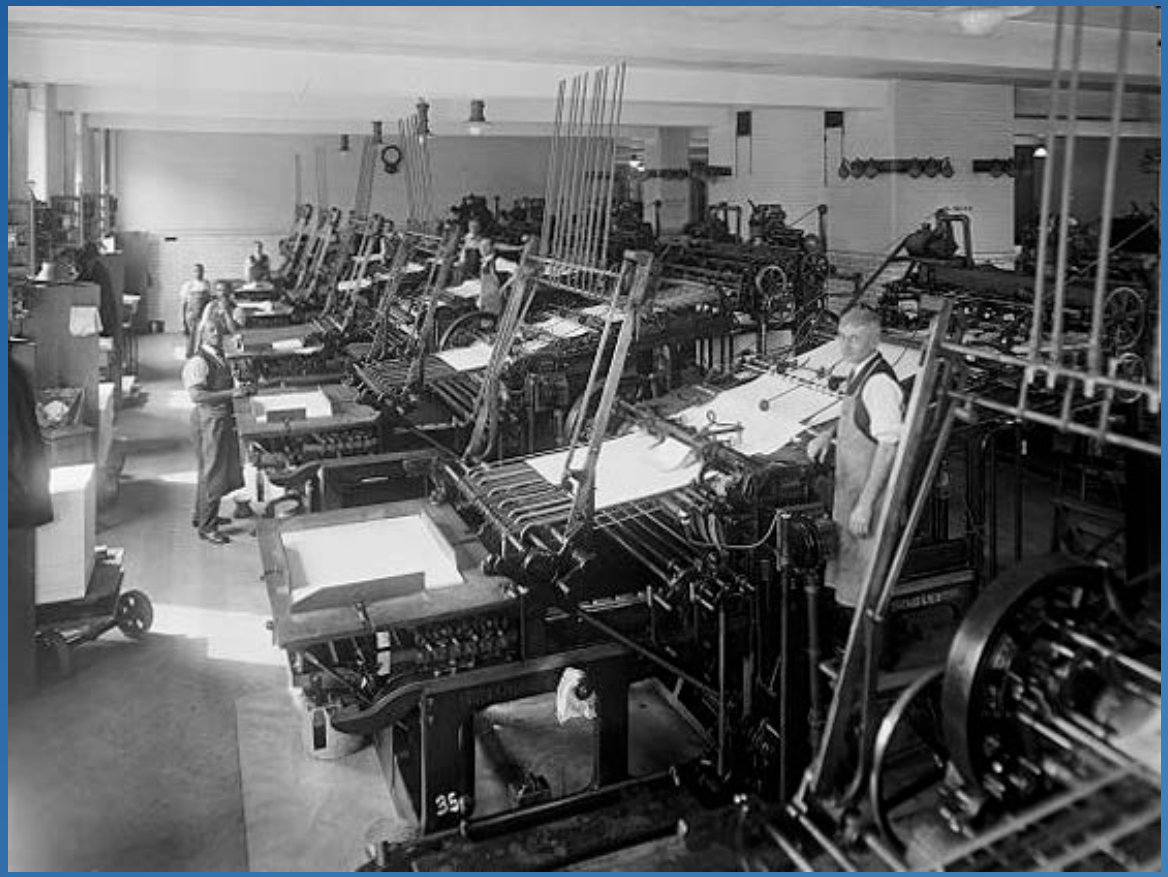
Bulk Data Task Force Meeting

February 8, 2018

Public Galleries
available in April, 1789



PREPARE

http://www.loc.gov/rr/main/govdocsguide/images/GPOPresses.jpg

MANAGE

https://blogs.loc.gov/law/2010/09/do-you-remember-how-to-use-a-card-catalog/

DISTRIBUTE

http://i.iheart.com/v3/url/aHR0cDovL2NvbnRlbnQuY2xlYXJjaGFubmVsLmNvbS9jYy1jb21tb24vbWxpYi82MTYvMDIvNjE2XzE0MjI4OTY3MDUxLmpwZw==?ops=maxcontain(750,473),quality(80)

ARCHIVE

https://www.archives.gov/files/publications/prologue/2008/spring/images/frc-lenexa-l.jpg

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 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".

