Coding for Efficiencies in Cataloging & Metadata: Practical Applications of XML, XSLT, XQuery, and PyMarc for Library Data

Timothy Cole, University of Illinois, Urbana-Champaign
Heidi Frank, New York University
Myung-Ja (“MJ”) Han, University of Illinois, Urbana-Champaign
Christine Schwartz, Princeton Theological Seminary Library

June 25, 2015
8:30 a.m. – 4:30 p.m.

Sponsor: Technical Services Workflows Efficiency Interest Group
Co-Sponsors:
LITA Program Planning Committee
Online Audiovisual Catalogers

alctsAC15
Schedule

Thursday, June 25, 2015

8:00 - 8:30 Registration and continental breakfast

8:30 - 8:45 Welcome and introductions: Annie Glerum, moderator

8:45 - 10:00 Timothy W. Cole & Myung-Ja "MJ" Han
   XML Technologies for Catalogers and Metadata Librarians: Part 1

10:00 - 10:15 Break

10:15 - 11:30 Timothy W. Cole & Myung-Ja "MJ" Han
   XML Technologies for Catalogers and Metadata Librarians: Part 2

11:30 - 1:00 Lunch (on your own)

1:00 - 2:30 Christine Schwartz
   Using XQuery as a Library Metadata Tool

2:30 - 2:45 Break

2:45 - 4:15 Heidi Frank
   Python/PyMARC for MARC Record Analysis and Manipulation

4:15 - 4:30 Questions and Wrap-Up
XML Technologies for Catalogers and Metadata Librarians

Timothy W. Cole & Myung-Ja Han
University of Illinois at Urbana-Champaign

2015 ALA Annual Conference
ALCTS Preconference
25 June 2015
San Francisco, California
1. A few introductory remarks about library metadata
2. eXtensible Markup Language (XML) – the basics
   hands-on exercise & break
3. XML validation, schema languages, etc.
4. Metadata crosswalking
   hands-on exercise & break
5. Extensible Stylesheet Language for Transformation (XSLT)
   hands-on exercise & break
A Few Introductory Remarks about Library Metadata
What is Metadata?
Definitions from First Principles

• Data about data
• Informative (useful) data about data
• Structured ("constructed"?) data about data
• Machine-readable structured data about data
• Machine-actionable structured data about data
• Surrogates for information resources (i.e., data)
• Structured data about data that facilitates use of data
How metadata will be used helps define what metadata is...

Who’s Here?

• Show of hands –
  
  
  • How many work in Technical Services unit or division?
  
  • How many are looking to improve or revamp workflows for main Library catalog? For content managements systems (e.g., CONTENTdm)? For Web Scale Discovery Systems (e.g., PRIMO)? For Institutional Repositories?
  
  • Other?
The MARC Metadata Format

MARC Format Records -

• Has been a cataloging tool in the library community since the 1960’s
• Optimized for streaming media (magnetic tape)
• Standard for library catalogs
• Used in conjunction with AACR2, the content standard, to facilitate bibliographic control
• Has more than 1900 fields
A Little Bit of a MARCXML Record

