
XML FOR INFORMATION MANAGEMENT

**WORKSHOP BACKGROUND, ACKNOWLEDGEMENTS,
AND CONTACT INFORMATION**

Workshop Background, Acknowledgements, and Contact Information

This workshop and all related materials are the direct result of a two-year grant to the State Archives Department of the Minnesota Historical Society (MHS) from the National Historical Publications and Records Commission (NHPRC). Work on the “Educating Archivists and Their Constituencies” project began in January 2001 and was completed in May 2003.

The project sought to address a critical responsibility that archives have discovered in their work with electronic records: the persistent need to educate a variety of constituencies about the principles, products, and resources necessary to implement archival considerations in the application of information technology to government functions. Several other goals were also supported:

- raise the level of knowledge and understanding of essential electronic records skills and tools among archivists,
- help archivists reach the electronic records creators who are their key constituencies,
- provide the means to form with those constituencies communities of learning that will support and sustain collaboration, and
- raise the profile of archivists in their own organizations and promote their involvement in the design and analysis of recordkeeping systems.

MHS administered the project and worked in collaboration with several partners: the Delaware Public Archives, the Indiana University Archives, the Ohio Historical Society, the San Diego Supercomputer Center, the Smithsonian Institution Archives, and the State of Kentucky. This list represents a variety of institutions, records environments, constituencies, needs, and levels of electronic records expertise. At MHS, Robert Horton served as the Project Director, Shawn Rounds as the Project Manager, and Jennifer Johnson as the Project Archivist.

MHS gratefully acknowledges the contribution of Advanced Strategies, Inc. (ASI) of Atlanta, Georgia, and Saint Paul, Minnesota, which specializes in a user-centric approach to all aspects of information technology planning and implementation. MHS project staff received training and guidance from ASI in adult education strategies and workshop development. The format of this course book is directly based on the design used by ASI in its own classes. For more information about ASI, visit <http://www.advstr.com/>

For more information regarding the workshop, contact MHS staff or visit the workshop web site at <http://www.mnhs.org/preserve/records/workshops/edarchivists.html>

Robert Horton: robert.horton@mnhs.org / 651-215-5866
Shawn Rounds: shawn.rounds@mnhs.org / 651-296-7953

INTRODUCTION

Introduction

This unit includes:

- Course objectives
- Course approach
- Course schedule
- Administrative items
- Participant introductions

Introduction

Course objectives

Upon completion of this course, you will be able to:

- understand basic information technology concepts and terminology
- understand what XML is and why it is useful
- understand the reasons for the development of XML
- comprehend the relationship between metadata and XML
- recognize XML markup
- identify other components of the XML standard
- identify XML tools and editors
- understand how XML is being implemented in various projects

Introduction

Course approach

- Spiral approach
- About the exercises
 - o The exercises in this course will be “live”
 - o The exercises will be limited in number
 - o The exercises may require some oral participation from students
 - o There is not a 100% correct answer; it is impossible to get an A, because there are many “right” answers
 - o Today our exercises will focus on a recipe
 - Recipes are familiar to everyone
 - There are various things you are able to do with a recipe, which fit nicely with the variety of things you are able to do with XML
 - Recipes have obvious, tangible results
- Parking lot
- Team teaching



Introduction

Course schedule

Introduction

Unit I: Defining XML

Break

Unit II: What Does XML Look Like?

Break

Unit III: Presenting XML

Lunch

Unit IV: XML Tools and Editors

Unit V: The Family of XML Standards

Break

Unit VI: Using XML

Unit VII: Encoded Archival Description

Break

Unit VIII: Putting It All Together

Appendices

Very flexible - driven by your needs.



Introduction

Administrative items

- Class will begin promptly at the scheduled time.
- Need a break: squirm, lie down, tell me!
- If something is not clear, if you have a question, or if I say something that does not make any sense: Stop me!
- Discussion is always better than lecture. Success of this workshop is dependent on all participants.
- Facilities (Parking, Restroom, Telephones, Exits)
- We will proceed informally.

Introduction

Participant introductions

NAME and ORGANIZATION: _____

Tell us a little bit about your program (think of what you wrote in your self-evaluation):

What experience have you had using XML?

What projects are you participating in that currently use XML, or do you know of any projects using XML?

Who do you know that is using XML?

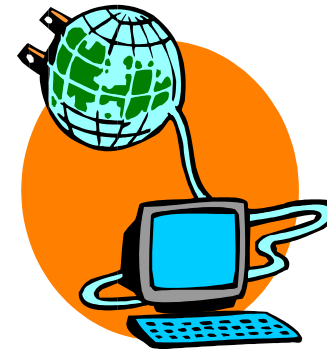
Who might you partner with for future XML projects?

UNIT I: DEFINING XML

Defining XML

This unit includes:

- What do we mean by information resources, digital objects, and electronic records?
- Defining digital objects.
- One person or organization cannot do it all.
- What is eXtensible Markup Language (XML)?
- Why XML?
- A quick definition of XML.
- Explaining the XML standard.
- Comparing XML to Standard Generalized Markup Language and Hypertext Markup Language.
- Defining well-formed and valid XML.



Defining XML

What do we mean by information resources, digital objects, and electronic records?

Information resources: The content of your information technology projects (data, information, records, images, digital objects, etc.)

Digital object: Information that is inscribed on a tangible medium or that is stored in an electronic or other medium and is retrievable in perceivable form. An object created, generated, sent, communicated, received, or stored by electronic means.¹

An electronic record is a specific type of digital object with unique characteristics described by archivists and records managers.

Types of digital objects:

- e-mail
- web pages
- databases
- spreadsheets
- word processing documents
- Portable Document Format (PDF) files
- PowerPoint presentations
- digital images
- ...and many more

¹ Electronic Signatures in Global and National Commerce Act (E-Sign). <http://thomas.loc.gov/cgi-bin/query/z?c106:S.761>:

Defining XML

Digital objects have three components:

Content: Informational substance of the object.

Structure: Technical characteristics of the objects (e.g., presentation, appearance, display).

Context: Information outside the object which provides illumination or understanding about it, or assigns meaning to it.

Defining XML

Defining information objects

Pittsburgh Project Definition

Content
Structure
Context

Order of Values

Data
Information
Knowledge

Information Technology Architecture

Data
Format
Application



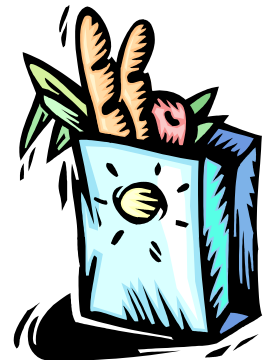
Defining XML

One person or organization cannot do it all

Components: $X + Y + Z =$ resources, tools, standards/methodologies, education, technology, partners

- Resources: skilled and informed staff, organizational funding, grants, ...
- Tools: metadata, markup languages, ...
- Standards/methodologies: Dublin Core, XML, ...
- Education: formal/informal education, workshops, conferences, ...
- Technology: TagGen, XMetal, ...
- Partners: stakeholders [record creators, agencies, users], other archivists, professional organizations [Midwest Archives Conference (MAC), Society of American Archivists (SAA), National Association of Government Archives and Records Administrators (NAGARA), Association of Records Managers and Administrators (ARMA)], ...

This class will cover one tool and standard: XML



Defining XML

Exercise: What do you think eXtensible Markup Language (XML) is?

Defining XML

Language means communication and communication leads to understanding

What makes understanding possible?

- vocabulary
- dictionary
- grammar

It's not just semantics. This is the structure of an "unstructured" text. It is executable knowledge.

*"When I say a word, it
means exactly what I
want it to mean."*

Humpty Dumpty –
Alice in Wonderland

Defining XML

What does eXtensible Markup Language mean?

- eXtensible:** In XML, you create the tags you want to use. XML extends your ability to describe a document, letting you define meaningful tags for your applications. For example, if your document contains many glossary terms, you can create a tag called <glossary> for those terms. If it contains employee identification numbers, you could use an <employeeid> tag. You can create as few or as many tags as you need.
- Markup:** Any means of making explicit an interpretation of a text. In this instance, a notation for writing text with tags. The tags may indicate the structure of the text, they may have names and attributes, and they enclose a part of the text.
- Language:** XML is designed to facilitate communication. It follows a firm set of rules that allow you to say what you want in a way that others will understand. It may let you create an extensible set of markup tags, but its structure and syntax remain firm and clearly defined.

Defining XML

Why XML?

- Share data** Different organizations rarely use the same tools to create and read data. XML can be used to store any kind of structured information, and to enclose or encapsulate it in order to pass the information between different computing systems which would otherwise be unable to communicate.
- Reuse data** XML documents can be moved to any format on any platform - without the elements losing their meaning. This means you can publish the same information to a web browser, or a personal digital assistant (PDA), and each device would use the information appropriately. XML can be designed in such a way that fragments or chunks can be pulled out of any given context and reused. So, when a chunk is updated, the resources that use the chunk are updated also.
- Customize data** XML allows for the development of user-defined document types. Users define the XML tags they want to encapsulate their data. XML also allows groups of people or organizations to create their own customized markup languages for exchanging information in their domain.