○

The paper version is the official document of record.

Today's Challenge:

How do parliaments modernize

their systems to utilize today's newer technologies to help citizens access

free & timely

legislative & legal information?

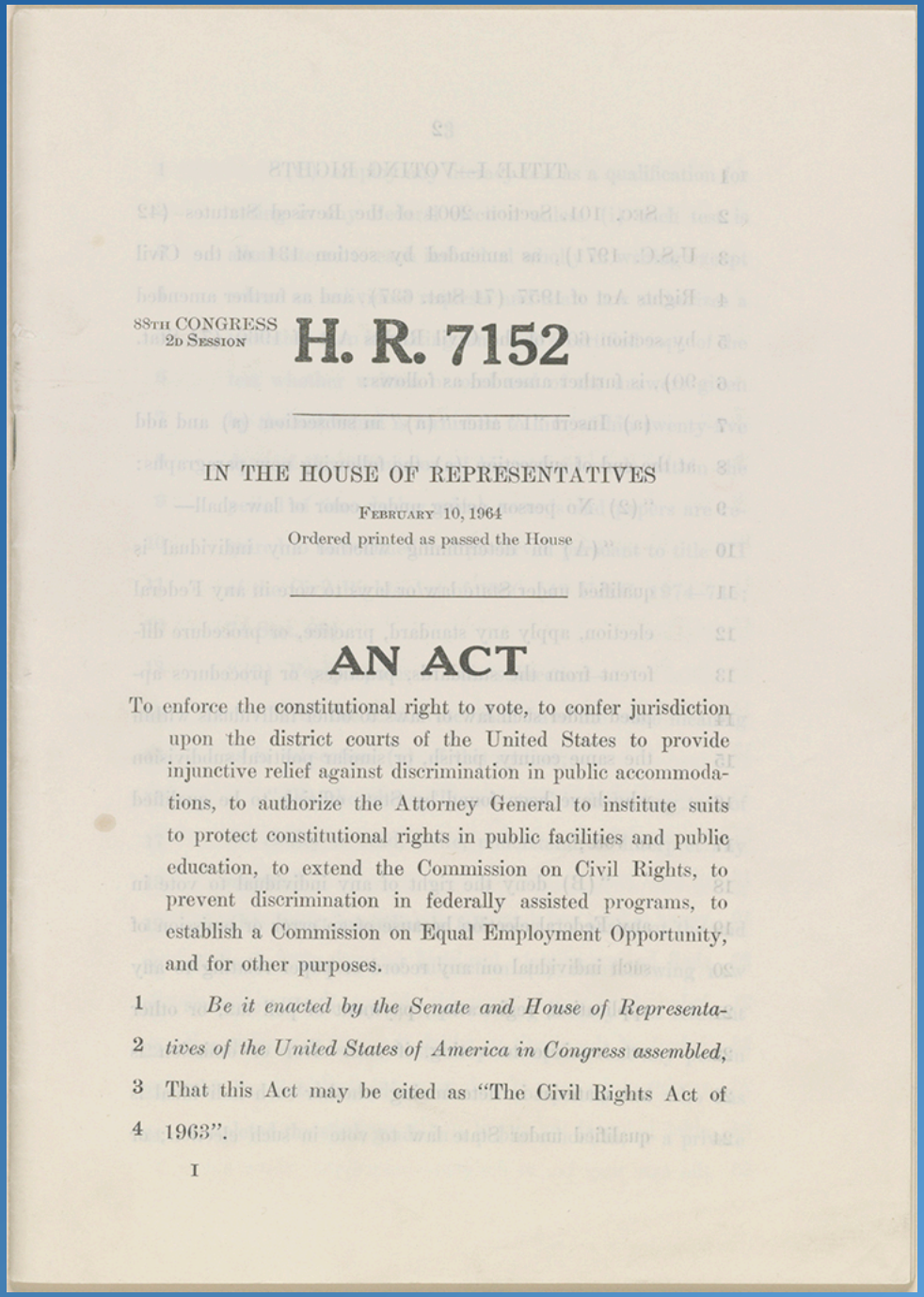
Opening the doors of Congress

http://history.house.gov/Collection/Detail/15032419525

https://geobrava.files.wordpress.com/2015/12/8787-mobile_2d00_device_2d00_security.jpg