<?xml version="1.0" encoding="UTF-8"?>
<record xmlns="http://www.loc.gov/MARC21/slim" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  <leader>00761cam a2200241i 4500</leader>
  <controlfield tag="001">2624108</controlfield>
  <controlfield tag="005">20081104163033.0</controlfield>
  <controlfield tag="008">900530s1841 eng 000 1 eng d</controlfield>
  <datafield tag="035" ind1=" " ind2="">
    <subfield code="a">(OCoLC)ocm03535290</subfield>
  </datafield>
  <datafield tag="040" ind1=" " ind2="">
    <subfield code="a">SUC</subfield>
    <subfield code="c">OCL</subfield>
    <subfield code="d">UIU</subfield>
  </datafield>
  <datafield tag="035" ind1=" " ind2="">
    <subfield code="a">ARY-4606</subfield>
  </datafield>
  <datafield tag="035" ind1=" " ind2="">
    <subfield code="a">UC 12037257</subfield>
  </datafield>
  <datafield tag="090" ind1=" " ind2="">
    <subfield code="a">PR4058</subfield>
    <subfield code="b">.365 1841</subfield>
  </datafield>
  <datafield tag="049" ind1=" " ind2="">
    <subfield code="a">UIUU</subfield>
  </datafield>
  <datafield tag="100" ind1="1" ind2="">
    <subfield code="a">Ingoldsby, Thomas,</subfield>
    <subfield code="d">1788-1845.</subfield>
  </datafield>
  <datafield tag="245" ind1="1" ind2="0">
    <subfield code="a">Some account of my cousin Nicholas</subfield>
    <subfield code="c">by Thomas Ingoldsby, esq. ; to which is added, The rubber of life.</subfield>
  </datafield>
  <datafield tag="260" ind1=" " ind2="">
    <subfield code="a">London : Richard Bentley,</subfield>
    <subfield code="b">1841.</subfield>
  </datafield>
  <datafield tag="300" ind1=" " ind2="">
    <subfield code="a">3 v. ;</subfield>
    <subfield code="c">19 cm.</subfield>
  </datafield>
</record>
Metadata Standards Used in Libraries

- MARC (MAchine Readable Cataloging)
- Metadata Object Description Schema (MODS)
- Metadata Authority Description Standard (MADS)
- Metadata Encoding & Transmission Standard (METS)
- Encoded Archival Description (EAD)
- NISO Metadata for Images in XML (MIX)
- Preservation Metadata (PREMIS)
- Dublin Core, VRA Core, PB Core, CDWA-Lite, IEEE-LOM, ODRL, ...

Show of hands: who has created or edited metadata in MARC? MODS? DC?
Housekeeping & logistics

• Website for AM: http://quest.library.Illinois.edu/ala2015/ALCTS-XML/
  • Includes updated slides
  • Includes examples used in lecture
  • Includes solutions to the exercises
  • Includes a zip of the site

• Time this morning allotted for 3 hands-on exercises
  • use these opportunities also as break time
  • About the software – Web browsers, plain text editors, oXygen, an XML editor, validating parser and XPath debugger
eXtensible Markup Language (XML) – the basics
What is XML?

Discussion points:

• First principles: Ordered Hierarchy Content of Objects (OHCO)

• Example: A simple XML fragment

• Compare/contrast: HTML, XHTML

• Elements, attributes, ... [the gory details]
Ordered Hierarchies of Content Objects (1)

A metadata record can be considered a form of a text document

• OHCO Premise: A text is the sum of its component parts
  • A `<Book>` might be defined as containing:
    `<FrontMatter>`, `<Chapter>s`, `<BackMatter>`
  • `<FrontMatter>` might contain:
    `<BookTitle>`, `<Author>s`, `<PubInfo>`
  • A `<Chapter>` might contain:
    `<ChapterTitle>`, `<Paragraph>s`
  • A `<Paragraph>` might contain:
    `<Sentence>s` or `<Table>s` or `<Figure>s` ...

• Components chosen should reflect anticipated use
• OHCO is a useful, albeit imperfect, model
  • Exposes an object’s intellectual structure
  • Supports reuse & abstraction of components
  • Better than a bit-mapped page image
  • Better than a model of text as a stream of characters plus formatting instructions
• Data management system for document-like objects
• Does not allow overlapping content objects
• Incomplete; requires infrastructure
Content Objects in a Catalog Card

Card
- CallNumber
- MainEntry
- TitleStatement
  - TitleProper
  - StatementOfResponsibility
- Imprint
- SummaryNote
- AddedEntrySubject(s)
- AddedEntryPersonalName(s)
How Does XML Work?

In XML, there’s content and there’s markup.

