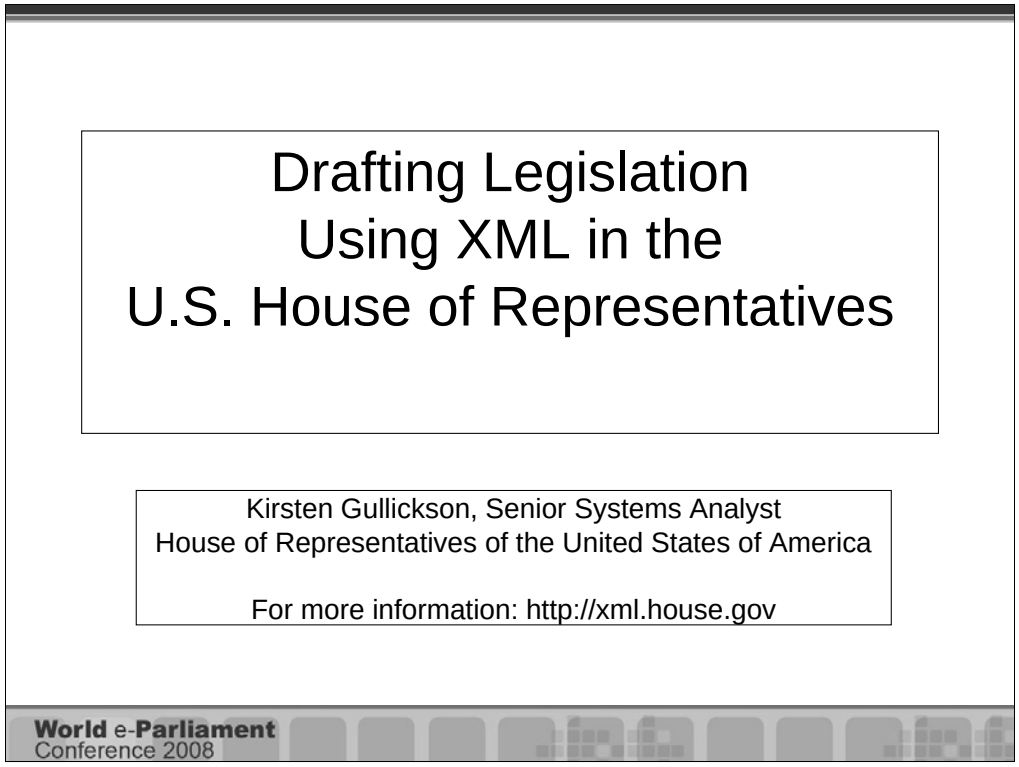

United Nations
EUROPEAN PARLIAMENT
Global Centre for ICT in Parliament
World e-Parliament
Conference 2008
25 - 26 November 2008 // European Parliament // Brussels



The United States Congress is engaged in an on-going effort to author and publish legislative documents in XML. The goals of this effort are to make improvements in the cost, accuracy, timeliness and efficiency of the legislative document creation and revision process, as well as establish common data standards for the exchange of legislative information with the Legislative Branch organizations. This presentation will provide you with some information on how this is being accomplished.

Following a feasibility study, the United States House and Senate extensively customized separate XML editors for drafting bills, resolutions and amendments. Organizations within the Legislative Branch exchange the XML documents using an agreed-upon Document Type Definition (DTD) for each document type. Our DTDs are schema-compliant.

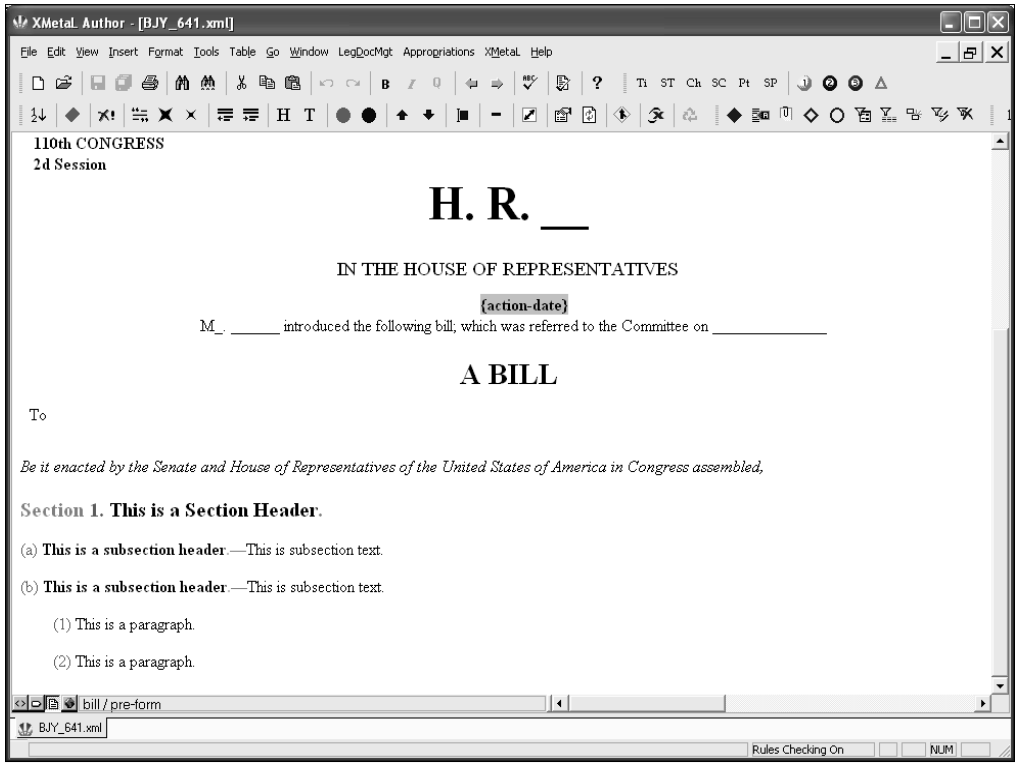
The production and exchange of legislative documents is a joint effort between the House, the Senate, the Library of Congress and the Government Printing Office.

The Government Printing Office prints and publishes the Congressional documents and the Library of Congress is responsible for posting the information for public and internal Legislative Branch consumers.

In addition, a Legislative Branch XML Technical Working Group was established to coordinate and manage the DTDs, shared tools, and the common tag library across the Legislative Branch of government. Collaboration on common standards and efforts to use some of the same tools is on-going. In the Senate and House XML editors, the same table tool component, the same composition system and the same translations are used. And in 2009, the same value-added script will be run on the XML files that are posted on the web.

The establishing of a common tag library is key to managing the XML elements. While a number of a DTDs (or schemas) may be managed, the common tag library always ensures reuse of the elements across data and other data sources.

The Office of the Clerk in the U.S. House provides the XML editor to a number of offices for their use.



[Slide illustrates XMetaL in Tags-Off view]

What is our XML editor?