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

https://www.loc.gov/exhibits/civil-rights-act/images/cr0162_enlarge.jpg



88TH CONGRESS
2D SESSION

H. R. 7152

IN THE HOUSE OF REPRESENTATIVES

FEBRUARY 10, 1964

Ordered printed as passed the House

AN ACT

To enforce the constitutional right to vote, to confer jurisdiction upon the district courts of the United States to provide injunctive relief against discrimination in public accommodations, to authorize the Attorney General to institute suits to protect constitutional rights in public facilities and public education, to extend the Commission on Civil Rights, to prevent discrimination in federally assisted programs, to establish a Commission on Equal Employment Opportunity, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled, That this Act may be cited as "The Civil Rights Act of 1963".

I

The Bulk Data Task Force

- created in 2012 to focus on increased dissemination of congressional information via bulk data

- partnership of staff of various Legislative Branch Agencies that work in a cooperative manner to coordinate transparency goals
    - House of Representatives
    - Senate
    - Library of Congress,
    - Government Publishing Office

- meets with civil society organizations

Establishment of exchange standards

§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 .)

2 U.S.C. 181

uscode.house.gov

Additional Policy Changes

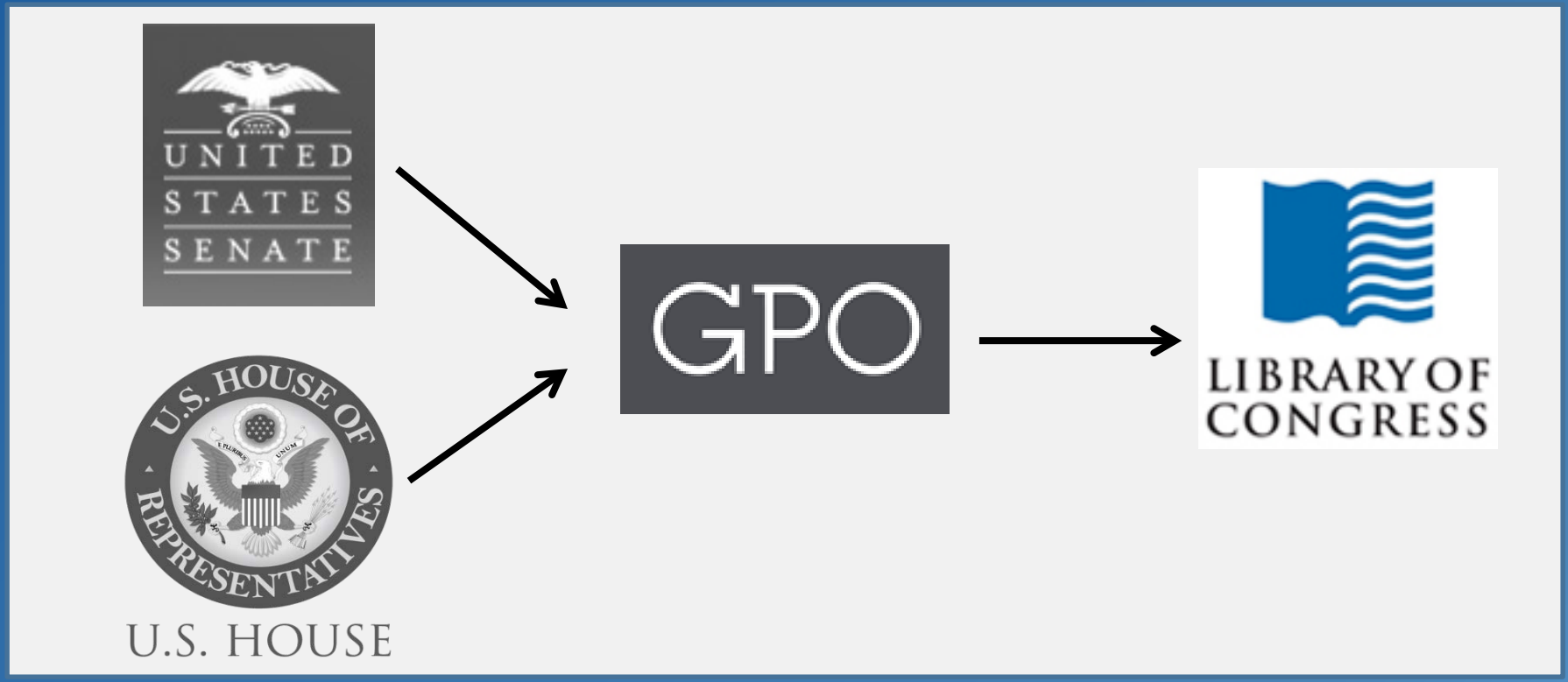
- **In 2011, the rules of the US House of Representatives for the 112th Congress called for greater transparency and accessibility to legislative documents.**

