

# The Bulk Data Task Force

Public Meeting  
Wednesday, July 14, 2021

# AGENDA

- 1. Welcome**
- 2. Introductions**
3. Reports/Updates from LOC, GPO, Clerk, Secretary
4. Reports/Updates from civil society organizations
5. Time for Questions and Answers
6. Announcements/Closing

*Meeting ends at 12:00 pm EST*

# The Bulk Data Task Force

The Bulk Data Task Force is a partnership of representatives from various Legislative Branch Agencies including the—

- 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 Bulk Data Task Force

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

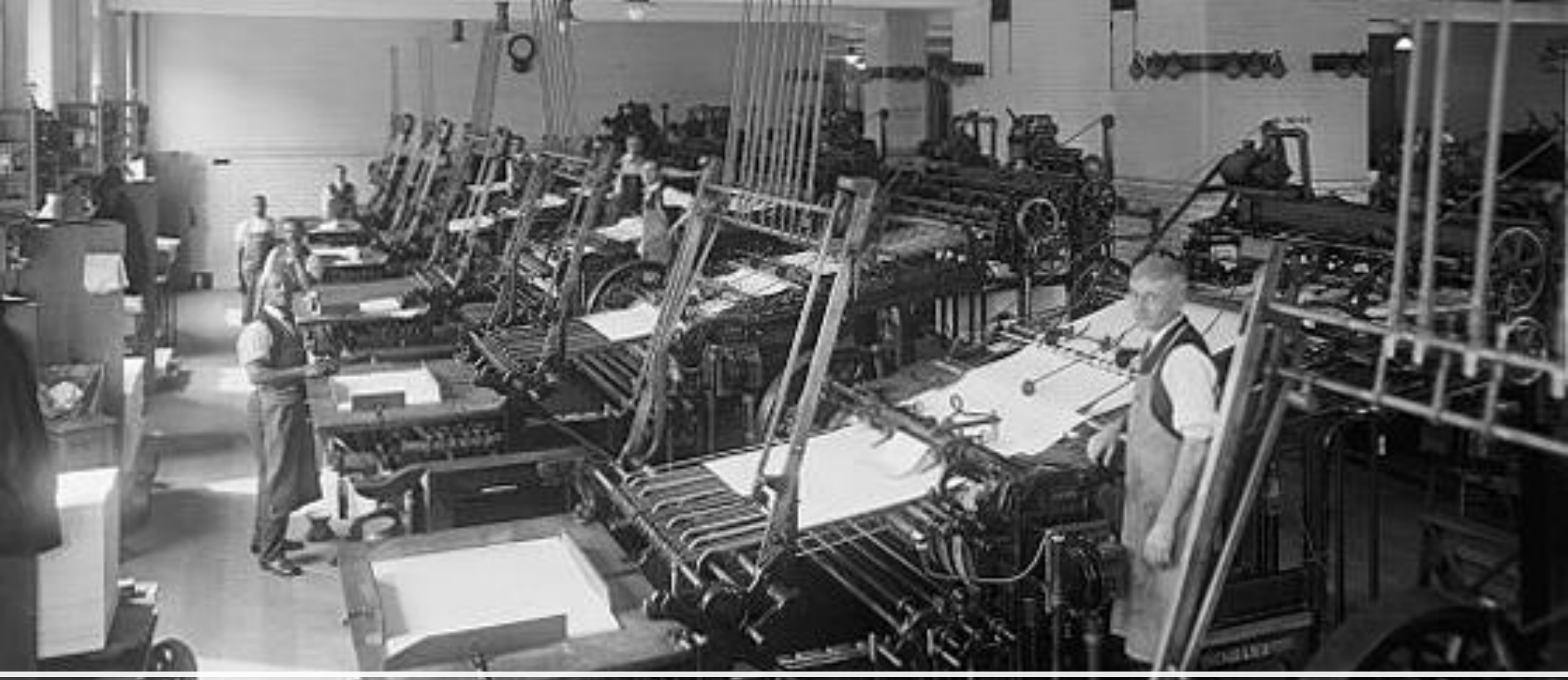
# The Bulk Data Task Force

Recommendation by the Select Committee on the Modernization of Congress to—

- Make permanent the Bulk Data Task Force and rename it the Congressional Data Task Force.