- **Markup** (within angle brackets, ‘<‘ & ‘>’)
  - Elements
  - Attributes
  - Comments
  - Processing instructions

- **Content**
  - Text (Unicode characters)
  - Entities
A Simple XML Fragment

<Book>
  <FrontMatter>
    <BookTitle>XML for Catalogers and Metadata Librarians</BookTitle>
    <Author>Timothy Cole</Author>
    <Author>MJ Han</Author>
    <PubInfo>Libraries Unlimited</PubInfo>
  </FrontMatter>
  <Chapter>
    <ChapterTitle>XML: What is it?</ChapterTitle>
    <Paragraph>
      The authoritative document defining the ...
    </Paragraph>
  </Chapter>
</Book>
This is NOT XML

<PoemFragment>

<Stanza>

(Line)<Sentence>It was six men of Indostan</Sentence>
(Line)To learning much inclined,
(Line)Who went to see the Elephant
(Line)(Though all of them were blind),
(Line)That each by observation
(Line)Might satisfy his mind.

</Sentence></Line>

</Stanza>

</PoemFragment>
XML, HTML, & XHTML

• HTML— a display-oriented, scheme for making Web pages
  • Syntax & allowed elements (semantics) are fixed

• XML— set of rules for defining markup schemes
  • Element set is fully extensible
  • Syntax is fixed

• XHTML— HTML modified to be XML-compliant
Markup Languages Compared

• XML vs. HTML
  • XML: no fixed tags, no predefined style
  • XML is extensible;
    Fixed elements (HTML) vs. rules (XML)
  • HTML elements describe how to present content
  • XML elements can describe the content itself

• XML syntax is stricter than HTML
  • Must explicitly close all elements
  • Attributes must be enclosed in quotes
  • All markup is case-sensitive

Questions?
Syntax – Elements

Elements are markup that enclose content
• <element_name>…</element_name>
  or <element_name />

• Element content models describe what elements may contain / where may occur
  • Parsed Character Data Only
  • Child Elements Only
  • Mixed (Elements & PC Data)

PC Data element:
  <author>Cole, T</author>
Syntax - Attributes

Associate a name-value pair with an element

• `<tag name1="value1" name2='value2'>...<tag>`
  • Can be used to embellish content…
  • or to associate content with an ‘empty’ element

`<author order='1'>Cole, T</author>
`<author name='Habing, T' />`
Syntax - Comments

Human-readable annotations
• Can be inserted anywhere after headers
• Not part of the document structure (is part of DOM)
• Usually ignored by XML parsers
• Does not have to be passed to application

<!-- This is a comment -->
Syntax - Processing Instructions

Machine-readable & application-specific
• Must be passed through by XML Parsers
• XML Declaration is a special PI
• XML Declaration is always first line in file

<?xml version='1.0' encoding='UTF-8' ?>
<?MyApp indent='on' linefeeds='off' ?>
Syntax - Entities

• Placeholders for internal or external content
  • Placeholder for a single character...
  • or string of text...
  • or external content (images, audio, etc.)
• Implementation specifics may vary

<!ENTITY copyright "&#169;" >
  means &copyright; should be replaced by Unicode code point 169 (base 10), i.e., ©
<!ENTITY pic SYSTEM "mugshot.gif" NDATA gif >
&pic; is replaced by graphic image
Character Encoding Issues

• XML Parsers must accept Unicode:
  • UTF-8 & UTF-16
  • Also &###nnnn; or &###xhhhh;

• MARC-8 encodings must be converted to Unicode for use in XML

http://lcweb.loc.gov/marc/specifications/specchartables.html
Why XML?

- Machine Processing
  - XML records easy to validate
  - XML records easy to transform in batches
  - XML records easy to ingest into databases

- Web services (http://www.w3.org/2002/ws/)
  - Facilitating machine-to-machine communications via XML
  - Simple Object Access Protocol (SOAP)
  - W3C XML Protocol Working Group

- Semantic Web (http://www.w3.org/2001/sw/)
  - Abstract knowledge representation on the Web
  - XML and databases
XML Terminology

• Document instance
• Well-formed XML
• Document class
• Document Type Definition (DTD), or schema
• Valid XML

• Stylesheets
  • Cascading Style Sheets (CSS)
  • XSL Transformations & Formatting Objects (XSL-T, XSL-FO)

• Document Object Model (DOM)
The Infrastructure of XML

• Required to make it work...

