



# Clerk of the House

Bob Reeves

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

Kirsten Gullickson

Principal Analyst, Legislative Computer Systems  
Office of the Clerk



CLERK



2019 HOUSE  
LEGISLATIVE DATA  
AND TRANSPARENCY  
CONFERENCE

# Legislative Data Transparency Conference

October 17, 2019

# The House Bulk Data Task Force

The Bulk Data Task Force is a partnership of representatives from the various Legislative Branch Agencies such as:

- House of Representatives,
- Senate,
- Library of Congress,
- Government Publishing Office,
- Government Accounting Office

that work in a cooperative manner to coordinate and accomplish openness and transparency goals.

# The House Bulk Data Task Force

- created in 2012 to focus on increased dissemination of congressional information via bulk data
- meets internally with Legislative Branch organizations to coordinate projects and other transparency initiatives
- meets with civil society organizations



**Public Galleries: available in April, 1789**



**PREPARE**



**MANAGE**



**DISTRIBUTE**





**ARCHIVE/PRESERVE**

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 COURT- HOUSE.

(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”.

The paper  
version is the  
official document  
of record.



**Today's House Chamber**



## Kiosks

<http://themedalofhonor.com/-/media/images/moh/exhibits/kiosk-1-new.ashx?h=500&w=640&la=en&hash=F428CEB5C31148825F4A60405DE5E4D42DAF9EED>



## **Digital displays**

<http://history.house.gov/mediacontent/Video/capitol/capitol-house-chamber-symbols.jpg>





[https://geobrava.files.wordpress.com/2015/12/8787-mobile\\_2d00\\_device\\_2d00\\_security.jpg](https://geobrava.files.wordpress.com/2015/12/8787-mobile_2d00_device_2d00_security.jpg)

# Opening the doors of Congress



---

## Today's Challenge:

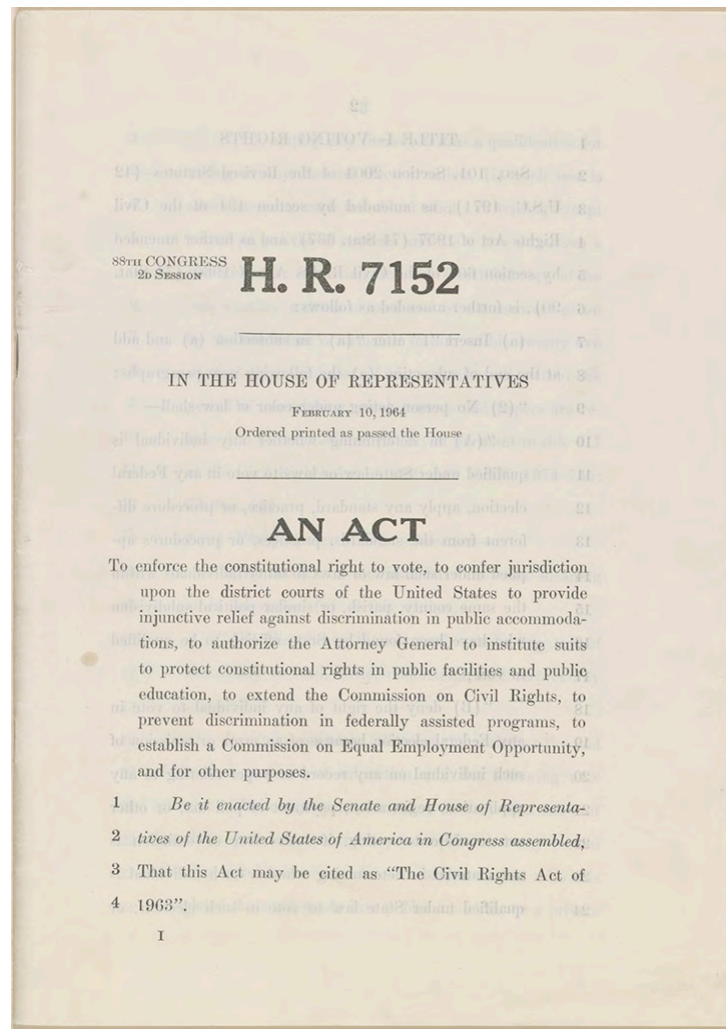
How do parliaments modernize  
their systems to utilize today's newer technologies to  
help citizens access  
free & timely  
legislative & legal information?