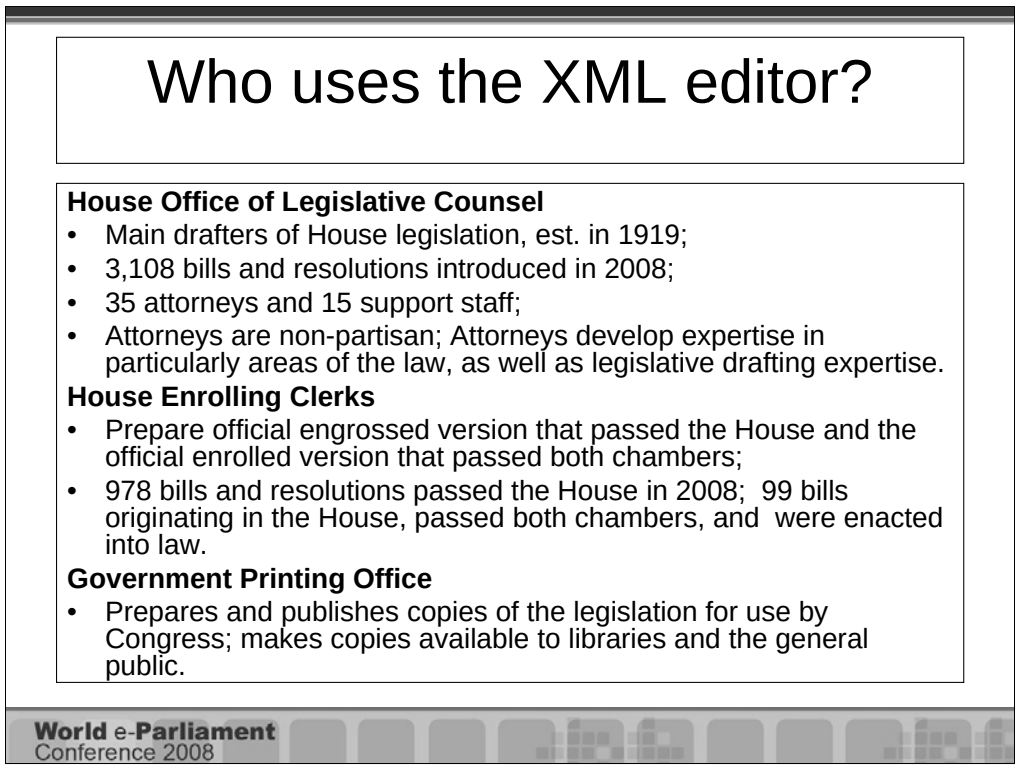
- The House purchased an off-the-shelf software product by JustSystems called XMetaL.

It was–

- Affordable;
- Used common development languages, and;
- Provided fully featured tags-off editing view.

The main customizations are written in–

- VBScript,
- Perlscript, and
- Visual Basic Dynamic Link Libraries.



The House Office of Legislative Counsel (HOLC) is the legislative drafting service of the House of Representatives.

- They are the chief drafters of House legislation.
- The U.S. Senate also has an Office of Legislative Counsel.

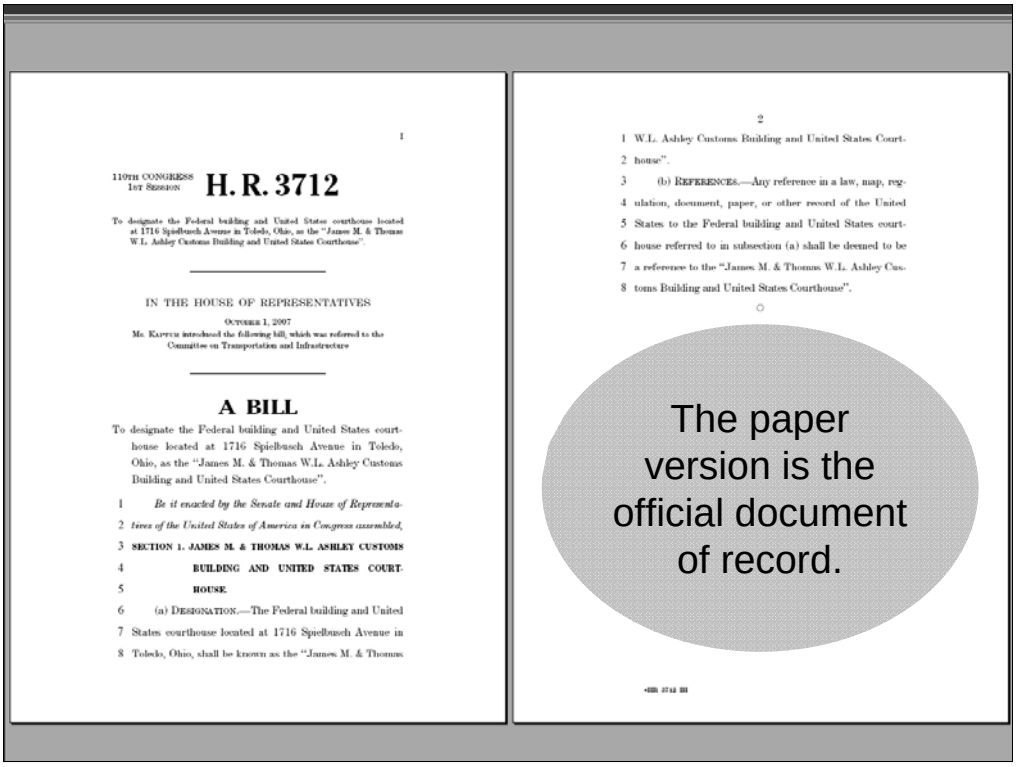
These Offices provide drafting and related assistance to their respective legislators and chamber committees. Although the Members and committees are not required to use the Office, most legislation in the House is worked on by attorneys in the Office.

The signed paper version submitted to the Clerk of the House on the House Floor is the official document of record.

HOLC consists of approximately 35 attorneys and a support staff of about 15 individuals, and is headed by the Legislative Counsel of the House who is appointed by the Speaker of the House. The Office is impartial as to issues of legislative policy and does not advocate the adoption or rejection of any proposal or policy. Although the attorneys in the Office naturally develop special expertise in certain areas of the law, they also assist each other in handling the workload of the Office.

In addition, the Government Printing Office, as well as the House Enrolling Clerks, uses XMetaL.

The Government Printing Office prepares and publishes the paper copies of legislation for use by Congress.



1

110TH CONGRESS
1ST SESSION

H. R. 3712

To designate the Federal building and United States courthouse located at 1716 Spielbusch Avenue in Toledo, Ohio, as the "James M. & Thomas W.L. Ashley Customs Building and United States Courthouse".

IN THE HOUSE OF REPRESENTATIVES

OCTOBER 1, 2007

Ms. KAPTUR introduced the following bill; which was referred to the Committee on Transportation and Infrastructure

A BILL

To designate the Federal building and United States courthouse located at 1716 Spielbusch Avenue in Toledo, Ohio, as the "James M. & Thomas W.L. Ashley Customs Building and United States Courthouse".

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. JAMES M. & THOMAS W.L. ASHLEY CUSTOMS BUILDING AND UNITED STATES COURTHOUSE.

(a) DESIGNATION.—The Federal building and United States courthouse located at 1716 Spielbusch Avenue in Toledo, Ohio, shall be known as the "James M. & Thomas

2

W.L. Ashley Customs Building and United States Courthouse".