  • DTDs & schemas: defining document classes

  • Validating XML parsers

  • Reusing & integrating schemas (using namespaces)

  • Stylesheets for presentation & transformation
Well-Formed XML

• No schema or DTD required
  – must follow XML syntax
  – no constraints on element or attribute names,
    use of hierarchy, element content models, etc.

• Well-Formed XML is easier, less overhead, good for rapid prototyping, good enough for:
  – Small team, short-duration ad-hoc projects
  – Highly homogeneous, simple records
  – Minimal expectations for reuse or retention of records
Illustrations of well-formed XML

• View well-formed XML in your Browser: illustration1.xml
  http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/Well-Formed-Metadata/illustration1.xml

• View well-formed XML in oXygen: illustration2.xml
  http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/Well-Formed-Metadata/illustration2.xml

• Not well-formed XML in your Browser: illustNotWF1.xml
  http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/Not-Well-Formed-Metadata/illustNotWF1.xml

• Not well-formed XML in oXygen: illustNotWF2.xml
  http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/Not-Well-Formed-Metadata/illustNotWF2.xml
Exercise 1 – Create well-formed XML
http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/

• Objective: create well-formed XML records describing the following resources or resources of your own choosing:
  • http://imagesearchnew.library.illinois.edu/cdm/ref/collection/koopman/id/777
  • http://imagesearchnew.library.illinois.edu/cdm/ref/collection/chicago/id/17
  • https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&v1=1&BBRecID=7285233
  • https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&v1=1&BBRecID=4554404

For additional details see: Exercise1-Description.pdf in the Exercise1 folder

Demonstrations using Template4Exercise1.xml
http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/Template4Exercise1.xml

• Open and save to you computer using Web Browser, edit in Notepad++
• Open and edit directly in oXygen
Exercise 1 -- Solutions

*Our Solutions:*

- sample1.xml
- sample2.xml
- sample3-1.xml, sample3-2.xml
- sample4.xml

*Incorrect, not Well-Formed:*

- sample1Not.xml (Mismatch between open & close tags)
- sample3-2Not.xml (Attribute value not in quotes)

*Discussion: What was easiest / hardest to remember?*
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
 <!-- Incomplete metadata record illustrating well-formedness -->
 <title>William T. Young Library, University of Kentucky</title>
 <type>Image</type>
 <publisher>ALA Archives - University of Illinois at Urbana-Champaign</publisher>
 <theme>Library Postcards</theme>
 <medium>Paper</medium>
 <resourceIdentifier>ALA0000483</resourceIdentifier>
 <repository>University of Illinois at Urbana-Champaign. University Archives</repository>
</metadata>
<xml version="1.0" encoding="UTF-8"/>
<metadata>
  <!-- Incomplete metadata record illustrating well-formedness -->
  <editor>Tennant, Roy.</editor>
  <publisher>New York : Neal-Schuman Publishers</publisher>
  <subjectLCSH>Libraries -- Data processing.</subjectLCSH>
  <persistentLinkToThisPage>https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&amp;v1=1&amp;BBRecID=4554404</persistentLinkToThisPage>
  <institution>University of Illinois at Urbana-Champaign Library</institution>
  <location>Shelved At Main Stacks</location>
</metadata>
<xml version="1.0" encoding="UTF-8"?>
<metadata>
<!-- Incomplete metadata record illustrating NOT well-formedness -->
<title>William T. Young Library, University of Kentucky</title>
<type>Image</type>
<publisher>ALA Archives - University of Illinois at Urbana-Champaign</publisher>
<theme>Library Postcards</theme>
<medium>Paper</medium>
<resourceIdentifier>ALA0000483</resourceIdentifier>
<repository>University of Illinois at Urbana-Champaign. University Archives</repository>
</metadata>

CASE OF ELEMENT NAME IN OPENING AND CLOSING TAG MUST MATCH!
Not Well-Formed XML
Illustration #2