\*Page 35, item #47 of the final report for the 116<sup>th</sup> Congress

<https://www.govinfo.gov/content/pkg/GPO-CRPT-116hrpt562/pdf/GPO-CRPT-116hrpt562.pdf>



**PREPARE**







**MANAGE**





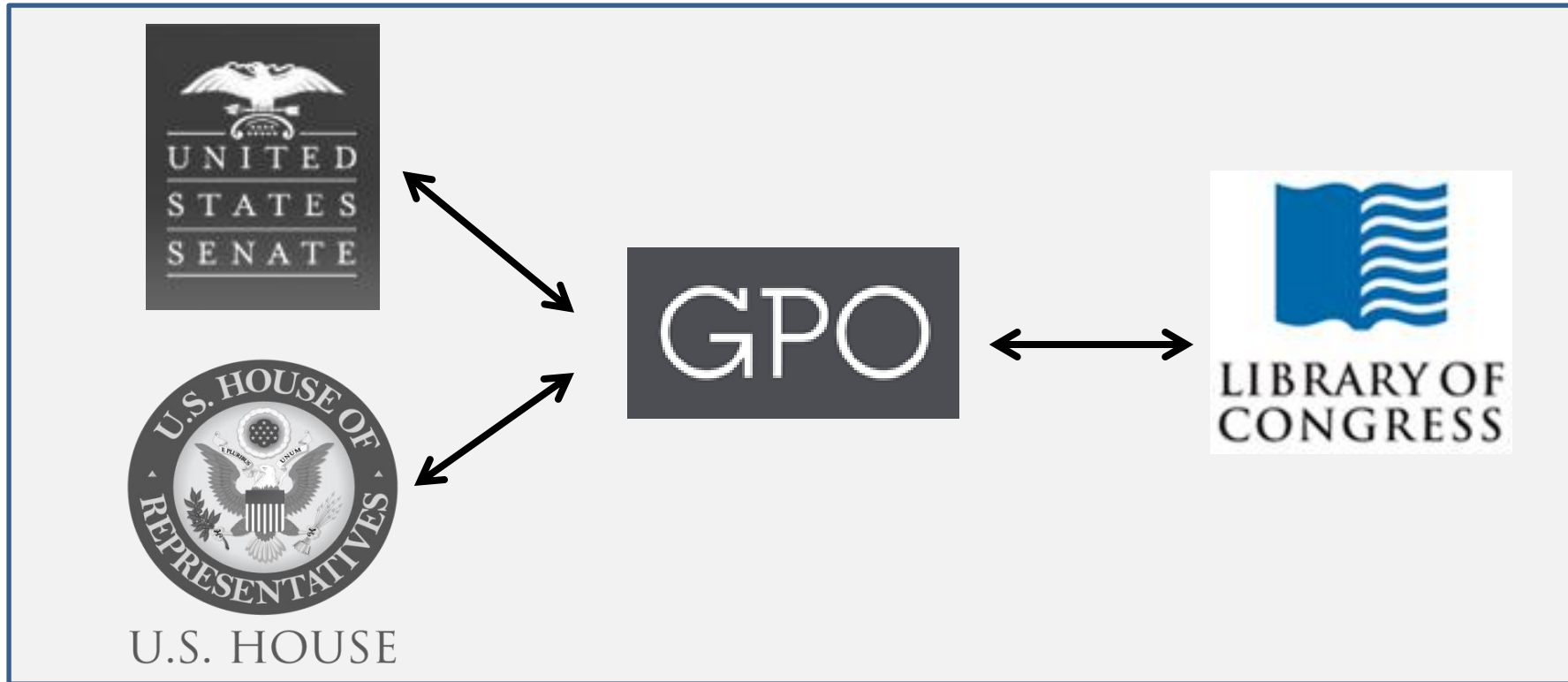
**DISTRIBUTE**





**ARCHIVE/PRESERVE**

# Data Exchange: Simple



More comprehensive workflows can be found on the Innovation Hub:  
<https://usgpo.github.io/innovation/resources/>

