



Legislative Data: A Technical Discussion

2015 Legislative Data and
Transparency Conference



Resources:

XML.House.Gov/Resources.htm

#LDTC15



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

2

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.



The Challenge

Systems to create, manage and distribute legislative data and documents must-

- be quick and flexible;
- have no interruptions, no failures, no risk, and no to little cost;
- meet the changing needs of the members and the public.

Well-equipped and well-trained staff can and do handle these constraints well.



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 entirety 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?



<http://5stardata.info/>

- ★ make your stuff available on the Web (whatever format) under an open license
- ★★ make it available as structured data (e.g., Excel instead of image scan of a table)
- ★★★ use non-proprietary formats (e.g., CSV instead of Excel)
- ★★★★ use URIs to denote things, so that people can point at your stuff
- ★★★★★ link your data to other data to provide context

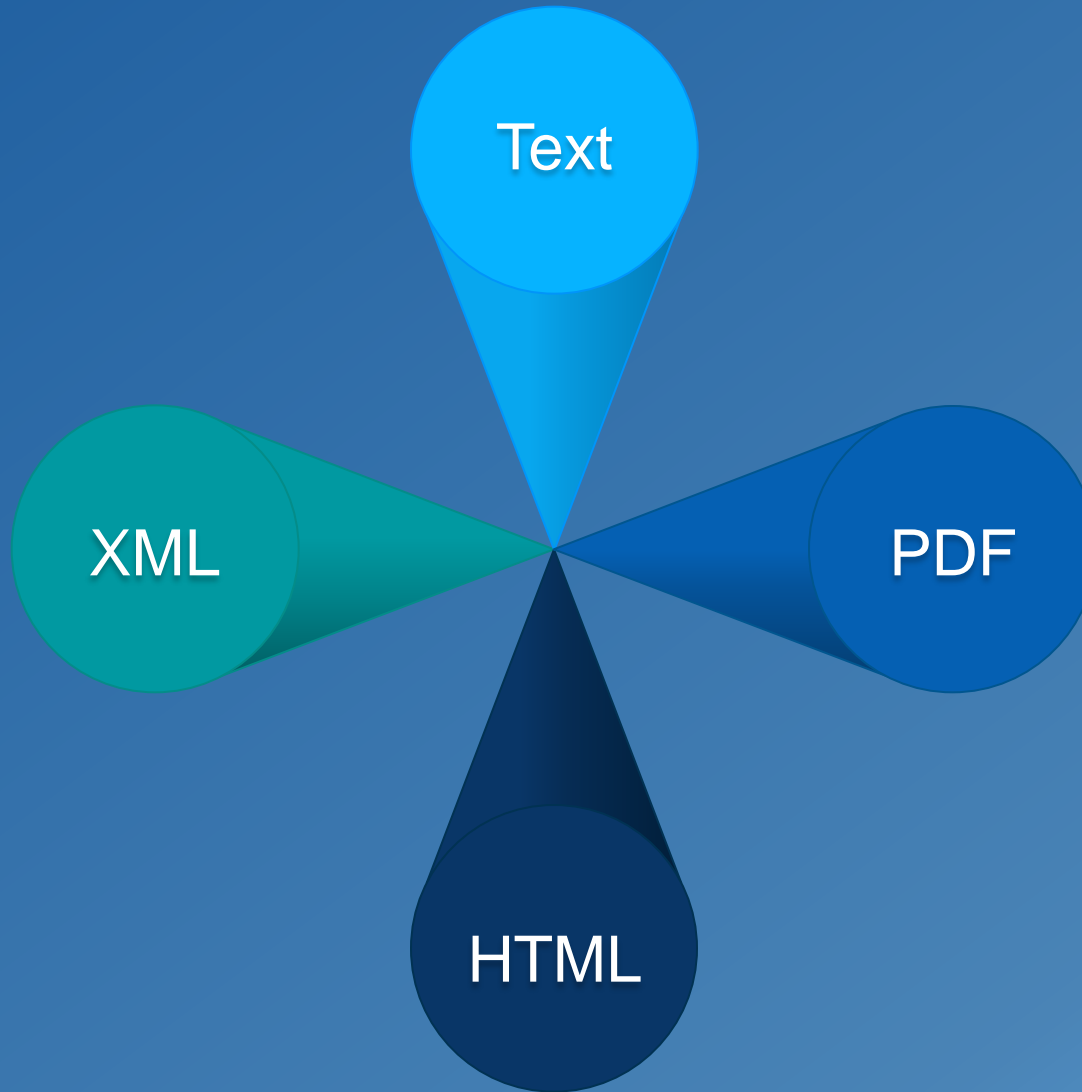


Meeting the Challenge:
How does a parliament meet the challenges
of **accountability, accessibility** and
empowerment of the citizen?