---



**Online means more than  
just the electronic  
representation of the  
paper document.**

**Public institutions need  
to deliver machine-  
readable, digital  
formats.**



# Hopper



# Clerk Seal





# World e-Parliament Report 2018

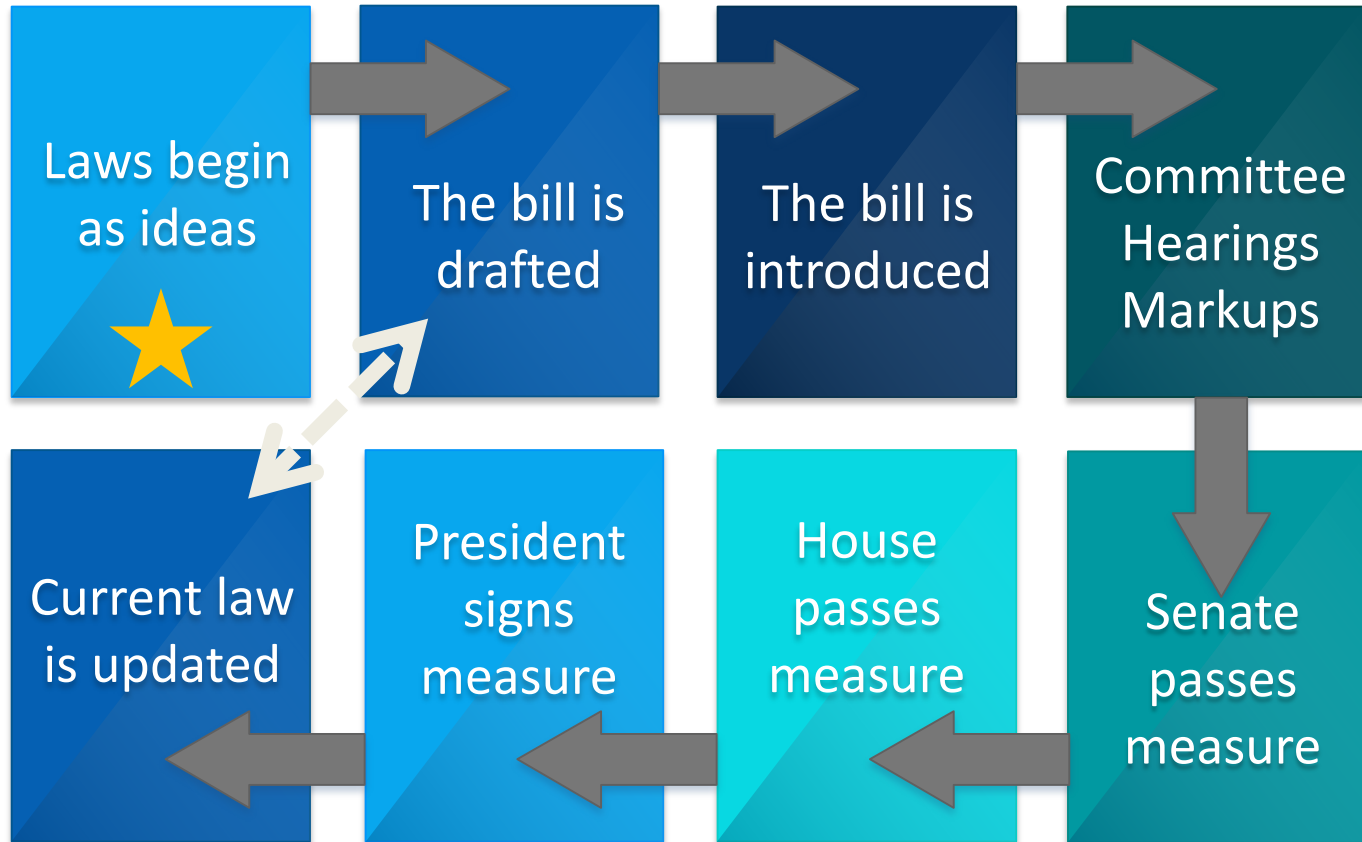
The key trends identified in the report:

- Digital technologies are now firmly embedded with clearly identified governance and technology practices in most parliaments.
- The use of instant messaging has increased significantly.
- Digital broadcasting and video streaming have overtaken traditional broadcasting.
- Barriers to greater use of ICT include training and skill deficits for staff and members.
- Knowledge of how parliaments work is seen as the biggest barrier to greater citizen engagement.
- Over a third of the parliaments now collaborate directly with PMOs.