(b) REFERENCES.—Any reference in a law, map, regulation, document, paper, or other record of the United States to the Federal building and United States courthouse referred to in subsection (a) shall be deemed to be a reference to the "James M. & Thomas W.L. Ashley Customs Building and United States Courthouse".

The paper version is the official document of record.

As stated before, the paper version is the document of record. Because the paper is the final product and not an electronic rendition, the drafters provide their clients with typeset drafts or PDF files that can be printed in the legislator's office.

The paper version of legislation is currently created by:

•Creating an XML file using a customized version of JustSystems XMetaL® product.

•This XML file is then converted to a U.S. Government Printing Office proprietary typesetting coded file.

•This file is then processed through a composition (or typesetting) program developed by the Government Printing Office. The output of the typesetting program is an Adobe® PostScript® file that is subsequently printed or distilled into a PDF file.

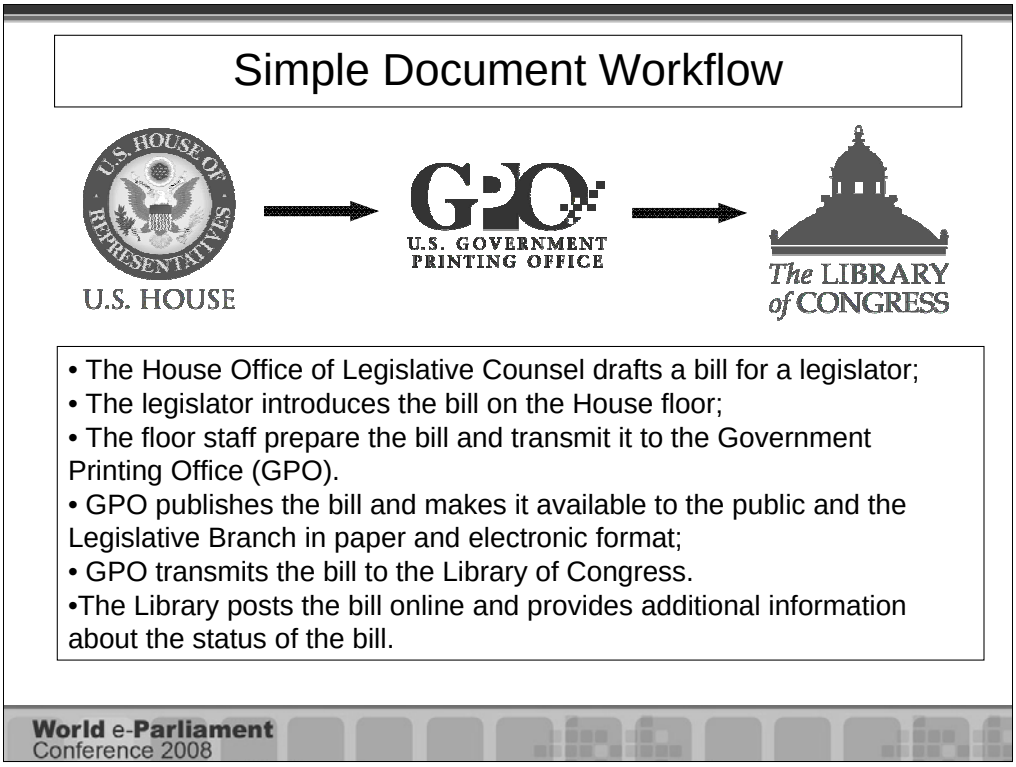
The conversion from the XML to the proprietary typesetting codes is challenging. As XSLT and other XML transformation tools mature, it is a goal to compose the PostScript or PDF file directly from the XML.

The House has achieved XML document creation and exchange by:

•Creating schema-compliant Document Type Definitions (DTD) that are used to exchange bills, resolutions, and amendments in XML between the House and the Senate, as well as support agencies (The Government Printing Office and The Library of Congress);

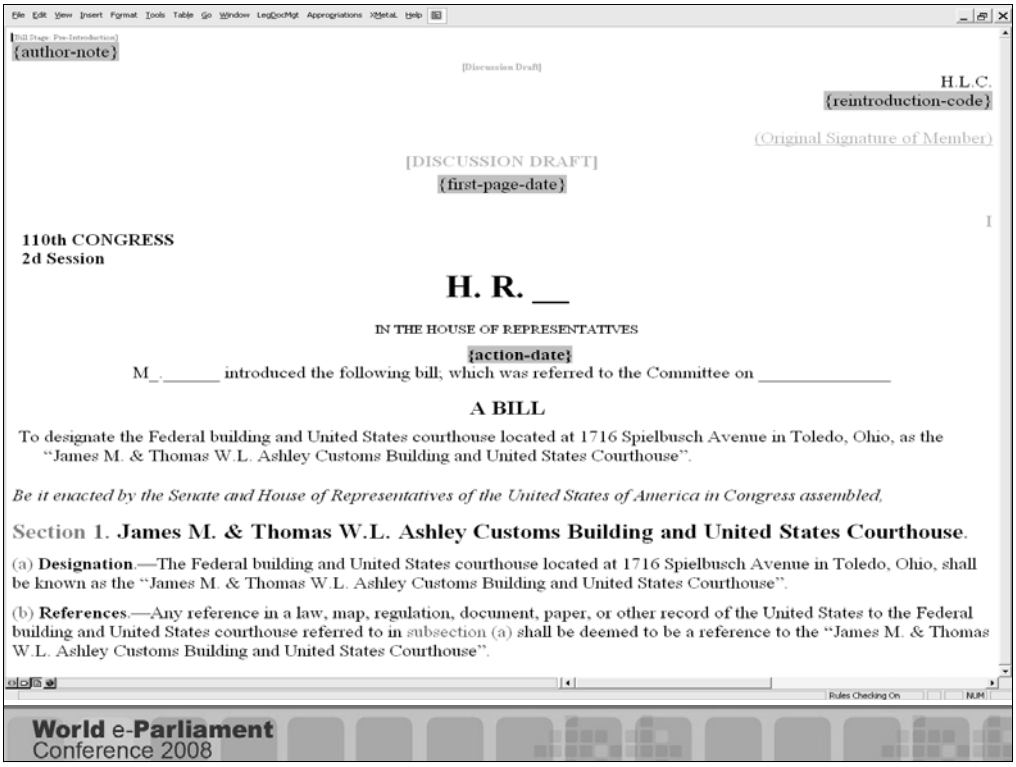
•Managing a common tag library to ensure that shared XML element names and attributes are the same across the various legislative documents;

•Highly customizing JustSystems XMetaL ® to meet the needs of the drafters in the Office of Legislative Counsel and our other users.



As with many legislative bodies, the document workflow can be very complex or very simple.

For the publication of an introduced bill, the electronic file starts in the House, and is transmitted to the Government Printing Office, then to the Library of Congress.



