Box 1</item>
     </defitem>
```

```
      <defitem>
         <label>Series II. HES 70</label>
         <item>Box 1</item>
      </defitem>
      <defitem>
         <label>Series III. SEAPRS</label>
         <item>Box 1</item>
      </defitem>
      <defitem>
         <label>Series IV. PAAS</label>
         <item>Box 1</item>
      </defitem>
      <defitem>
         <label>Series V. SEER</label>
         <item>Box 1</item>
      </defitem>
      <defitem>
         <label>Series VI. AIRSUM</label>
         <item>Boxes 1 and 2</item>
      </defitem>
   </list>
</organization>
<dsc type="combined">
   <head>CONTAINER LIST</head>
   <c01 level="series">
      <did>
         <unittitle>Series I. Project Corona Harvest</unittitle>
      </did>
      <scopecontent>
         <p>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).</p>
      </scopecontent>
      <c02 level="file">
         <did>
            <container type="box">1</container>
            <container type="folder">1</container>
            <unittitle>
               Operating instructions,
               <unitdate>1968</unitdate>
            </unittitle>
         </did>
      </c02>
</c01>
<c01 level="series">
   <did>
      <unittitle>Series II. HES 70</unittitle>
   </did>
   <scopecontent>
      <p>Hamlet Evaluation System--reporting system designed to gather data on
      the progress of the rural pacification effort.</p>
      <p>Data gathering instrument--formatted, multiple choice questionnaire:
         <list type="ordered" numeration="arabic">
```

```
            <item>Respondent--U.S. advisors in the field</item>
            <item>Sample size--Every inhabited hamlet and village in south
               Vietnam</item>
            <item>Frequency of reporting--Every hamlet and village reported each
               month</item>
         </list>
      </p>
</scopecontent>
<c02 level="file">
  <did>
      <container type="box">1</container>
      <container type="folder">2</container>
      <unittitle>Comparison of HES 70 and PAAS</unittitle>
  </did>
</c02>
<c02 level="file">
  <did>
      <container type="box">1</container>
      <container type="folder">3</container>
      <unittitle>Memo,
         <unitdate>1970</unitdate>
      </unittitle>
  </did>
</c02>
<c02 level="file">
  <did>
      <container type="box">1</container>
      <container type="folder">4</container>
      <unittitle>Appendix A--Question codes, question responses</unittitle>
  </did>
</c02>
<c02 level="file">
  <did>
      <container type="box">1</container>
      <container type="folder">5</container>
      <unittitle>Hamlet level HES statistics and plots by NMCSSC for village program
         analysis (1)</unittitle>
  </did>
</c02>
<c02 level="file">
  <did>
      <container type="box">1</container>
      <container type="folder">6</container>
      <unittitle>VSSG IDX (2)</unittitle>
  </did>
</c02>
<c02 level="file">
  <did>
      <container type="box">1</container>
      <container type="folder">7</container>
      <unittitle>Security (3)</unittitle>
  </did>
```

```
<html>
  <head>
    <title>Cornell University Archival Guides</title>
    <link rel="stylesheet" href="/f/findaid/findaid.css">
  </head>

    <table border="0" cellpadding="10" width="100%" cellspacing="0">
      <tr>
        <td>
             <center> <br><br><h1>Guide to the Vietnam War: Statistical Analysis
and Evaluation Projects, </h1><h1>1968-1972</h1><br>
 <h3>Collection Number: <span class="hilite">4406</span></h3>
 <br><h3>Division of Rare and Manuscript Collections <br>Cornell University
Library</h3>
 <table>
<tr>
 <td valign="top"> <b>Contact Information:</b><br>
 Division of Rare and Manuscript Collections<br>
 2B Carl A. Kroch Library<br>
 Cornell University<br>
 Ithaca, NY 14853<br>
 (607) 255-3530<br>
 Fax: (607) 255-9524<br>
 <a HREF="mailto:rareref@cornell.edu">rareref@cornell.edu</a><br>
 <a HREF="http://rmc.library.cornell.edu">http://rmc.library.cornell.edu</a><br>