<https://www.ipu.org/resources/publications/reports/2018-11/world-e-parliament-report-2018>



# How a Bill becomes the Law

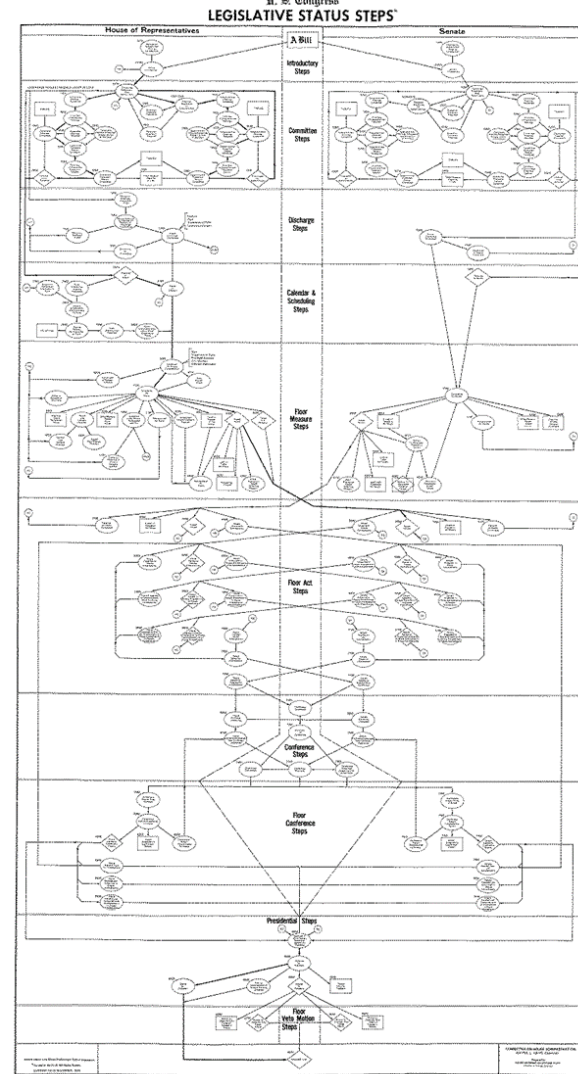


---

**The Legislative Process is complicated and rarely predictable.**

---

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





---

The Solution: Structured Documents  
Formats: XML, JSON, HTML

---

# Documents are not databases

## Database

- Predictable
- Relational (often)
- Homogeneous
- Scoped
- System provides meaning

## Documents

- Structured and un-structured
- Related (often) but not relational
- Heterogeneous
- Variable scope
- Intrinsic meaning

# Documents

- Documents have three characteristics:
  - Presentation: how it looks
  - Structure: how is it organized
  - Semantics: what does it mean

