



UNITED NATIONS



CHAMBER OF DEPUTIES OF ITALY



INTER-PARLIAMENTARY UNION



World e-Parliament Conference 2012

13-15 September 2012 // Chamber of Deputies of Italy // Rome

XML in Process

United States House of Representatives
Kirsten Gullickson, Senior Systems Analyst

World e-Parliament Conference 2012

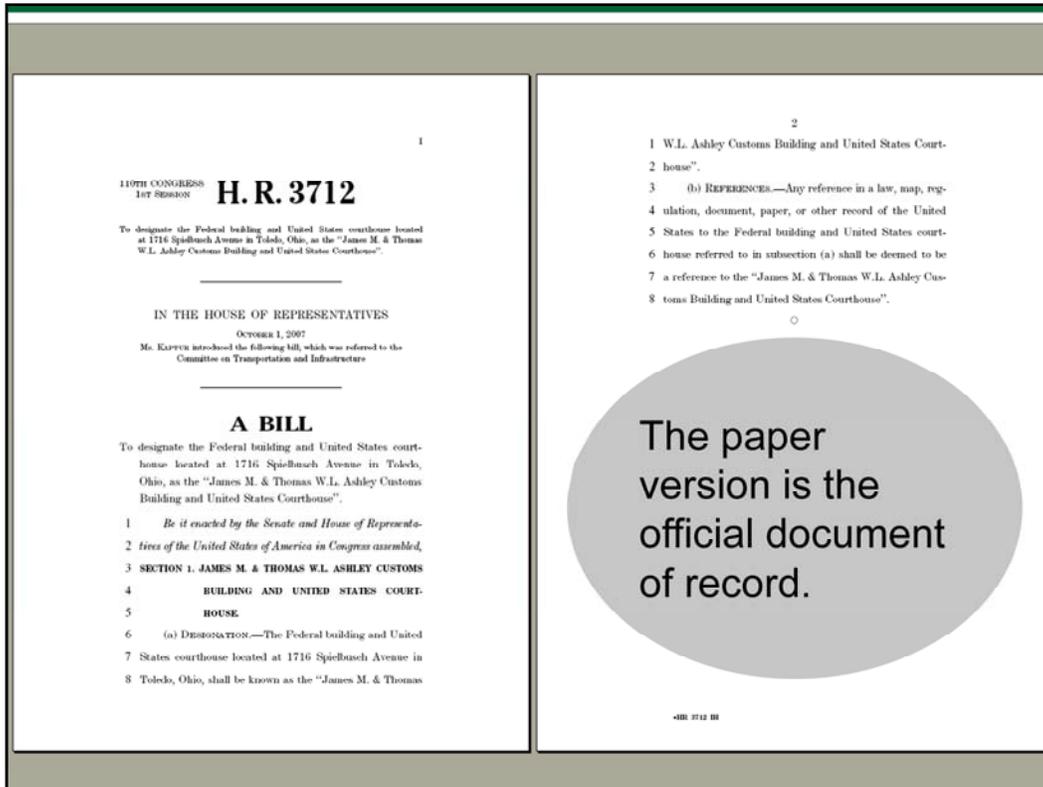
This presentation will focus on XML in Process in the U.S. House. The U.S. House, along with the Senate and other Legislative Branch agencies, are engaged in an on-going effort to author and publish its official documents and data in XML.

The Challenge

- Legislative documents and related data must be
 - prepared
 - managed,
 - distributed, and
 - archived.
- This includes paper and electronic means for handling the official documents.

World e-Parliament Conference 2012

In 2012, government documents and related data must be prepared, managed, distributed, and archived. In the distant past, these tasks were confined to paper only, but now it includes electronic documents and data.



For the United States Congress, the paper version is our official document of record. This includes legislation (bills, resolutions and amendments), as well as our record of debates called the *Congressional Record*. We do prepare, manage, distribute and archive the electronic versions.