Defining XML

A Quick definition of XML

“Extensible Markup Language (XML): An initiative from the W3C defining an “extremely simple” dialect of SGML suitable for use on the World-Wide Web.”

“XML: A metalanguage written in SGML that allows one to design a markup language, used to allow for the easy interchange of documents on the World Wide Web.”

(Dictionary.com)

“XML (Extensible Markup Language) is a flexible way to create common information formats and share both the format and the data on the World Wide Web, intranets, and elsewhere.”

(Whatis.com)

“The Extensible Markup Language (XML) is the universal format for structured documents and data on the Web.”

(World Wide Web Consortium (W3C))

Defining XML

The XML standard ^{2 3 4}

Ten Design Goals for XML:

1. XML shall be straightforwardly usable over the Internet.
2. XML shall support a wide variety of applications.
3. XML shall be compatible with SGML.
4. It shall be easy to write programs which process XML documents.
5. The number of optional features in XML is to be kept to the absolute minimum, ideally zero.
6. XML documents should be human-legible and reasonably clear.
7. The XML design should be prepared quickly.
8. The design of XML shall be formal and concise.
9. XML documents shall be easy to create.
10. Terseness in XML markup is of minimal importance.

In other words, XML is easy to create, easy to read, and designed for use over the Internet.

² World Wide Web Consortium (W3C). <http://www.w3.org/>

³ W3C Extensible Markup Language (XML). <http://www.w3.org/XML/>

⁴ Extensible Markup Language (XML) 1.0 (Second Edition). W3C Recommendation 6 October 2000. <http://www.w3.org/TR/REC-xml>

Defining XML

Comparing SGML, HTML, and XML

	Standard Generalized Markup Language (SGML)	Hypertext Markup Language (HTML)	eXtensible Markup Language (XML)
Standard	ISO standard 1986 [ISO 8879:1985]. Approximately 500 pages.	Current version HTML 4.01. Recommendation first released on 18th December 1997. 389 pages. Created by Tim Berners-Lee and Robert Caillau at CERN in 1991.	eXtensible Markup Language (XML) 1.0 (Second Edition). W3C Recommendation 6 October 2000. 59 pages.
Def.	An international documentation standard for defining descriptions of the structure of different types of electronic documents.	Web-based document presentation standard.	Document processing standard.
Use	Used for describing thousands of different document types in many fields. Used primarily for technical documentation and government applications, and in industries with huge documents, such as medical records, company databases, and aircraft parts catalogs.	Used worldwide by anyone creating web pages.	Gaining use.
Description	A method for creating interchangeable, structured documents. Content markup.	HTML is a SGML application used on the web. Markup for presentation or format. Most HTML browsers do not support basic SGML constructions, but nearly all SGML authoring tools are capable of producing good HTML documents.	XML is not a SGML application, rather a set of simple conventions for using SGML without some of the more esoteric features. Known as simplified SGML, XML is a lightweight cut-down version of SGML which keeps enough of its functionality to make it useful but removes all the optional features, and the more complex and less-used parts of SGML, in return for being easier to write applications for, easier to understand, and more suited to delivery and interoperability over the Web.

Defining XML

Comparing XML and HTML

Similarities:

- Both markup languages use tags `<xxx> yyy </xxx>`
- Web browsers display both markup languages
- Both markup languages have W3C working groups

Differences

- HTML tags describe how a document looks
XML tags describe what a document means
- HTML files can be sloppy
XML files have to be precise and exact – they have to be well-formed and valid

Defining XML

Well-formed and valid XML

Well-formed XML: Documents that follow XML tag rules, and each document is a complete, self-contained object.

XML tag rules:

- A root element is required
- Case matters: `<recipe>`, `<Recipe>`, `<RECIPE>` are not the same tag

- No unclosed tags

Incorrect: `<title>Original Nestle Toll House Chocolate Chip Cookies Recipe`

Correct: `<title>Original Nestle Toll House Chocolate Chip Cookies Recipe</title>`

- No overlapping tags, they must be properly nested

Incorrect: `<background><author>Ruth Wakefield</background></author>`

Correct: `<background><author>Ruth Wakefield</author></background>`

- Attribute values must be enclosed in quotes

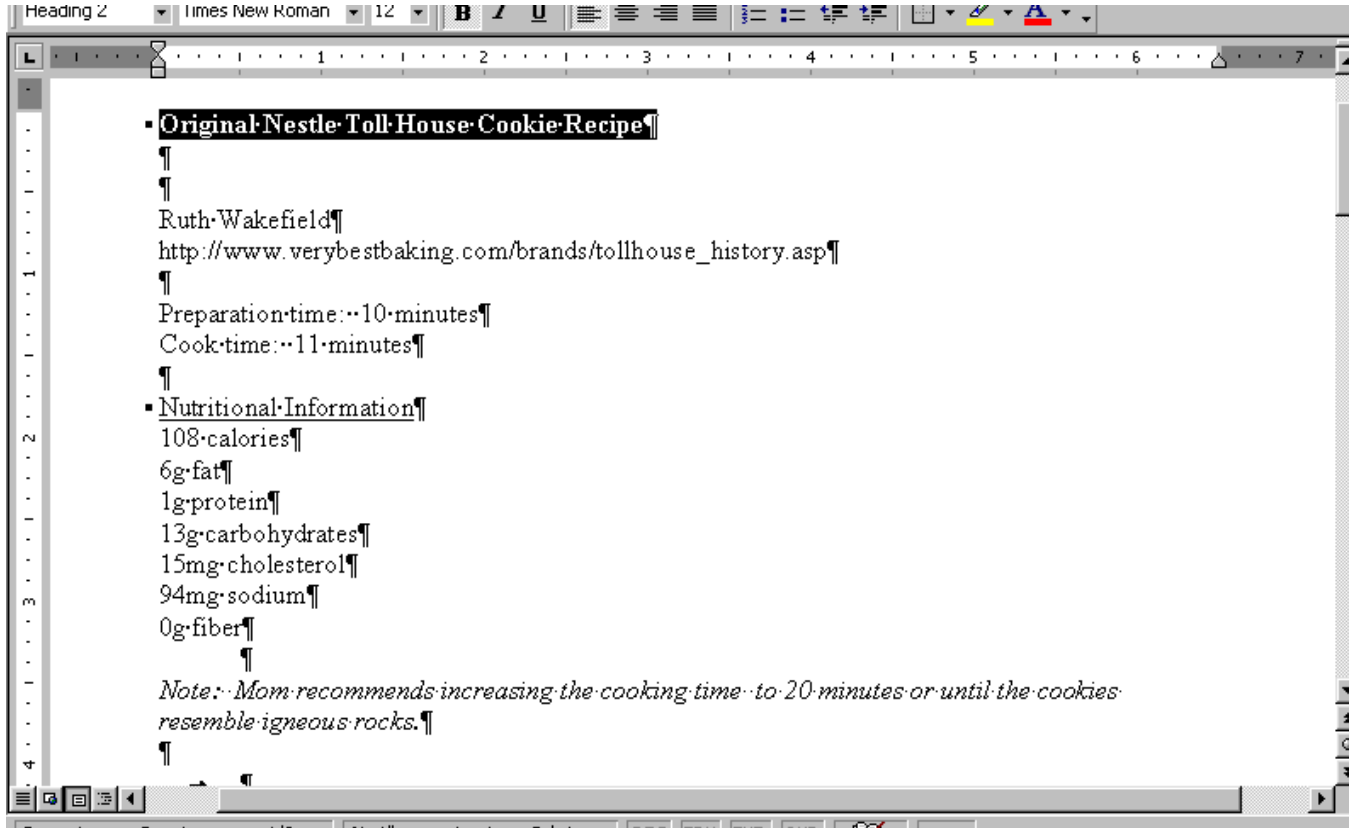
`<item quantity="2" unit="cup 12 oz pkg.">Nestle Toll House semi-sweet chocolate morsels</item>`

Valid XML: When it obeys the rules – the words are in the dictionary and the format is grammatical. In other words, a document that follows both the XML tag rules and the rules defined in its DTD or XML Schema.