---

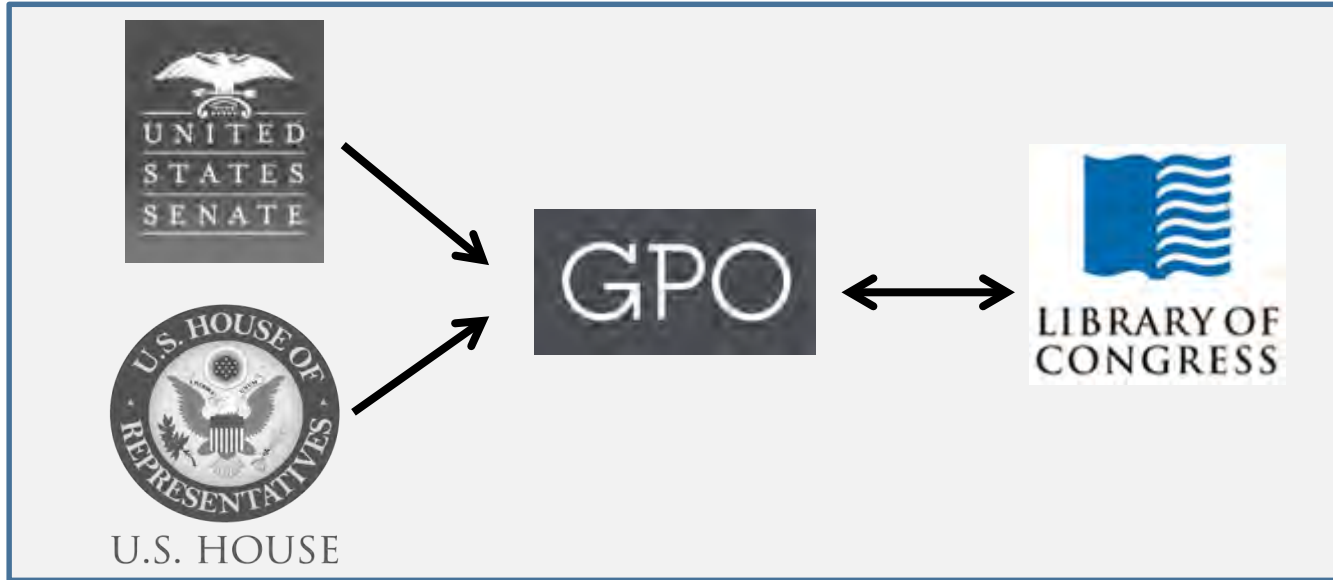
XML is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.

---

# XML – snippet from S. 100

```
8 <legis-num display="yes">S. 100</legis-num>
9 <current-chamber display="yes">IN THE SENATE OF THE UNITED
10 STATES</current-chamber>
11 <action display="yes">
12   <action-date date="20070104">January 4, 2007</action-date>
13   <action-desc><sponsor name-id="S223">Mrs. Boxer</sponsor> introduced
14     the following bill; which was read twice and referred to the
15     <committee-name committee-id="SSF100">Committee on
16       Finance</committee-name></action-desc>
17 </action>
18 <legis-type display="yes">A BILL</legis-type>
19 <official-title display="yes">To encourage the health of children in
20   schools by promoting better nutrition and increased physical activity, and for
21   other purposes.</official-title>
22 </form>
23 <legis-body display-enacting-clause="yes-display-enacting-clause" style="OLC">
24   <section commented="no" display-inline="no-display-inline" id="S1" section-type="section-one"><enum>1.</enum><header disp:
25     <quote><short-title>Healthy Students Act of
26       2007</short-title></quote>.</text>
27   </section><section commented="no" display-inline="no-display-inline" id="idE055AAD810E54A94B899EF7BC4BB22F2" section-type=
28     and breakfasts</header>
29   <subsection commented="no" display-inline="no-display-inline" id="id94A593417F1746548BE34A69949E0D7A"><enum>(a)</enum>
30     National School Lunch Act (42 U.S.C. 1758(a)) is amended by striking paragraph
31     (4) and inserting the following:</text>
```

# Simple Workflow



# Finding Legislative Information

## United States Congress

SENATE.GOV  
*Senate information*

CLERK.HOUSE.GOV & clerkpreview.house.gov  
*House information*

CONGRESS.GOV  
*House and Senate data and documents*

GOVINFO.GOV  
*all government documents*

---

[www.github.com/usgpo](https://www.github.com/usgpo)

[www.govinfo.gov/bulkdata](https://www.govinfo.gov/bulkdata)