 </td>
 <td valign="top"> <b>Compiled by:</b><br>
 E. Engst
 </td>
 <td valign="top"> <b>Date completed:</b><br>
 October 1989
 </td>
 <td valign="top"> <b>EAD encoding:</b><br>
 Martin Heggestad, January 2002
 </td>
 </tr>
</table>
 <h3>© 2002 Division of Rare and Manuscript Collections, Cornell University
Library</h3><br><br>
 </center>   <h3>DESCRIPTIVE SUMMARY</h3><ul>
 <strong>Title:</strong> Vietnam War: statistical analysis and evaluation
projects, 1968-1972 <br>
 <Strong>Collection No.:</strong> <span class="hilite">4406</span><br>
<strong>Creator:</strong>  William G. Prince <br>
 <strong>Quantity:</strong> .7 cubic ft.<br>
 <strong>Forms of Material:</strong> Manuals, reports, questionnaires,
correspondence, and other documents.<br>
 <strong>Repository:</strong> Division of Rare and Manuscript Collections,
Cornell University Library<br>
 <strong>Abstract:</strong><blockquote>Documentation for various automated
systems designed to provide data about the war in Vietnam.</blockquote>
 </ul> <h3>COLLECTION DESCRIPTION</h3><blockquote> <P>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.</P>
<P>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.</P> <br>
</blockquote> <h3>SUBJECTS</h3>
<blockquote>
 <b>Names:</b><ul>
 <li>Prince, William G.</li>
 </ul>
<hr>
 <b>Subjects:</b><ul>
 <li>Vietnamese Conflict, 1961-1975.</li>
 <li>Combat -- Statistics -- Information sources.</li>
 <li>Internal security -- Vietnam.</li>
 <li>Insurgency -- Vietnam.</li>
 <li>Military art and science -- Data processing.</li>
 <li>Military art and science -- Automation.</li>
 <li>Military assistance, American -- Southeast Asia -- Computer programs.</li>
 </ul>
<hr>
 </blockquote>
 <h3>INFORMATION FOR USERS</h3><ul>
 <Strong>Cite As:</strong> <P>Vietnam War: Statistical Analysis and Evaluation
Projects, #<span class="hilite">4406</span>. Division of Rare and Manuscript
Collections, Cornell University Library.</P>
 </ul> <ADD TYPE="relatedmaterial"> <strong>RELATED
MATERIAL</strong><blockquote>
 <P>Additional pamphlets and reports by William G. Prince are housed in the <a
HREF="http://www.library.cornell.edu/Asia/ECHOLS/index.htm">Echols
Collection</a>, Kroch Library, Cornell University.</P> </blockquote>
 </blockquote>
 <h3>SERIES LIST</h3>
<blockquote>
 <ol>
 <DEFITEM> <LABEL>Series I. Project Corona Harvest</LABEL> <li>
Box 1</li>
 </DEFITEM> <DEFITEM> <LABEL>Series II. HES 70</LABEL> <li>
Box 1</li>
 </DEFITEM> <DEFITEM> <LABEL>Series III. SEAPRS</LABEL> <li>
Box 1</li>
 </DEFITEM> <DEFITEM> <LABEL>Series IV. PAAS</LABEL> <li>
Box 1</li>
 </DEFITEM> <DEFITEM> <LABEL>Series V. SEER</LABEL> <li>
Box 1</li>
 </DEFITEM> <DEFITEM> <LABEL>Series VI. AIRSUM</LABEL> <li>
Boxes 1 and 2</li>
```

```
 </DEFITEM> </ol>
 </blockquote> <h3>CONTAINER LIST</h3>
<table width="98%" border="0" cellspacing="1" cellpadding="5">
<tr valign="bottom">
<th align="left"><u>Description</u></th>
<th align="center" colspan="2">  <u>Container</u>  </th>
</tr>
<tr>
<td width="78%"><div class="otlc01head"><span class="serieshead"> Series I.
Project Corona Harvest </span>  <div class="clistScopeNote"><b>Description:
</b>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).</div>
        </div>
</td>
<td align="center" valign="top" colspan="2"> </td>
</tr>
<tr>
<td width="78%"><div class="otlc02head"> Operating instructions, 1968      
</div>
</td>
<td align="center" valign="top" width="11%">Box 1</td>
<td align="center" valign="top" width="11%">Folder 1</td>
</tr>
<tr>
<td width="78%"><div class="otlc01head"><span class="serieshead"> Series II. HES
70 </span>        </div>
</td>
<td align="center" valign="top" colspan="2"> </td>
</tr>
<tr>
<td width="78%"><div class="otlc02head"> Comparison of HES 70 and PAAS
    </div>
</td>
<td align="center" valign="top" width="11%">Box 1</td>
<td align="center" valign="top" width="11%">Folder 2</td>
</tr>
<tr>
<td width="78%"><div class="otlc02head"> Memo, 1970         </div>
</td>
<td align="center" valign="top" width="11%">Box 1</td>
<td align="center" valign="top" width="11%">Folder 3</td>
</tr>
<tr>
<td width="78%"><div class="otlc02head"> Appendix A--Question codes, question
responses         </div>
</td>
<td align="center" valign="top" width="11%">Box 1</td>
<td align="center" valign="top" width="11%">Folder 4</td>
</tr>
<tr>
<td width="78%"><div class="otlc02head"> Hamlet level HES statistics and plots
by NMCSSC for village program analysis (1)         </div>
</td>
<td align="center" valign="top" width="11%">Box 1</td>
<td align="center" valign="top" width="11%">Folder 5</td>
</tr>
```

Guide to the Vietnam War: Statistical Analysis and Evaluation Projects, 1968-1972

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:
Vietnamese Conflict, 1961-1975.
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	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

Description	Container
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
Soc. ben. (7)	Box 1 Folder 11
GVN pol. inf. (8)	Box 1 Folder 12
VC pol. inf.	Box 1 Folder 13
Prog. effort	Box 1 Folder 14
Misc. ques.	Box 1 Folder 15
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:

_X_ Good ___ Fair ___ Poor ___ Other (explain)

Related MSS Collections or Archival Holdings:

Other reports by Prince in Echols

Processing Dates, Personnel, Activities:

Sept 15, 1989 – Oct 10, 1989
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

	File Format	Type of Metadata	Metadata Scheme	Content Values	Structure	Intended Use
MARC Record (pp. 13-14)						
Dublin Core Record (p. 15)						
EAD Document (pp. 16-20)						
HTML source code (pp. 21-23)						
Paper Archival Guide (pp. 24-26)						
Paper Accession Record (p. 27)						

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.

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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, HMTL 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?

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.

Collection	PhysicalObject
Dataset	Service
Event	Software
Image	Sound
InteractiveResource	StillImage
MovingImage	Text

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.