Defining XML

Exercise: Comparing XML and HTML using a recipe

Part I: What can we do with a recipe? ⁵



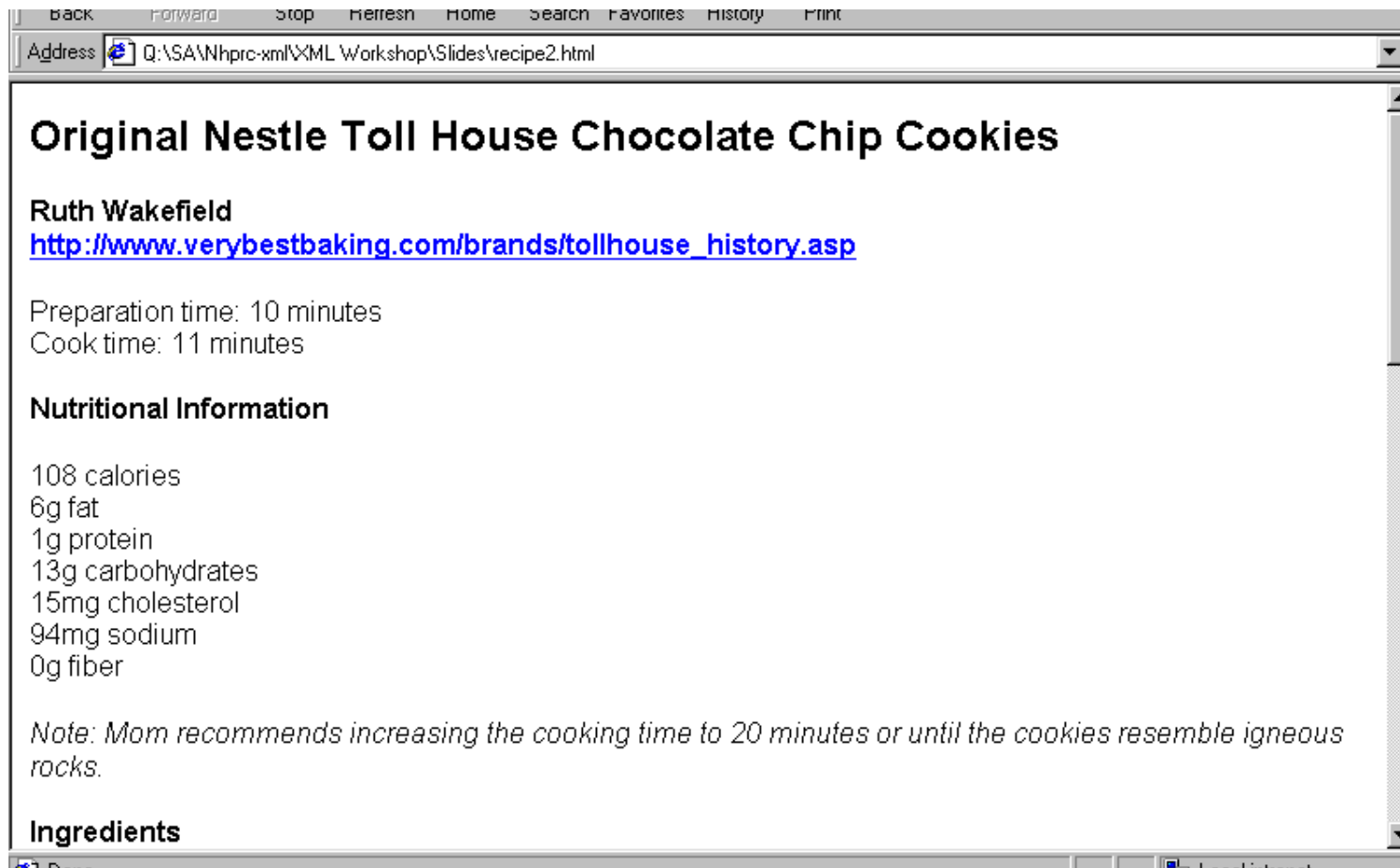
⁵ http://www.verybestbaking.com/recipes/recipe_template.asp?id=614

Defining XML

Exercise: Comparing XML and HTML using a recipe

Part II: What does a recipe look like in HTML and XML?

HTML as viewed in a browser



Defining XML

Exercise: Comparing XML and HTML using a recipe

Part II: What does a recipe look like in HTML and XML?

XML as viewed in a browser

```
Address | U:\SA\Nhpoc-xml\XML Workshop\Slides\recipe2.xml
<?xml version="1.0" ?>
- <recipe>
  <title>Original Nestle Toll House Cookies</title>
  - <background>
    <author>Ruth Wakefield</author>
    <history>http://www.verybestbaking.com/brands/tollhouse\_history.asp</history>
  </background>
  - <recipe_info>
    <prep_time>10 minutes</prep_time>
    <cook_time>11 minutes</cook_time>
  </recipe_info>
  - <nutritional_info>
    <calories>108</calories>
    <fat>6g</fat>
    <protein>1g</protein>
    <carbohydrates>13g</carbohydrates>
    <cholesterol>15 mg</cholesterol>
    <sodium>94 mg</sodium>
    <fiber>0 g</fiber>
  </nutritional_info>
  <comments>Mom reccomends increasing the cooking time to 20 minutes or until the cookies resemble igneous rocks.</comments>
  - <ingredients>
    <item quantity="1" unit="cup">butter</item>
    <item quantity="3/4" unit="cup">brown sugar</item>
```

Defining XML

Exercise: A Markup exercise

A joke.

Two North Dakotans come into a bar, slapping each other on the back, laughing, clearly happy as clams. One says to the bartender, "We're celebrating! Give everybody a round on us!"

The bartender says, "So what's the big deal? What are you celebrating?"

And the North Dakotan says, "We just finished a jigsaw puzzle and it only took us four days."

The bartender says, "A jigsaw puzzle? Two people? Four days? That doesn't sound like much reason to celebrate."

And the other North Dakotan says, "Are you kidding? The box said '2-3 Years.'"

Defining XML

A markup exercise example

A joke.

```
<?xml version="1.0?">
<text>
  <paragraph>
    <sentence type="expository">Two North Dakotans come into a bar, slapping each other on the back, laughing, clearly happy as
      clams.</sentence>
    <sentence type="exclamation">One says to the bartender,<quotation> "We're celebrating! Give everybody a round on
      us!"</quotation></sentence>
  </paragraph>
  <paragraph>
    <sentence type="question">The bartender says, <quotation>"So what's the big deal? What are you
      celebrating?"</quotation></sentence>
  </paragraph>
  <paragraph>
    <sentence type="expository">And the North Dakotan says, <quotation>"We just finished a jigsaw puzzle and it only took us four
      days."</quotation></sentence>
  </paragraph>
  <paragraph>
    <sentence type="other">The bartender says, <quotation>"A jigsaw puzzle? Two people? Four days? That doesn't sound like
      much reason to celebrate."</quotation></sentence>
  </paragraph>
  <paragraph>
    <sentence type="other">And the other North Dakotan says, <quotation>"Are you kidding? The box said '2-3
      Years.'"</quotation></sentence>
  </paragraph>
</text>
```

Defining XML

A markup exercise example

A joke.

```
<?xml version="1.0?>
```

```
<story>
```

```
<setting>Two North Dakotans come into a bar, slapping each other on the back, laughing, clearly happy as clams.
```

```
</setting>
```

```
<dialogue>
```

```
<character1>One</character1> says to the <character2>bartender</character2>, "We're celebrating! Give everybody a round on us!"
```

```
The bartender says, "So what's the big deal? What are you celebrating?"
```

```
And the North Dakotan says, "We just finished a jigsaw puzzle and it only took us four days."
```

```
The bartender says, "A jigsaw puzzle? Two people? Four days? That doesn't sound like much reason to celebrate."
```

```
And the <character3>other North Dakotan</character3> says, "Are you kidding? The box said '2-3 Years.'"
```

```
</dialogue>
```

```
</story>
```

Defining XML

A markup exercise example

A joke.

```
<?xml version="1.0?">  
<humor>
```

```
  <joke taste="questionable">
```

```
    Two <ethnic subject>North Dakotans</ethnic subject> come into a bar, slapping each other on the back, laughing, clearly happy  
    as clams. One says to the <ethnic subject>bartender</ethnic subject>, "We're celebrating! Give everybody a round on us!"
```

```
    The <ethnic subject>bartender</ethnic subject> says, "So what's the big deal? What are you celebrating?"
```

```
    And the <ethnic subject>North Dakotan</ethnic subject> says, "We just finished a jigsaw puzzle and it only took us four days."
```

```
    The <ethnic subject>bartender</ethnic subject> says, "A jigsaw puzzle? Two people? Four days? That doesn't sound like much  
    reason to celebrate."
```

```
    <punchline>And the other <ethnic subject>North Dakotan</ethnic subject> says, "Are you kidding? The box said '2-3 Years.'"  
  </punchline>
```

```
</joke>
```

```
</humor>
```

Defining XML

Key Messages

- A digital object may be created, generated, sent, communicated, received, or stored by electronic means.
- An electronic record is a specific type of digital object with unique characteristics described by archivists and records managers.
- In order to preserve content, context, and structure in digital form, we are going to have to move to capturing data, format, and application.
- One person or organization cannot do it all. You are going to have to make choices.
- XML is meant to be easy to create, easy to read, and is designed for use over the Internet.
- eXtensible Markup Language is a means of marking up data, using a specific syntax that you define.
- XML allows you to share, reuse, and customize data.
- XML is an international standard supported by the World Wide Web Consortium.
- Both HTML and XML were derived from SGML.
- HTML and XML are similar in that they both use markup tags; however, HTML's markup defines a document's format, whereas XML's markup defines a document's content.
- Well-formed XML documents follow XML tag rules, and each document is a complete, self-contained object.
- Valid XML documents obey the rules – the words are in the dictionary and the format is grammatical. In other words, a document that follows both the XML tag rules and the rules defined in its DTD or XML Schema.

UNIT II: WHAT DOES XML LOOK LIKE?

What Does XML Look Like?

This unit includes:

- What is an XML document?
- The metadata and XML relationship.
- Document Type Definition (DTD).
- XML Schema.
- Namespaces.
- What is involved in using XML?

What Does XML Look Like?

What is an XML document?

