

Government Publishing Office

Lisa LaPlant

Program Manager, Govinfo

Matt Landgraf

Lead Program Planner

The GPO logo consists of the letters 'GPO' in a white, sans-serif font, centered within a dark blue circle. This circle is surrounded by a white ring, which is itself set within a larger, light gray ring.

GPO

Bulk Data Task Force Update

Government Publishing Office

Lisa LaPlant, Govinfo Program Manager
Matt Landgraf, XPub Program Manager

October 17, 2019

Topics

- XPub (formerly Composition System Replacement)
 - Background and Overview
 - U.S. Code in Production
 - Release Roadmap
- Govinfo
 - Highlights
 - Statute Compilations in USLM
 - H. Rept. 115-696

Introducing XPub

- Formerly known as GPO's Composition System Replacement (CSR)
- XML-based composition system to replace legacy Microcomp system and locator-coded text format

XPub Goals

- Accept content in any form
- Simultaneously generate content in a variety of print and digital formats, including USLM
- Create templates so content can be easily repurposed and reproduced
- Produce PDF files and digital products that are Section 508-compliant

XPub Process Modernization

- Comprehensive environment for managing GPO's publishing operations
 - Content intake and job monitoring
 - Tools for integrated work assignment, receipt, submission, and approval
 - Proof reading/correction
 - Structural and semantic XML markup
 - Job assembly and production of print and digital products

U.S. Code in Production via XPub

- 2018 Main Edition is being produced using XPub
- First large scale production job using the new system
- GPO plans to complete the publishing of the 2018 Main Edition in less than a year

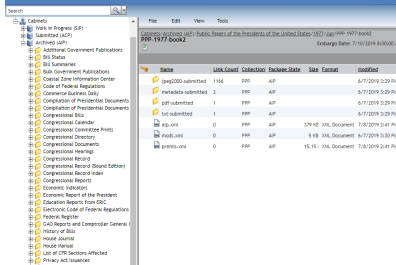
XPub Release Targets

- Complete U.S. Code (end of 2019)
- Congressional Bills, Public Laws, Statutes at Large (for the 117th Congress)
- House and Senate Calendars (2020)

Trustworthy Digital Repository

govinfo

content management system



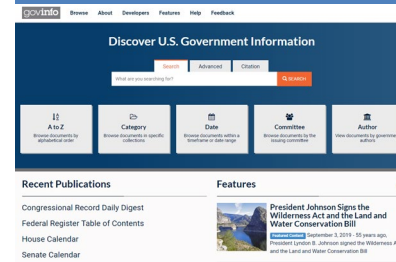
manages digital content to ensure authenticity and

preservation repository

```
<?xml version="1.0" encoding="UTF-8"?>
<meta:meta xmlns:xsi="http://www.loc.gov/METS/
http://www.loc.gov/standards/mets/mets.xsd"
xmlns:dc="http://www.w3.org/2001/XMLSchema-instance"
xmlns:meta="http://www.loc.gov/METS/"
xmlns:mets="http://www.loc.gov/METS/"
CREATEDATE="2019-06-07T15:37:49-04:00">
  <meta:agent ROLE="CREATOR" TYPE="ORGANIZATION">
    <meta:agentName>United States Government Publishing Office</meta:agentName>
  </meta:agent>
  <meta:metaIdentifier>
    <meta:metaIdentifierType>ACIP</meta:metaIdentifierType>
  </meta:metaIdentifier>
  <meta:object>
    <meta:objectID>PMD_OTHER</meta:objectID>
    <meta:objectID>P0002e61a151694</meta:objectID>
    <meta:objectName>Title</meta:objectName>
    <meta:objectName>http://www.w3.org/1999/xlink LOCTYPE="URL"
    HREF="METS" />
  </meta:object>
  <meta:amfID>
    <meta:amfID>AMM_OTHER</meta:amfID>
    <meta:amfID>M0902e61a151927-EEC</meta:amfID>
    <meta:amfID>M0902e61a151927-48V</meta:amfID>
    <meta:amfID>http://www.w3.org/1999/xlink LOCTYPE="URL"
    HREF="METS" />
  </meta:amfID>
  <meta:diagnosticID>
    <meta:diagnosticID>D0902e61a151927-DIG</meta:diagnosticID>
    <meta:diagnosticID>M0902e61a151927-48V</meta:diagnosticID>
    <meta:diagnosticName>Title</meta:diagnosticName>
    <meta:diagnosticName>http://www.w3.org/1999/xlink LOCTYPE="URL"
    HREF="METS" />
  </meta:diagnosticID>
  </meta:diagnosticID>
</meta:meta>
```

follows preservation standards enabling access to digital content

public website



combines extensive metadata, robust search, and modern

Notable Content

- Congressional Record (1873 to present)
- Federal Register (1936 to present)
- Code of Federal Regulations (1996 to present)
- Bills, Slip Laws, and United States Code (1994 to present)
- United States Statues at Large (1951 to present)
- Public Papers of the Presidents of the United States (1929 to present)
- Budget of the United States Government (1996 to present)
- Federal District, Bankruptcy, and Appellate Opinions (1.2M cases from 132 courts)
- House and Senate Rules, Procedures, Precedents, Calendars, and Journals
- House and Senate Committee Hearings, Prints, Reports, and Documents
- Directories of Organizations and Officials in the U.S. Federal Government
- Current and Historical Publications from Federal Agencies and Commissions

Websites



Govinfo Highlights

- FDsys retirement and redirects to govinfo.
- ISO 16363 Trustworthy Digital Repository certification.
- USLM XML for Enrolled Bills, Public Laws, and Statutes at Large.
- API for Bill Status and eCFR Bulk Data.
- Major upgrades to key system components including the CMS, search engine, databases, application servers, and frameworks.
- Over 1,500 digitized Congressional hearings from various House and Senate committees, over 2,000 digitized executive agency and commission publications, and the digitized Public Papers of the Presidents back to Hoover in 1929.
- See also www.govinfo.gov/features/release-notes.

Statute Compilations in USLM

Goal is to provide a uniform set of laws in USLM to enable downstream processes and increase efficiencies.

Phase 1 – 319 COMPS Available on govinfo

- <https://www.govinfo.gov/app/collection/comps>
- In coordination with HOLC/SOLC/Clerk/Secretary, make select Statute Compilations in PDF available on govinfo as a Beta.
- Ingest legacy COMP DTD XML & locator files into govinfo CMS.

Phase 2 – Procurement In Process

- Convert legacy COMP DTD XML or locator files into USLM XML.
- Provide access to COMPS in USLM XML as Bulk Data.

H. Rept. 115-696

- In H. Rept. 115-696, the report that accompanied the FY19 House legislative branch appropriation bill, the House Appropriations “Committee direct[ed] the GPO to assess the costs associated with converting Statutes at Large from 1789-2002 into USLM XML format.”
- In FY20, GPO plans to initiate a project to digitize the remaining Statutes at Large volumes back to Volume 1 (1789-1799)
- As part of the project, GPO plans to prototype the conversion of a subset of digitized Statutes at Large into USLM XML in order to assess the technical feasibility and determine future budget requirements.
- Funding for the digitization and prototyping efforts is included in GPO’s proposed FY20 budget.

Contacts and Sites

llaplant@gpo.gov

@lisalaplant

mlandgraf@gpo.gov

github.com/usgpo

usgpo.github.io/innovation

github.com/usgpo/uslm

govinfo.gov

api.govinfo.gov