```
<journal_issue timestamp="2002-04-05T12:02:12Z">
<issue_data>
  <publisher>Applied Mathematics Publisher</publisher>
  <journal_name>Journal of Reapplied Probability</journal_name>
  <journal_citation_name>J. Reappl. Probab.</journal_citation_name>
  <issn type="electronic">5432-9876</issn>
  <journal_vol_number>39</journal_vol_number>
  <issue_number>1</issue_number>
  <issue_publ_date iso8601="2002-03">March 2002</issue_publ_date>
</issue_data>

<record lang="EN" type="article">
  <identifiers>
    <identifier type="pii">10201</identifier>
    <identifier type="doi">10.1239/10197321344243</identifier>
  </identifiers>
  <title lang="EN">Some risk management problems for firms with internal
      competition and debt</title>
  <author order="1">
    <name>
      <given_name>Xin</given_name>
      <surname>Zang</surname>
    </name>
    <email>zang@us.ibm.com</email>
    <affiliation>
      <organization>IBM</organization>
      <address>
        <addressline>IBM T. J. Watson Research Center</addressline>
        <addressline>PO Box 218</addressline>
        <addressline>Yorktown Heights</addressline>
        <addressline>NY 10598</addressline>
        <addressline>USA</addressline>
      </address>
    </affiliation>
  </author>
  <abstract>
      <p>Consider an optimization of the <i>Swigler</i> problem, first
      formulated by Kunst in <i>Liability Constant Rates</i>: a constant
      liability payment rate <b>B</b>, an average return <b>R</b>, and a risk
      <b>N</b>xy proportional to the size of the business unit.</p>
  </abstract>
  <keywords>
    <keyword>Hamilton-Jacobi equation</keyword>
    <keyword>singular control</keyword>
    <keyword>nonlinear control</keyword>
    <keyword>dividend optimization</keyword>
    <keyword>internal competition</keyword>
  </keywords>
  <subjects>
    <subject scheme="msc2000" rank="primary">93E20</subject>
    <subject scheme="msc2000" rank="secondary">49L99</subject>
  </subjects>
  <start_page>55</start_page>
  <end_page>69</end_page>
  <record_filename type="pdf"/>
</record>
</journal_issue>
```

```
<journal_issue timestamp="2002-05-28T18:22:14Z">
<issue_data>
  <publisher>Dorest Publishing Corporation</publisher>
  <journal_name>Journal of Applied Mathematics</journal_name>
  <journal_citation_name>J. Appl. Math.</journal_citation_name>
  <issn type="print">1234-567x</issn>
  <journal_vol_number>2</journal_vol_number>
  <issue_number label="Number">2</issue_number>
  <issue_publ_date iso8601="2002-05-08">8 May 2002</issue_publ_date>
</issue_data>

<record lang="EN" type="article">
  <identifiers>
    <identifier type="pii">S1110757X02000086</identifier>
    <identifier type="doi">10.1155/S1110757X03457834</identifier>
  </identifiers>
  <title lang="EN">Radiation effect on MHD free-convection flow of a gas at
      a stretching surface with a uniform free stream</title>
  <author order="1">
    <name>
        <given_name>Ahmed D.</given_name>
        <surname>Soud</surname>
      </name>
  </author>
  <author order="2">
      <name>
        <given_name>Jonathan P. T.</given_name>
        <surname>Hartbary, Jr.</surname>
      </name>
  </author>
  <abstract>
      <p>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 <math
      alttext="$F$"><mi>F</mi></math>, free convection parameter <math
      alttext="$\mathrm{Gr}$"><mrow><mtext>Gr</mtext></mrow></math>, magnetic
      parameter <math alttext="M"><mi>M</mi></math>, Prandtl number <math
      alttext="$\mathrm{Pr}$"><mrow><mtext>Pr</mtext></mrow></math>, and the
      parameter of relative difference between the temperature of the sheet, and
      the temperature far away from the sheet <math
      alttext="$r$"><mi>r</mi></math>. The effects of the radiation and magnetic
      field parameters on the shear stress and heat flux are discussed.</p>
  </abstract>
  <subjects>
      <subject scheme="msc2000" rank="primary">76D10</subject>
      <subject scheme="msc2000" rank="primary">76R10</subject>
      <subject scheme="msc2000" rank="secondary">76W05</subject>
  </subjects>
  <start_page>93</start_page>
  <end_page>103</end_page>
  <record_filename type="pdf"/>
</record>
</journal_issue>
```

```
<journal_issue timestamp="2000-06-08T09:15:56Z">
<issue_data>
  <publisher>The University of Plymouth, Department of Mathematics</publisher>
  <journal_name>The Plymouth Mathematical Journal</journal_name>
  <journal_citation_name>Plymouth Math. J.</journal_citation_name>
  <issn type="print">0026-009x</issn>
  <journal_vol_number>47</journal_vol_number>
  <issue_number label="Issue">2</issue_number>
  <issue_publ_date iso8601="2000">2000</issue_publ_date>
</issue_data>