---



---

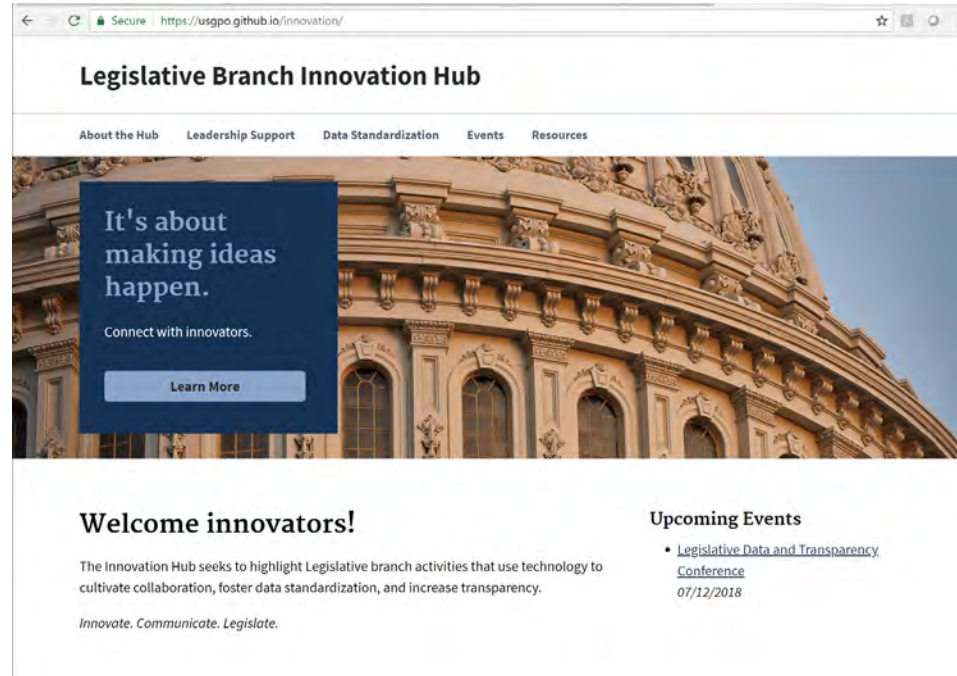
## Technology Timeline

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

---

# MORE INFORMATION

- [USGPO.GITHUB.IO/INNOVATION/](https://usgpo.github.io/innovation/)
  - Code is available on github; pull requests taken from all
  - Uses the United States Web Design System
    - [DESIGNSYSTEM.DIGITAL.GOV](https://designsystem.digital.gov)
    - Responsive design removes the needs for a mobile app
- [XML.HOUSE.GOV](https://xml.house.gov)
  - Legacy website, circa 1998
  - Manually updated infrequently using basic HTML coding



116<sup>th</sup> Congress

# **OFFICE OF THE CLERK UPDATES**

# Websites Updates

## **LIVE.HOUSE.GOV**

- Old service offline in July, 2019
- Streaming using Azure Media Services
- Consuming a Clerk API from XML sources
- Consuming GPO/LOC bill data
- Responsive Design

## **CLERKPREVIEW.HOUSE.GOV**

- Beta version
- Streaming using Azure Media Services
- Consuming a Clerk API from XML sources
- Consuming GPO/LOC bill data
- Responsive Design

## Other updates

- New version of BioGuide.Congress.gov
  - New CMS back-end
  - New public facing site by the end of the calendar year
- Standardized committee witness forms
  - Used by committees; posted on docs.house.gov
- Managing cosponsor information

## Posey Comparative Prints

# 115<sup>th</sup> Congress House Rules Change

- Subsection (s) of H. Res. 5, Adopting the rules of the 115<sup>th</sup> Congress added Clause 12 to House Rule XXI.
- Deadline of 12/31/2017 was met by
  - Creating a new website for doc-to-doc comparisons
  - Augmenting the existing HOLC Ramseyer tool for changes in existing law

(s) COMPARATIVE PRINTS FOR BILLS OR JOINT RESOLUTIONS CONSIDERED ON FLOOR.—Effective December 31, 2017, in rule XXI, add at the end the following new clause:

“12.(a)(1) Before a bill or joint resolution proposing to repeal or amend a statute or part thereof may be considered, there shall be made available on a publicly available website of the House an easily searchable electronic comparative print that shows how the bill or joint resolution proposes to change current law, showing (to the greatest extent practicable) by appropriate typographical devices the omissions and insertions proposed.

“(2) Before an amendment in the nature of a substitute may be considered if the amendment proposes to repeal or amend a statute or part thereof, there shall be made available on a publicly available website of the House an easily searchable electronic comparative print that shows (to the greatest extent practicable) how the amendment proposes to change current law, showing by appropriate typographical devices the omissions and insertions proposed.

“(b) If a committee reports a bill or joint resolution, before the bill or joint resolution may be considered with text different from the text reported, there shall be made available on a publicly available website of the House a document that shows, by appropriate typographical devices, the differences between the text of the bill or joint resolution as proposed to be considered and the text of the bill or joint resolution as reported.”.

