

---

# XML in the U.S. House of Representatives

---

**Kirsten Gullickson**

Office of the Clerk, U.S. House of Representatives

---

# The Challenge

---

# Data and Documents

Documents are not databases.

Documents have three characteristics:

- **Presentation:** how it looks
- **Structure:** how is it organized
- **Semantics:** what does it mean

Raw data is not always raw. Look at it in context

# XML Basics

Element

Attribute

```
<recorded-vote>  
  <legislator name-id="C001101" party="D">  
    Ms. Clark</legislator>  
    <vote>Yea</vote>  
  </recorded-vote>  
  
<recorded-vote>  
  <legislator name-id="B001250" party="R">  
    Mr. Bishop</legislator>  
    <vote>Yea</vote>  
  </recorded-vote>
```

# XML Basics

section enum **Sec. 2.** /enum header **Training and employment for veterans and members of armed forces in curation and historic preservation.** /header

subsection enum **(a)** /enum header **Training and Employment.**— /header text The Secretary of the Army, acting through the Chief of Engineers, shall carry out a Veterans' Curation Program to hire veterans and members of the Armed Forces to assist the Secretary in carrying out curation and historic preservation activities. /text /subsection  
/section

**SEC. 2. TRAINING AND EMPLOYMENT FOR VETERANS AND MEMBERS OF ARMED FORCES IN CURATION AND HISTORIC PRESERVATION.**

(a) TRAINING AND EMPLOYMENT.—The Secretary of the Army, acting through the Chief of Engineers, shall carry out a Veterans' Curation Program to hire veterans and members of the Armed Forces to assist the Secretary in carrying out curation and historic preservation activities.

(b) AUTHORIZATION.—The Secretary shall carry out this section.

(1) \$5,000

(2) \$6,000

(3) \$7,000

**17 SEC. 2. TRAINING AND EMPLOYMENT FOR VETERANS AND**  
**18 MEMBERS OF ARMED FORCES IN CURATION**  
**19 AND HISTORIC PRESERVATION.**

20 (a) TRAINING AND EMPLOYMENT.—The Secretary of  
21 the Army, acting through the Chief of Engineers, shall  
22 carry out a Veterans' Curation Program to hire veterans  
23 and members of the Armed Forces to assist the Secretary  
24 in carrying out curation and historic preservation activi-  
25 ties.

---

# History of XML in the U.S. Congress

---

# History of XML in U.S. Congress

1997

**Establishment of exchange standards:** During the 105th Congress, the Clerk's Office started an effort to establish a document exchange standard and the transformation, creation, and exchange of House documents using XML. This included an effort to customize an XML editor for the creation of bills and resolution. The initiative is rooted in a **1996 directive** from the chairmen of the Committee on House Oversight (now known as the Committee on House Administration) and the Senate Committee on Rules and Administration to the Clerk of the House and Secretary of Senate, respectively, to work together toward establishing common data standards for the exchange of legislative information. (**2 U.S.C. 181**) (See also Public Law 104-197)



[\[Print\]](#) [\[Print selection\]](#)

[\[Close\]](#) [He](#)

## **§181. Program for exchange of information among legislative branch agencies**

(a) On September 16, 1996, there shall be established a program for providing the widest possible exchange of information among legislative branch agencies with the long-range goal of improving information technology planning and evaluation. The Committee on House Oversight of the House of Representatives and the Committee on Rules and Administration of the Senate are requested to determine the structure and operation of this program and to provide appropriate oversight. All of the appropriate offices and agencies of the legislative branch as defined below shall participate in this program for information exchange, and shall report annually on the extent and nature of their participation in their budget submissions to the Committee on Appropriations of the House of Representatives and the Committee on Appropriations of the Senate.

(b) As used in this section-

(1) the term "offices and agencies of the legislative branch" means, the office of the Clerk of the House, the office of the Secretary of the Senate, the office of the Architect of the Capitol, the Government Accountability Office, the Government Publishing Office, the Library of Congress, the Congressional Research Service, the Congressional Budget Office, the Chief Administrative Officer of the House of Representatives, and the Sergeant at Arms of the Senate; and

(2) the term "technology" refers to any form of computer hardware and software; computer-based systems, services, and support for the creation, processing, exchange, and delivery of information; and telecommunications systems, and the associated hardware and software, that provide for voice, data, or image communication.