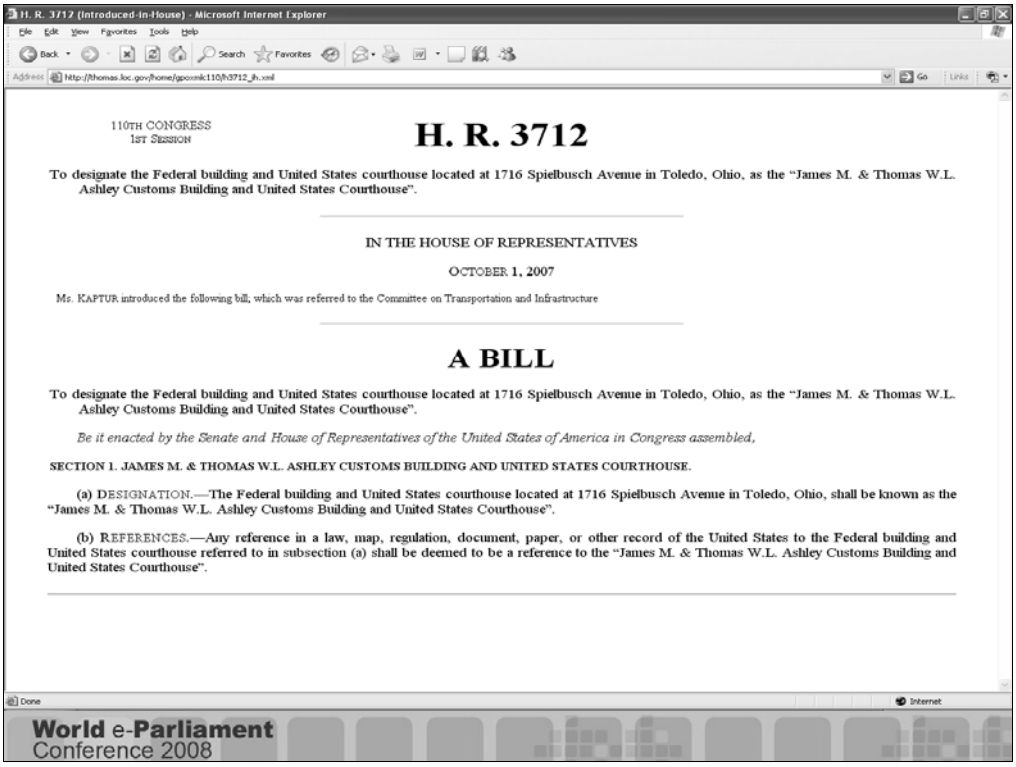
[Slide illustrates XMetaL in Tags-Off view]

- This is what the bill looks like in the XML Editor, XMetaL.

After the Member has filed the paper copy on the House floor, the Office of the Clerk updates the document by adding the sponsor name, date of introduction and the committee referral information.

The XML file is transmitted to the Government Printing Office, which produces the paper copy. The paper copy is made available to the public.

[A paper copy is shown during presentation.]



[Slide illustrates XML files in web browser]

•The XML file is transmitted to the Library of Congress for Internet display. This is the original source document created by the author.

•Links to public law and U.S. Code citations are added to the XML document before posting to the web.

•In addition, dublinCore metadata is added to the document.

•There are over 22,000 bills and resolutions posted online and available to the public.

SAMPLE:

<dublinCore>

<dc:title>110 HR 3712 IH: To designate the Federal building and United States</dc:title>

<dc:publisher>U.S. House of Representatives</dc:publisher>

<dc:date>2007-10-01</dc:date>

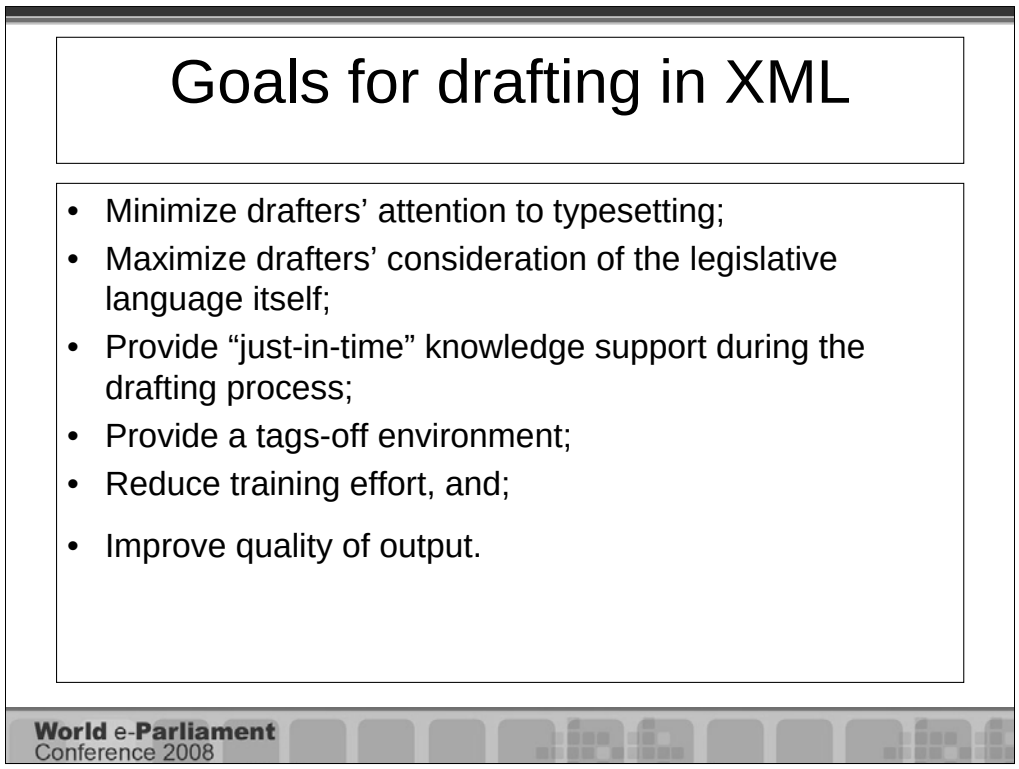
<dc:format>text/xml</dc:format>

<dc:language>EN</dc:language>

<dc:rights>Pursuant to Title 17 Section 105 of the United States Code, this file is not subject to copyright protection and is in the public domain.</dc:rights>

</dublinCore>

Additional information about the metadata added in the Legislative documents (not included in presentation): In November 2004, the Legislative Branch XML Working



When discussing goals of authoring in XML, the following goals are often mentioned:

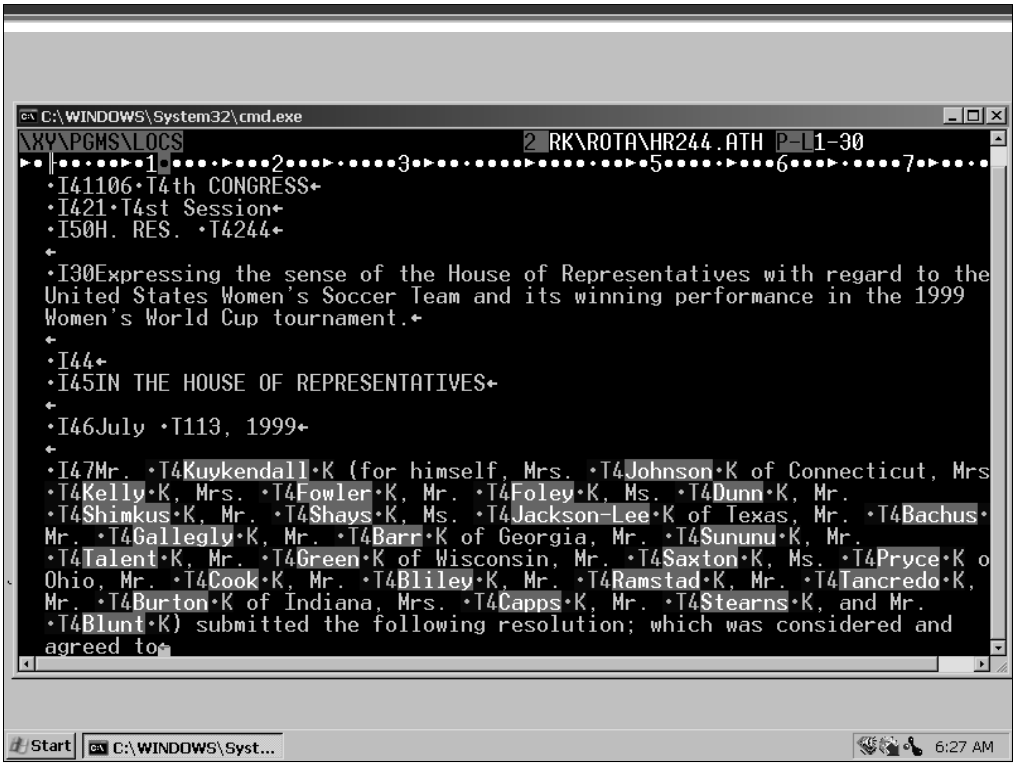
- **modularity,**
- **version control,**
- **consistent formatting,**
- **search ability, and**
- **the ability to publish in multiple formats.**

The House was interested in achieving those goals too, but we also wanted to provide a better authoring environment for our drafters.

Prior to XML, our authors were using XyWrite, a DOS-based text editor and inserting proprietary typesetting codes into their text file. It took months for the drafters to the learn the codes and after becoming proficient, the end result was not a proficient drafter, but a proficient typesetter.

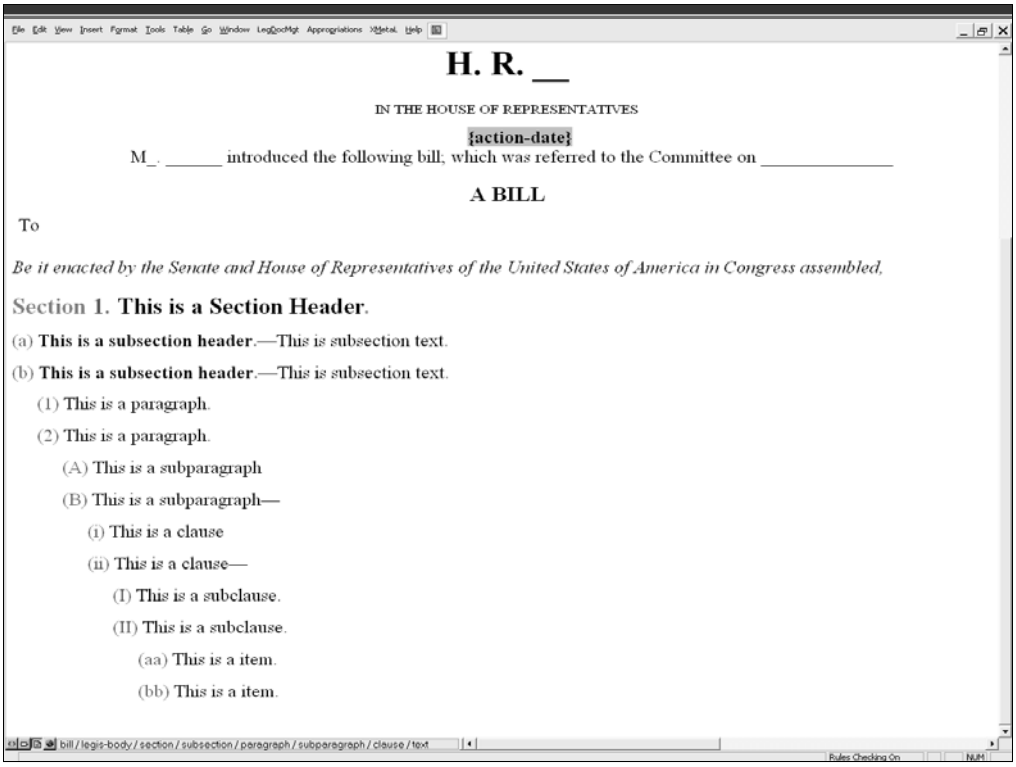
With this, we focused on providing the authors a "smart" authoring environment that would:

- Minimize drafters' attention to the typesetting product;
- Maximize drafters' consideration of the legislative language itself;
- Provide "just-in-time" knowledge support during the drafting process;
- Provide all this within a WYSIWYG (what you see is what you get) environment – one where the user does not see XML tags;
- Reduce training effort, and;
- Improve quality of output.



[Illustrates the old environment]

XyWrite, a DOS-based text editor with propriety typesetting codes.



[Slide illustrates XMetaL in Tags-Off view]

This is the XMetaL environment.

•This slide illustrates some of the legislative structure found in U.S. Federal legislation.

- •The basic unit of legislation is the section.
- •Sections can contain seven (7) levels of hierarchy within them (subsection, paragraph, subparagraph, clause, subclause, item, and subitem). Sections can also be within eight (8) higher levels (division, subdivision, title, subtitle, chapter, subchapter, part, and subpart).

•In the DTD/Schema, we do name each hierarchy by the name it is known as. We felt that it was important to use element names that the drafters recognized.