Declaration: Declares what version of XML you are using. Appears first in an XML document. Also called a processing instruction.

```
<?xml version="1.0" standalone="yes"?>
```

Elements: The most basic unit of an XML document. The name of the element (defined by you) should assign some meaning to the content.

```
<recipe>
  <title>Original Nestle Toll House Chocolate Chip Cookies</title>
  <background>
    <author>Ruth Wakefield</author>
  </background>
</recipe>
```

Attributes: Additional data elements that help to more accurately describe an element. Attributes have quotation-mark delimited values that further describe the purpose and content of an element. Information contained in an attribute is generally considered metadata.

```
<ingredients>
  <item quantity="1" unit="12 oz pkg.">Nestle Toll House semi-sweet chocolate morsels</item>
</ingredients>
```

What Does XML Look Like?

The metadata and XML relationship

XML is a meta language. XML is really about adding layers of information to your data, so that the data can be processed, used, and transferred between applications.

Different definitions of metadata:

- data about data
- information about information
- descriptive information which facilitates management of, and access to, other information
- evaluation tool

Metadata assists you with the discovery, description, evaluation, and management of records

The XML and metadata connection

- Elements and attributes are all metadata. They assign meaning to the text within the tags.
- The decision of whether to present your information as attributes or sub-elements will depend on your business needs.

What Does XML Look Like?

Document Type Definition (DTD)

The document which holds the rules that govern what makes an XML document valid. A standard mechanism for defining what elements and attributes may be used in an XML document, where they may appear, and indicating their relationship to one another within the document. In other words, a DTD is the grammar of an XML document.

- A DTD may be internal to an XML document, or external.

Internal

```
<?xml version="1.0" standalone="yes"?>
```

External

```
<?xml version="1.0" URI HERE>
```

- A DTD may be public or private. Private meaning that you or your organization has created it, whereas a public DTD is defined by a standards body.

What Does XML Look Like?

Exercise: Document Type Definition of a recipe

```
<!DOCTYPE recipe[
<!ELEMENT recipe (title, background, recipe_info, nutritional_info, comments, ingredients, directions)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT background (author, history)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT history (#PCDATA)>
<!ELEMENT recipe_info (prep_time, cook_time)>
<!ELEMENT cook_time (#PCDATA)>
<!ELEMENT prep_time (#PCDATA)>
<!ELEMENT nutritional_info (calories, fat, protein, carbohydrates, cholesterol, sodium, fiber)>
<!ELEMENT protein (#PCDATA)>
<!ELEMENT calories (#PCDATA)>
<!ELEMENT carbohydrates (#PCDATA)>
<!ELEMENT sodium (#PCDATA)>
<!ELEMENT cholesterol (#PCDATA)>
<!ELEMENT fat (#PCDATA)>
<!ELEMENT comments (#PCDATA)>
<!ELEMENT fiber (#PCDATA)>
<!ELEMENT ingredients (item+)>
<!ELEMENT directions (directions_standard, directions_variation+)>
<!ELEMENT item (#PCDATA)>
<!ELEMENT directions_standard (step+)>
<!ELEMENT directions_variation (variation_name+, step+, variation_comment?)>

<!ATTLIST item
    quantity CDATA #REQUIRED
    unit CDATA #REQUIRED>
<!ELEMENT step (#PCDATA)>
<!ELEMENT variation_comment (#PCDATA)>
<!ELEMENT variation_name (#PCDATA)>
]>
```

What Does XML Look Like?

XML Schema ⁶

Specifies the structure of an XML document and constraints on its content. A schema defines the grammar of an XML document and is for validation.

What are the downsides of using DTDs?

- a. DTDs do not follow XML syntax and semantics
- b. DTDs are mixed into the XML 1.0 specification
- c. No support for Namespaces
- d. DTDs are difficult to extend
- e. No support for schema evolution, extension, or inheritance of declarations
- f. No embedded, structured self-documentation
- g. Defaults cannot be specified separate from the declarations
- h. DTDs cannot specify data types

⁶ XML Schema. <http://www.w3.org/XML/Schema>

What Does XML Look Like?

XML Schema continued

What are the benefits of XML Schemas?

- XML Schema is expressed in well-formed XML. DTDs are not expressed in XML language.
 - XML Schema offers better control over grouping of elements and attributes.
 - Allows you to define global element (those that must be used in the same way throughout the XML document) and local elements (those that can have a particular meaning in a particular context).
 - Offers an extensive system of datatypes that you can specify for a given element. For example, an element may be an integer, contain a period of time, contain a string, boolean, a language code, etc. DTDs are unable to restrict character data to a pattern.
 - XML Schema supports the use of namespaces.
- See the following web site for a list of Document Type Definitions and XML Schemas promoted by every conceivable industry:
http://www.xml.org/xml/registry_searchresults.jsp

What Does XML Look Like?

Namespaces ⁷

XML namespace: In order for XML documents to be able to use elements and attributes that have the same name but come from different sources, there must be a way to differentiate between the markup elements that come from the different sources.

- XML namespaces are used in XML documents so that elements with the same name, but with different purposes, can be used in the same document.

For example, one instance of a <table> tag may refer to a data structure with rows and columns while another instance of a <table> tag may refer to a four-legged piece of furniture.

- Appears as “prefix”:"element name”

For example, f:table or h:table

- Namespaces can only assure that names are unique and unambiguous. They have nothing to do with document validity.

⁷ Namespaces in XML. World Wide Web Consortium, 14 January 1999. A W3C recommendation. <http://www.w3.org/TR/REC-xml-names/>

What Does XML Look Like?

Exercise: XML Schema of a recipe

What Does XML Look Like?

What is involved in using XML?

Equipment

- Parser: Program or class that can read any well-formed XML at its input.
 - o Non-validating parser: All XML parsers check the well-formedness of documents.
 - o Validating parser: Validating parsers also confirm whether the document is *valid*; that is, that the structure and number of tags make sense.
- Browser capable of reading/working with XML
 - o Internet Explorer 5.0 (and higher) ⁸
 - o Mozilla 1.0 ⁹
 - o Opera 5 (and higher) ¹⁰
 - o Netscape 7.0 ¹¹

⁸ <http://www.microsoft.com/windows/ie/default.asp>

⁹ <http://www.mozilla.org/>

¹⁰ <http://www.opera.com/>

¹¹ <http://www.netscape.com/>

What Does XML Look Like?

Key messages

- An XML document is made up of a declaration, elements, and attributes.
- A DTD holds the rules that govern what makes an XML document valid.
- An XML Schema specifies the structure of an XML document and constraints on its content.
- DTDs and XML schemas have similar functions. DTDs are more widely used, since the XML schema specification is still new.
- In order to use an XML document, you may need to have the following:
 - o a parser
 - o a browser capable of reading/working with XML

UNIT III: PRESENTING XML

Presenting XML

This unit includes:

- eXtensible Stylesheet Language (XSL)
- XSL Transformations (XSLT)
- XHTML

Presenting XML

eXtensible Stylesheet Language (XSL) ¹² ¹³

A language for expressing stylesheets.

Stylesheet: A definition of a document's appearance or layout in terms of such elements as default typeface, size, and color of headings and body text, how sections are laid out in terms of space, line spacing, margin widths on all sides, spacing between headings, etc. Typically expressed at the beginning of an electronic document. May be embedded in or linked to a document.

XSL is a family of applications:

- XSL Transformations (XSLT)
- XML Path Language (XPath)
- XSL Formatting Objects (XSL-FO)

¹² The Extensible Stylesheet Language. <http://www.w3.org/Style/XSL/>

¹³ Extensible Stylesheet Language (XSL), Version 1.0. W3C Recommendation 15 October 2001. <http://www.w3.org/TR/xsl/>

Presenting XML

XSL Transformations (XSLT) ¹⁴

A language for transforming XML documents. A tool which uses XSL to act on XML documents. XSLT is used to transform XML document contents into something else more suitable for a particular task.

Why would we want to transform a document from one format into another?

- store in one format, display in another
- convert to a more useful format

Example of implementation: Outputting Encoded Archival Description (EAD) documents to HTML for presentation on the World Wide Web.