Solution: Structured Data

- Publishing well-formed XML can help meet the goals of
 - accountability of the government,
 - transparency,
 - accessibility,
 - empowerment of the public,
 - higher quality of information, and
 - a way to make technology more human.

Document Formats



Structured data
provides a means to publish
raw data and for others to **add value**.

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 Basics

Element

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

Attribute

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

Selected List of Legislative Branch Data

[http://xml.house.gov/resources/
SelectedList-
LegisBranchData.htm](http://xml.house.gov/resources/SelectedList-LegisBranchData.htm)

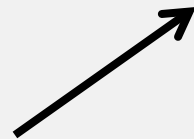
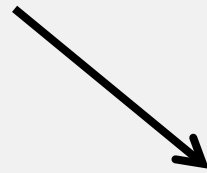
Data and Documents

- 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

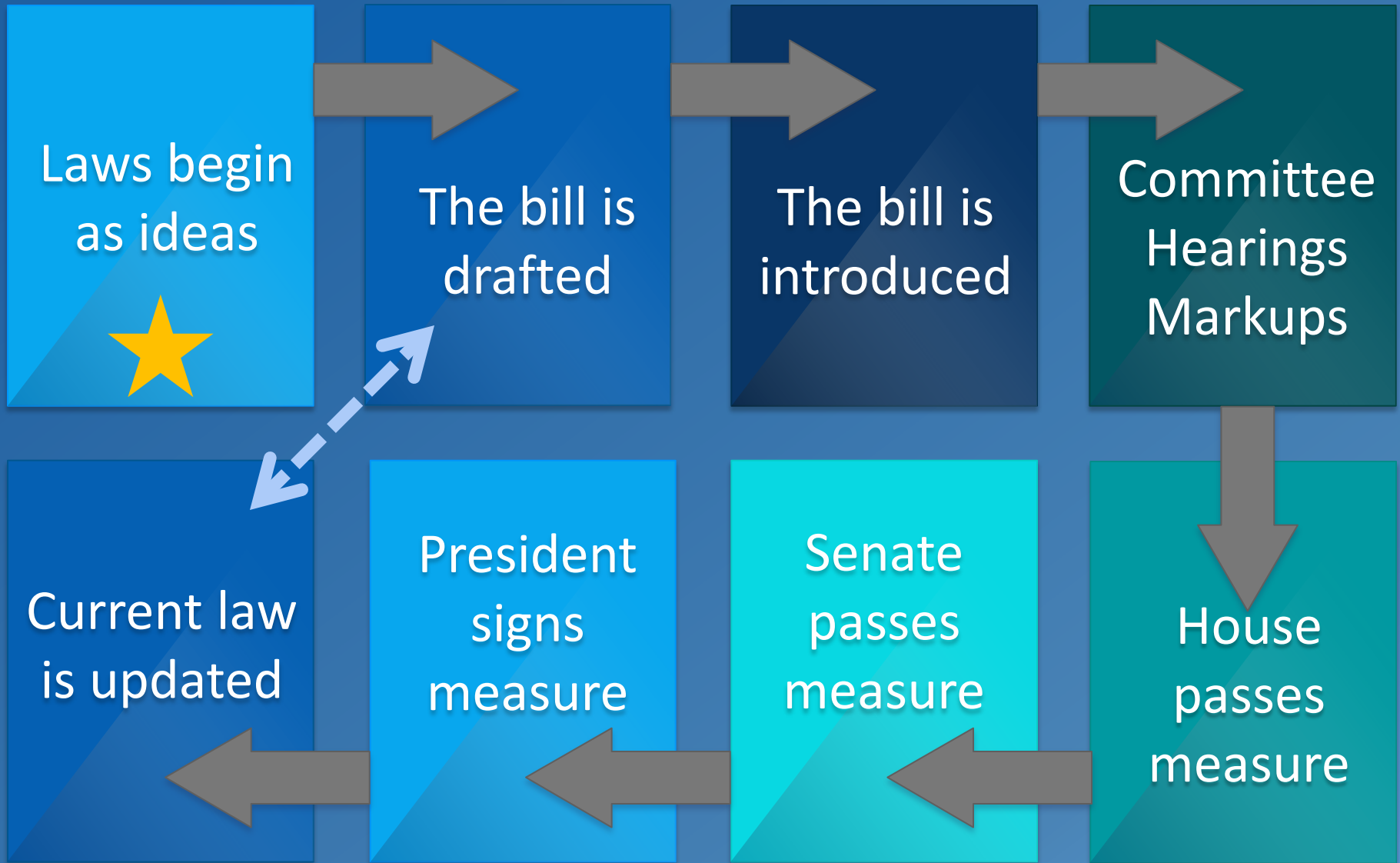
Simple Workflow



U.S. HOUSE

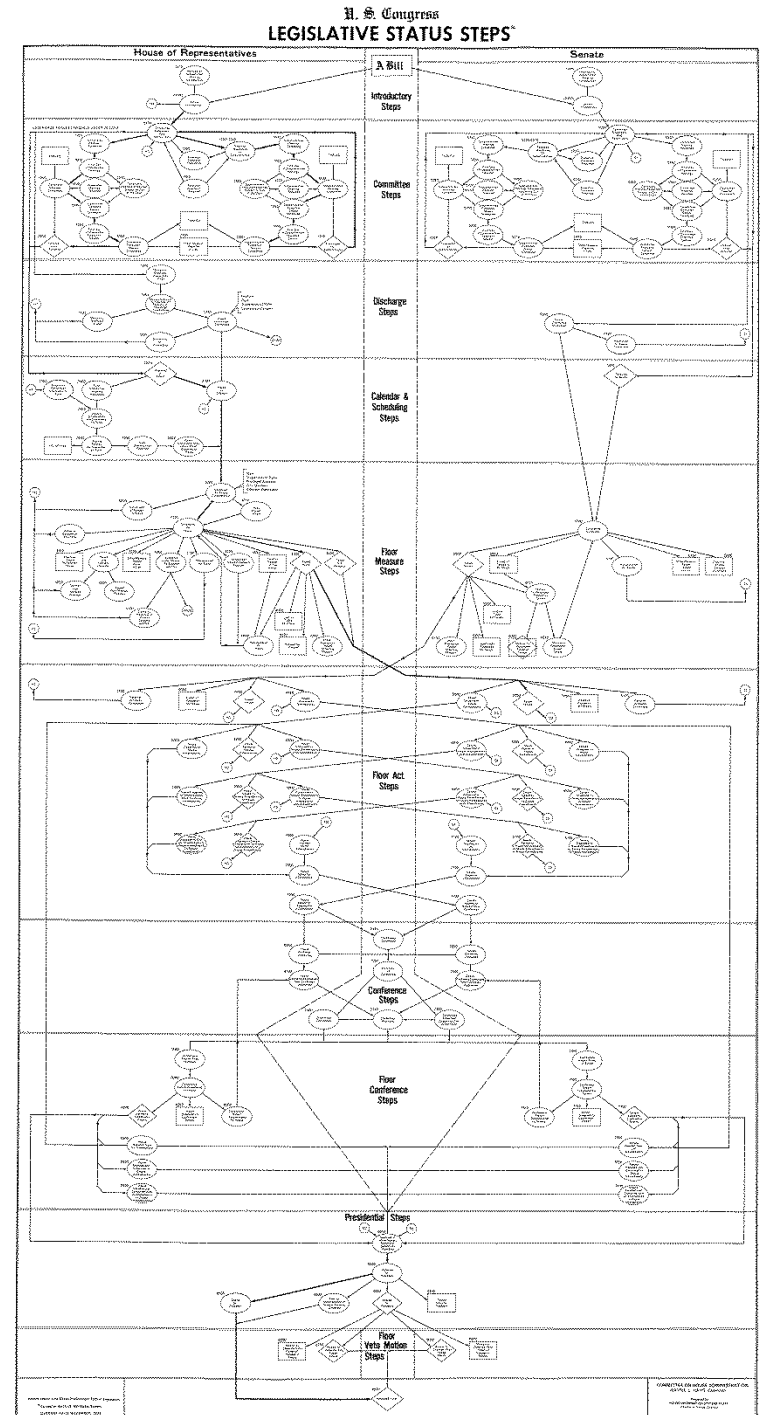


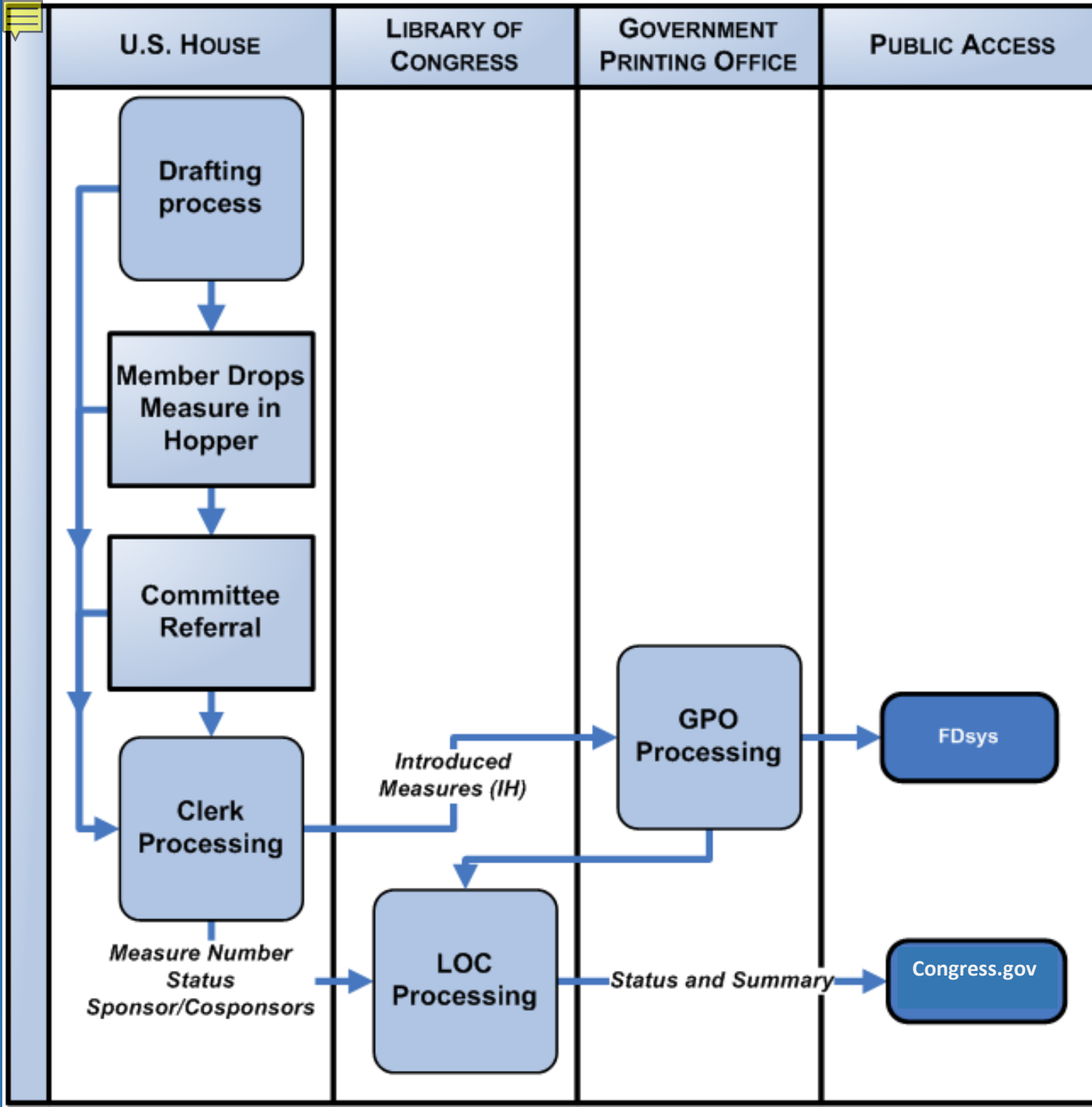