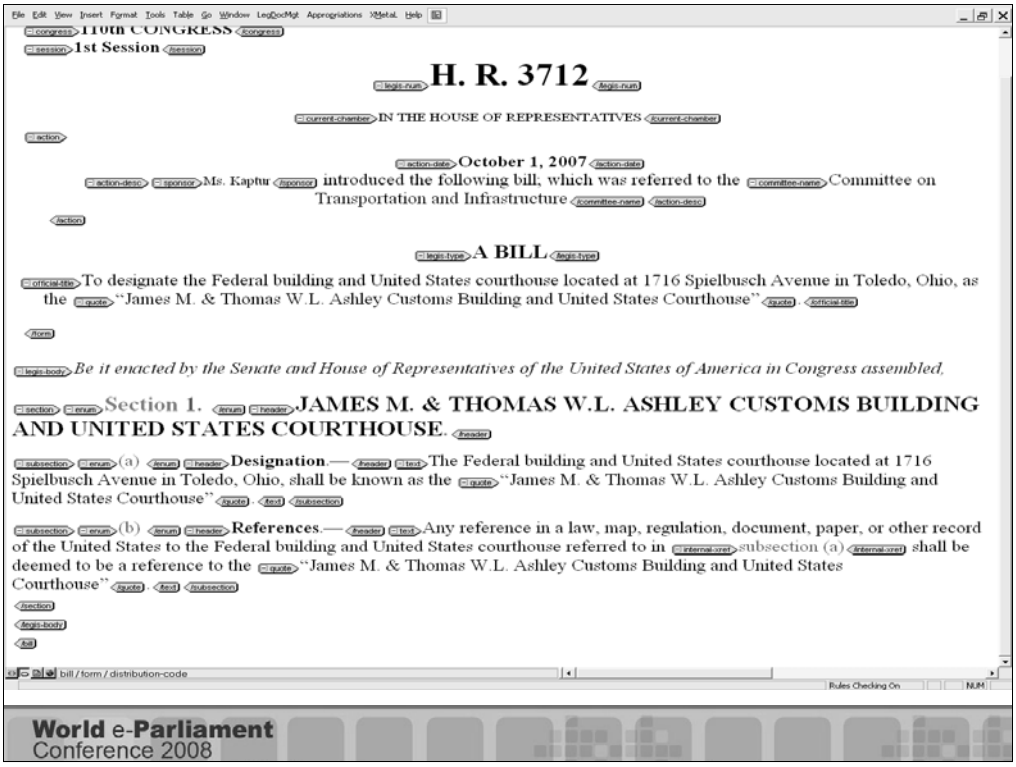
•The levels are nested in a parent/child relationship.

•It was important in the development of XML authoring environment for drafting legislation to create a "smart" editing environment based on context.

•We did not want to exchange our old typesetting codes for a set of XML tags.

•We ensured that editing environment provided the tags based on the content structure instead of the authors picking the tag.

In addition, legislation is drafted using various styles. The styles are a byproduct of previously passed legislation and sometimes individual draft style preferences. Over the years, there have been efforts to standardized the styles of the legislation. However, because Congress continues to amend prior enacted laws, the various styles must be available to legislative drafters.



[During presentation, this will be illustrated in the XMetaL application.]

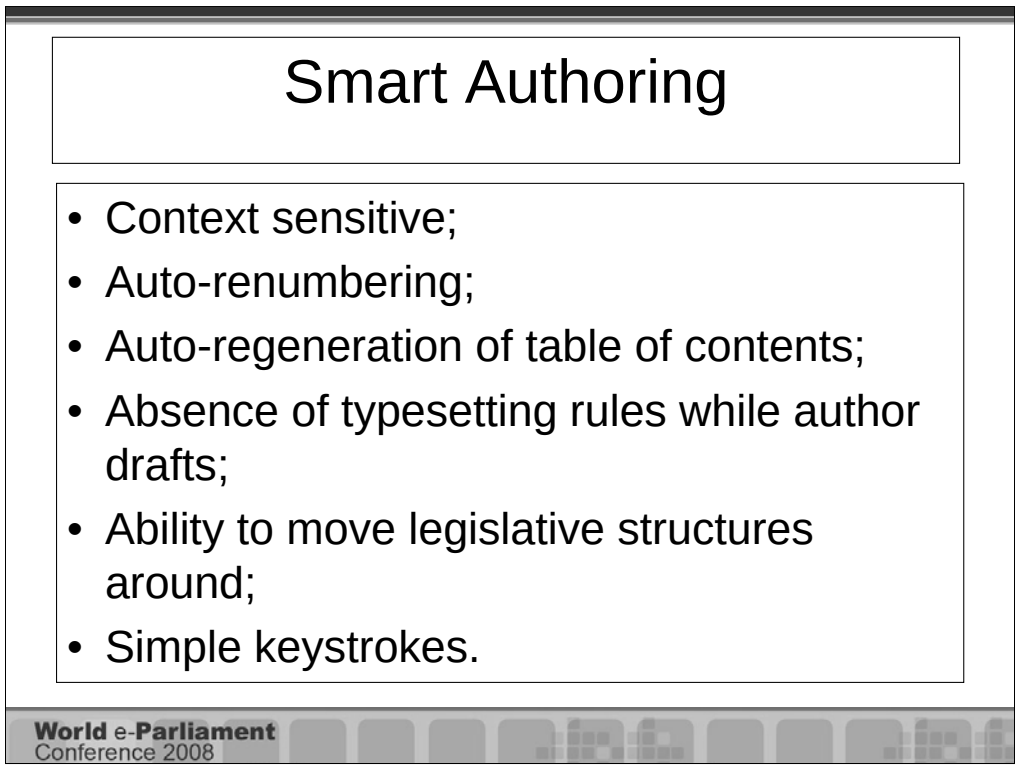
[Slide illustrates XMetaL in Tags-On view]

The users have access to the tags-on environment, however, very little manipulation of the document is allowed in this view.

The elements or tags illustrate the nested structure of the DTD/Schema.

[During presentation, this will be illustrated in the XMetaL application.]

[Slide illustrates XMetaL in Plain View. This is the “native” XML file.]



What is Smart Authoring?

•Context sensitive.

•Helps the drafter improve the document and focus on the content and not the typesetting or the style.

Key Features In The Editor

•Auto-renumbering

•Absence of typesetting rules while author drafts – Example: headers

•Navigation

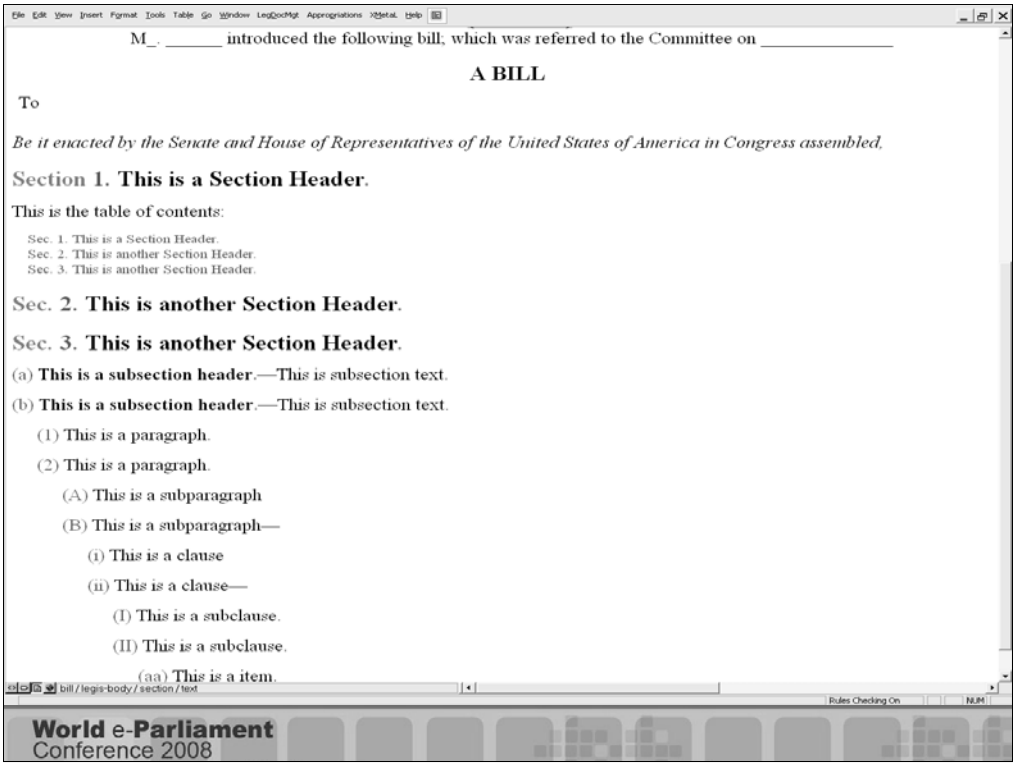
•Adherence to legislative drafting rules

•The ability to move legislative structures around

•Auto-generation of table of contents.