( Pub. L. 104–197, title III, §314, Sept. 16, 1996, 110 Stat. 2415 ; Pub. L. 108–271, §8(b), July 7, 2004, 118 Stat. 814 ; Pub. L. 113–235, div. H, title I, §1301(b), Dec. 16, 2014, 128 Stat. 2537 .)

# History of XML in U.S. Congress

**2004:** House and Senate Bills available in XML: The Library of Congress began publicly posting House and Senate bills in XML format. It was agreed that the same style sheet used for the web display should be used by all Legislative Branch web sites displaying XML bill format.

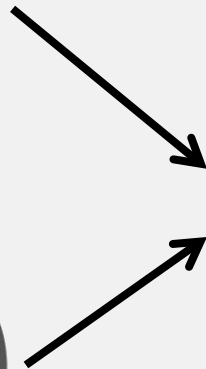
**2003:** House votes available in XML: During the 108th Congress, House Roll Call votes were made available in XML on the House Clerk's website. Votes were available online, but in HTML format, since 1997.

**2001:** During the 107th Congress, the House Office of Legislative Counsel began to use an XML editor to draft legislation.

**2000:** The House Clerk presented the SGML/XML Feasibility Study Final Report to the House Committee on Administration.

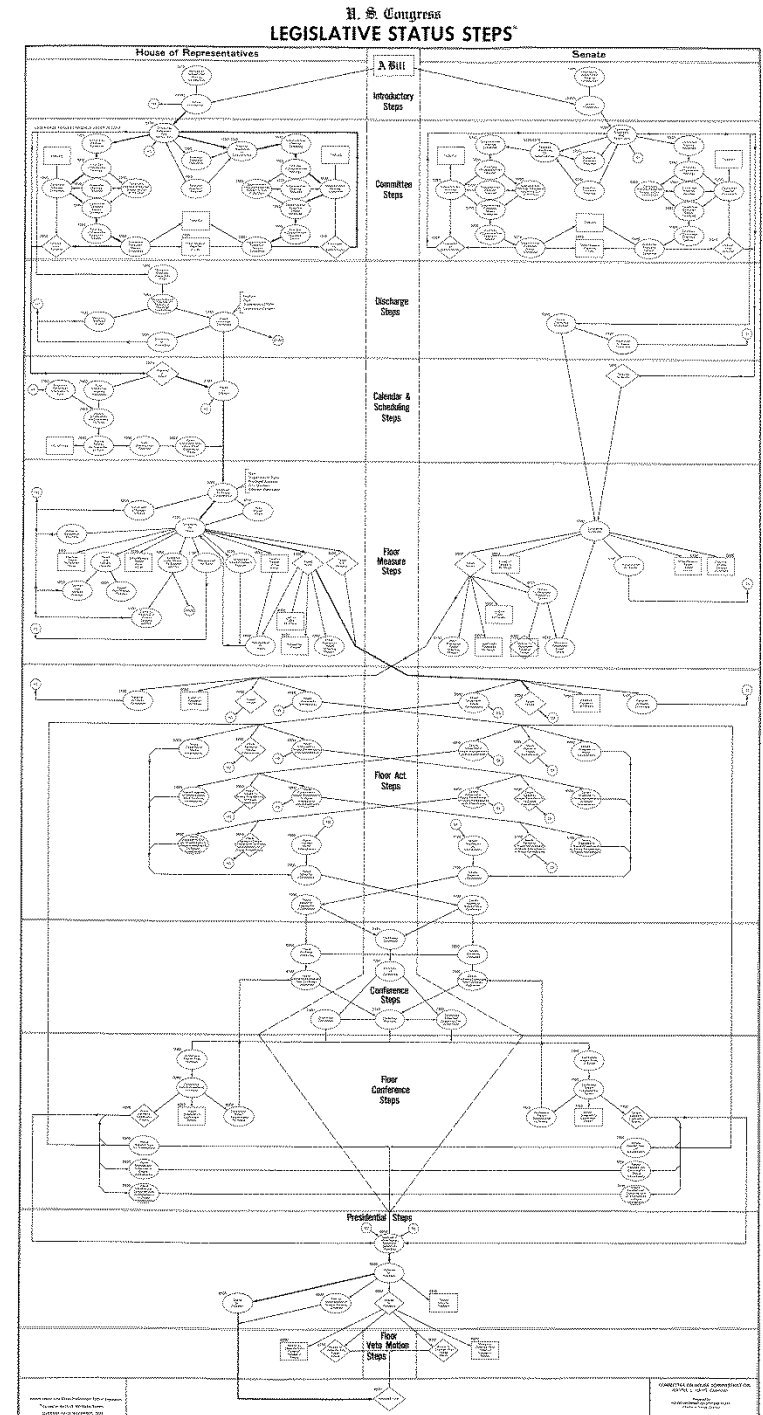
<http://xml.house.gov/resources/TechTimeline.htm>