¹⁴ XSL Transformations (XSLT), version 1.0. W3C Recommendation 16 November 1999. <http://www.w3.org/TR/xslt>

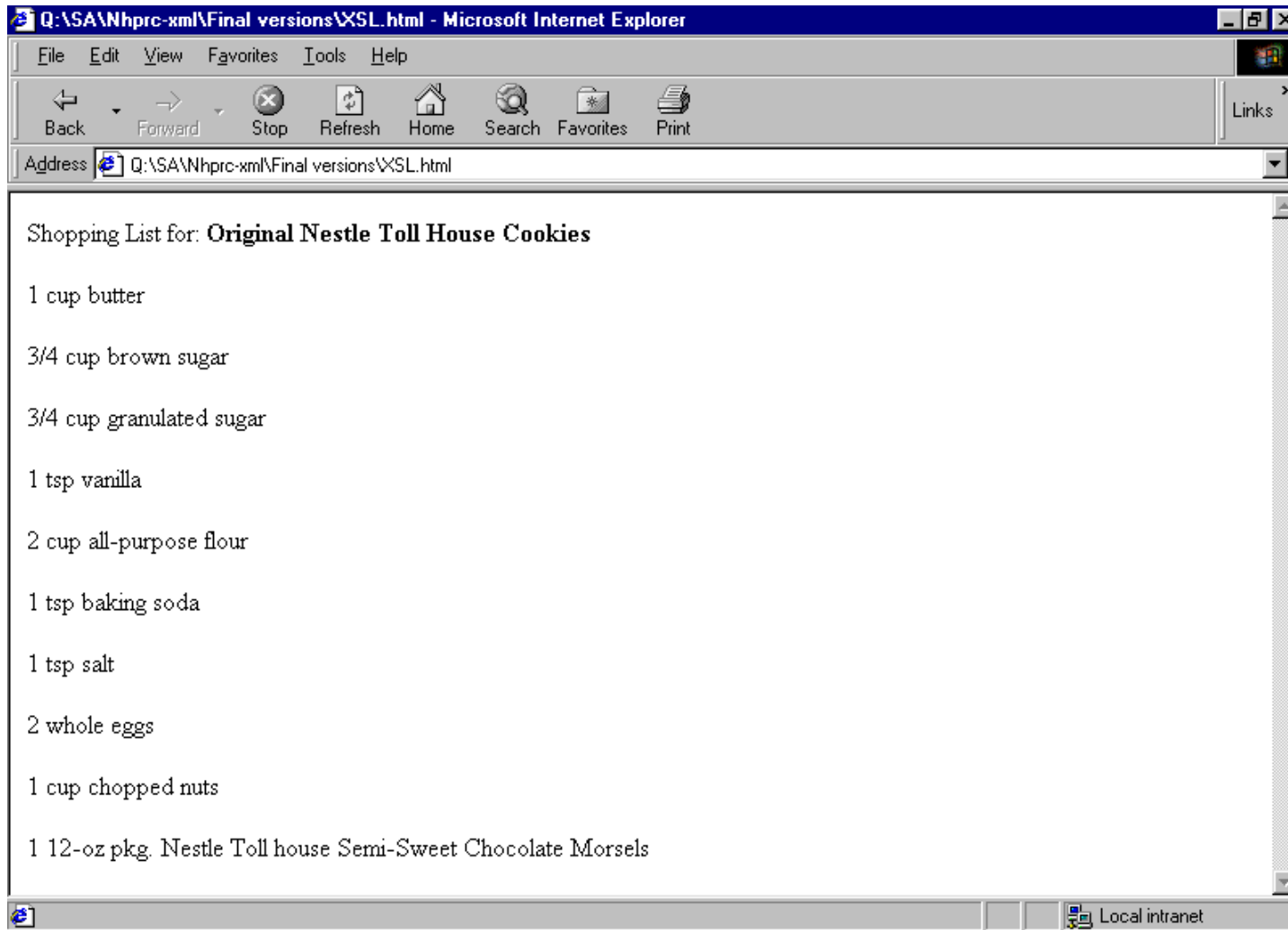
Presenting XML

Example: eXtensible Stylesheet Language (XSL) of our recipe

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:template match="/">
    <html>
      <head/>
      <body>
        <p>Shopping List for: <b><xsl:value-of select="recipe/title"/></b></p>
        <xsl:for-each select="recipe/ingredients/item">
          <p>
            <xsl:value-of select="@quantity"/>
            <xsl:text> </xsl:text>
            <xsl:value-of select="@unit"/>
            <xsl:text> </xsl:text>
            <xsl:value-of select="."/>
          </p>
        </xsl:for-each>
      </body>
    </html>
  </xsl:template>
</xsl:stylesheet>
```

Presenting XML

Example: XSL of a recipe. How does the above style sheet display in a browser?



Presenting XML

XHTML¹⁵

XHTML: a reformulation of HTML 4 as an XML application.

- HTML tags written with XMLs strict syntactic rules

Why should you use XHTML over HTML?

- XHTML is an XML conforming standard. XHTML documents can be used with any general-purpose XML editor, validator, browser, or other program to work on XML docs.
- Documents that follow the stricter XML rules are cleaner, more predictable, and better-behaved in browsers and XML software
- The extensible qualities of XML will benefit XHTML in the long run, making it easier to add new elements and functionality.

Three flavors of XHTML

- Strict: a clean break from current HTML, many HTML elements deprecated, greater use of Cascading Stylesheets (CSS)
- Transitional: if you want your pages to remain compatible with older browsers that do not support stylesheets, retains the elements and attributes of HTML
- Frameset: like strict XHTML with the ability to use frames

¹⁵ XHTML™ 2.0: The Extensible HyperText Markup Language. A reformulation of HTML 4 in XML 1.0. W3C Working Draft 5 August 2002. <http://www.w3.org/TR/2002/WD-xhtml2-20020805/>

Presenting XML

Key messages

- XSL, XSLT, and XHTML are extensions of XML used for presentation of documents.
- XSL uses stylesheets to control how XML documents are presented
- XSLT is a language for transforming XML documents from one format into another
- XHTML is a transition from HTML to XML. XHTML is HTML tags written in compliance with XML's strict syntactic rules

UNIT IV: XML TOOLS AND EDITORS

XML Tools and Editors

This unit includes:

- A discussion of why you need tools and editors
- A brief look at XML tools and editor that are free

XML Tools and Editors

A discussion about why you need tools and editors

- To create XML tags
- To create different types of XML documents
- To validate XML documents

XML Tools and Editors

Free tools and editors

XML Editor:

- XML Cooktop
<http://www.xmlcooktop.com/>

Online tools

Parser

- Userland Frontier
[http://frontier.userland.com/stories/storyReader\\$1092](http://frontier.userland.com/stories/storyReader$1092)

Well-formedness and validator

- <http://www.cogsci.ed.ac.uk/~richard/xml-check.html>

Other

- The Apache XML Project
<http://xml.apache.org>
- netbeans.org
<http://www.netbeans.org/nonav/index2.html>

XML Tools and Editors

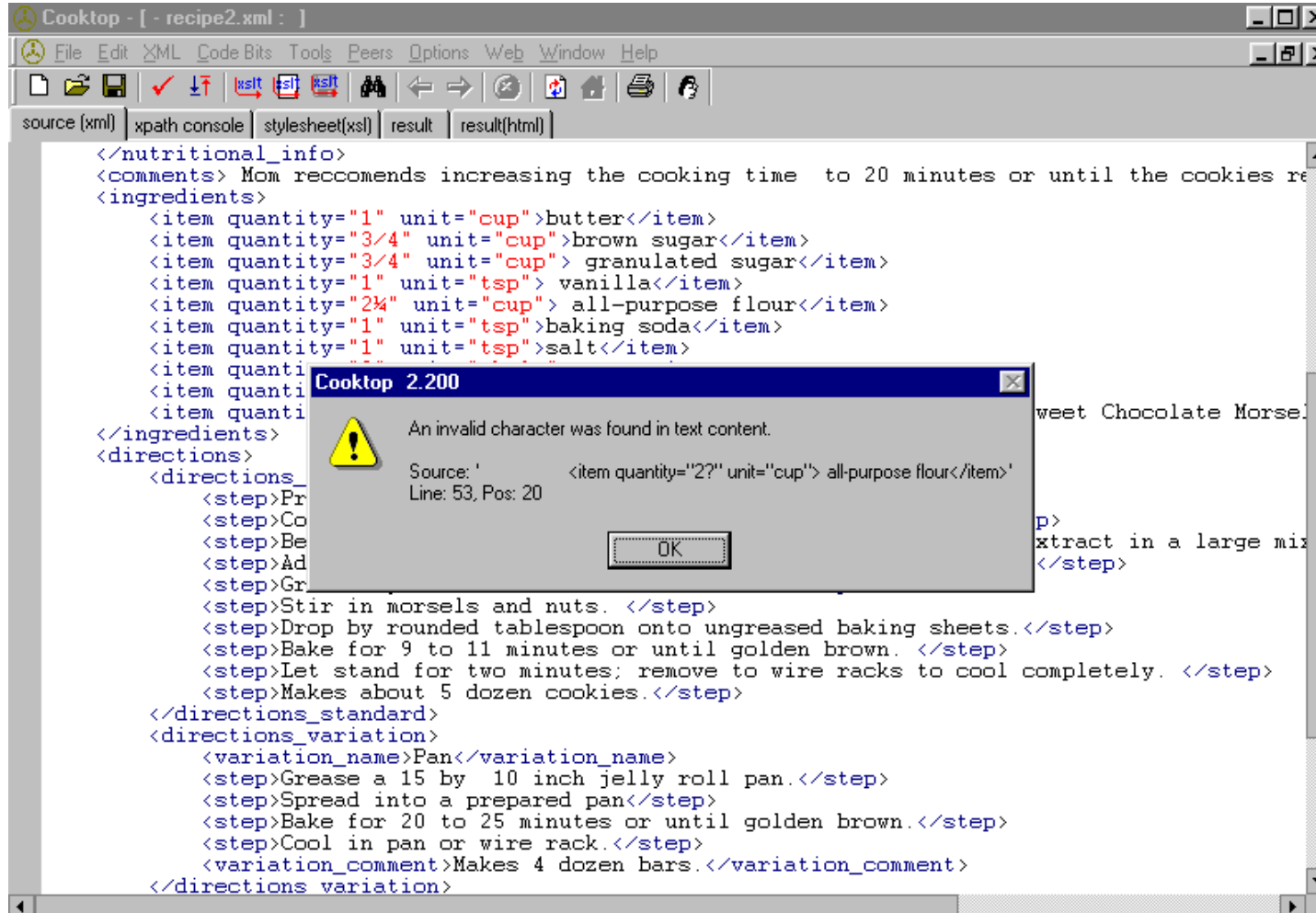
Free tools and editors continued

URLs with access to many tools:

- XML.com
<http://www.xml.com/buyersguide/>
- XML Cover Pages
<http://www.oasis-open.org/cover/publicSW.html#xmlTools>
- <http://www.xmlsoftware.com>
- Information by Lars M. Garshol
<http://www.garshol.priv.no/download/>
- sourceforge.net
<http://sourceforge.net/>

XML Tools and Editors

Screenshot of XML Cooktop



XML Tools and Editors

Key Messages

- XML tools and editors facilitate your work
- There are free tools and editors out there

UNIT V: THE FAMILY OF XML STANDARDS

The Family of XML Standards

This unit includes:

- XPath
- XPointer
- XLink
- XSL Formatting Objects (XSL-FO)
- XForms
- SVG
- XQuery

The Family of XML Standards

XPATH ¹⁶

The primary purpose of XPath is to address and access parts of an XML document.

- XPath models XML as a tree of nodes. XPath helps you get to a point on the tree
- Used to identify particular parts of XML documents.
- XPath functions as part of XSLT, addressing the parts of an XML document that an author wishes to transform.



¹⁶ XML Path Language (XPath) Version 1.0. W3C recommendation, 16 November 1999. <http://www.w3.org/TR/xpath>

The Family of XML Standards

XPOINTER ¹⁷

XPointer supports addressing into the internal structures of XML documents.

XPointer specifies a mechanism for pointing to arbitrary chunks (fragments) of a target document.

XPointer enables you to target a given element by number, name, type, or relation, to other elements in the document.



¹⁷ XML Pointer Language (XPointer) Version 1.0. W3C recommendation 27 June 2001. <http://www.w3.org/TR/xptr>

The Family of XML Standards

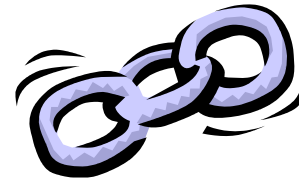
XLINK¹⁸

XLink allows elements to be inserted into XML documents in order to create and describe links between resources.

XLink works by providing global attributes you can use to mark your elements as *linking elements*.

XLink provides a framework for creating both basic unidirectional links and more complex linking structures. It allows XML documents to:

- Assert linking relationships among more than two resources
- Associate metadata with a link
- Express links that reside in a location separate from the linked resources



¹⁸ XML Linking Language (XLink) Version 1.0. W3C recommendation, 27 June 2001. <http://www.w3.org/TR/xlink>

The Family of XML Standards

XSL Formatting Objects (XSL-FO)

XSL Formatting Objects: an XML vocabulary for specifying formatting semantics.

- Formatting is the process of turning the result of an XSL Transformation into a tangible form for the reader or listener.

XSL-FO is a pagination markup language describing a rendering vocabulary capturing the semantics of formatting information for paginated presentation. Essentially, it describes how pages will look when presented to a reader.

The Family of XML Standards

XForms^{19 20}

XForms is comprised of separate sections that describe what the form does, and how the form looks. It separates *purpose* from the *presentation*.

In the XForms approach, forms are comprised of a section that describes what the form does, called the XForms Model, and another section that describes how the form is to be presented.

XForms uses XML for data transport and HTML for data display. With XForms, the data that are displayed in a form, and the data that are submitted from the form, are transported over the net using XML.

¹⁹ <http://www.w3.org/MarkUp/Forms/>

²⁰ XForms 1.0, W3C Working Draft, 18 January 2002. <http://www.w3.org/TR/2002/WD-xforms-20020118/>

The Family of XML Standards

SVG Scalable Vector Graphics ^{21 22}

Scalable Vector Graphics (SVG): a language for describing two-dimensional vector and mixed vector/raster graphics in XML.

SVG is a language for describing two-dimensional graphics in XML. SVG allows for three types of graphic objects:

- vector graphic shapes (e.g., paths consisting of straight lines and curves)
- images
- text

SVG sends instructions for drawing lines or curves (vectors), and filling these shapes

Some benefits of SVG:

- Can be printed with high quality at any resolution, without the “staircase” effects you see when printing bitmapped images (e.g., GIF or JPEG)
- Text in SVG is selectable and searchable. For example, you can search for specific text strings, like city names on a map.
- Non-proprietary

²¹ <http://www.w3.org/Graphics/SVG/Overview.htm#>