# document to document comparison

- Clause 12(b) calls for a document-to-document comparison between different versions of bill language.
- Built tool – [billcompare.house.gov](https://billcompare.house.gov).
- Currently maintaining and updating tool

COMPARATIVE PRINT

Comparing the base document H.R. 2203 as reported, with the Rules Committee Print 116-27

This document was computer-generated in an effort to show how two versions of the same legislative text differ **and may not have been reviewed for accuracy**. This document may comply with clause 12(b) of rule XXI of the Rules of the House. This document does not represent an official expression by the House and should not be relied on as an authoritative delineation of the proposed change(s) to existing law.

Omitted text is shown ~~stricken~~, new matter that is proposed is in underlined italics, and existing text in which no change is being proposed is shown in regular roman. Typesetting and stylistic characteristics, particularly in the headings and indentations, may not conform to how the text, if adopted, would be illustrated in subsequent versions of legislation or public law.

---

## Sec. 1. Short title; table of contents

(a) SHORT TITLE.— This Act may be cited as the “Homeland Security Improvement Act”.

(b) TABLE OF CONTENTS.—

The table of contents for this Act is as follows:

Sec. 1. Short title; table of contents.

### TITLE I—NATIONAL COMMISSION TO INVESTIGATE THE TREATMENT OF THE MIGRANT FAMILIES AND CHILDREN BY THE TRUMP ADMINISTRATION

Sec. 101. Establishment of commission.

Sec. 102. Purposes.

Sec. 103. Composition of the Commission; qualifications.

Sec. 104. Functions of commission.

Sec. 105. Powers of commission.

Sec. 106. Public meetings and release of public version of reports.

Sec. 107. Staff of commission.

Sec. 108. Compensation and travel expenses.

Sec. 109. Security clearances for commission members and staff.

Sec. 110. Reports of commission; termination.

Sec. 111. Funding.

### TITLE II—OMBUDSMAN, BORDER COMMUNITY LIAISONS, TRAINING AND MANAGEMENT

Sec. 201. Establishment of the Office of the Ombudsman for Border and Immigration Related Concerns.

Sec. 202. Establishment of border community stakeholder advisory committee.

Sec. 203. Training and continuing education.

Sec. 204. Body-worn cameras.

Sec. 205. Limitation on the separation of families.

Sec. 206. Prohibition on exceeding 72 hours in short-term detention.

Sec. 207. Electronic tracking.

Sec. 208. ~~Preventing implementation of anti-asylum access interim final rule~~; Termination of ~~§ 6(a)(1) in Mexico~~ protocol and metering policy.

Sec. 209. Surge support for care of families and unaccompanied children.

Sec. 210. Port of entry temporary duty assignments.

Sec. 211. Prohibition on physical restraints for women who are pregnant or post-delivery.

### TITLE III—OVERSIGHT

Sec. 301. GAO report on the extent of CBP activities, operations, and claimed authority.

Sec. 302. Migrant deaths reporting.

Sec. 303. Review of the use of force.

Sec. 304. Accountability and transparency within the border zone.



# changes to existing law

- Clause 12(a) calls for a comparative print that illustrates changes made by the proposal to current law.
- augmented existing tool in HOLC to meet early deadline
- building a more modern tool now.

## COMPARATIVE PRINT

In compliance with clause 12(a) of rule XXI of the Rules of the House of Representatives, this document was generated to show how legislative text that may be considered by the House proposes to change existing law. It does not represent an official expression by the House and it should not be relied on in lieu of the legislative proposal to understand the proposed change(s).

Omitted text is enclosed in black brackets, new matter is proposed in italics, and existing law in which no change is proposed is shown in roman. Typesetting and stylistic characteristics, particularly in the headings and indentations, may not conform to how the proposed text, if adopted, would be executed in current law. This comparative print may not illustrate changes to tables of contents if the legislative text is proposing such a change(s).

## CONTINUING APPROPRIATIONS ACT, 2018

### DIVISION D—CONTINUING APPROPRIATIONS ACT, 2018

SEC. 106. Unless otherwise provided for in this Act or in the applicable appropriations Act for fiscal year 2018, appropriations and funds made available and authority granted pursuant to this Act shall be available until whichever of the following first occurs:

- (1) the enactment into law of an appropriation for any project or activity provided for in this Act;
- (2) the enactment into law of the applicable appropriations Act for fiscal year 2018 without any provision for such project or activity; or
- (3) **January 19, 2018** *February 16, 2018.*