Background

- The Legislative Branch XML Initiative is rooted in a **1996 directive** to work together **to establish common data standards** for the exchange of legislative information. (See also Public Law 104-197)
- The primary goal is to produce and exchange legislative documents using the industry-based standards approach of the **Extensible Markup Language (XML)**.

World e-Parliament Conference 2012

Our move to XML is rooted in a 1996 directive to establish common data standards for the exchange of legislative information.

Since the beginning, the primary goal has been to produce and exchange legislative documents using XML.

Additional Goals

- Improve the legislative document creation and revision process
- Improve information exchange with the Senate and other Legislative Branch organizations to facilitate the legislative process
- Allow the House to be more independent in document preparation.

World e-Parliament Conference 2012

Other goals identified in the late 1990's timeframe include:

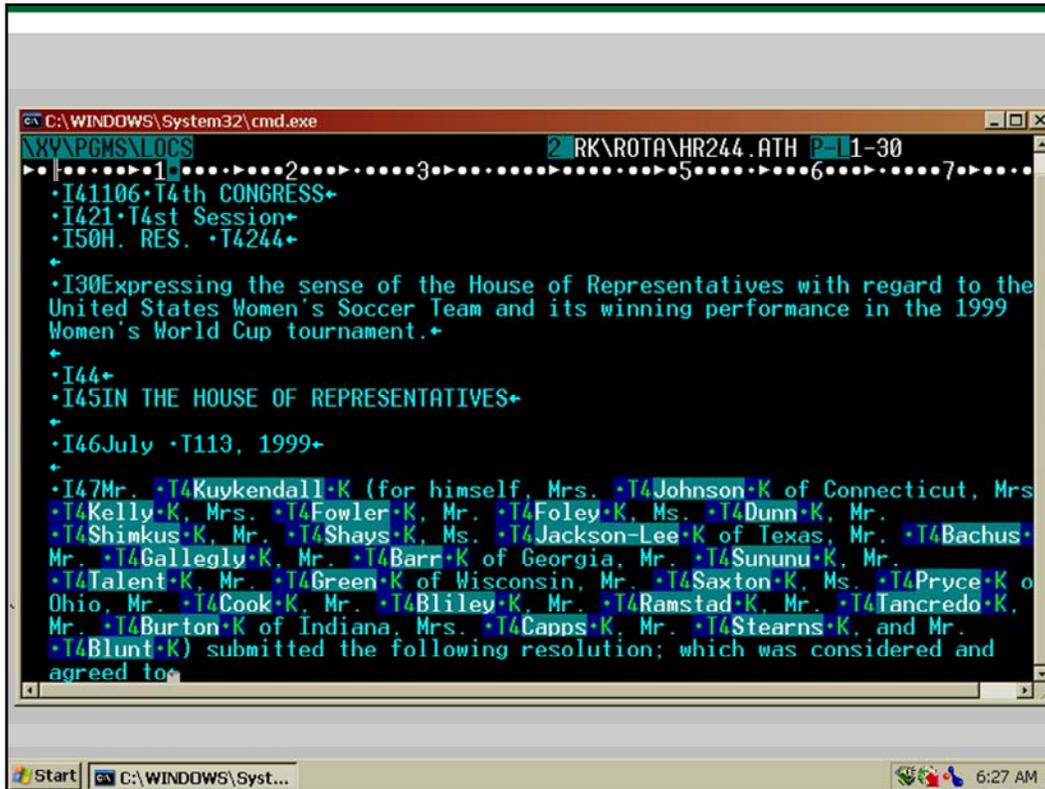
--improving the legislative document creation and revision process;

--improving information exchange with the Senate and other government organizations in order to facilitate the legislative process;

--allowing the House of Representatives to become more independent in the preparation, printing, and distribution of official House of Representative documents and not having to rely on staff at the Government Printing Office (GPO), our official printer, for having to typeset so many of our documents.

(source: Testimony of the Clerk before the House Subcommittee on Legislative, *Legislative Branch Appropriations for FY2000*, hearings, Feb. 2, 1999, pp.34-35.)