²² Scalable Vector Graphics (SVG) 1.0 Specification. W3C Recommendation, 04 September 2001. <http://www.w3.org/TR/SVG/>

The Family of XML Standards

XQuery Language ²³

The XQuery language is designed to be a small, easily implementable language in which queries are concise and easily understood. It is also flexible enough to query a broad spectrum of XML information sources, including both databases and documents.

XML querying is relevant for:

- human-readable documents: to retrieve individual documents, to perform context-sensitive searching,
- data-oriented documents: to query XML representations of databases,
- mixed-model documents: to perform queries on documents with embedded data

In short, XML querying is relevant for information retrieval.



²³ XQuery 1.0: An XML Query Language. W3C Working Draft, 30 April 2002. <http://www.w3.org/TR/xquery/>

The Family of XML Standards

Key messages

- This unit introduced you to some of the standards in the XML family. Other recommendations are constantly emerging.
- These trends extend XML to perform like other applications.

UNIT VI: USING XML

Using XML

This unit includes:

- XML initiatives in a variety of communities ^{24 25}
 - o Legal XML
 - o MathML
 - o Theological ML
 - o AgXML
 - o Real estate/mortgage XML
 - o Wireless ML
 - o Minimal ML

²⁴ XML Cover Pages. <http://www.oasis-open.org/cover/sgml-xml.html>

²⁵ XML.gov. <http://xml.gov/>

Using XML

XML initiatives in a variety of communities

Legal XML ²⁶

- Legal and technical experts
- Organization for the Advancement of Structured Information Standards (OASIS) framework
- National consensus
- Court filings, court documents, legal citations, transcripts
- Criminal Justice Information Systems



²⁶ <http://www.legalxml.org/>

Using XML

XML initiatives in a variety of communities

MathML ²⁷ ²⁸

- W3C working group
- Expressing math equations in web pages
- Graphic representation and calculation



²⁷ <http://www.w3.org/Math/>

²⁸ Mathematical Markup Language (MathML™) Version 2.0. W3C recommendation 21 February 2001. <http://www.w3.org/TR/MathML2/>

Using XML

Non-archival XML initiatives continued

Theological Markup Language (ThML) ²⁹

- Christian Classics Ethereal Library
- Subject and Scripture references
- Intelligent searching
- XSLT: Automatic conversion, variable formats

²⁹ <http://www.ccel.org/ThML/>

Using XML

Non-archival XML initiatives continued

AgXML ³⁰

- Grain and oilseed processing
- International markets
- Improving business functions
- Data models
- Promoting adoption of schemas



³⁰ <http://www.agxml.org/>

Using XML

Non-archival XML initiatives continued

Real Estate/Mortgage XML

- Real estate recording
- Property transactions, mortgages, secondary mortgage market
- MISMO (Mortgage Industry Standards Maintenance Organization)
- PRIJTF (Property Records Industry Joint Task Force)



Using XML

Non-archival XML initiatives continued

Wireless Markup Language (WML)³¹

- Same content, many devices
- Different applications
- Broadband and narrowband



³¹ <http://www.oasis-open.org/cover/wap-wml.html>

Using XML

Non-archival XML initiatives continued

Minimal XML (MinML) ³²

- Subset of XML
- Only essential features for data sharing
- Faster parsers
- Simpler information models
- Easier to learn

³² <http://www.docuverse.com/smldev/minxmlspec.html>

Using XML

Key Messages

- XML has a rich potential
- XML's potential is realized through the work of specific communities
- It takes a lot of work
- And there are many communities

UNIT VII: ENCODED ARCHIVAL DESCRIPTION

Encoded Archival Description

This unit includes:

- What is EAD?
- What does EAD do?
- What are the practical components of EAD?
- What makes EAD work?

Encoded Archival Description

What is EAD?

Encoded Archival Description (EAD)^{33 34}

The EAD Document Type Definition (DTD) is a standard for encoding archival finding aids. EAD is a set of rules for designating the intellectual and physical parts of archival finding aids so that the information contained therein may be searched, retrieved, displayed, and exchanged in a predictable platform-independent manner.

Project Timeline

- Berkeley Finding Aid Project 1993
- Bentley Fellowship 1995
- Prototype DTD (SGML) 1996
- XML compliant DTD 1998
- Society of American Archivists (SAA), Library of Congress (LoC) support ongoing

³³ <http://www.loc.gov/ead/>

³⁴ Metadata Made Simpler. http://www.niso.org/news/Metadata_simpler.pdf

Encoded Archival Description

What does EAD do?

EAD provides the:

1. Ability to present extensive and interrelated descriptive information found in archival finding aids
2. Ability to preserve the hierarchical relationships existing between levels of description
3. Ability to represent descriptive information that is inherited by one hierarchical level from another
4. Ability to move within a hierarchical informational structure and
5. Support for element-specific indexing and retrieval.

Encoded Archival Description

What are the practical components of EAD?

Standard

DTD

Mechanism for standard refinement and maintenance

EAD Tag Library ³⁵

EAD Cookbook ³⁶

EAD workshop ³⁷

XSL workshop

XML editor

Access mechanism - e.g., printed finding aids, Web delivery, portals

³⁵ EAD Tag Library for Version 1.0. Encoded Archival Description (EAD) Document Type Definition (DTD), Version 1.0 Technical Document No. 2. June 1998. <http://lcweb.loc.gov/ead/tglib/tlhome.html>

³⁶ <http://www.iath.virginia.edu/ead/cookbookhelp.html>

³⁷ Society of American Archivists, Upcoming Professional Education Offerings. http://www.archivists.org/prof-education/seasonal_schedule.asp

Encoded Archival Description

What makes EAD work?

Business purpose

Community of interests

Constant promotion and education

Encoded Archival Description

Key Messages

- EAD is a practical tool for the standardized creation and presentation of finding aids.
- Archivists use EAD because it fulfills a real business need.
- But using EAD is demanding. Creating and supporting the EAD standard was even more demanding.

UNIT VIII: PUTTING IT ALL TOGETHER

Putting It All Together

This unit includes:

- Learning to speak the same language
- Business needs: processes and functions
- Identifying and forming communities
- Developing the application
- Case study: Minnesota's Electronic Real Estate Recording Task Force
- Implementing your own solution

Putting It All Together

Learning to speak the same language

A common language needs a:

- Vocabulary
- Dictionary
- Grammar
- And an educational system

A successful XML project needs a:

- Compelling business need
- Collaborative community
- Practical application
- And a very large up-front investment in people, time, money and knowledge

Putting It All Together

Business needs

Data sharing

Infrastructure independent applications

Web-based transactions

Improved business processes

Legal mandates

Preservation

The **first** concern is having a real application or business need that XML may help fulfill. The second step is developing the appropriate XML language.

Introduction to XML for Decision-Makers

Legal mandates

E-Government Act of 2002 ³⁸

“4) enterprise architecture

(A) means

- (i) a strategic information asset base, which defines the mission;
- (ii) the information necessary to perform the mission;
- (iii) the technologies necessary to perform the mission;
- (iv) the transitional processes for implementing new technologies in response to changing mission needs”

“(6) interoperability means the ability of different operating and software systems, applications, and services to communicate and exchange data in an accurate, effective, and consistent manner;”