# Simple Workflow



The shortest  
success path for a  
bill to become law  
is twenty decision  
points

<http://www.loc.gov/pictures/resource/ppmsca.33996/>



# Finding Legislative Information

## United States House

DOCS.HOUSE.GOV

*Text of legislation being considered now*

HOUSE.GOV

*House information*

CONGRESS.GOV

*House and Senate data and documents*

GPO.GOV

*all government documents*



## A Selected List of Legislative Branch Data

This document was compiled by staff in the Legislative Computer Systems division of the House Office of the Clerk for the 2015 Legislative Data and Transparency Conference (4/29/2015). Staff at the Library of Congress, the Government Publishing Office and the Senate Office of the Secretary assisted in the preparation of this document.

Additional government publications and publications related to Congress can be found at GPO.gov. The list of collections can be found: <http://www.gpo.gov/fdsys/browse/collectiontab.action>

Legislative Branch-wide DTDs, Schemas and other resources can be found at <http://xml.house.gov>.

## U.S. House and U.S. Senate

What	Location(s)		Format(s)
Chamber activity	House and Senate Calendars	<a href="http://www.gpo.gov/fdsys/browse/collection.action?collectionCode=CCAL">http://www.gpo.gov/fdsys/browse/collection.action?collectionCode=CCAL</a>  <i>Links available on Congress.gov home page.</i>	PDF, TXT
	Congressional Record	<a href="https://www.congress.gov/congressional-record">https://www.congress.gov/congressional-record</a> (searchable)  <a href="https://www.congress.gov/advanced-search/command-line#toggler-congressional-record">https://www.congress.gov/advanced-search/command-line#toggler-congressional-record</a> (describes searchable data fields)  Daily: <a href="http://www.gpo.gov/fdsys/browse">http://www.gpo.gov/fdsys/browse</a>	PDF, TXT

# Drafting Legislation with an XML editor

**“The less the people know  
about how sausages and  
laws are made, the better  
they sleep in the night.”**

Unknown phrase,  
often attributed to  
Otto von Bismarck

# Who uses the XML editor?

## House Office of Legislative Counsel

- Main drafters of House legislation, est. in 1919;
- 35 attorneys and 15 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;

## Government Printing Office

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



# Original 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.

2 RK\ROTA\HR244.ATH P-11-30



←

•T45IN THE HOUSE OF REPRESENTATIVES•



100

[Bill Stage: Pre-Introduction]  
{author-note}

[Discussion Draft]

H.L.C.  
{reintroduction-code}

(Original Signature of Member)

[DISCUSSION DRAFT]  
{first-page-date}

I

110th CONGRESS  
2d Session

H. R. \_\_\_\_

IN THE HOUSE OF REPRESENTATIVES

{action-date}

M\_\_\_\_ introduced the following bill; which was referred to the Committee on \_\_\_\_\_

**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 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”.

# 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.

---

# Lessons Learned: The Editor

---

# Original Lessons Learned

## (from 2005 to 2008)

- 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.

# Smart Authoring For the Future

- 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.
- Expand “just-in-time” knowledge support during the drafting process
- Provide an editor that does both things – keeps the editor in a valid document, but provides flexibility to draft “free-form”
- Easy methods to insert added value tags and maintain unique identifiers

---

# **Lessons Learned: The DTDs**

---



---

## **The U.S. Code**

the general and permanent laws of  
the United States, organized into  
titles based on subject matter

---

# Lesson Learned:

## Model the data as it appears

1 *Be it enacted by the Senate and House of Representa-*  
2 *tives of the United States of America in Congress assembled,*  
3 **SECTION 1. SHORT TITLE.**

4 This Act may be cited as the “National Park Service  
5 100th Anniversary Commemorative Coin Act”.

6 **SEC. 2. FINDINGS.**

7 The Congress

```
1 <section id="H354CA4636A1B41A693F90E736E818B0A"  
  section-type="section-one"><enum>1.</enum><header>Short  
  Title</header><text display-inline="yes-display-inline">  
  This Act may be cited as the <quote>National Park  
  Service 100th Anniversary Commemorative Coin Act</quote>  
  . </text></section>  
2 <section id="H5181530F2C244C558618C5BD142456CD"><enum>2.  
  </enum><header>Findings</header><text display-inline=  
  "yes-display-inline">The Congress finds the following:  
  </text> </section>
```