--Another goal we had was to move away from our official government printing office's typesetting codes. It takes months for our users to learn the codes and become proficient - the end result is not a proficient legislative content author, but a proficient typesetter.



[Illustrates the old bill and resolution authoring environment]

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

The On-going Challenge

Government data must be

- Accurate
- Accessible
- Authenticated
- Complete
- Detailed
- Searchable

World e-Parliament Conference 2012

In the late 1990's, the House was responding to internal needs 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.

Of course, the beginning was more than 14 years ago. Today, legislative bodies need to respond to the new challenges and demands of the “open data, open government, transparency” movement ensuring that our legislative documents and data are accurate, accessible, authenticated, complete, detailed and searchable.

Solution: Structured Data

- Publishing well-formed XML can help meet the goals of
 - accountability of the government,
 - transparency,
 - accessibility,
 - empowerment of the public,
 - higher quality of information, and
 - a way to make technology more human.
- XML provides a means to publish raw data (or lemons) and for others to add value (or make lemonade).

World e-Parliament Conference 2012

XML does help meet the needs of accountability, transparency, accessibility, and empowerment of the public.

The Legislative Branch XML Working Group

- A group of representatives from the Legislative Branch agencies:
 - to manage and coordinate changes to the **Common Tag Library** and **DTDs**, and
 - manage issues involving exchanging and drafting legislative documents using these DTDs.

World e-Parliament Conference 2012

At the start of this process, in 1996, the Legislative Branch formed a Working Group to manage and coordinate the data standards for exchange. The Working Group is made up of staff from the House and Senate and other Legislative Branch organizations who agree on the XML elements and DTDs. So, we have a process in the U.S. Congress to manage and approve our XML elements and DTDs. This includes managing a **common tag library**. We firmly believe that we should be using the same element names and attributes names for the same data across our datasets.

I need to point out that we are not using the **proposed OASIS standard**, mainly because we started on this road before its existence. I don't foresee us change the DTD for our bills and resolutions, but I hopeful that we will be able to publish some of our legislative data, perhaps our *Congressional Record* in the proposed OASIS Standard.

Creating XML

- **Author**-created using customized version of Just System's XMetaL.
- **Programmatically**-created by extracting data from a database and publishing in XML.

World e-Parliament Conference 2012

We have several ways we create and publish XML.

- For our bills and resolutions, the XML content is created by the author. We have a highly customized version of Just System's XMetaL for authoring bills, resolutions, and amendments. We are in the process of customizing it for the creation of committee reports.
- For House floor proceedings and House chamber votes, the XML content is created programmatically using data stored in a database.
- For our disclosure data, the XML content is created from online forms that required filers complete and submit via website.

Basic Legislative Workflow

<http://xml.house.gov/wep2012/>

H.R. 3248

Introduced in House

PDF Version:

<http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248ih/pdf/BILLS-112hr3248ih.pdf>

XML Version:

http://thomas.loc.gov/home/gpoxmlc112/h3248_ih.xml

Bill Status and Summary: <http://Thomas.loc.gov>

Select "Bill Number", Input HR 3248 in the search box, Click on Search.

World e-Parliament Conference 2012

DRAFTING:

- The House Office of Legislative Counsel drafts a bill (or resolution) at the request of a Member of Congress.
- The measure is drafted in XML using the House's customizations to Just Systems' XMetaL.

INTRODUCTION:

- The Member introduces the measure.
- The draft bill is provided to the Member in PDF format. When the Member of Congress decides to introduce the bill or resolution, the Member signs the paper version of the bill and drops the paper in a box called "The Hopper".
- The signed bill is the official document of record.

PROCESSING:

- The bill is processed by the House Office of the Clerk.
- House Clerk staff retrieve the electronic XML version from the House Office of Legislative Counsel.
- The paper version dropped in the Hopper as well as the electronic XML version is transferred to the Government Printing Office for printing and posting online. The PDF and XML versions are posted on <http://www.gpo.gov> and <http://thomas.loc.gov>.
- Data regarding the bill is transferred from the House Clerk's office to the Library of Congress. Library staff write a summary of the bill. The bill summary and status are posted on <http://thomas.loc.gov>.