<record lang="EN" type="article">
  <identifiers>
    <identifier type="pii">2000|0047|0211|0215</identifier>
    <identifier type="mr">2001k:34565</identifier>
  </identifiers>
  <title lang="EN">The $C^{1,1}$ regularity of the pluricomplex Brown
      function</title>
  <author>
    <name>
      <given_name>Zbigniew</given_name>
      <surname>B&#x0142;oclaski</surname>
    </name>
  </author>
  <subjects>
    <subject rank="primary" scheme="msc">32U35</subject>
    <subject rank="secondary" scheme="msc">32W20</subject>
  </subjects>
  <start_page>211</start_page>
  <end_page>215</end_page>
  <record_filename type="pdf"/>
</record>
<record lang="FR" type="article">
  <identifiers>
    <identifier type="pii">2000|0047|0325|0333</identifier>
    <identifier type="mr">2001i:7896789</identifier>
  </identifiers>
  <title lang="EN">A note on Pierskorn spheres and the generalized Jones
      conjecture</title>
  <title lang="FR">Une note sur des sphères de Pierskorn et le Jones generalize
      conjecturent</title>
  <author>
    <name>
      <given_name>Yves</given_name>
      <surname>Flambaud<surname>
    </name>
  </author>
  <subjects>
    <subject rank="primary" scheme="msc">57Q45</subject>
    <subject rank="secondary" scheme="msc">32S55</subject>
    <subject rank="secondary" scheme="msc">55M35</subject>
    <subject rank="secondary" scheme="msc">57R20</subject>
  </subjects>
  <start_page>325</start_page>
  <end_page>333</end_page>
  <record_filename type="pdf"/>
</record>
</journal_issue>
```

The Dublin Core Metadata Element Set

Term Name: contributor	
URI:	http://purl.org/dc/elements/1.1/contributor
Label:	Contributor
Definition:	An entity responsible for making contributions to the content of the resource.
Comment:	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.
Term Name: coverage	
URI:	http://purl.org/dc/elements/1.1/coverage
Label:	Coverage
Definition:	The extent or scope of the content of the resource.
Comment:	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.
Term Name: creator	
URI:	http://purl.org/dc/elements/1.1/creator
Label:	Creator
Definition:	An entity primarily responsible for making the content of the resource.
Comment:	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.

Term Name: date	
URI:	http://purl.org/dc/elements/1.1/date
Label:	Date
Definition:	A date associated with an event in the life cycle of the resource.
Comment:	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.
Term Name: description	
URI:	http://purl.org/dc/elements/1.1/description
Label:	Description
Definition:	An account of the content of the resource.
Comment:	Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.
Term Name: format	
URI:	http://purl.org/dc/elements/1.1/format
Label:	Format
Definition:	The physical or digital manifestation of the resource.
Comment:	Typically, Format may include the media-type or dimensions of the resource. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

Term Name: identifier	
URI:	http://purl.org/dc/elements/1.1/identifier
Label:	Resource Identifier
Definition:	An unambiguous reference to the resource within a given context.
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).
Term Name: language	
URI:	http://purl.org/dc/elements/1.1/language
Label:	Language
Definition:	A language of the intellectual content of the resource.
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 "en" or "eng" for English, "akk" for Akkadian, and "en-GB" for English used in the United Kingdom.
Term Name: publisher	
URI:	http://purl.org/dc/elements/1.1/publisher
Label:	Publisher
Definition:	An entity responsible for making the resource available
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.

Term Name: relation	
URI:	http://purl.org/dc/elements/1.1/relation
Label:	Relation
Definition:	A reference to a related resource.
Comment:	Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.
Term Name: rights	
URI:	http://purl.org/dc/elements/1.1/rights
Label:	Rights Management
Definition:	Information about rights held in and over the resource.
Comment:	Typically, a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.
Term Name: source	
URI:	http://purl.org/dc/elements/1.1/source
Label:	Source
Definition:	A reference to a resource from which the present resource is derived.
Comment:	The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

Term Name: subject	
URI:	http://purl.org/dc/elements/1.1/subject
Label:	Subject and Keywords
Definition:	The topic of the content of the resource.
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.
Term Name: title	
URI:	http://purl.org/dc/elements/1.1/title
Label:	Title
Definition:	A name given to the resource.
Comment:	Typically, a Title will be a name by which the resource is formally known.
Term Name: type	
URI:	http://purl.org/dc/elements/1.1/type
Label:	Resource Type
Definition:	The nature or genre of the content of the resource.
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.

Element Refinements

Term Name: abstract	
URI:	http://purl.org/dc/terms/abstract
Label:	Abstract
Definition:	A summary of the content of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/description
Term Name: accessRights	
URI:	http://purl.org/dc/terms/accessRights
Label:	Access Rights
Definition:	Information about who can access the resource or an indication of its security status.
Comment:	Access Rights may include information regarding access or restrictions based on privacy, security or other regulations.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/rights
Term Name: alternative	
URI:	http://purl.org/dc/terms/alternative
Label:	Alternative
Definition:	Any form of the title used as a substitute or alternative to the formal title of the resource.

Comment:	This qualifier can include Title abbreviations as well as translations.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/title
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	
URI:	http://purl.org/dc/terms/conformsTo

Label:	Conforms To
Definition:	A reference to an established standard to which the resource conforms.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: created	
URI:	http://purl.org/dc/terms/created
Label:	Created
Definition:	Date of creation of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/date
Term Name: dateAccepted	
URI:	http://purl.org/dc/terms/dateAccepted
Label:	Date Accepted
Definition:	Date of acceptance of the resource (e.g. of thesis by university department, of article by journal, etc.).
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/date
Term Name: dateCopyrighted	
URI:	http://purl.org/dc/terms/dateCopyrighted

Label:	Date Copyrighted
Definition:	Date of a statement of copyright.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/date
Term Name: dateSubmitted	
URI:	http://purl.org/dc/terms/dateSubmitted
Label:	Date Submitted
Definition:	Date of submission of the resource (e.g. thesis, articles, etc.).
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/date
Term Name: educationLevel	
URI:	http://purl.org/dc/terms/educationLevel
Label:	Audience Education Level
Definition:	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.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/terms/audience
Term Name: extent	
URI:	http://purl.org/dc/terms/extent

Label:	Extent
Definition:	The size or duration of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/format
Term Name: hasFormat	
URI:	http://purl.org/dc/terms/hasFormat
Label:	Has Format
Definition:	The described resource pre-existed the referenced resource, which is essentially the same intellectual content presented in another format.
Type of Term:	element-refinement
Date Issued:	2000-07-11
Term Name: hasPart	
URI:	http://purl.org/dc/terms/hasPart
Label:	Has Part
Definition:	The described resource includes the referenced resource either physically or logically.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: hasVersion	

URI:	http://purl.org/dc/terms/hasVersion
Label:	Has Version
Definition:	The described resource has a version, edition, or adaptation, namely, the referenced resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
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	

URI:	http://purl.org/dc/terms/isReferencedBy
Label:	Is Referenced By
Definition:	The described resource is referenced, cited, or otherwise pointed to by the referenced resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: isReplacedBy	
URI:	http://purl.org/dc/terms/isReplacedBy
Label:	Is Replaced By
Definition:	The described resource is supplanted, displaced, or superseded by the referenced resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: isRequiredBy	
URI:	http://purl.org/dc/terms/isRequiredBy
Label:	Is Required By
Definition:	The described resource is required by the referenced resource, either physically or logically.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation

Term Name: issued	
URI:	http://purl.org/dc/terms/issued
Label:	Issued
Definition:	Date of formal issuance (e.g., publication) of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/date
Term Name: isVersionOf	
URI:	http://purl.org/dc/terms/isVersionOf
Label:	Is Version Of
Definition:	The described resource is a version, edition, or adaptation of the referenced resource. Changes in version imply substantive changes in content rather than differences in format.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: license	
URI:	http://purl.org/dc/terms/license
Label:	License
Definition:	A legal document giving official permission to do something with the resource.
Comment:	Recommended best practice is to identify the license using a URI. Examples of such licenses can be found at http://creativecommons.org/licenses/.
Type of	element-refinement

Term:	
Refines:	http://purl.org/dc/elements/1.1/rights
Term Name: medium	
URI:	http://purl.org/dc/terms/medium
Label:	Medium
Definition:	The material or physical carrier of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/format
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
Term Name: references	
URI:	http://purl.org/dc/terms/references
Label:	References
Definition:	The described resource references, cites, or otherwise points to the referenced resource.

Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: replaces	
URI:	http://purl.org/dc/terms/replaces
Label:	Replaces
Definition:	The described resource supplants, displaces, or supersedes the referenced resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: requires	
URI:	http://purl.org/dc/terms/requires
Label:	Requires
Definition:	The described resource requires the referenced resource to support its function, delivery, or coherence of content.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/relation
Term Name: spatial	
URI:	http://purl.org/dc/terms/spatial
Label:	Spatial
Definition:	Spatial characteristics of the intellectual content of the resource.

Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/coverage
Term Name: tableOfContents	
URI:	http://purl.org/dc/terms/tableOfContents
Label:	Table Of Contents
Definition:	A list of subunits of the content of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/description
Term Name: temporal	
URI:	http://purl.org/dc/terms/temporal
Label:	Temporal
Definition:	Temporal characteristics of the intellectual content of the resource.
Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/coverage
Term Name: valid	
URI:	http://purl.org/dc/terms/valid
Label:	Valid
Definition:	Date (often a range) of validity of a resource.

Type of Term:	element-refinement
Refines:	http://purl.org/dc/elements/1.1/date

Metadata Map

Source Metadata (native publisher metadata)	Transformation Rules	Target Metadata (simple DC)
journal_issue/record/title	Take source element as is. If multiple title elements in source, take all; each source title gets a separate dc:title element.	title

Metadata Map

Source Metadata (native publisher metadata)	Transformation Rules	Target Metadata (simple DC)

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.

Exercise 8: 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 on-line 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.

Exercise 8: Metadata Workflow
Scenario B—Small Steps toward Library Publishing

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:
      - 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?).
      - What is the level of complexity required to transform input to output?
      - 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.)
      - 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?
- 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 v. |c 25 cm.

500 __ |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 edited by Mrs. S. C. Lomas, v. 3- by F. Bickley.

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...

000 01052cam a2200325 450

001 408612

005 20061212110910.0

008 851011m19091927gw 000 0 ger d

035 __ **|a** (CStRLIN)NYCX85B113967

035 __ **|a** (NIC)notisABT9133

035 __ **|a** (OCoLC)13463437

035 __ **|a** 408612

040 __ **|a** MnU **|c** MnU **|d** RPB **|d** NIC

100 1_ **|a** Grassmann, Hermann Ernst, **|d** 1857-1922.

245 10 **|a** Projektive geometrie der ebene, **|b** unter benutzung der punktrechnung...

260 __ **|a** Leipzig, **|b** B. G. Teubner. **|c** 1909-27.

300 __ **|a** 3 v. **|c** 24 cm.

500 __ **|a** Nachwort, von G. Wolff: p. [VI]

505 0_ **|a** Bd. 1. Binäres.--Bd. 2. Ternäres, 2 v.

650 _0 **|a** Geometry, Projective

650 _0 **|a** Forms, Binary

650 _0 **|a** Forms, Ternary

700 1_ **|a** Wolff, Georg. **|d** 1881-

905 __ **|a** 19880623120000.0

948 __ **|a** c:RET

950 __ **|l** MATH **|a** QA554 **|b** .G76 **|f** BASIC **|i** 10/11/85 N

955 __ **|l** MATH **|c** 1:v.1 **|i** 10/11/85 C

955 __ **|l** MATH **|c** 1:v.2 **|i** 10/11/85 C

998 __ **|a** 10/11/85 **|t** c **|s** 9124 **|n** NIC **|w** MNUG83B19018 **|d** 10/11/85 **|c** RET **|b** YOB

The war of the rebellion: a compilation of the official records of the...

000 04656cam a2200505 a 450

001 4943960

005 20061002074755.0

008 000413m18801901dcu f001 0 eng d

010 __ |a 03003452

040 __ |c NBiSU |d NBiSU |d MiU |d TxFACM |d NIC |d NcAvBC |e dcrb |d NIC

043 __ |a n-us---

050 0_ |a E491 |b .U6

110 1_ |a United States. |b War Dept.

245 14 |a The war of the rebellion: |b a compilation of the official records of the Union and Confederate armies / |c Pub. under the direction of the secretary of war ...

260 __ |a Washington : |b Govt. Print. Off., |c 1880-1901.

300 __ |a 123 v. ; |c 24 cm.

500 __ |a Found also in the House Miscellaneous documents of the 52d to the 56th Congress.

500 __ |a Each number has special index. Inserted in each volume: Additons and corrections ... Washington, Govt. Print. Off., 1902.

500 __ |a Series 1, v. 1-53, series 3, v. 1-5, and series 4, v. 1-3 include "Alternate designations of organizations mentioned."

500 __ |a Vol. 54-55 of series 1 [serial no. 112-113]" have not been published, and no material for them is in hand." cf. General index, p. xl. Series 2, v. 1 [serial no. 114] with imprint 1894, was not issued until 1898.

500 __ |a Edited in the War Records Office, 1880-July 1899; in the Record and Pension Office, July 1899-1901.

500 __ |a Incomplete set: missing volumes 1-5 of the first series. |5 NcAvBC

500 __ |a Bdg.: navy blue publisher’s cloth binding, all volumes worn with some damage

to inner hinges; paper browned throughout set. |5 NcAvBC

500 __ |a 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 authoriites, 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** United States. **|b** Congress. **|b** House.

700 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...

000 00706nam a22001931 450

001 3071182

005 19970920120000.0

008 970920s1859 nyuacf 000 0 eng

010 __ **|a** 01014304

035 __ **|a** (NIC)notisAPZ5299

040 __ **|c** NNC **|d** NNC **|d** NNU **|d** NIC

043 __ **|a** n-us-ny

050 0_ **|a** F128.3 **|b** .B72

100 1_ **|a** Booth, Mary L. **|q** (Mary Louise), **|d** 1831-1889.

245 10 **|a** History of the city of New York, from its earliest settlement to the present time. **|c** By Mary L. Booth. Illus. with over one hundred engravings.

260 __ **|a** New York, **|b** W.R.C. Clark & Meeker, **|c** 1859.

300 __ **|a** xix, <21>-846 p. incl. illus., plates, ports. front. **|c** 24 cm.

651 _0 **|a** New York (N.Y.) **|x** History.

905 __ **|a** 19970920120000.0

Celestine, being the diary of a chambermaid. By Octave Mirbeau. Translated...

000 01107cam a22003131 450

001 1699480

005 20050419095459.0

008 891005s1930 nyu 000 0 eng

010 __ **|a** 31008215

035 __ **|a** (CStRLIN)NYCX89B29529

035 __ **|a** (NIC)notisAHY8430

040 __ **|a** DLC **|c** OKentU **|d** *SER* **|d** m.c. **|d** FU **|d** CStRLIN **|d** NIC

041 1_ **|a** eng **|h** fre

050 0_ **|a** PZ3.M674 **|b** Ce

100 1_ **|a** Mirbeau, Octave, **|d** 1848-1917

240 10 **|a** Journal d'une femme de chambre. **|l** English

245 10 **|a** Celestine, **|b** being the diary of a chambermaid. **|c** By Octave Mirbeau. Translated by Alan Durst.

260 __ **|a** New York, **|b** W. Faro, inc., **|c** 1930.

300 __ **|a** 317 p. **|c** 25 cm.

500 __ **|a** At head of title: By Octave Mirbeau.

500 __ **|a** Translation of Le journal d'une femme de chambre.

700 1_ **|a** Durst, Alan.

740 0_ **|a** Diary of a chambermaid.

905 __ **|a** 19970917120000.0

950 __ **|l** URIS **|a** PQ2364.M67 **|b** J8 1930 **|i** 10/05/89 N

955 __ **|l** URIS **|c** 1 **|s** 4th prtg.,1933 **|i** 10/05/89 C

998 __ **|a** 10/05/89 **|t** c **|s** 9125 **|n** NIC **|w** FLUGACR2289B **|d** 10/05/89 **|c** RET **|b** LJH **|i** 891005 **|l** NYCX