“(7) integrated service delivery means the provision of Internet-based Federal Government information or services integrated according to function or topic rather than separated according to the boundaries of agency jurisdiction”

Electronic Signatures in Global and National Commerce Act (E-Sign) ³⁹

“A Federal regulatory agency shall not adopt any regulation, order, or guidance described in paragraph, and a State regulatory agency is preempted by section 101 from adopting any regulation, order, or guidance described in paragraph, unless--

- (iii) the methods selected to carry out that purpose do not require, or accord greater legal status or effect to, the implementation or application of a specific technology or technical specification for performing the functions of creating, storing, generating, receiving, communicating, or authenticating electronic records or electronic signatures.”

³⁸ H.R. 2458, E-Government Act of 2002. <http://thomas.loc.gov/cgi-bin/query/D?c107:2:./temp/~c107yKQjZ9:>

³⁹ Electronic Signatures in Global and National Commerce Act (E-Sign). <http://thomas.loc.gov/cgi-bin/query/z?c106:S.761:>

Putting It All Together

Project Plan

Collaborative Community

Agreed upon business need

Project Sponsor

Education

Practical Application

Design

- Data modeling
- Process modeling

Business process re-engineering

Test

Assessment

Delivery

Refinement

Resources

People

Time

Money

Knowledge

Putting It All Together

Case study: Minnesota Electronic Real Estate Recording Task Force ⁴⁰

Task force formed 2000

Project to end 2004

Funded by filing fee surcharge

Private-public partnership

Entirely voluntary

⁴⁰ <http://www.commissions.leg.state.mn.us/lcc/erertf.htm>

Putting It All Together

Case study: Minnesota Electronic Real Estate Recording Task Force

What do we mean by recording and electronic recording?

Recording:

- Recording is the act of entering deeds, mortgages, easements, and other written instruments that affect title to real property into the public record.
- The purpose of recording is to give notice, to anyone who is interested, of the various interests that parties hold in a particular tract of land. Recording determines the legal priority of instruments that affect title to a particular tract of land.

Electronic recording:

- A publicly owned and managed county system, defined by statewide standards, that does not require paper or “wet” signatures, and under which real estate documents may be electronically:
 - o Created, executed, and authenticated;
 - o Delivered to and recorded with, as well as indexed, archived, and retrieved by, county recorders and registrars of title; and
 - o Retrieved by anyone from both on- and off-site locations.

Putting It All Together

Case study: Minnesota Electronic Real Estate Recording Task Force

What are the business needs?

- Huge and increasing volume of filings
- Highly inefficient paper workflow between automated activities
- Secondary mortgage market demand for digital records
- Increasing complexity of property rights and descriptions
- Legislative mandate to develop common technical and information architectures

Putting It All Together

Case study: Minnesota Electronic Real Estate Recording Task Force

What are the options?

- Level 1: images and minimal metadata (scan paper documents and send them in an email message that is manually processed)
- Level 2: images, metadata, digital or digitized signature (scan documents with some substantive metadata that can be automatically processed)
- Level 3: so-called “smart” documents in XML format, following recognized standards (create documents in an XML format which can be entirely processed automatically, populating tract and grantor/grantee indexes, calculating taxes and fees, validating legal descriptions, forwarding approvals and authorizations to different offices etc.)

Putting It All Together

Case study: Minnesota Electronic Real Estate Recording Task Force

What are the problems?

- Ordinary challenges to re-engineering
- Extraordinary political challenges to re-engineering
- Connectivity to existing systems
- Resources
- Setting standards
- Role of the archivist

Putting It All Together

Case study: Minnesota Electronic Real Estate Recording Task Force

How are we using XML?

- Business and workflow analysis
- Data and process models
- National standards
- Schemas ⁴¹
- XSLT

⁴¹ <http://www.commissions.leg.state.mn.us/lcc/erertfdrafts/>

Putting It All Together

What about your situation?

You will have to make choices.

The first choice is XML right for me?

Variables for the Electronic Real Estate Recording Task Force:

- Resources: consultants with business analysis and XML expertise, a special surcharge on filing to generate sufficient funds, volunteers with subject matter expertise.
- Tools: business analysis, communication, negotiation, web sites, facilitated meetings, pilots, subcommittees.
- Standards/methodologies: XML, XML Schemas, national standard with local extensions.
- Education: not enough of it, no common basis of understanding of the choices.
- Technology: web-based transactions, digital signatures, imaging, XML as middleware between legacy applications.
- Partners: banks, realtors, title companies, Fannie Mae, an archivist, county officials, state agencies, legislators ...



Putting It All Together

Key Messages

- In order to speak the same language, we need to have a common vocabulary, dictionary, and grammar. We will also need an educational system to establish, understand, and learn these common factors.
- A successful XML project needs a:
 - o Compelling business need
 - o Collaborative community
 - o Practical application
 - o And a very large investment in people, time, money and knowledge
- In Minnesota, the Minnesota Electronic Real Estate Recording Task Force is currently evaluating XML as a solution to automating real estate recording processes.
- In order to implement your own XML project, you will need to consider the following variables:
 - o Resources
 - o Tools
 - o Standards/Methodologies
 - o Education
 - o Technology
 - o Partners

Putting It All Together

Course summary and conclusions

- A review of XML and its applications
 - o Digital objects are a fact of life and we need a tool to manage them.
 - o Extensible Markup Language is a means of marking up data, using a specific syntax that you define.
 - o XML is made up of a declaration, elements, and attributes, and is verified against a DTD or XML Schema.
 - o XSL, XSLT, and XHTML are extensions of XML used for presentation of documents.
 - o Free XML tools and editors are available to assist you in your work.
 - o There is an evolving family of XML-related standards which extend XML's functionality.
 - o XML's potential is realized through the work of specific communities.
 - o EAD is a practical tool for the standardized creation and presentation of finding aids using XML.
 - o You must make choices in order to implement XML within your situation.

APPENDIX A

Appendix A

List of Acronyms

ARMA	Association of Records Managers and Administrators
ASI	Advanced Strategies, Inc.
CDATA	Character Data
CSS	Cascading Stylesheets
DTD	Document Type Definition
EAD	Encoded Archival Description
E-SIGN	Electronic Signatures in Global and National Commerce Act
GIF	Graphics Interchange Format
HTML	Hypertext Markup Language
ISO	International Organization for Standardization
IT	Information Technology
JPEG	Joint Photographic Experts Group
LoC	Library of Congress
MAC	Midwest Archives Conference

Appendix A