Available online
http://quest.library.illinois.edu/ala2015/ALCTS-XML/Exercises/Exercise1/Not-Well-Formed-Metadata/illustNotWF2.xml

<?xml version="1.0" encoding="UTF-8"?>
<metadata>
<!-- Incomplete metadata record illustrating NOT well-formedness -->
<editor>Tennant, Roy.</editor>
<publisher>New York : Neal-Schuman Publishers</publisher>
<subject authority="lcsh">XML (Document markup language)</subject>
<subject authority="lcsh">Cataloging --Data processing.</subject>
<persistentLinkToThisPage>https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&amp;l=v1&amp;BBRecID=4554404</persistentLinkToThisPage>
<institution>University of Illinois at Urbana-Champaign Library</institution>
<location>Shelved At Main Stacks</location>
</metadata>

ATTRIBUTE VALUES MUST BE ENCLOSED IN QUOTES!
Exercise 1
Creating well-formed XML documents

1. Click on the following URLs. These URLs show two digital images (with descriptions) and descriptions of two print monographs.

   •  http://imagesearchnew.library.illinois.edu/cdm/ref/collection/koopman/id/777
   •  http://imagesearchnew.library.illinois.edu/cdm/ref/collection/chicago/id/17
   •  https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&v1=1&BBRecID=7285233
   •  https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&v1=1&BBRecID=4554404

2. Please examine the information displayed at each URL about the resource. Then create a well-formed XML document describing each resource.

3. Please view each of the well-formed XML documents you create in oXygen and/or your Web browser to verify that they are well-formed (or to find your errors and fix).

A template / incomplete well-formed metadata record to get you started:

   •  http://quest.library.illinois.edu/ALA2015/ALCTS-XML/Exercises/Exercise1/Template4Exercise1.xml
<xml version="1.0" encoding="UTF-8"?>
<metadata>
  <title>...put title text here...</title>
  <publisher>...put publisher information here...</publisher>
  <!-- and so on. -->
</metadata>
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
    <title>William T. Young Library, University of Kentucky</title>
    <type>Image</type>
    <digitizedMaterial>black and white postcard</digitizedMaterial>
    <coverageSpatial>Lexington, Kentucky</coverageSpatial>
    <subject>Library buildings Kentucky</subject>
    <subject>Public libraries</subject>
    <subject>Academic libraries</subject>
    <description>On Front: William T. Young Library, University of Kentucky; On Back: The William T. Young Library opened on the University of Kentucky campus in April 1998. It is the third central library in the University's history, replacing the Margaret T. King Library which opened in 1931 and the Carnegie Library constructed in 1909. Michael McKinnell of Boston was the architect and Bovis the construction management firm. The building has over 350,000 square feet of space and cost $58 million. President Wethington's commitment to the project was critical as the facility was constructed with private funds by the UK Alumni Association and bonds were sold to finance part of the construction expenses. William T. Young of Lexington was the leading contributor to the Library.</description>
    <publisher>ALA Archives - University of Illinois at Urbana-Champaign</publisher>
    <theme>Library Postcards</theme>
    <subTheme>Kentucky</subTheme>
    <language>English</language>
    <workDimensionsMeasurement>14.0x10.8 cm</workDimensionsMeasurement>
    <medium>Paper</medium>
    <format>TIFF</format>
    <resourceIdentifier>ALA0000483</resourceIdentifier>
    <dateCreated>October 21 2008</dateCreated>
    <copyright>Unknown</copyright>
    <repository>University of Illinois at Urbana-Champaign. University Archives</repository>
    <collection>Sjoerd Koopman Library Postcard Collection</collection>
</metadata>
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <title>50-50: Fighting Chicago's Crime Trusts</title>
  <creator>French, George W.</creator>
  <description>Cover of the book "50-50": Fighting Chicago's Crime Trusts.</description>
  <date>1916</date>
  <type>Book Covers</type>
  <type>Drawings</type>
  <type>Cartoons</type>
  <identifier>5050fightingchic00cham_09.jp2</identifier>
  <sourcePage>http://www.archive.org/stream/5050fightingchic00cham#page/n2/mode/1up</sourcePage>
  <linkTotheFullText>http://www.archive.org/stream/5050fightingchic00cham#page/n2/mode/1up</linkTotheFullText>
  <physicalLocation>University of Illinois at Chicago</physicalLocation>
  <format>image/jp2</format>
  <language>English</language>
  <subject>Book covers</subject>
  <subject>Covers (Illustration)</subject>
  <coverageSpatial>Chicago (Ill.)</coverageSpatial>
  <coverageTemporal>1910-1919</coverageTemporal>
  <publisher>University of Illinois at Urbana-Champaign</publisher>
  <publisher>University of Illinois at Chicago</publisher>
  <rights>Images in this collection were digitized through the University of Illinois Library's participation in the Open Content Alliance and may be used freely. Attribution to the University of Illinois is appreciated. High-resolution images can be downloaded from the Internet Archive at www.archive.org. For further information, contact dcc@library.uiuc.edu.</rights>
  <subCollection>people</subCollection>
  <subCollection>Buildings</subCollection>
  <bookTitle>50-50: fighting Chicago's crime trusts</bookTitle>
  <authorOftheBook>Chamberlin, Henry Barrett, b. 1867.</authorOftheBook>
  <collection>Picture Chicago</collection>
</metadata>
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <author>Cole, Timothy W.</author>
  <author>Han, Myung-Ja K.</author>
  <title>XML for catalogers and metadata librarians</title>
  <series>Third millennium cataloging</series>
  <physicalDescription>xiv, 388 pages : illustrations ; 26 cm.</physicalDescription>
  <includes>Includes bibliographical references and index.</includes>
  <summary>"This book provides a foundation of knowledge for catalogers, metadata librarians, and library school students on the Extensible Markup Language (XML)--one of the most commonly listed qualifications in today’s cataloger and metadata librarian job postings. How are today’s librarians to manage and describe the ever-expanding volumes of resources, in both digital and print formats? The use of XML in cataloging and metadata workflows can improve metadata quality, the consistency of cataloging workflows, and adherence to standards. This book is intended to enable current and future catalogers and metadata librarians to progress beyond a bare surface-level acquaintance with XML, thereby enabling them to integrate XML technologies more fully into their cataloging workflows....."--Publisher’s website.</summary>
  <subjectLCSH>XML (Document markup language)</subjectLCSH>
  <subjectLCSH>Cataloging --Data processing.</subjectLCSH>
  <subjectLCSH>Metadata.</subjectLCSH>
  <subjectOther>LANGUAGE ARTS &amp; DISCIPLINES / Library &amp; Information Science / General.</subjectOther>
  <subjectOther>LANGUAGE ARTS &amp; DISCIPLINES / Library &amp; Information Science / Cataloging &amp; Classification.</subjectOther>
  <isbn>9781598845198 (pbk. : acid-free paper)</isbn>
  <isbn>1598845195 (pbk. : acid-free paper)</isbn>
  <isbn>9781610692915 (ebook)</isbn>
  <lcCardNumber>2013006978</lcCardNumber>
  <persistentLinkToThisPage>https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&amp;Dv1=1&amp;DBBRecID=7285233</persistentLinkToThisPage>
  <institution>University of Illinois at Urbana-Champaign Library</institution>
  <location>Shelved At Social Sciences, Health, and Education</location>
  <callNumber>Z678.93.X54 C65 2013</callNumber>
  <copyNumber>1</copyNumber>
  <status>Available</status>
</metadata>
Well-Formed XML Exercise
Solution for Resource #3, version 2

Available online
http://quest.library.illinois.edu/ALA2015/ALCTS-ExerciseSolutions/Exercise1/Well-Formed-Metadata/sample3-2.xml