Updates from

# **THE OFFICE OF THE CLERK**

# Response to Public Health Emergency

## Electronic submission of legislative documents

- Bills and resolutions (eHopper)
- Committee reports
- Other material

## Proxy voting

- <https://clerk.house.gov/ProxyLetter>
- To date, over 1400 letters are posted



# New item due to House Rules change

CONGRESS.GOV Advanced Searches Browse Search Tools Support

Legislation Examples: hr

Home > Legislation > 117th Congress > S.578

## S.578 - FASTER Act of 2021

117th Congress (2021-2022) | [Get alerts](#)

**LAW** Hide Overview ✕

**Sponsor:** [Sen. Scott, Tim \[R-SC\]](#) (Intro)

**Latest Action:** 04/23/2021 Became Public La

**Roll Call Votes:** There has been [1 roll call vote](#)

**Tracker:**

Introduced > Passed Senate > Passed Ho

### Supporting House Members for S.578 - FASTER Act of 2021 [117th]

The following Members of the House of Representatives have indicated their support for this Senate measure pursuant to section 3(x) of H. Res. 8:

- [Rep. Taylor, Van \[R-TX-3\]](#) - 03/16/2021

#### About Supporting House Members

On January 4, 2021, the House of Representatives issued orders regarding [Support For Senate Measures](#). See also [REGULATIONS FOR SUPPORTING SENATE MEASURES PURSUANT TO HOUSE RESOLUTION 8, 117TH CONGRESS](#) in the February 2, 2021 Congressional Record.

Citation Subscribe Share/Save S

**For Congressional Users**

[Search CRS.gov](#)

**More on This Bill**

- [CBO Cost Estimates \[0\]](#)
- [Supporting House Members \[1\]](#)

**Subject — Policy Area:**  
Health

# Deployment of new BioGuide site

The screenshot shows a web browser displaying the BioGuide website. The URL is <https://bioguide.congress.gov/search/bio/R000082>. The page features the logo of the Biographical Directory of the United States Congress and a search bar containing "rayburn, s". Navigation buttons for "Back" and "Download" are visible. The main content area is titled "RAYBURN, Samuel Taliaferro 1882 – 1961" and includes a portrait of the subject. A sidebar on the left provides key statistics: 25 US Congresses Served (63rd – 87th, 1913 – 1961), House Years of Service (1913 – 1961), State / Territory (Texas), 3 Positions (Majority Leader, Representative, Speaker Of The House), and Party (Democrat). The main text is a detailed biography of Samuel Taliaferro Rayburn, starting with "RAYBURN, Samuel Taliaferro, a Representative from Texas; born near Kingston, Roane County, Tenn., January 6, 1882; moved to Fannin County, Tex., in 1887 with his parents who settled near Windom; attended the rural schools and was graduated from the East Texas Normal College, Commerce, Tex., in 1903; studied law at the University of Texas at Austin; was admitted to the bar in 1908 and commenced practice in Bonham, Fannin County, Tex.; member of the State house of representatives 1907-1913, and served as speaker during the last two years; elected as a Democrat to the Sixty-third and to the twenty-four succeeding Congresses and served from March 4, 1913, until his death; chairman, Committee on Interstate and Foreign Commerce (Seventy-second, Seventy-third, and Seventy-fourth Congresses); majority leader (Seventy-fifth and Seventy-sixth Congresses), minority leader (Eightieth and Eighty-third Congresses); elected Speaker of the House of Representatives September 16, 1940, to fill the vacancy caused by the death of Speaker William B. Bankhead; reelected Speaker in the Seventy-seventh, Seventy-eighth, Seventy-ninth, Eighty-first, Eighty-second, Eighty-fourth, Eighty-fifth, Eighty-sixth, and Eighty-seventh Congresses; died in Bonham, Tex., November 16, 1961; interment in Willow Wild Cemetery." The footer contains contact information for the House and Senate Historians, the Office of Art & Archives, and the Office of the Clerk, along with a "Resources" section listing links to the Retro Biographical Directory, Privacy Policy, and Copyright Information.