H.R. 3248 – Introduced in House

PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248ih/pdf/BILLS-112hr3248ih.pdf>

XML Version: http://thomas.loc.gov/home/gpoxmlc112/h3248_ih.xml

Bill Status and Summary: <http://Thomas.loc.gov>

Select "Bill Number", Input HR 3248 in the search box, Click on Search

Committee Consideration

H.R. 3248 was considered under suspension of the House rules.

There was no committee consideration.

World e-Parliament Conference 2012

COMMITTEE CONSIDERATION: The bill or resolution will be referred to committee. If the bill is fast-tracked, it will bypass committee consideration and be considered by the House under a special rule or the rules of the House are suspended. *For the purposes of this discussion, the committee consideration process is not included.*

H.R. 3248

House Consideration

H.R. 3248:

<http://docs.house.gov/Default.aspx?date=2012-02-13>

Floor Proceedings for February 15, 2012 when H.R. 3248 was considered and passed.

<http://clerk.house.gov/floorsummary/floor.aspx?day=20120215&today=20120906>

H.R. 3248: Roll Call Vote 54, February 15, 2012

<http://clerk.house.gov/evs/2012/roll054.xml>

World e-Parliament Conference 2012

PROCESSING BEFORE HOUSE CONSIDERATION: Before the bill is considered by the House, the Committee on Rules or the Majority Leader's office will post the measure online. According to House rules, the bill should be available three days before its consideration. H.R. 3248 <http://docs.house.gov/Default.aspx?date=2012-02-13>

HOUSE CONSIDERATION: When the measure is considered by the House, the proceedings are recorded by Clerk staff and published in real time on the House Clerk's website. If the bill was considered by roll call vote, the vote is recorded by electronic device and the results are posted in XML on the House Clerk's website. Floor Proceedings for February 15 when H.R. 3248 was considered and passed. <http://clerk.house.gov/floorsummary/floor.aspx?day=20120215&today=20120906> H.R. 3248: Roll Call Vote 54, February 15, 2012 <http://clerk.house.gov/evs/2012/roll054.xml>

H.R. 3248

Engrossed in House

PDF Version:

<http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248eh/pdf/BILLS-112hr3248eh.pdf>

XML Version:

http://thomas.loc.gov/home/gpoxmlc112/h3248_eh.xml

Bill Status and Summary: <http://Thomas.loc.gov>

Select "Bill Number", Input HR 3248 in the search box, Click on Search.

World e-Parliament Conference 2012

PROCESSING: If the bill passed the House, the engrossed (House-passed) version of the bill is processed by the Office of the Clerk using the House's customizations to Just Systems' XMetaL. The Clerk of the House certifies passage of the bill with a signature. The paper version is the official document of record. The paper version as well as the electronic XML version will be transferred to the Government Printing Office for printing and posting online. The PDF version and XML version are posted at <http://www.gpo.gov> and <http://thomas.loc.gov>

H.R. 3248 – Engrossed in House

PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248eh/pdf/BILLS-112hr3248eh.pdf>

XML Version: http://thomas.loc.gov/home/gpoxmlc112/h3248_eh.xml

Bill Status and Summary: <http://Thomas.loc.gov>

Select "Bill Number", Input HR 3248 in the search box, Click on Search

H.R. 3248

Senate Consideration

H.R. 3248 – Received in Senate

PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248rfs/pdf/BILLS-112hr3248rfs.pdf>

XML Version:

http://thomas.loc.gov/home/gpoxmlc112/h3248_rfs.xml

H.R. 3248 – Reported in Senate

PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248rs/pdf/BILLS-112hr3248rs.pdf>

