Using XML to Share Statistics Education Materials in PAPIRIS

H. Edward Donley, Mathematics Department, Indiana University of Pennsylvania, Indiana, PA, 15701, USA.

Introduction to PAPIRIS

Principles and Practices: Internet Resources in Statistics (PAPIRIS) will be a Web-based statistics reference library for teachers, students, and lifelong learners, consisting of an encyclopedia of statistical terms, definitions, and examples; annotated links to external Web sites for statistics and statistics education; a collection of lessons; and critiques of those lessons, for students aged 6 through 18 years old. PAPIRIS will be seeded with approximately 200 lessons developed by teachers who participated in the Statistics Education through Quantitative Literacy (SEQuaL) workshops since 1992 throughout the state of Pennsylvania in the United States.

This paper focuses on the technical aspects of PAPIRIS' implementation, in particular, its use of Extensible Markup Language (XML). Another paper in this conference, "Simple Linear Regression: A PAPIRIS Example", by Maher Shawer, John Baker, Michael Bossé, and Fred Morgan, more fully describes PAPIRIS' educational uses.

Introduction to XML

Extensible Markup Language (XML) is a set of rules for organizing information [1]. XML documents are easy for computer programs to read and process, so XML provides a convenient mechanism for computers to share data. XML documents look like HTML, because both specifications are derived from the Standard General Markup Language (SGML) specification, developed in the early 1980s. XML documents use tags surrounded by angle brackets, < and >, and attributes of the form name="value". Whereas, HTML has a fixed set of tag names, such as <body>, <pre>, and , the authors of XML documents can create their own tag names. XML tag names can describe the contents of the elements. For example, the abstract for a research article can be enclosed in <abstract> and </abstract> tags, rather than using HTML's and tags. Consequently, a computer program that reads the XML document containing the research article can extract the abstract. Research articles might also contain <title>, <author>, <body>, and <reference> tags. Computer programs can easily construct tables of contents, author indices, and bibliographies for collections of research articles. If the computer programs are running on Web servers, then the Web server can dynamically create up-to-date indices for visitors to the Web site. These collections of articles can even reside on several servers, provided that all authors agree to use the same tag names. This agreement would be specified in an XML Schema document [2, 3], which describes the required structure of an XML document. An XML Schema for research articles might state that every article must have one title, one or more authors, one abstract, one body, and zero or more references. Many industries and professional organizations have published XML Schemas for sharing industry-specific information. For example, MathML is a specification for mathematical expressions [4].

Extensible Stylesheet Language (XSL) is a related standard for specifying how an XML document is to be transformed into another format. Most often, XSL is used to transform XML documents into HTML documents. For the research article example, an XSL stylesheet could state that article titles should be displayed in HTML <h1> header tags, authors should be in italics, abstracts should be in paragraphs and italics, and references should be displayed as ordered lists.

Another related standard, XML Query [5], which is still in Working Draft stage, will enable Web site creators to extract portions of XML documents using URLs. No computer programming will be necessary. In contrast, if information were stored in a database rather than in XML, then a computer program would have to retrieve the information from the database.

An ideal strategy for this research article publication system would be to store the articles in XML format, so that computer programs could easily extract portions of the articles — possibly from different servers — and an XSL stylesheet so that visitors to the Web site can view the articles in HTML format in their Web browsers.

A large number of XML tools are available [6], ranging from XML editors to software development tools for almost every popular computer programming language.

XML Specifications for Educational Resources

Several XML specifications have been published for education resources. The Dublin Core [7] is a widely accepted metadata standard supporting a broad range of applications. The schema is quite generic, including such elements as Description, Creator, and Date. The Dublin Core is often used by libraries to describe their diverse resources in card catalogs. The U.S. National Science Foundation has adopted the Dublin Core for its National Science Digital Library.