How a Bill Becomes the Law

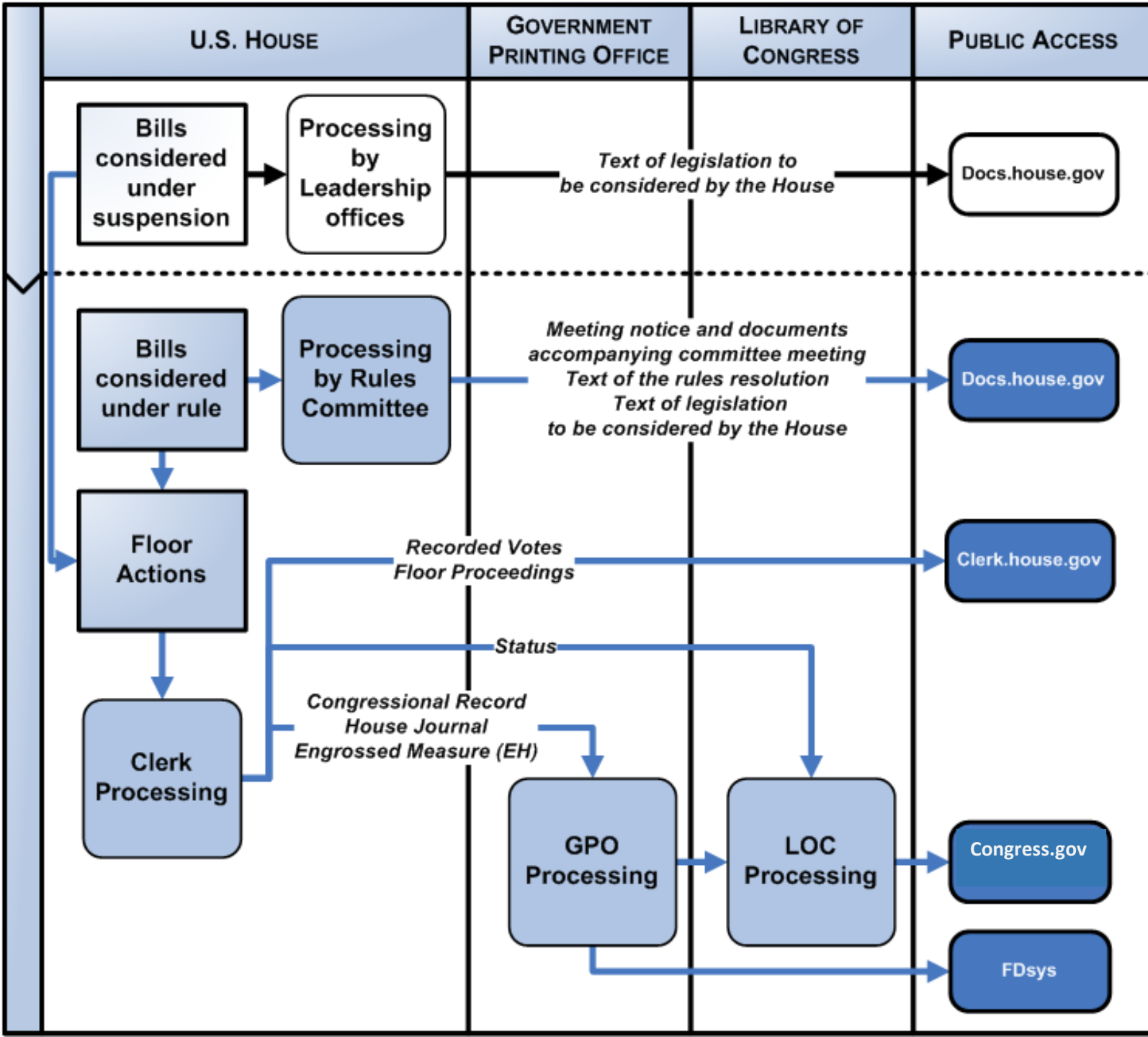


The shortest success path for a bill to become law is twenty decision points

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







Thank you

Kirsten Gullickson

Office of the Clerk
U.S. House of Representatives

Kirsten.Gullickson@mail.house.gov