- **The intention of the change was to place electronic distribution on par with traditional printing.**

Technology Timeline

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

Simple Workflow



www.githhub.com/usgpo
www.govinfo.gov/bulkdata

Finding Legislative Information

- DOCS.HOUSE.GOV
  - Text of legislative being considered this week
- CLERK.HOUSE.GOV
  - House information
- CONGRESS.GOV
  - House and Senate data and documents
- GOVINFO.GOV
  - All government publications

Schemas in use by the U.S. House

- Bill DTD
  - Generation one
  - created in 1998 by Congress during the feasibility study
  - still in use for bills and resolutions
- USLM schema
  - Generation two
  - created in 2013 by Congress
  - in use for the US Code
  - uses AKN and HTML standards

---

xml.house.gov

LegalDocML is the next version of the Akoma Ntoso Legislative XML

Originally developed for the United Nations Department of Economic & Social Affairs (UNDESA)

Defined at the University of Bologna, Italy

http://www.akomantoso.org/

Pronounced “A-coma In-tozo”

Roadmap for More Docs in USLM

Completed	U.S. Code
Current Stage	Enrolled bills, Public Laws, and Statutes at Large
A	HOLC/SOLC 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*

*excerpt from Clerk Karen Haas testimony before the Committee on House Administration, October 11, 2017

Qualities of Government Data

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **ACCESSIBLE** also includes **available, discoverable, and online**

Where do members of parliament, staff, and the public access legislative information? It is available? Easy to find? How and where? Accessible data also implies that the data is online and discoverable by humans and computers.

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **ACCURATE** also includes **trusted**

Are the documents and data accurate? Do they tell the story of the legislative process in manner that is understandable to a broad audience? Can it be trusted? How is it authenticated? Parliaments should be the authoritative source for their records and ensure that information released to the public is accurate and trusted.

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **COMPLETE** also includes **detailed**

It is complete? Detailed? Legislative processes are complex. Legislative information needs to be as complete as possible, reflecting the entirely of the legislative actions, subject only to narrowly defined exceptions.

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **DESCRIBED**

Is the legislative information described for all parties using the information? When publishing structured formats, such as XML, a user guide and/or data dictionary defining the formats is a must.

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **FREE**

Data should be available free of charge and in a non-discriminatory manner.

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **MACHINE-READABLE**

Can a computer read and analyze the data using data models that are understood by the widest possible audience? Are best practices being used?

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **PERMANENT**

Is the legislative information available permanently? Are URLs and URIs durable? Predictable? Permanent? Persistent?

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **SEARCHABLE**

Can the data be searched? And to what degree?

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **TIMELY**

Is it timely? Can it be provided in real time? Legislative information should be released publicly as quickly as it is available internally to members and staff.

Government Data Should Be:

- **ACCESSIBLE**
- **ACCURATE**
- **COMPLETE**
- **DESCRIBED**
- **FREE**
- **MACHINE-READABLE**
- **PERMANENT**
- **SEARCHABLE**
- **TIMELY**
- **USABLE**

- **USABLE**

Is it usable? Has the legislative data been published in a manner that is usable by all parties using the information?

Government Data Should Be:

- ACCESSIBLE
- ACCURATE
- COMPLETE
- DESCRIBED
- FREE
- MACHINE-READABLE
- PERMANENT
- SEARCHABLE
- TIMELY
- USABLE

Bulk Data Task Force Meeting

February 8, 2018