•Simple keystrokes to create legislative structures (ENTER; CTRL+ENTER; CTRL+SHIFT+ENTER)

Additional Information on Development and Deployment: We first deployed the customized XMetaL editor to the Legislative Counsel in December of 2001. A phased-in deployment approach has been used. In the earlier years, the editor could only be used to draft simple House bills and resolutions. As the project progressed, more complex legislative structures were added. Today, our authors can draft nearly 99% of all the legislative structures that exist in Federal law. All XML documents are exchanged and processed by the Government Printing Office and Library of Congress. When the XML editor cannot be used, documents are drafted or edited in the former GPO proprietary typesetting system.



[During presentation, these features will be illustrated in the XMetaL application.]

[Slide illustrates XMetaL in Tags-OFF view]

Features of "Smart Authoring"

•WYSIWYG-Like XML Authoring (What-you-see-is-what-you-get)

Authoring in a structured XML environment is different than a free-form word processor, but XML editors are beginning to see the value of providing a CSS-based authoring environment.

•Minimize drafters' attention to the typesetting product

Auto-numbering: Automatic numbering of levels has freed the drafter from this administrative chore. When new levels are inserted, moved, or deleted, the renumbering occurs automatically. This reduces the attention needed to number the levels correctly and increases the attention to legislative language. Drafters can also turn this feature off if desired.

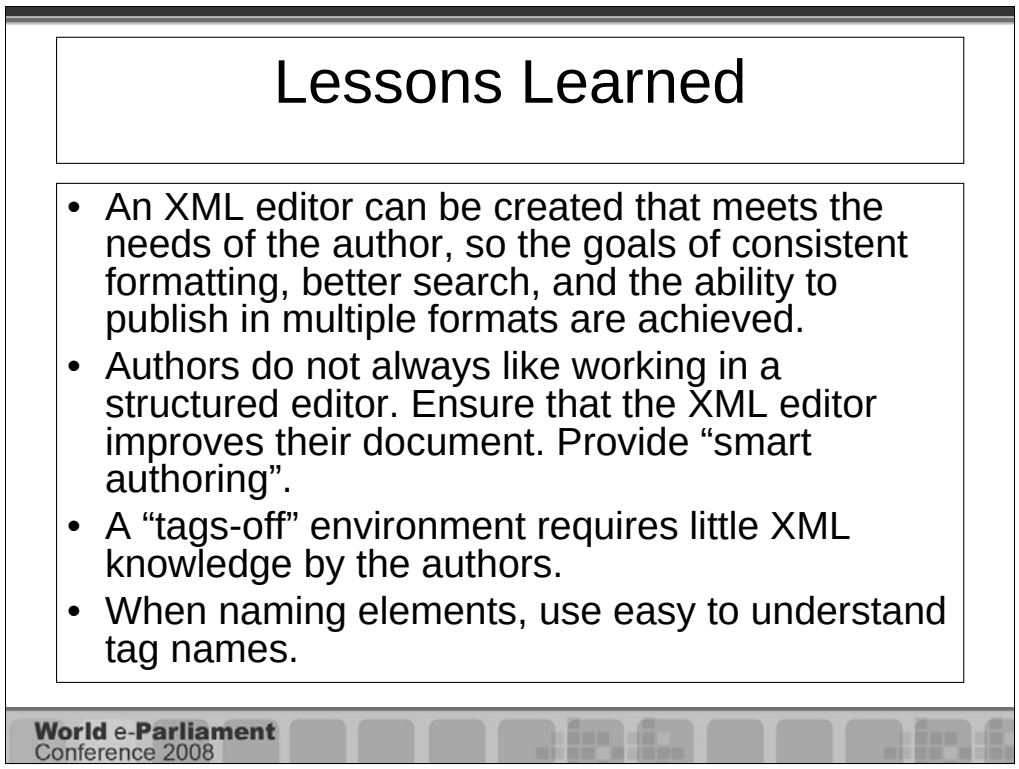
Heading Casing and Table of Contents Generation. In the DOS environment, drafters needed to use the appropriate upper and lower casing for each level's heading based on the required appearance when typeset. This created several drafting inefficiencies: (1) headings needed to be retyped with the correct casing when levels were demoted or promoted; (2) drafters needed to learn the casing rules for each level; and (3) additional work was needed to create tables of contents because the casing of headings within tables of contents use sentence style casing. The solution was to draft headers using sentence style casing. The appropriate casing for each level is then applied automatically by the rendering mechanism. **Although this has improved efficiency and quality, the main benefit is that drafters have one less typesetting requirement to think about in the drafting process.**

•Maximize drafters' consideration of the legislative language itself

Moving levels/Promote and Demote. While there is no arbitrary cut and paste in XMetaL, the XML allows us to create tools that enable the draft to move whole legislative provisions. For example, a subsection can be promoted to a section, or a paragraph can be demoted to a subparagraph. Levels, like a section and its children, can be moved up or down in the document. This re-organization of the document by moving whole legislative provisions is extremely useful to the drafters.

These are just a few of the smart editing features built into the application.

These improve the quality of the drafting and enable drafters to focus on content, not style or typesetting.



Lessons Learned

•An XML editor can be created that meets the needs of the author, so the goals of consistent formatting, better search, and the ability to publish in multiple formats are achieved.

•Drafters do not always like drafting in a structured editor. To overcome that, ensure that there are enough features that help the drafter improve their drafting process and create a document with less errors.

•It is important to know the structure of your legislation before creating the DTD/Schema.

