
XML in the U.S. House of Representatives

Kirsten Gullickson

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

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.

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”.

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. JAMES M. & THOMAS W.L. ASHLEY CUSTOMS**
4 **BUILDING AND UNITED STATES COURT-**
5 **HOUSE**

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

1 W.L. Ashley Customs Building and United States Court-
2 house”.

3 (b) REFERENCES.—Any reference in a law, map, reg-
4 ulation, document, paper, or other record of the United
5 States to the Federal building and United States court-
6 house referred to in subsection (a) shall be deemed to be
7 a reference to the “James M. & Thomas W.L. Ashley Cus-
8 toms Building and United States Courthouse”.



The paper
version is the
official
document
of record.

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

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)

§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.

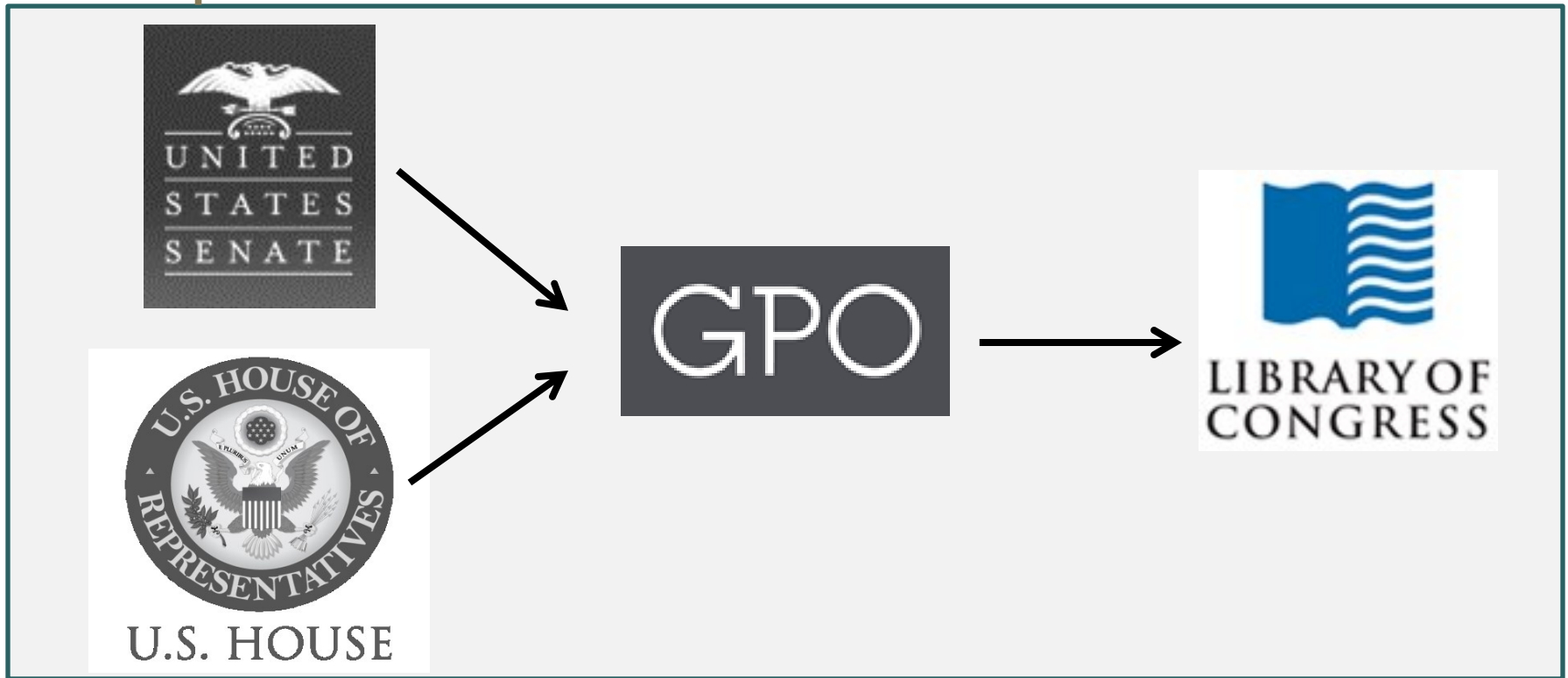
2003: House votes available in XML: During the 108th Congress, House Roll Call votes were made available in XML on the House Clerk's website. Votes were available online, but in HTML format, since 1997.

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.