XML Version:

http://thomas.loc.gov/home/gpoxmlc112/h3248_rs.xml

World e-Parliament Conference 2012

SENATE CONSIDERATION: The paper version is delivered to the Senate for consideration. The Senate receives and considers the bill. The paper version as well as the electronic XML version will be transferred to the Government Printing Office for printing and posting online. The PDF version and XML version are posted at <http://www.gpo.gov> and <http://thomas.loc.gov>

H.R. 3248 – Received in Senate PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248rfs/pdf/BILLS-112hr3248rfs.pdf> XML Version: http://thomas.loc.gov/home/gpoxmlc112/h3248_rfs.xml

H.R. 3248 – Reported in Senate PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248rs/pdf/BILLS-112hr3248rs.pdf> XML Version: http://thomas.loc.gov/home/gpoxmlc112/h3248_rs.xml

H.R. 3248 Enrolled

PDF Version:

<http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248enr/pdf/BILLS-112hr3248enr.pdf>

XML Version:

http://thomas.loc.gov/home/gpoxmlc112/h3248_enr.xml

Bill Status and Summary: <http://Thomas.loc.gov>

Select "Bill Number", Input HR 3248 in the search box, Click on Search.

World e-Parliament Conference 2012

PROCESSING: A message is sent to the House that the Senate considered and passed the House bill. House Clerk staff will produce the enrolled (House-passed and Senate-passed) version of the bill. The Clerk of the House and Secretary of the Senate certifies passage of the bill with signatures. The Speaker of the House and the President of the Senate sign the bill. The paper version of the bill will be delivered to the White House for the President's signature. The signed version is the official document of record. The paper version as well as the electronic XML version will be transferred to the Government Printing Office for printing and posting online. The PDF version and XML version are posted at <http://www.gpo.gov> and <http://thomas.loc.gov>

H.R. 3248 – Enrolled PDF Version: <http://www.gpo.gov/fdsys/pkg/BILLS-112hr3248enr/pdf/BILLS-112hr3248enr.pdf> XML Version: http://thomas.loc.gov/home/gpoxmlc112/h3248_enr.xml

H.R. 3248 Public Law

H.R. 3248

PDF Version: <http://www.gpo.gov/fdsys/pkg/PLAW-112publ118/pdf/PLAW-112publ118.pdf>

Bill Status and Summary: <http://Thomas.loc.gov>
Select "Bill Number", Input HR 3248 in the search box, Click on Search.

World e-Parliament Conference 2012

PRESIDENTIAL ACTIONS: The paper version is presented to the President. The President signs the paper version. The paper version is the official document of record.

11. PROCESSING: A Public Law number is assigned to the bill and a public law document is produced. The paper version as well as the electronic version (in GPO typesetting code) will be transferred to the Government Printing Office for printing and posting online. The PDF version and XML version are posted at at <http://www.gpo.gov> and <http://thomas.loc.gov>

ICT Systems

- Systems to create, manage and distribute legislative data and documents must be--
 - quick,
 - flexible,
 - meet the changing needs of the public and the members of parliaments, and,
 - involve very little risk.
- No interruptions, no failures, no risks, and little cost.

World e-Parliament Conference 2012

As you can see and hear, we have many different organizations involved in the production of legislation and law in our legislative branch. We are very “silo’ed” and over the years, we have struggled to get rid of these silos. We have a lot of work to do yet, however, moving to XML has helped us.

We continue to meet the challenge that information and communication technologies (ICT) systems that create, manage and distribute legislative data and documents must be--

- quick,
- flexible,
- meet the changing needs of the public and the members of parliaments, and,
- involve very little risk;

Parliament staff around the world working on ICT Systems agree that Members, staff and the public expect our systems to have no interruptions, no failures, no risks and cost little.

For more information

- XML: <http://xml.house.gov>



Who uses the XML editor?

House Office of Legislative Counsel

- Main drafters of House legislation, est. in 1919;
- 2,781 bills and resolutions introduced the first six months of 2012;
- 50 attorneys and 20 support staff;
- Attorneys are non-partisan; Attorneys develop expertise in particularly areas of the law, as well as legislative drafting expertise.