The IMS Learning Resource Meta-data Information Model [8] grew from an IEEE working group and builds on the Dublin Core. Since it is specifically for learning resources, it is more detailed than the Dublin Core. The schema includes language, learner's age range, type of learning resource (including Exercise, Simulation, Questionnaire, Narrative Text, and Problem Statement), prerequisites, typical learning time, copyright restrictions, and numerous other education-related elements. The IMS Project [9] publishes other educational technical standards, such as interoperability standards for sharing quizzes and tests.

Sharable Courseware Object Reference Model (SCORM) [10] integrates the IMS model with other related specifications to allow Web-based learning management systems (LMS), such as Blackboard and WebCT, to import, share, combine, reuse and export learning modules and track students' progress. The goal is to allow a learner to enroll in a course in one LMS, use learning modules from various LMS's and the course's LMS should track the learner's progress, including time spent on task and scores on exams.

XML and PAPIRIS

PAPIRIS will store in XML format its taxonomy of about 150 statistical terms, its term definitions, its collection of links to external resources, and information about its examples and lessons.

PAPIRIS will use XML for several reasons.

- Other Web servers can drill down into PAPIRIS' taxonomy of statistics terms and retrieve lessons on specific topics. For example, textbook publishers can correlate PAPIRIS lessons with their statistics textbooks.
- Other Web servers can create indices of age-appropriate PAPIRIS encyclopedia terms and lessons. For example, an educational Web site for students aged 6 to 12 can link to all of PAPIRIS' elementary content.
- PAPIRIS' partner Web sites will use XSL to present PAPIRIS content using their own Web site's look and feel. So PAPIRIS' content will appear to be the partner's content.
- PAPIRIS lesson authors can create lists of their lessons.

The IMS Model is the appropriate XML specification for PAPIRIS. It includes education-specific elements, but does not include the additional complexity of SCORM. Since PAPIRIS is not using learning management systems, it does not need to use SCORM. A sample IMS document for a simple linear regression lesson is listed in Appendix 1. This is a fictitious example, since PAPIRIS is still in the design phase.

XML and the Supercourse

The Supercourse would benefit from storing learning resource metadata in XML format. XML can be the glue that combines the various Supercourse projects. XML can:

- aggregate Supercourse learning modules across all of the national Supercourse projects,
- extract all Supercourse learning modules written in a particular language, and
- create a subject index combining all Supercourse projects.

The IMS Model is the appropriate XML specification for the Supercourse or the same reasons that IMS is appropriate for PAPIRIS. It includes education-specific elements, but does not include the additional complexity of SCORM. Since the Supercourse is not using learning management systems, it does not need to use SCORM.

References

1. XML in 10 Points, World Wide Web Consortium, http://www.w3.org/XML/1999/XML-in-10-points, 1999.
2. XML Schema Part 1: Structures, World Wide Web Consortium, http://www.w3.org/TR/xmlschema-1/, May 2001.
3. XML Schema Part 2: Data Types, World Wide Web Consortium, http://www.w3.org/TR/xmlschema-2/, May 2001.
4. W3C Math Home, World Wide Web Consortium, http://www.w3.org/Math/, February 2001.
5. XML Query, World Wide Web Consortium, http://www.w3.org/XML/Query, February 2001.
6. Dr. Tom's Guide to XML, http://www.imsglobal.org/drtomxml.html#Tools.

7. Dublin Core Metadata Initiative, http://dublincore.org/.
8. IMS Learning Resource Meta-data Information Model, IMS Global Learning Consortium, Inc., http://www.imsproject.org/metadata/imsmdv1p2p1/imsmd_infov1p2p1.html, September 2001.
9. IMS Global Learning Consortium, Inc., http://www.imsproject.org/.
10. Sharable Content Object Reference Model, Version 1.2, Advanced Distributed Learning, http://www.adlnet.org/index.cfm?fuseaction=scormabt, May 2002.