# Not this...

1 *Be it enacted by the Senate and House of Representa-*  
2 *tives of the United States of America in Congress assembled,*

3 SHORT TITLE

4 SECTION 1. This Act may be cited as the “National  
5 Park Service 100th Anniversary Commemorative Coin  
6 Act”.

7 FINDINGS

8 SEC. 2. The Congress finds the following:

```
1 <section id="H354CA4636A1B41A693F90E736E818B0A"  
  section-type="section-one"><enum>1.</enum><header>Short  
  Title</header><text display-inline="yes-display-inline">  
  This Act may be cited as the <quote>National Park  
  Service 100th Anniversary Commemorative Coin Act</quote>  
  . </text></section>  
2 <section id="H5181530F2C244C558618C5BD142456CD"><enum>2.  
  </enum><header>Findings</header><text display-inline=  
  "yes-display-inline">The Congress finds the following:  
  </text> </section>
```

# Lesson Learned:

## Avoid generated text

3 SECTION 1. FINDINGS AND DEFINITIONS.

4 (a) FINDING.—Congress finds it in the public interest  
5 to authorize the sale of certain federally owned land in

```
<section id="H171EBBB26FB14E62A44B0B19946A7C98"  
  section-type="section-one"><enum>1.</enum><header>  
  Findings and definitions.</header>  
<subsection id="HC3CC357C072B42A0B016AF9F81724346"  
  ><enum>(a)</enum><header>Finding.—</header><text  
  display-inline="yes-display-inline">Congress finds it  
  in the public interest to authorize the sale of  
  certain federally owned land in
```

# Not this...

## 3 SECTION 1. FINDINGS AND DEFINITIONS.

4 (a) FINDING.—Congress finds it in the public interest  
5 to authorize the sale of certain federally owned land in

```
<section id="H171EBBB26FB14E62A44B0B19946A7C98"  
section-type="section-one"><enum>1.</enum><header>  
Findings and definitions</header>  
<subsection id="HC3CC357C072B42A0B016AF9F81724346"  
><enum>(a)</enum><header>Finding</header><text  
display-inline="yes-display-inline">Congress finds it  
in the public interest to authorize the sale of  
certain federally owned land in
```

# Lesson Learned:

No attribute text in online or printed presentation

## SECTION 1. FINDINGS AND DEFINITIONS.

(a) FINDING.—Congress finds it in the public interest to authorize the sale of certain federally owned land in

```
<section id="H171EBBB26FB14E62A44B0B19946A7C98"
section-type="section-one"><enum>1.</enum><header>
Findings and definitions.</header>
<subsection id="HC3CC357C072B42A0B016AF9F81724346"
><enum>(a)</enum><header>Finding.—</header><text
display-inline="yes-display-inline">Congress finds it
in the public interest to authorize the sale of
certain federally owned land in
```

# Not this...

## SECTION 1. FINDINGS AND DEFINITIONS.

(a) FINDING.—Congress finds it in the public interest to authorize the sale of certain federally owned land in

```
<section id="H171EBBB26FB14E62A44B0B19946A7C98"
  section-type="section-one" enum="1"><header>Findings
and definitions.</header>
<subsection id="HC3CC357C072B42A0B016AF9F81724346"
  enum="(a)"><header>Finding.—</header><text
display-inline="yes-display-inline">Congress finds it
in the public interest to authorize the sale of
certain federally owned land in
```



# Lesson Learned:

## Reuse elements and data models from other standards

Akoma  
Ntoso

XHTML

```
<paragraph style="-uslm-lc:I11" class=
"indent0" id=
"id1c9b745f-76d8-11e3-b262-e79c7393386e"
identifier="/us/usc/t28/s332/e/10">
<num value="10">(10)</num><content>
Preparing an annual report to the
circuit and to the Administrative
Office of the United States Courts for
the preceding calendar year, including
recommendations for more expeditious
disposition of the business of the
circuit.
<p style="-uslm-lc:I11" class="indent0">
All duties delegated to the circuit
executive shall be subject to the
general supervision of the chief judge
of the circuit.</p>
</content>
```



# **Best Practice:**

## User Manual

**<http://uscode.house.gov>**

Thank you  
Merci  
Gracias

**Kirsten Gullickson**

Office of the Clerk  
U.S. House of Representatives

Kirsten.Gullickson@mail.house.gov  
@GullicksonK