<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <author>Cole, Timothy W.</author>
  <author>Han, Myung-Ja K.</author>
  <title>XML for catalogers and metadata librarians</title>
  <series>Third millennium cataloging</series>
  <publisher>Santa Barbara, California : Libraries Unlimited, an imprint of ABC-CLIO, LLC</publisher>
  <dateOfPublication>2013</dateOfPublication>
  <physicalDescription>xiv, 388 pages : illustrations ; 26 cm.</physicalDescription>
  <note>Includes bibliographical references and index.</note>
  <summary>"This book provides a foundation of knowledge for catalogers, metadata librarians, and library school students on the Extensible Markup Language (XML)--one of the most commonly listed qualifications in today’s cataloger and metadata librarian job postings..." --Publisher’s website.</summary>
  <subject authority="lcsh">XML (Document markup language)</subject>
  <subject authority="lcsh">Cataloging --Data processing.</subject>
  <subject authority="lcsh">Metadata.</subject>
  <subject authority="other">LANGUAGE ARTS &amp; DISCIPLINES / Library &amp; Information Science / General.</subject>
  <subject authority="other">LANGUAGE ARTS &amp; DISCIPLINES / Library &amp; Information Science / Cataloging &amp; Classification.</subject>
  <identifier type="ISBN">9781598845198 (pbk. : acid-free paper)</identifier>
  <identifier type="ISBN">1598845195 (pbk. : acid-free paper)</identifier>
  <identifier type="ISBN">9781610692915 (ebook)</identifier>
  <identifier type="LCCN">2013006978</identifier>
  <persistentLinkToThisPage>https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&amp;Dv1=1&amp;DBBRecID=7285233</persistentLinkToThisPage>
  <institution>University of Illinois at Urbana-Champaign Library</institution>
  <location>Shelved At Social Sciences, Health, and Education</location>
  <callNumber>Z678.93.X54 C65 2013</callNumber>
  <copyNumber>1</copyNumber>
  <status>Available</status>
</metadata>
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <title>XML in libraries</title>
  <otherFormatsOfTitle>Online version: XML in libraries. New York : Neal-Schuman Publishers, c2002</otherFormatsOfTitle>
  <editor>Tennant, Roy.</editor>
  <publisher>New York : Neal-Schuman Publishers</publisher>
  <copyrightdate>2002</copyrightdate>
  <physicalDescription>xi, 213 p. : ill. ; 23 cm.</physicalDescription>
  <includes>Includes bibliographical references and index.</includes>
  <contents>Updating MARC records with XMLMARC / Kevin S. Clarke -- Searching and retrieving XML records via the Web / Theo van Veen -- Improving interlibrary loan with XML / Kyle Banerjee -- Harnessing Oracle and XT for finding aid dissemination and search / Leslie Myrick -- Creating a unified e-government portal using XML / Lloyd Sokvitne and Jan Lavelle -- Expediting the work of the indexer with XML / Walter Lewis, Gail Richardson, and Geoff Cannon -- Using XML to federate collections: the Legacy Tobacco Documents Library / Heidi Schmidt -- Publishing books online at eScholarship / Roy Tennant -- Building XML databases with Zope and Castor / Art Rhyno -- Migrating native law cases from HTML to XML / Darlene Fichter -- Transforming word processing documents into XML: electronic scholarly publishing at the University of Michigan / Brian Rosenblum -- Encoding digital objects with METS / Jerome McDonough -- Integrating systems with XML-based Web services / Don Gourley.</contents>
  <subjectLCSH>XML (Document markup language)</subjectLCSH>
  <subjectLCSH>Libraries -- Data processing.</subjectLCSH>
  <subjectOther>XML</subjectOther>
  <isbn>1555704433 (pbk. : alk. paper)</isbn>
  <isbn>9781555704438 (pbk. : alk. paper)</isbn>
  <lcCardNumber>2002005807</lcCardNumber>
  <otherIdentifyingNumber>20020143478</otherIdentifyingNumber>
  <persistentLinkToThisPage>https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&amp;v1=1&amp;BBRecID=4554404</persistentLinkToThisPage>
  <institution>University of Illinois at Urbana-Champaign Library</institution>
  <location>Shelved At Main Stacks</location>
  <callNumber>005.72 X62</callNumber>
  <copy>1</copy>
  <status>Available</status>
</metadata>
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <title>William T. Young Library, University of Kentucky</title>
  <type>Image</type>
  <digitizedMaterial>black and white postcard</digitizedMaterial>
  <coverageSpatial>Lexington, Kentucky</coverageSpatial>
  <subject>Library buildings Kentucky</subject>
  <subject>Public libraries</subject>
  <subject>Academic libraries</subject>
  <description>On Front: William T. Young Library, University of Kentucky; On Back: The William T. Young Library opened on the University of Kentucky campus in April 1998. It is the third central library in the University's history, replacing the Margaret T. King Library which opened in 1931 and the Carnegie Library constructed in 1909. Michael McKinnell of Boston was the architect and Bovis the construction management firm. The building has over 350,000 square feet of space and cost $58 million. President Wethington's commitment to the project was critical as the facility was constructed with private funds by the UK Alumni Association and bonds were sold to finance part of the construction expenses. William T. Young of Lexington was the leading contributor to the Library.</description>
  <publisher>ALA Archives - University of Illinois at Urbana-Champaign</publisher>
  <theme>Library Postcards</theme>
  <subTheme>Kentucky</subTheme>
  <language>English</language>
  <workDimensionsMeasurement>14.0x10.8 cm</workDimensionsMeasurement>
  <medium>Paper</medium>
  <format>TIFF</format>
  <resourceIdentifier>ALA0000483</resourceIdentifier>
  <dateCreated>October 21 2008</dateCreated>
  <copyright>Unknown</copyright>
  <repository>University of Illinois at Urbana-Champaign. University Archives</repository>
  <collection>Sjoerd Koopman Library Postcard Collection</collection>
</metadata>
Well-Formed XML Exercise
An INCORRECT Solution for Resource #3