Appendix 1: Sample IMS Model for Simple Linear Regression Lesson

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE lom SYSTEM "imsmd_v1p2p2.dtd">
<lom xmlns="http://www.imsglobal.org/xsd/imsmd_v1p2"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.imsglobal.org/xsd/imsmd_v1p2 imsmd_v1p2p2.xsd">
    <general>
        <identifier>2005-PAPIRIS</identifier>
        <language>en-US</language>
        <title>
            <langstring xml:lang="en">Simple Linear Regression</langstring>
            <langstring xml:lang="de">Einfache Lineare Regression</langstring>
        </title>
        <catalogentry>
            <catalog>URI</catalog>
            <entry>
                <langstring xml:lang="en">
http://www.math.iup.edu/papiris/regression/simple.html</langstring>
            </entry>
        </catalogentry>
        <catalogentry>
            <catalog>PAPIRIS</catalog>
            <entry>
                <langstring xml:lang="en">2005-PAPIRIS#</langstring>
            </entry>
        </catalogentry>
        <description>
            <langstring xml:lang="en">This lesson introduces students to simple linear regression
through exploration of examples.</langstring>
        </description>
        <structure>Collection</structure>  <!-- choose from Collection, Mixed, Linear, Heirarchical,
Networked, Branched, Parceled, Atomic -->
        <aggregationlevel>2</aggregationlevel>
        <contribute>
            <role>
                <source>
                    <langstring xml:lang="x-none">LOMv1.0</langstring>
                </source>
                <value>
                    <langstring xml:lang="x-none">Author</langstring>
                </value>
            </role>
            <centity>
                <vcard>
BEGIN:vCard
Shawer;Maher
END:vCard
```

```
</vcard>
            </centity>
        </contribute>
        <contribute>
            <role>
                <source>
                    <langstring xml:lang="x-none">LOMv1.0</langstring>
                </source>
                <value>
                    <langstring xml:lang="x-none">Publisher</langstring>
                </value>
            </role>
            <centity>
                <vcard>
BEGIN: vCard
ORG: Indiana University of Pennsylvania
ADR: Mathematics Department;210 South Tenth Street;Indiana;Pennsylvania;15701;USA
END: vCard
</vcard>
            </centity>
            <date>
                <datetime>2002-06-30</datetime>
            </date>
        </contribute>
    </general>
    <technical>
        <format>text/html</format>
        <location type="URI">
http://www.math.iup.edu/papiris/regression/simple.html
</location>
        <otherplatformrequirements>
            <langstring xml:lang="en">Statistical software package or graphing calculator</langstring>
        </otherplatformrequirements>
    </technical>
    <educational>
        <intendedenduserrole>
            <source>
                <langstring xml:lang="x-none">LOMv1.0</langstring>
            </source>
            <value>
                <langstring xml:lang="x-none">Learner</langstring>
            </value>
        </intendedenduserrole>
        <context>
            <source>
                <langstring xml:lang="x-none">LOMv1.0</langstring>
            </source>
            <value>
                <langstring xml:lang="x-none">Secondary Education</langstring>
            </value>
        </context>
        <typicalagerange>
```

<langstring xml:lang="en">14-18</langstring>
</typicalagerange>
<interactivitytype>Active</interactivitytype>
<difficulty>medium</difficulty>
<typicallearningtime>02:00</typicallearningtime>
<cost>Active</interactivitytype>
</educational>
<rights>
<cost>
<source>
<langstring xml:lang="x-none">LOMv1.0</langstring>
</source>
<value>
<langstring xml:lang="x-none">no</langstring>
</value>
</cost>
<copyrightandotherrestrictions>
<source>
<langstring xml:lang="x-none">LOMv1.0</langstring>
</source>
<value>
<langstring xml:lang="x-none">yes</langstring>
</value>
</copyrightandotherrestrictions>
<description>
<langstring xml:lang="en">Cannot redistribute without author's permission. Contact PAPIRIS, someone@somewhere.edu</langstring>
</description>
</rights>
<classification>
<taxonpath>
<source>
<langstring xml:lang="en">PAPIRIS</langstring>
</source>
<taxon>
<entry>
<langstring xml:lang="en">Regression</langstring>
</entry>
<taxon>
<entry>
<langstring xml:lang="en">Simple Linear</langstring>
</entry>
</taxon>
</taxon>
</taxonpath>
</classification>
</lom>