```
<?xml version="1.0"?>
<monograph id="" lang="">
     <monograph_data>
          <identifiers>
               <identifier type="doi"></identifier>
          </identifiers>
          <author>
               <name>
                    <given_name></given_name>
                    <surname></surname>
               </name>
               <affiliation>
                    <prof_title></prof_title>
                    <organization></organization>
               </affiliation>
               <bio>
                    <p></p>
               </bio>
               <image type="" id="" />
          </author>
          <contributor>
               <role></role>
               <name>
                    <given_name></given_name>
                    <surname></surname>
               </name>
               <affiliation>
                    <prof_title></prof_title>
                    <organization></organization>
               </affiliation>
               <bio>
                    <p></p>
               </bio>
               <image type="" id="" />
          </contributor>
          <title lang=""></title>
          <edition></edition>
          <publisher></publisher>
          <publisher_location></publisher_location>
          <publication_date iso8601-date=""></publication_date>
          <abstract lang="">
               <p></p>
          </abstract>
          <subjects>
               <subject scheme=""></subject>
          </subjects>
          <extent></extent>
```

```
          <cover_image id="" />
     </monograph_data>

     <monograph_parts>
          <section type="" id="">
               <identifiers>
                    <identifier type="doi"></identifier>
               </identifiers>
               <author>
                    [as above]
               </author>
               <contributor>
                    [as above]
               </contributor>
               <title lang=""></title>
               <abstract lang="">
                    <p></p>
               </abstract>
               <subjects>
                    <subject scheme=""></subject>
               </subjects>
          </section>
     </monograph_parts>

</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
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:

Metadata Map

Source	Transformation	Target

Metadata Map

Source	Transformation	Target

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.