Available online
http://quest.library.illinois.edu/ALA2015/ALCTS-ExerciseSolutions/Exercise1/Not-Well-Formed-Metadata/Sample3-2Not.xml

<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <author>Cole, Timothy W.</author>
  <author>Han, Myung-Ja K.</author>
  <title>XML for catalogers and metadata librarians</title>
  <series>Third millennium cataloging</series>
  <publisher>Santa Barbara, California : Libraries Unlimited, an imprint of ABC-CLIO, LLC</publisher>
  <dateOfPublication>2013</dateOfPublication>
  <physicalDescription>xiv, 388 pages : illustrations ; 26 cm.</physicalDescription>
  <note>Includes bibliographical references and index.</note>
  <summary>This book provides a foundation of knowledge for catalogers, metadata librarians, and library school students on the Extensible Markup Language (XML)--one of the most commonly listed qualifications in today’s cataloger and metadata librarian job postings..."--Publisher’s website.</summary>
  <subject authority="lcsh">XML (Document markup language)</subject>
  <subject authority="lcsh">Cataloging --Data processing</subject>
  <subject authority="lcsh">Metadata</subject>
  <subject authority="other">LANGUAGE ARTS &amp; DISCIPLINES / Library &amp; Information Science / General</subject>
  <subject authority="other">LANGUAGE ARTS &amp; DISCIPLINES / Library & Information Science / Cataloging & Classification</subject>
  <identifier type="ISBN">9781598845198 (pbk. : acid-free paper)</identifier>
  <identifier type="ISBN">1598845195 (pbk. : acid-free paper)</identifier>
  <identifier type="ISBN">9781610692915 (ebook)</identifier>
  <identifier type="LCCN">2013006978</identifier>
  <persistentLinkToThisPage>https://i-share.carli.illinois.edu/uiu/cgi-bin/Pwebrecon.cgi?DB=local&amp;Dv1=1&amp;DBBRecID=7285233</persistentLinkToThisPage>
  <institution>University of Illinois at Urbana-Champaign Library</institution>
  <location>Shelved At Social Sciences, Health, and Education</location>
  <callNumber>Z678.93.X54 C65 2013</callNumber>
  <copyNumber>1</copyNumber>
  <status>Available</status>
</metadata>