SEC. 148. *Funds appropriated by the Department of Defense Missile Defeat and Defense Enhancements Appropriations Act, 2018 (division B of Public Law 115-96) may be obligated and expended notwithstanding section 504(a)(1) of the National Security Act of 1947 (50 U.S.C. 3094(a)(1)).*

SEC. 149. *Amounts made available by section 101 for "Department of Agriculture—Food and Nutrition Service—Child Nutrition Programs" to carry out section 749(g) of the Agriculture Appropriations Act of 2010 (Public Law 111-80) may be apportioned up to the rate for operations necessary to ensure that the program can be fully operational by May 2018.*

SEC. 150. *Amounts made available by section 101 for "National Aeronautics and Space Administration—Exploration" may be apportioned up to the rate for operations necessary to maintain the planned launch capability schedules for the Space Launch System launch vehicle, Exploration Ground Systems, and Orion Multi-Purpose Crew Vehicle programs.*

SEC. 151. *Amounts made available by section 101 for "Department of Energy—Energy Programs—Office of the Inspector General" may be apportioned up to the rate for operations necessary to sustain staffing levels achieved on June 30, 2017.*

SEC. 152. *Amounts made available by section 101 for "Small Business Administration—Business Loans Program Account" may be apportioned up to the rate for operations necessary to accommodate increased demand for commitments for general business loans authorized under section 7(a)*

# Three types of comparisons



Document to document



Amendments contained in a  
legislative proposal to current law

Changes in existing law  
(Ramseyer/Cordon)



Amendments to legislative  
proposal

Committee amendments  
Floor amendments

# Current Project Goal

- **Deploy a** robust, scalable, and secure **tool so all House staff** (and others) **can create on-demand point-in-time comparative prints** that illustrate the changes between:
  - Two versions of a bill, resolution, or amendment (document to document comparison)
  - Current law and current law as proposed to be changed by amendments contained in a bill, resolution, or amendment to current law (positive and non-positive law)
  - A bill or resolution and the bill or resolution as proposed to be modified by amendments (amendment impact)

# Current work – clause 12(a)

Changes to existing law

# What do the Natural Language Processors Need to Do?

- Recognize
- Interpret
- Retrieve
- Execute

## Section 704 of the Homeland Security Act of 2002

(6 U.S.C. 344) is amended—

(1) in subsection (b)—

(A) in paragraph (1)—

(i) by inserting “, including with respect to leader development and employee engagement,” after “policies”;

(ii) by striking “and in line” and inserting “, in line”; and

(iii) by inserting “and informed by best practices within the Federal government and the private sector,” after “priorities,”;

## Section 2(a)(6) of H.R.6901 RH of the 115<sup>th</sup> Congress

(6) E-GOVERNMENT REPORT.—Section 3606 of title 44, United States Code, is amended—

(A) in the heading, by striking “E-Government” and inserting “Annual”; and

(B) in subsection (a), by striking “an E-Government status report to the Committee on Governmental Affairs of the Senate and the Committee on Government Reform of the House of Representatives” and inserting “a report to the Committee on Homeland Security and Governmental Affairs of the Senate and the Committee on Oversight and Government Reform of the House of Representatives”.

## Section 3606 of chapter 36 of title 44, United State Code, showing proposed amendments

### § 3606. ~~E-Government~~ Annual report

(a) Not later than March 1 of each year, the Director shall submit ~~an E-Government status report to the Committee on Governmental Affairs of the Senate and the Committee on Government Reform of the House of Representatives~~ a report to the Committee on Homeland Security and Governmental Affairs of the Senate and the Committee on Oversight and Government Reform of the House of Representatives.

(b) The report under subsection (a) shall contain—

- (1) a summary of the information reported by agencies under section 202(f) of the E-Government Act of 2002;
- (2) the information required to be reported by section 3604(f); and
- (3) a description of compliance by the Federal Government with other goals and provisions of the E-Government Act of 2002.

## Welcome

Use this tool to create on-demand static comparative prints that illustrate changes between:

- Two versions of a bill, resolution, or amendment
- Current law and current law as proposed to be changed by amendments contained in a bill, resolution or amendment
- A bill or resolution and the bill or resolution as proposed to be modified by amendments

[Learn More](#)


### Please Sign In

Log in using your Active Directory username and password.

☐ Remember Me

Login

# New Report

 What would you like to view?