Biographical Directory of the United States Congress

rayburn, s

RAYBURN, Samuel Taliaferro  
1882 – 1961

Biography Congresses Bibliography Research Collections

Collection of the U.S. House of Representatives  
[About this object](#)

25 US Congresses Served  
63rd – 87th (1913 – 1961)

House Years of Service  
1913 – 1961

State / Territory  
Texas

3 Positions  
Majority Leader  
Representative  
Speaker Of The House

Party  
Democrat

RAYBURN, Samuel Taliaferro, a Representative from Texas; born near Kingston, Roane County, Tenn., January 6, 1882; moved to Fannin County, Tex., in 1887 with his parents who settled near Windom; attended the rural schools and was graduated from the East Texas Normal College, Commerce, Tex., in 1903; studied law at the University of Texas at Austin; was admitted to the bar in 1908 and commenced practice in Bonham, Fannin County, Tex.; member of the State house of representatives 1907-1913, and served as speaker during the last two years; elected as a Democrat to the Sixty-third and to the twenty-four succeeding Congresses and served from March 4, 1913, until his death; chairman, Committee on Interstate and Foreign Commerce (Seventy-second, Seventy-third, and Seventy-fourth Congresses); majority leader (Seventy-fifth and Seventy-sixth Congresses), minority leader (Eightieth and Eighty-third Congresses); elected Speaker of the House of Representatives September 16, 1940, to fill the vacancy caused by the death of Speaker William B. Bankhead; reelected Speaker in the Seventy-seventh, Seventy-eighth, Seventy-ninth, Eighty-first, Eighty-second, Eighty-fourth, Eighty-fifth, Eighty-sixth, and Eighty-seventh Congresses; died in Bonham, Tex., November 16, 1961; interment in Willow Wild Cemetery.

**CONTACT**

Office of the House Historian: [history@mail.house.gov](mailto:history@mail.house.gov)  
Office of the Senate Historian: [historian@sec.senate.gov](mailto:historian@sec.senate.gov)  
Office of Art & Archives, Office of the Clerk: [art@mail.house.gov](mailto:art@mail.house.gov), [archives@mail.house.gov](mailto:archives@mail.house.gov)

**Resources**

[Retro Biographical Directory](#)  
[Privacy Policy](#)  
[Copyright Information](#)

# Reports Due to the Adoption of H. Res. 756 on March 10, 2020

- <https://cha.house.gov/modernization-reports>  
– Under the committee activity tab

## TITLE V—IMPROVING ACCESS TO DOCUMENTS AND PUBLICATIONS

Sec. 501. Adopting standardized format for legislative documents.

Sec. 502. Legislation comparison project.

Sec. 503. Database of information on expiration of authorizations of programs.

Sec. 504. Database of votes taken in committees.

Sec. 505. Assignment of unique identifiers for reports filed by registered lobbyists.

Sec. 506. Public availability of reports.

# Large bills and the ecosystem

## H.R. 133, Consolidated Appropriations, FY 2021

- Engrossed Amendment in House (EAH): 5,432 pages
- Enrollment (ENR): 2,124 pages

## H.R. 8352 116<sup>th</sup>

- Introduced House (IH) : 4,435 pages

## H.R. 6395, NDAA, FY 2021

- Engrossed House (EH) – 3,470 pages
- Placed on the Calendar Senate (PCS) – 3,470 pages