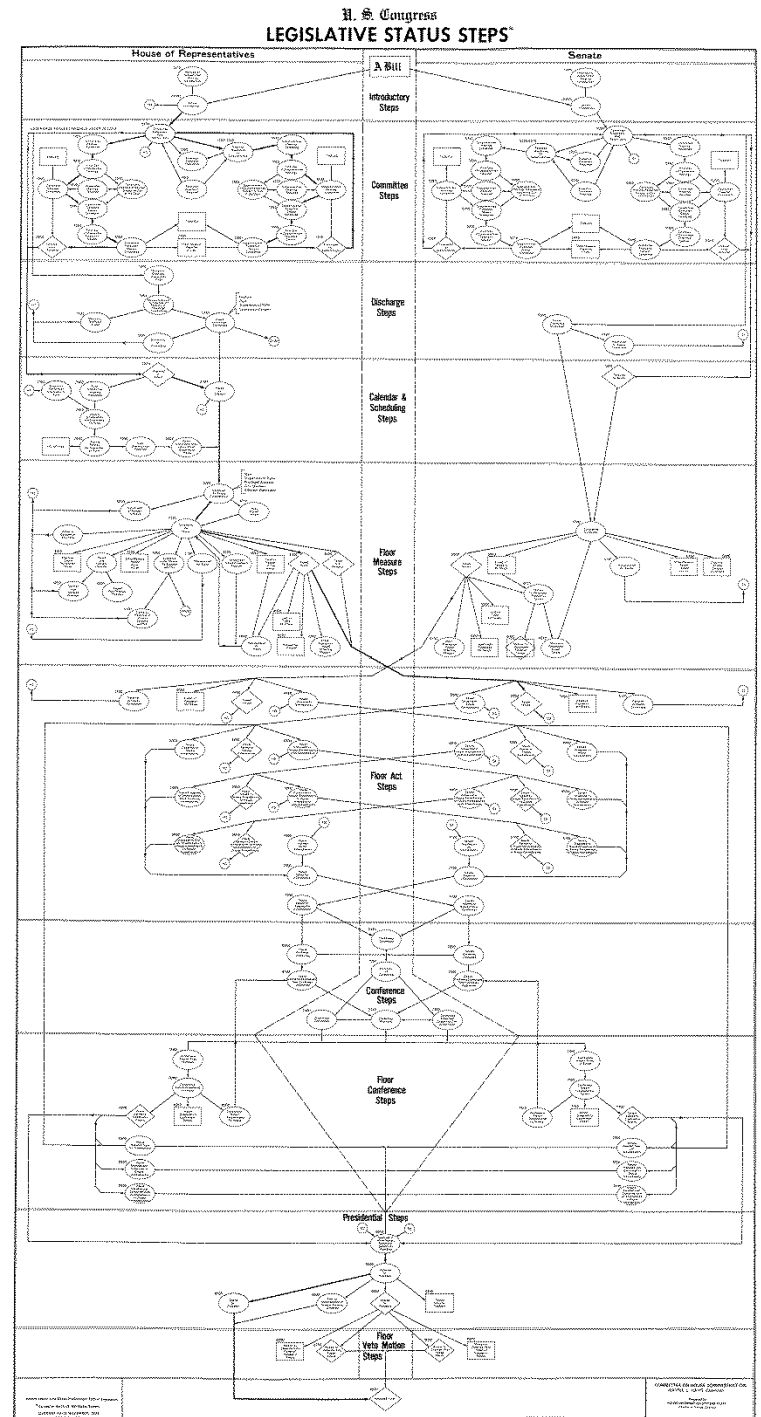
<http://xml.house.gov/resources/TechTimeline.htm>

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](https://docs.house.gov)

Text of legislation being considered now

[HOUSE.GOV](https://www.house.gov)

House information

[CONGRESS.GOV](https://www.congress.gov)

House and Senate data and documents

[GPO.GOV](https://www.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 <i>Links available on Congress.gov home page.</i>	PDF, TXT
	Congressional Record https://www.congress.gov/congressional-record (searchable) https://www.congress.gov/advanced-search/command-line#toggler-congressional-record (describes searchable data fields) Daily: http://www.gpo.gov/fdsys/browse	PDF, TXT

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

The diagram consists of two blue circles on the left. The top circle is split diagonally and contains the text 'Akoma Ntoso'. The bottom circle is also split diagonally and contains the text 'XHTML'. Two white arrows point from the right side of each circle towards the right, pointing to two separate blocks of XML code. The top arrow points to a block of XML code starting with a paragraph tag, and the bottom arrow points to a block of XML code starting with a paragraph tag.

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