How Bill Changes Current Law

 Document 1

Reset

Congress:\*

116 (current)

Search By:

Bill Type:\*

House Bill (H.R.)

Bill Number:\*

####

Bill Stage:\*

Select...

Upload 

Cancel

Submit



## Sec. 2. Changes relating to electronic Government services

(a) CHANGE OF CERTAIN NAMES IN CHAPTER 36 OF TITLE 44.—

(6) E-GOVERNMENT REPORT.— Section 3606 of title 44, United States Code, is amended—

(A) in the heading, by striking “E-Government” and inserting “Annual” ; and

(A) assess the feasibility, scientific and technical merits, and estimated cost of each proposal, and

(B) submit each proposal, and the assessment of the proposal, to ~~the Administrator~~ the Federal Chief Information Officer.

(3) The technical assistance team shall not consider or evaluate proposals submitted in response to a solicitation for offers for a pending procurement or for a specific agency requirement.

(4) After receiving proposals and assessments from the technical assistance team, ~~the Administrator~~ the Federal Chief Information Officer shall consider recommending appropriate proposals for funding under the ~~E-Government Fund~~ Federal IT Fund established under section 3604 or, if appropriate, forward the proposal and the assessment of it to the executive agency whose mission most coincides with the subject matter of the proposal.

### § 3606. ~~E-Government~~ Annual report

(a) Not later than March 1 of each year, the Director shall submit ~~an E-Government status report to the Committee on Governmental Affairs of the Senate and the Committee on Government Reform of the House of Representatives~~ a report to the Committee on Homeland Security and Governmental Affairs of the Senate and the Committee on Oversight and Government Reform of the House of Representatives.

(b) The report under subsection (a) shall contain—

# Challenges: Finding provisions in Federal Law

No single unified code

Draft to several sources

- U.S. Code
- Statutes at Large
- Statute Compilations
- Other law (Revised Statutes of the United States, Uniform Code of Federal Evidence, Tariff Tables,...)

No single effective date; conditional effective dates

# Challenges: Interpreting Amendatory Instructions

strike “Chief” and insert “director”.

At the end of Title 15, add this new section...

Strike subsection (e) and insert a new (e) and (f), redesignate (f), (g), (h) as (g), (h), and (i).

modify subpart S of part 76 of the Commission’s rules by eliminating any requirements relating to network nonduplication and syndicated exclusivity for open video systems.

in paragraph (9), as so redesignated, by striking “325(b)(7)” and inserting “325”; and

in subsection (j), as so redesignated—

- in paragraph (3), as so redesignated, by striking “122(j)” and inserting “122(g)”;

# What would make building the 12(a) tool easier

- Data modernization
- Positive Law Codification
- Improvements in the Statute Compilations

---

Roadmap for Publishing Legislative Documents in USLM XML Format	
Completed	U.S. Code
Completed	Enrolled bills, Public Laws, and Statutes at Large
Current Stage	HOLC/SOLC statute compilations (non-positive law)
B	Remaining bill/resolution versions (introduced, reported, engrossed, all amendments, et. al.)
C	Committee Reports
D	House portions of the Congressional Record, including the Daily Digest
E	Hearing Record
F	House Calendar
G	Precedents of the U.S. House of Representatives
H	House Journal

---

## More information

Clerk testimony before Legislative Branch subcommittee (Committee on Appropriations):

*March 12, 2019*

<https://docs.house.gov/Committee/Calendar/ByEvent.aspx?EventID=109069>

Clerk testimony before Committee on House Administration

*April 9, 2019*

<https://docs.house.gov/Committee/Calendar/ByEvent.aspx?EventID=109268>

Deputy Clerk testimony before Committee on House Modernization

*May 10, 2019*

<https://docs.house.gov/Committee/Calendar/ByEvent.aspx?EventID=109468>

---

Demonstration of tool  
during the  
Vision of the Future session

---

# Thank you

## **Kirsten Gullickson**

Office of the Clerk  
U.S. House of Representatives

[Kirsten.Gullickson@mail.house.gov](mailto:Kirsten.Gullickson@mail.house.gov)

For more information:  
[XML.HOUSE.GOV](https://XML.HOUSE.GOV)  
[USGPO.GITHUB.IO/INNOVATION/](https://USGPO.GITHUB.IO/INNOVATION/)