List of Acronyms (continued)

MHS	Minnesota Historical Society
MinML	Minimal XML
MISMO	Mortgage Industry Standards Maintenance Organization
NAGARA	National Association of Government Archives and Records Administrators
NHPRC	National Historical Publications and Records Commission
NISO	National Information Standards Organization
OASIS	Organization for the Advancement of Structured Information Standards
PCDATA	Parsed Character Data
PDA	Personal Digital Assistant
PDF	Portable Document Format
PRIJTF	Property Records Industry Joint Task Force
SAA	Society of American Archivists
SGML	Standard Generalized Markup Language
SVG	Scalable Vector Graphics

Appendix A

List of Acronyms (continued)

ThML	Theological Markup Language
TIS	Trustworthy Information Systems
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
WML	Wireless Markup Language
XHTML	eXtensible Hypertext Markup Language
XLINK	XML Linking Language
XML	eXtensible Markup Language
XPATH	XML Path Language
XPOINTER	XML Pointer Language
XSL	eXtensible Stylesheet Language
XSL-FO	XSL Formatting Objects
XSLT	XSL Transformations

APPENDIX B

Appendix B

XML tools and editors that are not free

Editors

- XMetal – SoftQuad, Inc.
http://www.softquad.com/top_frame.sq
- Xeena
<http://www.alphaWorks.ibm.com/tech/xeena>
- XMLSpy
<http://www.xmlspy.com>
- ArborText – Epic
<http://www.arbortext.com/>

Parsers

- Xerces
<http://www.alphaworks.ibm.com/tech/xml4j>
- Expat
<http://expat.sourceforge.net>

Appendix B

XML tools and editors that are not free (continued)

Other tools

- TIBCO Extensibility
<http://www.tibco.com/products/extensibility/>
- XML Notepad – MSDN
<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnxml/html/xmlpaddownload.asp>

Tool Providers

- Inso Corporation
<http://www.inso.com/>
- Macromedia
<http://www.macromedia.com/>
- Microsoft
<http://www.microsoft.com/ms.htm>

APPENDIX C

Appendix C

Bibliography

Workshop web site

- Educating Archivists and their Constituencies
<http://www.mnhs.org/preserve/records/edarchivists.html>

Electronic Records

- Minnesota Historical Society, State Archives Department. *Electronic Records Management Guidelines*. 2002.
<http://www.mnhs.org/preserve/records/electronicrecords/erguidelines.html>
- DLM Forum. *Guidelines on Best Practices for Using Electronic Information*. 1997.
<http://europa.eu.int/ISPO/dlm/documents/gdlines.pdf>

Dollar, Charles M. *Authentic Electronic Records: Strategies for Long-Term Access*. Chicago: Cohasset Associates, Inc. 2000.

Saffady, W. *Managing Electronic Records*. 2nd ed. Prairie Village, Kan.: ARMA International, 1998.

- Electronic Signatures in Global and National Commerce Act (E-Sign)
[http://thomas.loc.gov/cgi-bin/query/z?c106:S.761:](http://thomas.loc.gov/cgi-bin/query/z?c106:S.761)
- Trustworthy Information Systems Handbook
<http://www.mnhs.org/preserve/records/tis/tis.html>
- Minnesota Technical Enterprise Architecture, Revision 1.0. 2001.
<http://www.ot.state.mn.us/architecture>

Appendix C

Bibliography (continued)

XML

- Why XML? by Simon St. Laurent. November 1999.
http://www.webdevelopersjournal.com/articles/why_xml.html
- XML in Ten Points from the W3C
<http://www.w3.org/XML/1999/XML-in-10-points>
- World Wide Web Consortium (W3C)
<http://www.w3.org/>
- W3C Extensible Markup Language (XML)
<http://www.w3.org/XML/>
- Extensible Markup Language (XML) 1.0 (Second Edition). W3C Recommendation 6 October 2000
<http://www.w3.org/TR/REC-xml>

DTDs and XML Schemas

- XML Schema
<http://www.w3.org/XML/Schema>

Samples of DTDs and XML Schemas

- http://www.xml.org/xml/registry_searchresults.jsp

Appendix C

Bibliography (continued)

Namespaces

- Namespaces in XML. World Wide Web Consortium, 14 January 1999. A W3C recommendation.
<http://www.w3.org/TR/REC-xml-names/>

Browsers compatible with XML

- Internet Explorer 5.0 (and higher)
<http://www.microsoft.com/windows/ie/default.asp>
- Mozilla 1.0
<http://www.mozilla.org/>
- Opera 5 (and higher)
<http://www.opera.com/>
- Netscape 7.0
<http://www.netscape.com/>

eXtensible Stylesheet Language (XSL)

- The Extensible Stylesheet Language
<http://www.w3.org/Style/XSL/>
- Extensible Stylesheet Language (XSL), Version 1.0. W3C Recommendation 15 October 2001.
<http://www.w3.org/TR/xsl/>

Appendix C

Bibliography (continued)

XSL Transformations (XSLT)

- XSL Transformations (XSLT), version 1.0. W3C Recommendation 16 November 1999.
<http://www.w3.org/TR/xslt>

XHTML

- XHTML™ 1.0: The Extensible HyperText Markup Language. A reformulation of HTML 4 in XML 1.0. W3C recommendation, 26 January 2000.
<http://www.w3.org/TR/xhtml1/>

XML Tools and Editors (see also Appendix B)

- The Apache XML Project
<http://xml.apache.org>
- netbeans.org
<http://www.netbeans.org/nonav/index2.html>

XML Editor:

- XML Cooktop
<http://www.xmlcooktop.com/>

Appendix C

Bibliography (continued)

XML Tools and Editors (see also Appendix B) (continued)

Online tools

Parser

- Userland Frontier
[http://frontier.userland.com/stories/storyReader\\$1092](http://frontier.userland.com/stories/storyReader$1092)

Well-formedness and validator

- <http://www.cogsci.ed.ac.uk/~richard/xml-check.html>>

URLs with access to many tools

- XML.com
<http://www.xml.com/buyersguide/>
- XML Cover Pages
<http://www.oasis-open.org/cover/publicSW.html#xmlTools>
- <http://www.xmlsoftware.com>
- Information by Lars M. Garshol
<http://www.garshol.priv.no/download/>
- sourceforge.net
<http://sourceforge.net/>

Appendix C

Bibliography (continued)

Future Trends

- XML Path Language (XPath) Version 1.0. W3C recommendation, 16 November 1999.
<http://www.w3.org/TR/xpath>
- XML Linking Language (XLink) Version 1.0. W3C recommendation, 27 June 2001.
<http://www.w3.org/TR/xlink>
- XML Pointer Language (XPointer) Version 1.0. W3C recommendation 27 June 2001.
<http://www.w3.org/TR/xptr>
- XForms
<http://www.w3.org/MarkUp/Forms/>
- XForms 1.0, W3C Working Draft, 18 January 2002.
<http://www.w3.org/TR/2002/WD-xforms-20020118/>
- Scalable Vector Graphics (SVG)
<http://www.w3.org/Graphics/SVG/Overview.htm8>
- Scalable Vector Graphics (SVG) 1.0 Specification. W3C Recommendation, 04 September 2001.
<http://www.w3.org/TR/SVG/>
- XQuery 1.0: An XML Query Language. W3C Working Draft, 30 April 2002.
<http://www.w3.org/TR/xquery/>

Appendix C

Bibliography (continued)

Non-archival XML initiatives

- XML Cover Pages
<http://www.oasis-open.org/cover/sgml-xml.html>
- XML.gov
<http://xml.gov/>
- LegalXML
<http://www.legalxml.org/>
- MathML
<http://www.w3.org/Math/>
- Mathematical Markup Language (MathML™) Version 2.0. W3C recommendation 21 February 2001.
<http://www.w3.org/TR/MathML2/>
- Theological Markup Language (ThML)
<http://www.ccel.org/ThML/>
- AgXML
<http://www.agxml.org/>
- Wireless Markup Language (WML)
<http://www.oasis-open.org/cover/wap-wml.html>
- Minimal XML (MinML)
<http://www.docuverse.com/smldev/minxmlspec.html>

Appendix C

Bibliography (continued)

Encoded Archival Description

- Encoded Archival Description (EAD)
<http://www.loc.gov/ead/>
- Metadata Made Simpler.
http://www.niso.org/news/Metadata_simpler.pdf
- EAD Tag Library for Version 1.0. Encoded Archival Description (EAD) Document Type Definition (DTD), Version 1.0 Technical Document No. 2. June 1998.
<http://lcweb.loc.gov/ead/tglib/tlhome.html>
- EAD Cookbook
<http://www.iath.virginia.edu/ead/cookbookhelp.html>
- Society of American Archivists, Upcoming Professional Education Offerings.
http://www.archivists.org/prof-education/seasonal_schedule.asp

Minnesota Electronic Real Estate Recording Task Force

- <http://www.commissions.leg.state.mn.us/lcc/erertf.htm>
- XML Schemas
<http://www.commissions.leg.state.mn.us/lcc/erertfdrafts/>