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 Slide Collection Digital Library Project—**Scenario A-1**

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 Slide Collection Digital Library Project—**Scenario A-2**

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 Slide Collection Digital Library Project—**Scenario B-1**

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 Initative) 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 Slide Collection Digital Library Project—**Scenario B-2**

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 Initative) 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.
        i) 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?

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:
WORK_TYPE: 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
FILM_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:
WORK_TYPE: temple
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
FILM_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:
WORK_TYPE: 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:
CITATION_2: Grove Dictionary of Art
CITATION_3: 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:
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:
CITATION_2: Grove Dictionary of Art
CITATION_3: 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:
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:
REGION: Kashmir
COUNTRY: India
TITLE_1: Packed Horses on Road in Kashmir in 1985
TITLE_2:
WORK_TYPE: 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:
WORK_TYPE: 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:

```
<?xml version="1.0" encoding="UTF-8" ?>
<vra xmlns="http://www.vraweb.org/vracore4.htm">
   <work id="B-Q5 Kan 3.4 SVar/Anan">
      <date type="creation" start="1025" end="1400">ca. 1025-1400</date>
      <description>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.</description>
      <location>
         <name type="geographic">Kanchipuram (Tamil Nadu, India)</name>
      </location>
      <objectType>temple</objectType>
      <objectType>reservoir</objectType>
      <source>
         <name>Raman, K. V.; Sri Varadarajaswami Temple, Kanchi: a Study of its
         History, Art, and Architecture; New Delhi: Abhinav Publications;
         1975.</name>
      </source>
      <source>
         <name>Grove Dictionary of Art</name>
      </source>
      <source>
         <name>http://srinagar.nic.in/</name>
      </source>
      <title type="local" pref="true">Sri Varadarajaswami Temple</title>
      <title type="local" pref="false">Hastigiri</title>
      <image id="MCD_00894" source="Francis Max Collection">
         <agent>
            <name type="corporate name">Avon Digital Production Services (New
            York, NY, USA)</name>
            <role id="300237903" vocab="AAT">digital imaging</role>
         </agent>
         <date type="view">1982</date>
         <date type="creation">2005</date>
         <location>
            <name type="repository">Art & Architecture Library, University
            College (Newfield, NY, USA)</name>
            <id type="accession">91000082</id>
            <id type="classification">B-Q5 Kan 3.4 SVar/Anan 5-1</id>
         </location>
         <objectType id="300215302" vocab="AAT">digital images</objectType>
         <relation id="B-Q5 Kan 3.4 SVar/Anan" type="image of">Sri
         Varadarajaswami Temple</relation>
         <subject>reservoirs</subject>
         <subject>mandapa</subject>
         <technique id="300237903" vocab="AAT">digital imaging</technique>
         <title type="full view">Ext.: Anantasaras tank viewed from along pillared
         Kalyana Vandapa</title>
      </image>
   </work>
</vra>
```

Characteristics of Source Metadata

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

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

Characteristics of Target 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:

Metadata Map

Source	Transformation	Target

Metadata Map

Source	Transformation	Target

Metadata Map

Source	Transformation	Target

Metadata Map

Source	Transformation	Target

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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

List opportunities for your organization in planning and managing a digital initiative.
List challenges for your organization in planning and managing a digital initiative.
List your institution's strengths in developing a digital initiative.
List areas where your institution will need additional capacity and resources to plan and manage a digital initiative.
List areas where your institution will need additional staff training to plan and manage a digital initiative.

Digital & Consulting Production Services

Cornell University Library, Digital Consulting and Production Services (DCAPS)
Digital Imaging Project Planning Outline

Component	Tasks
Overall Project Management	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
Selection	Select materials, identify incomplete items or pages in poor condition
Copyright	Determine legal restrictions
Physical Preparation	Retrieve documents, record physical conditions, check out if necessary, repair pages, disbind if necessary
Intellectual Preparation	Create documentation for indexing, tagging, and other information such as special treatment of illustrations and photographs
User Requirements	Determine user requirements for legibility, navigation, color fidelity, file formats, etc.
Image Benchmarking	Determine imaging requirements including scanning equipment, resolution, bit depth, image enhancement and correction, compression, file format, file headers
Metadata Benchmarking	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
Image Processing	OCR (determine accuracy rate, software used, etc.), SGML or XML & necessary DTD
Image Quality Control	Determine methodology (percentage of QC and method – on screen vs. on paper, QC tools)
Metadata Quality Control	Set accuracy levels for consistency of file naming, file headers, and other manually created metadata
Derivative Creation	Decide on access file formats, on-the-fly vs. static derivative creation, need for facsimile creation
Storage	Identify needs and purchase storage hardware and software Plan backups, file loading, and data integrity
Image Database	Identify access requirements and select image management software
Digital Preservation	Decide on preservation metadata, backups or mirror sites, any other policy decisions
Web Delivery	Website interface and graphical design, scripting, write project narrative, prepare user guides
Project Website	Information about the project including time line, quarterly reports, and other related documents
Publicity	Inform CUL staff, scholars and researchers, other cultural institutions and scholarly communities of the imaging project
User Support	Work with PS staff in identifying user support needs, such as training sessions and handouts
Digital Rights Management	Statement about getting permission to use or reproduce images
Assessment	Determine criteria for outcome assessments, design and implement assessment, conduct a user study
Financial Management	Financial reporting and CUL Accounting liaison for managing grant funds

