


SGML→XML →HTML for Displaying EAD Finding Aids



Susan Hamburger, Ph.D.

Penn State University

Forum on Digital Library Practices

July 17, 1999

Washington, D.C.



Delivery: Interleaf's Panorama

■ PRO

- Searchable
- Customizable display

■ CON

- Proprietary software
- No longer offers downloadable free viewer for remote patrons
- Not networkable for distributed campus system



Delivery: HTML

■ PRO

- Anyone can view
- Easy to create
- No special software needed

■ CON

- Lose navigation
- Lose sophisticated searchability
- Displays linear not hierarchical



Delivery: XML

■ PRO

- Cutting-edge technology
- Compatible with EAD dtd

■ CON

- Most browsers, except Internet Explorer 5.0, cannot read it



Solution: compromise

- Create finding aid in SGML
- Map SGML to XML
- Export XML files to HTML
 - PRO
 - Mimics Panorama display with frames and navigation
 - CON
 - Lose customized font color
 - When you choose one, all subheads in navigator frame display



Solution: alpha test conversion program

- Daniel Pitti wrote a script to convert SGML files into XML
 - SGML files must contain attributes XML can read:

```
<HEAD ID="dsc">Container List</HEAD>
<C01 LEVEL="series">
<HEAD ID="s1">Photographs Series</HEAD>
<DID><UNITTITLE>Photographs, <UNITDATE>1891-
1915.</UNITDATE></UNITTITLE></DID>
```
 - XML files contain java scripts



XML to HTML

■ XML to HTML script

– Creates five files to display one SGML finding aid

- [filename].html = frames
- [filename1].html = navigation frame for Table of Contents
- [filename2].html = headers in body linked from Table of Contents
- [filename3].html = headers in body linked to text
- [filename4].html = text of finding aid



How it works at Penn State

- Create finding aid using Author/Editor
- Use EditPad for parsing
- Save finding aids being worked on in
 - working/findaids/inprocess/ae
 - working/findaids/inprocess (for SGML version)

How it works at Penn State

- When finding aid done, export to XML
 - Parse in NSGMLS: **pwork** [filename]
 - Save to XML: **sx** [filename]
 - In **XMLIN folder** open [filename].xml and **system.txt** in EditPad
 - Replace **<!DOCTYPE EAD SYSTEM...** with string in **system.txt**: **<!DOCTYPE EAD SYSTEM**
"../..../shared/ead/eadx.dtd" [
 - Change **entityref=psuseal** to **href="psuseal.gif"**
 - Remove extraneous character **Å** at copyright symbol



How it works at Penn State

- In **Parse folder** run **xp1**

- `<?xml version="1.0" encoding="ISO-8859-1"?>`
`<!DOCTYPE xsl:stylesheet>`
`<xsl:stylesheet`
`xmlns:xsl='http://www.w3.org/TR/WD-xsl'`
`smlns:html="http://www.w3.org/TR/REC-`
`html40/"`
`results-ns="html">`
`<xsl:define-constant name='name'`
`value='nassh'/>`
`</xsl:stylesheet>`
- Change **value** to current [filename]



How it works at Penn State

- In **HTMLOUT** folder open [filename4].html and do a **find and replace** `
` with `
`
- Save in **DONE** folder:
 - catalog, entityrc, psuseal.gif
 - [filename].sgm, [filename].html (all five files for each finding aid)
- Move all but SGML file to **FindingAids** folder on libraries' website

Making of America II Testbed

Inventory of the Pennsylvania Bridges Collection, 1884-1915

Descriptive Summary

+ Administrative Information

Scope and Content

+ Index Terms

+ Container List

Inventory of the Pennsylvania Bridges Collection,
1884-1915

Accession number: 1991-0012H

Historical Collections and Labor Archives
Special Collections Department



University Libraries
The Pennsylvania State University

Contact Information:

Pennsylvania State University
University Libraries
Special Collections Department
W339 Pattee Library
University Park, PA 16802
814/863-2505
FAX 814/863-5775
E-mail: dmc@psulias.psu.edu

©1998 Pennsylvania State University.
All rights reserved.



Expanded navigator frame

Inventory of the
Pennsylvania
Bridges Collection,
1884-1915

Descriptive Summary

- Administrative
Information

Access

Preferred Citation

Scope and Content

- Index Terms

Corporate Subjects

Topical Subjects

Photographers

Corporate Creators

Form/Genre Terms

- Container List

Photographs Series

Technical Drawings Series

Lantern Slides Series



Links to individual digital objects

Box: 3

[Item 5241]

Simple truss bridge, Pittsburgh, Pa., built 1903. Pittsburgh, Ft. Wayne and Chicago Railroad bridge crosses Allegheny River. Photo by Pittsburgh, Ft. Wayne and Chicago Railroad. `<DAO SHOW="EMBED" ACTUATE="USER"></DAO>`

Box: 1

[Item 5242]

Three-hinged arch bridge, Schenley Park, Pittsburgh, Pa., built 1898? Bridge crosses Panther Hollow. Photo by Steiner and Co.