*H. R. 133 ENR on its way to the Senate (signature); 12/24/2020*

*Photo courtesy of the staff in Legislative Operations, Office of the Clerk*





# Comparative Print Project United States House of Representatives

---

# 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

Changes to existing law

# 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



# Data Standards

## *Roadmap for Publishing Legislative Documents in USLM XML Format*

### United States Legislative Markup (USLM)

- <https://github.com/usgpo/uslm>

### Oasis LegalDocML or Akoma Ntoso

- <http://www.akomantoso.org/>

Completed	U.S. Code
Completed	Enrolled bills, Public Laws, and Statutes at Large
IN PROGRESS <i>Nearing completion</i>	HOLC/SOLC statute compilations
IN PROGRESS <i>Initiated</i>	Remaining bill/resolution versions (introduced, reported, engrossed, all amendments, et. al.)
C	Committee Reports
D	House portions of the <i>Congressional Record</i> , including the <i>Daily Digest</i>
E	Hearing Record
F	House Calendar
G	<i>Precedents of the U.S. House of Representatives</i>
H	<i>House Journal</i>

## 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,”;

# 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).

Amended to read as follows:

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)”;

Bill to bill differences



“(b) Chair; Membership.—The Administrator shall designate a representative to serve as the Chair of the advisory board. In addition to the Chair, the membership of the advisory board shall include—

“(1) the Deputy Director for Management of the Office of Management and Budget;

“(2) the Director of the Office of Personnel Management;

“(3) the Administrator of the Office of Electronic Government of the Office of Management and Budget;

“(4) the Assistant to the President and Chief Technology Officer; and

“(5) other individuals as may be designated by the Administrator.

# Changes and Differences

Comparison variation  
two

“(b) Chair; Membership.—The Administrator shall designate a representative to serve as the Chair of the advisory board. In addition to the Chair, the membership of the advisory board shall include ~~the Director of the Office of Personnel Management;~~—

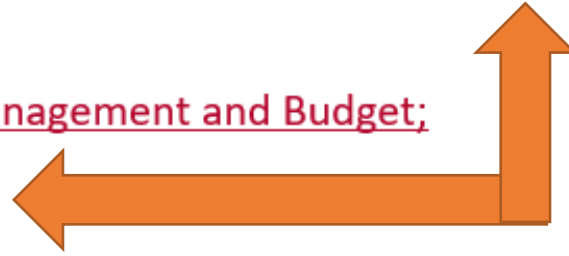
“(1) the Deputy Director for Management of the Office of Management and Budget;

“(2) the Director of the Office of Personnel Management;

“(3) the Administrator of the Office of Electronic Government of the Office of Management and Budget;

“(4) the Assistant to the President and Chief Technology Officer; and

“(5) other individuals as may be designated by the Administrator.



# Changes and Differences

Comparison variation one

# Trust: A Drafter's Viewpoint

---

Wade Ballou, Legislative Counsel  
U.S. House of Representatives

# Demonstration

*Happy path*

# 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:  
[USGPO.GITHUB.IO/INNOVATION/](https://www.usgpo.gov/github.io/innovation/)

# AGENDA

1. Welcome
2. Introductions
- 3. Reports/Updates from LOC, GPO, Clerk, Secretary**
4. Reports/Updates from civil society organizations and others
5. Time for Questions and Answers
6. Announcements/Closing

*Meeting ends at 12:00 pm EST*



# AGENDA

1. Welcome
2. Introductions
3. Reports/Updates from GPO, LOC, Clerk, Secretary
- 4. Reports/Updates from civil society organizations**
5. Time for Questions and Answers
6. Announcements/Closing

*Meeting ends at 12:00 pm EST*

# Other items

- **Job Openings with the House**

<https://www.house.gov/employment/positions-with-other-house-organizations>

## Clerk

- Software Engineer II (Full Stack Developer – Backend focus)
- Senior Database Administrator
- Software Engineer II (Full Stack Developer – Frontend focus)
- Senior IT Business and Systems Analyst

## CAO

- Deputy Chief Information Officer
- Technical Project Manager
- Senior Internet Systems Engineer (DevOps)
- Application Systems Administrator (PeopleSoft)

Resources and other information:

<https://usgpo.github.io/innovation/>

# Bulk Data Task Force

Bulk Data Task Force – Public Meeting  
Wednesday, July 14, 2021



# Thank you for attending!

Bulk Data Task Force – Public Meeting  
Wednesday, July 14, 2021