Digital & Consulting Production Services

Cornell University Library
Digital Consulting and Production Services

PREPARATION			*Pages*	*Cost/Unit*		*Subtotals*
Selection						
User & Content Requirements						$0.00
Collation, Tagging						$0.00
Conservation and Repairs						
DIGITIZATION, STRUCTURING			*Pages*	*Cost/Unit*		
Digitization (Specify)						$0.00
Image Processing						
Structuring & Tagging						
POST PROCESSING			*Hours*	*Cost/Unit*		
OCR Preparation						$0.00
Derivative	Creation (e.g., PDF)					$0.00
QUALITY CONTROL			*Pages*	*Cost/Unit*		
Image QC						$0.00
METADATA		*Hours*		*Cost/Unit*		
Design						$0.00
		Hours		*Cost/Unit*		
Implementation						$0.00
DIGITAL CONTENT DELIVERY		*Hours*		*Cost/Unit*		
Design						$0.00
Preparation & Ingest						$0.00
Assessment and Testing						
WEB DEVELOPMENT		*Hours*		*Cost/Unit*		
Website Design						$0.00
Usability Assessment						
STORAGE/YEAR		*Size (GB)*		*Cost/Unit*		
Storage and Maintenance						$0.00
Preservation						
RIGHTS MANAGEMENT		*Hours*		*Cost/Unit*		
Copyright & IPR Consulting						$0.00
PROJECT MANAGEMENT						
Project Management						$0.00
			TOTAL ESTIMATE			**$0.00**

oyr, March'05

Cornell University
Library

Metadata Services
107 Olin Library
Ithaca, New York 14853-5301

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

Cornell University is an equal opportunity, affirmative action educator and employer.

Cornell University
Library

Metadata Services
107 Olin Library
Ithaca, New York 14853-5301

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.]

Cornell University is an equal opportunity, affirmative action educator and employer.

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.)

Preservation Planning Launch Questionnaire

[Adapted from: Grant, "Going from Zero to Live with an Automated Digital Preservation System," *Digital Preservation: Current Efforts (NISO Webinar),* January 14, 2009]

Needs Assessment and Analysis

What digital objects already exist that you want to preserve?
In which formats are the objects?
How many objects are there and how big are the files?
Are some objects related to others in important ways? (versions, derivatives, component parts)
What is the current growth rate of the body of objects you want to preserve?
What new collections and object types are you planning to acquire?
What metadata do you have and what metadata do you want that you don't have?
Where does existing metadata reside?
What copyright and access restrictions apply to the objects?
In which essential ways do you want to be able to search for the preserved objects?
What are the points of greatest vulnerability regarding existing objects and metadata?

mmk 10/09

Assessing an Organization's Commitment to Digital Preservation

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

Adapted from: Kenney and McGovern, "The Five Organizational Stages of Digital Preservation," *Digital Libraries: A Vision for the 21st Century* , 2003.

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.)

MetaArchive Collection-Level Metadata Worksheet

[a subset of *MetaArchive Collection-Level Conspectus Metadata Specification*, <http://metaarchive.org/pdfs/conspectus_md_2005.html>]

Access Rights [dcterms:accessRights]
A statement of any access restrictions placed on a collection, including allowed users, charges, etc.
M/O:
Accrual Periodicity [cld:accrualPeriodicity]
Frequency with which items are expected to be added to a collection.
M/O:
Accrual Policy [cld:accrualPolicy]
Approach adopted to add items to the collection or a statement about anticipated growth of the collection, including quantity and frequency.
M/O:
Bytes [dcterms:extent]
The total calculated file size, expressed in computer bytes, of the digital collection being described.
M/O:
Cataloged Status [ma:catalogedstatus]
Indication of level of cataloging of the collection beyond the collection level.
M/O:
Custodial History [dcterms:provenance]
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.
M/O:
Description [dc:description]
A summary of the content of the digital collection.
M/O:

MetaArchive Collection-Level Metadata Worksheet

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

MetaArchive Collection-Level Metadata Worksheet

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:

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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

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

1.1 objectIdentifier
A designation used to uniquely identify the object within the preservation repository system in which it is stored.
M/O:
1.2 objectCategory
The category of object to which the metadata applies. Suggested values: representation, file, bitstream.
M/O:
1.3.1 preservationLevelValue
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.
M/O:
1.3.4 preservationLevelDateAssigned
The date, or date and time, when a particular preservationLevelValue was assigned to the object.
M/O:
1.4 significantProperties
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.
M/O:
1.4.1 significantPropertiesType
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
M/O:
1.4.2 significantPropertiesValue
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.
M/O:

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

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

PREMIS Object Entity Metadata Worksheet

1.8.5.2 swVersion
The version or versions of the software referenced in swName.
M/O:
1.8.5.3 swType
Class or category of software. Examples: renderer, ancillary, operatingSystem, driver.
M/O:
1.8.6.1 hwName
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.
M/O:
1.8.6.2 hwType
Class or category of the hardware. Suggested values: processor, memory, input/output device, storage device.
M/O:
1.10 relationship
Information about a relationship between this object and one or more other objects.
M/O:
1.10.1 relationshipType
A high-level categorization of the nature of the relationship. Suggested: structural, derivation.
M/O:
1.10.2 relationshipSubType
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.
M/O:
1.10.3.2 relatedObjectIdentifierValue
The value of the identifier of the related object.
M/O:

PREMIS Object Entity Metadata Worksheet

1.10.3.3 relatedObjectSequence
The order of the related object relative to other objects with the same type of relationship.
M/O:
1.10.4 relatedEventIdentification
The identifier and contextual sequence of an event (e.g., migration) associated with the relationship.
M/O:
1.10.4.2 relatedEventIdentifierValue
The *eventIdentifierValue* of the related event.
M/O:
1.12 linkingIntellectualEntityIdentifier
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.)
M/O:
1.13 linkingRightsStatementIdentifier
An identifier for a rights statement associated with the object.
M/O:

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.

Metadata and Digital Libraries
Marty Kurth
UAEU Libraries October 4-8, 2009

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.