

KoreaMed technical workshop, 2006

Time: August 11, Friday, 10:00 AM -2:00 PM

Place: Conference Room in the Second Floor, Sookmyung Women's University Library, Seoul, Korea

Subject: PubMed Central XML and XSLT Processor

Organizer: Sun Huh (Chair of the Committee on Information Management, KAMJE)

General Purpose: Attendants should be able to make PubMed Central (PMC) XML

Specific Goals:

- 1) Attendants should be able to explain the structure of PMC DTD
- 2) Attendants should be able to input the XML tags.
- 3) Attendants should be able to validate the PMC XML grammar.
- 4) Attendants should be able to transform the PMX XML to HTML file with XALAN-C program.
- 5) Attendants should be able to explain the process of PMC application.
- 6) Attendants should be able to find the list of medical journals from Korea eligible to be applied to PMC.

Schedule of workshop

1. See the example - 15 min
2. How to use drXML - 15 min
3. Table and reference is the most complicated ones -5 min
4. Exercise according to example with JKMS - 30 min
break -10 min
5. Each person makes one paper into PMC XML - 30 min
6. How to validate or check the validity of PMC XML files - 15 min
7. XML to html -XALAN-C -10 min (not necessary at now since JKMS has pdf files)
lunch - 1 hour

* Figure ID is not still solved. SQL db is required to match the ID and figure names. It is not necessary at now. It will be done by NLM PMC team

Plan after the workshop>

- 1) Let's try to make PMC XML files rapidly and send them to NLM to be checked for their validity.
- 2) Before Sep. 19, there should be PMC files of JKMS

How to use DrXML, automatic PMC XML tagging program?

1. Whole tag library was added.
2. How to input affiliation tag?
3. Precise name of journal-id is required.
4. Body section's meaning; Is it necessary to add sec in p?
5. Table in any paragraphs should be possible. Table label is table title. caption is title.
6. Reference is in the back matter.
7. Table and reference are the most important parts.
8. Format according to each journal is possible. Automatic inclusion of journal metadata is possible
9. Let's try and report the problems.
10. Ninety percent is already succeeded in DrXML

Files and sites for PMC XML

Archiving and Interchange: <http://dtd.nlm.nih.gov/2.2/>

Journal Publishing: <http://dtd.nlm.nih.gov/publishing/2.2/> Article

Authoring: <http://dtd.nlm.nih.gov/articleauthoring/2.2/>

W3C Schema Version pages: Archiving and Interchange:

<http://dtd.nlm.nih.gov/w3c-schema.html>

Journal Publishing: <http://dtd.nlm.nih.gov/publishing/w3c-schema.html>

New versions of the XML to HTML XSL converter and the corresponding CSS file have been released.

Both are now available through the FTP site:

ftp://ftp.ncbi.nih.gov/pub/archive_dtd/tools/.

As with the beta version, a copy of the CSS resides on the DTD site at

<http://dtd.nlm.nih.gov/tools/ViewNLM.css> for your use.

eXtyles -XML editor for ms word <http://inera.com>

Merops -XML editor for ms word http://www.shabash.net/merops_modules.htm

XMLprobe --XML-based content strategy & check XML, <http://www.XMLprobe.com/>

XML Validator:

<http://www.pubmedcentral.gov/utis/validate/XMLcheck.cgi> - This validator works with the Journal Publishing DTD.

PubMed Central Style Checker:

http://www.pubmedcentral.gov/utis/style_checker/stylechecker.cgi

PubMed Central XML Tagging Guidelines:

<http://dtd.NLM.nih.gov/publishing/coding/PMC/style.html> - The PMC tagging style is defined in this document.

The process is outlined in on the PMC web site here:

<http://www.pubmedcentral.gov/about/pubinfo.html>

Installation of XALAN for transformation of XML file to HTML file

1. Down load two programs.

XALAN-C downloads (For windows user) XALAN-C_1_10_0-win32-msvc_60.zip

Available from: <http://XML.apache.org/XALAN-C/>

Xerces-C downloads (For windows user)

Available from:

http://mirror.apache.or.kr/XML/xerces-c/binaries/xerces-c_2_7_0-windows_2000-msvc_60.zip)

2. Unzip the files.

3. Copy xerces-c_2_7.dll from XERECES/BIN folder to XALAN/BIN folder.

4. Save the XALAN-C_1-10_0-win32-msvc_60\BIN folder to environmental variables.

Method>

My computer -->Push right button of mouse -->High --> Environmental variables
-->System Variables

Add the new access path to the end of path. Add ' ; ' before access path.

For example>

;C:\Program Files\XALAN-C_1_10_0-win32-msvc_60\bin

5. Under the XALAN/bin directory, journalpublishing.dtd, dtd related files and NLMview-v2.xsl should be located together.

6. Transfer the XML file to html files at CMD window.

XALANTRANSFORM INPUT_filename.XML STYLE_filename.XSL
OUTPUT_filename.HTML

Example> XALANtransform 1.XML viewNLM-v2.xsl 1.html

7) In case that step 4 is not applied

Add XALAN files under the c:\XALAN-C\, also add xerces-c_2_7.dll, journalpublishing.dtd, dtd-related files, and NLMview-v2.xsl under the bin directory,

Type as follows:

```
c:\XALAN-C\bin> XALANtransform 1.XML viewNLM-v2.xsl 1.html
```

After that, XML file will be transformed to html file.

(End)