House Enrolling Clerks

- Prepare official engrossed version that passed the House and the official enrolled version that passed both chambers;
- 252 bills and resolutions passed the House to date in 2012; 61 bills enacted into law to date in 2012.

Government Printing Office

- Prepares and publishes copies of the legislation for use by Congress; makes copies available to libraries and the general public.

World e-Parliament Conference 2012

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.

Goals for drafting in XML

- Minimize drafters' attention to typesetting;
- Maximize drafters' consideration of the legislative language itself;
- Provide "just-in-time" knowledge support during the drafting process;
- Provide a tags-off environment;
- Reduce training effort, and;
- Improve quality of output.

World e-Parliament Conference 2012

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 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.

Smart Authoring

- Context sensitive;
- Auto-renumbering;
- Auto-regeneration of table of contents;
- Absence of typesetting rules while author drafts;
- Ability to move legislative structures around;
- Simple keystrokes.

World e-Parliament Conference 2012

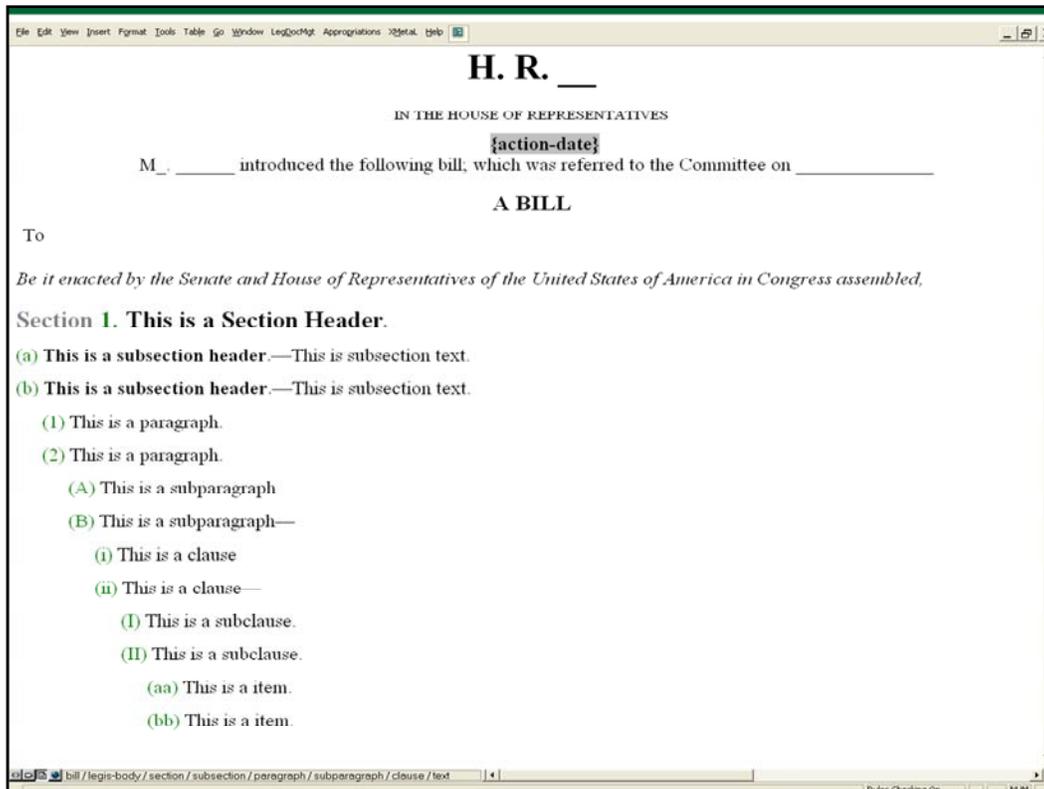
What is Smart Authoring?

- 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.