•An XML editing environment is excellent for sophisticated subject matter experts

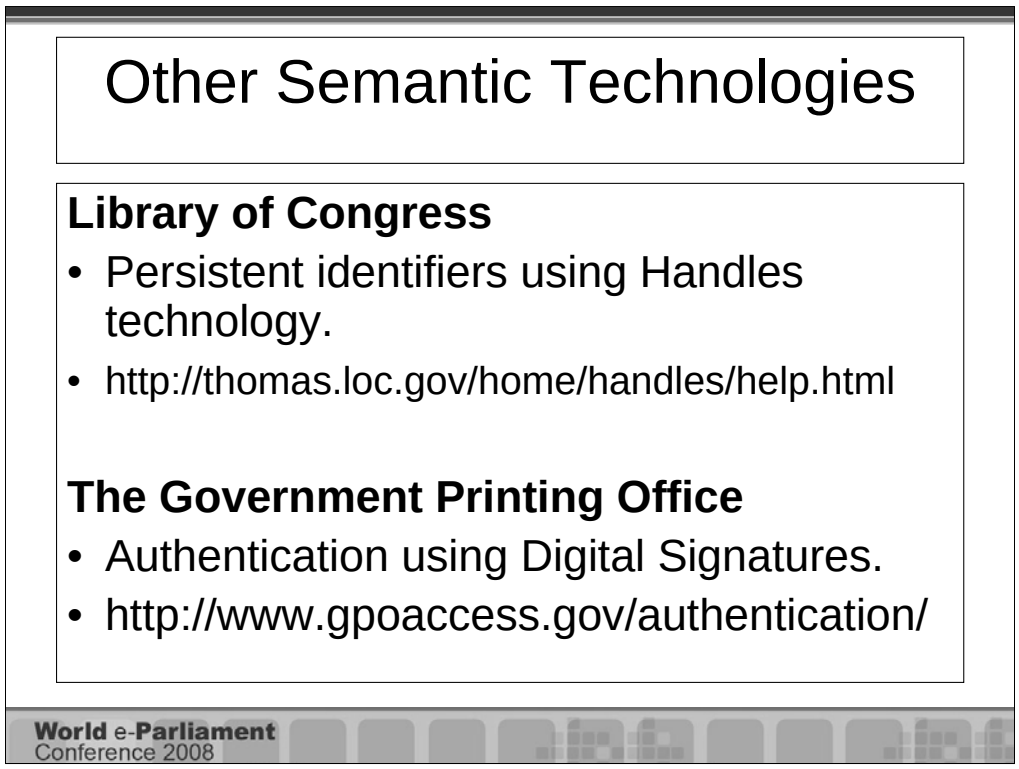
•A tags-off editing environment requires very little understanding of XML by the authors. Authors only need to know how legislative is structured.

•When naming element use easy to understand tag names.

In the end, the transition to XML for the drafting of legislation has been both challenging and highly rewarding for the House.

•The House has been using the XML authoring environment since January 2001. Through the Library of Congress’ Thomas website (http://thomas.loc.gov/) and GPO’s GPO Access website (http://www.access.gpo.gov/), the public accesses legislative documents.

•DTDs, schemas, and examples of XML bills and resolutions are available at http://xml.house.gov/.



Other Uses Of Semantic Technologies

In addition to XML, the Legislative Branch in the U.S. government is also taking advantage of digital signatures and the use of handles.

Handles

Persistent identifiers provide long-term access to a web page. Unlike URLs, persistent identifiers do not change when the underlying file or content moves. A handle is a type of persistent identifier for a digital object.

Using Handle technology, the Library of Congress is able to create a persistent identifier for each legislative document. This helps achieve continuity internally and externally.

A handle system helps minimize confusion as new technologies are implemented. Handles would provide a consistent approach to accessing data through the planned technology transitions.

In addition, a handle system would allow outside partners (or those groups) that link to the Library's resouces. With the handle system, our links would be easy to type into an address bar. Online content producers could easily reference content directly.

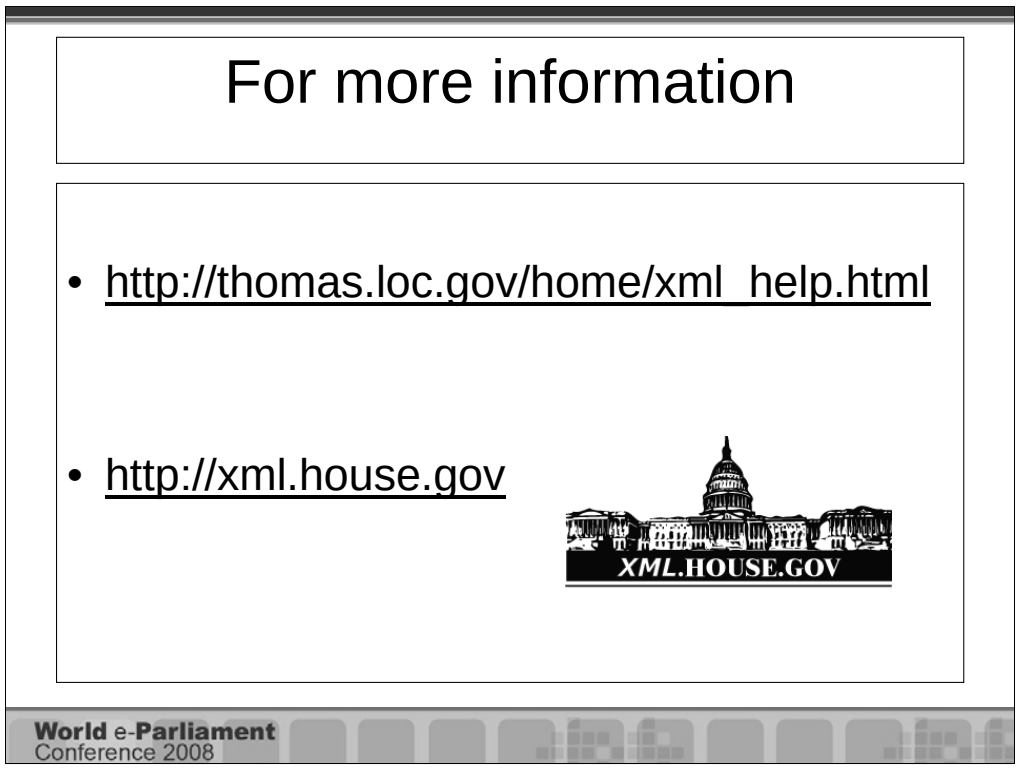
Digital Signatures

For almost 150 years, the U.S. Government Printing Office (GPO) has been the official disseminator of Government documents and has assured users of their authenticity.

In the 21st century, the increasing use of electronic documents poses special challenges in verifying authenticity, because digital technology makes such documents easy to alter or copy, leading to multiple non-identical versions that can be used in unauthorized or illegitimate ways.

To help meet the challenge of the digital age, GPO has begun implementing digital signatures to certain electronic documents on GPO Access that not only establish GPO as the trusted information disseminator, but also provide the assurance that an electronic document has not been altered since GPO disseminated it.

The visible digital signatures on online PDF documents serve the same purpose as handwritten signatures or traditional wax seals on printed documents. A digital signature, viewed through the GPO Seal of Authenticity, verifies document integrity and authenticity on GPO online Federal documents, at no cost to the customer.



Additional Information Not Covered in the Presentation

Exchange DTD/Schemas Versus Editing DTD/Schemas

The U.S. House and U.S. Senate do have different authoring environments.

The House's editor uses the agreed-upon exchange DTD/Schema. This DTD/Schema keeps the authors in the hierarchical structure of the legislations. The legislative structures are is a nested hierarchy (parent-child relationships). With the exchange DTD/Schema, the document is valid all the time and forces the author to think about structure while drafting.

The Senate uses an editing DTD/Schema. The hierarchical structure is flattened and offers the author a more flexible environment while drafting. The legislative structures are peers and are not contained by a parent element. The Senate application converts the authoring DTD/Schema to the exchange DTD/Schema after a validation routine is run to check for XML and business rule validity. Both approaches are valid and meet the needs of the author.