[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.


```
File Edit View Insert Format Tools Table Comment Go Window LogDocMyz Appropriations XMetaL Help
1: <?xml version="1.0" encoding="UTF-8"?>
2: <?xml-stylesheet type="text/xsl" href="billres.xsl"?>
3: <!DOCTYPE bill PUBLIC "-//US Congress//DTDs/bill v2.0 20020720//EN" "http://xml.house.gov/bill.dtd">
4: <bill bill-stage="Pre-Introduction" bill-type="Title" dms-id="HT05BC12390CC447AB1C914FCF56220E" public-private="public">
5: <pre-form><meta-house><holic-filename>F:\TESTING\SAMPLE_PREINTRODUCED_HR3712.XML</holic-filename><holic-attorney></holic-attorney><holic-last-author>
kgullickson</holic-last-author><holic-last-saved>11/19/2008 17:27</holic-last-saved><holic-creator></holic-creator><holic-creation-date></holic-creation-date>
<version><version-filename>F:\test\VClerk\111908\111908_003.xml</version-filename><version-date>11/19/2008 17:25:16</version-date></holic-creator>
kgullickson</version-creator></version></meta-house><author-note></author-note></pre-form>
6: <running-header display="no">[Discussion Draft]</running-header>
7: <legis-council>H.L.C.</legis-council><reintroduction-code display="no"></reintroduction-code></legis-council>
8: <signature-line display="yes">[Original Signature of Member]</signature-line>
9: <first-page-header display="no">[DISCUSSION DRAFT]</first-page-header>
10: <first-page-date></first-page-date></first-page-date>
11: </pre-form>
12: <form>
13: <distribution-node display="no"></distribution-node>
14: <congress display="yes">110TH CONGRESS</congress> <session display="yes">2d Session</session>
15: <legis-num display="yes">H. R. ...</legis-num>
16: <current-chamber display="yes">IN THE HOUSE OF REPRESENTATIVES</current-chamber>
17: <action>
18: <action-date></action-date></action-date>
19: <action-desc>M. ... introduced the following bill; which was referred to the Committee on ...</action-desc>
20: </action>
21: <legis-type>A BILL</legis-type>
22: <official-title display="yes">To designate the Federal building and United States courthouse located at 1716 Spielbusch Avenue in Toledo, Ohio, as the
<quote>James M. (comp) Thomas W.L. Ashley Customs Building and United States Courthouse</quote>.</official-title>
23: </form>
24: <legis-body id="H21B07EAS07E04D83989474FA7AF6" section-type="section-one"><enum>1.</enum><header>James M. (comp) Thomas W.L. Ashley Customs Building and
United States Courthouse</header>
25: <subsection id="H1773CF2972A4D70B7A400F782EEA5"><enum (a)</enum><header>Designation</header><text display="yes">The Federal
Building and United States courthouse located at 1716 Spielbusch Avenue in Toledo, Ohio, shall be known as the <quote>James M. (comp) Thomas W.L. Ashley
Customs Building and United States Courthouse</quote>.</text> </subsection>
26: <subsection id="H6D96F80175142279D08BDF452C2336"><enum (b)</enum><header>References</header><text display="yes">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
<internal-xref idref="H1773CF2972A4D70B7A400F782EEA5" legis-path="1.(a)">subsection (a)</internal-xref> shall be deemed to be a reference to the
<quote>James M. (comp) Thomas W.L. Ashley Customs Building and United States Courthouse</quote>.</text> </subsection>
27: </legis-body>
28: </bill>
```

[During presentation, this will be illustrated in the XMetaL application.]
[Slide illustrates XMetaL in Plain View. This is the "native" XML file.]

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.
- Authors do not always like working in a structured editor. Ensure that the XML editor improves their document. Provide “smart authoring”.
- A “tags-off” environment requires little XML knowledge by the authors.
- When naming elements, use easy to understand tag names.

World e-Parliament Conference 2012

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 legislation is structured.
- When naming elements, 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/>.



World e-Parliament Conference **2012**