

Access to Government Information

How libraries can continue to ensure sustained access to this information for researchers

Annelise Sklar



The Library
UC SAN DIEGO

How we got here: FDLP legacy print collection

- Estimated at 2.3 million items.
- As much as 1/3 is not catalogued.
- Most depositories joined the program after 1813
 - not required to acquire retrospectively
 - or retroactively in the event of loss



Congressional
Research
Service

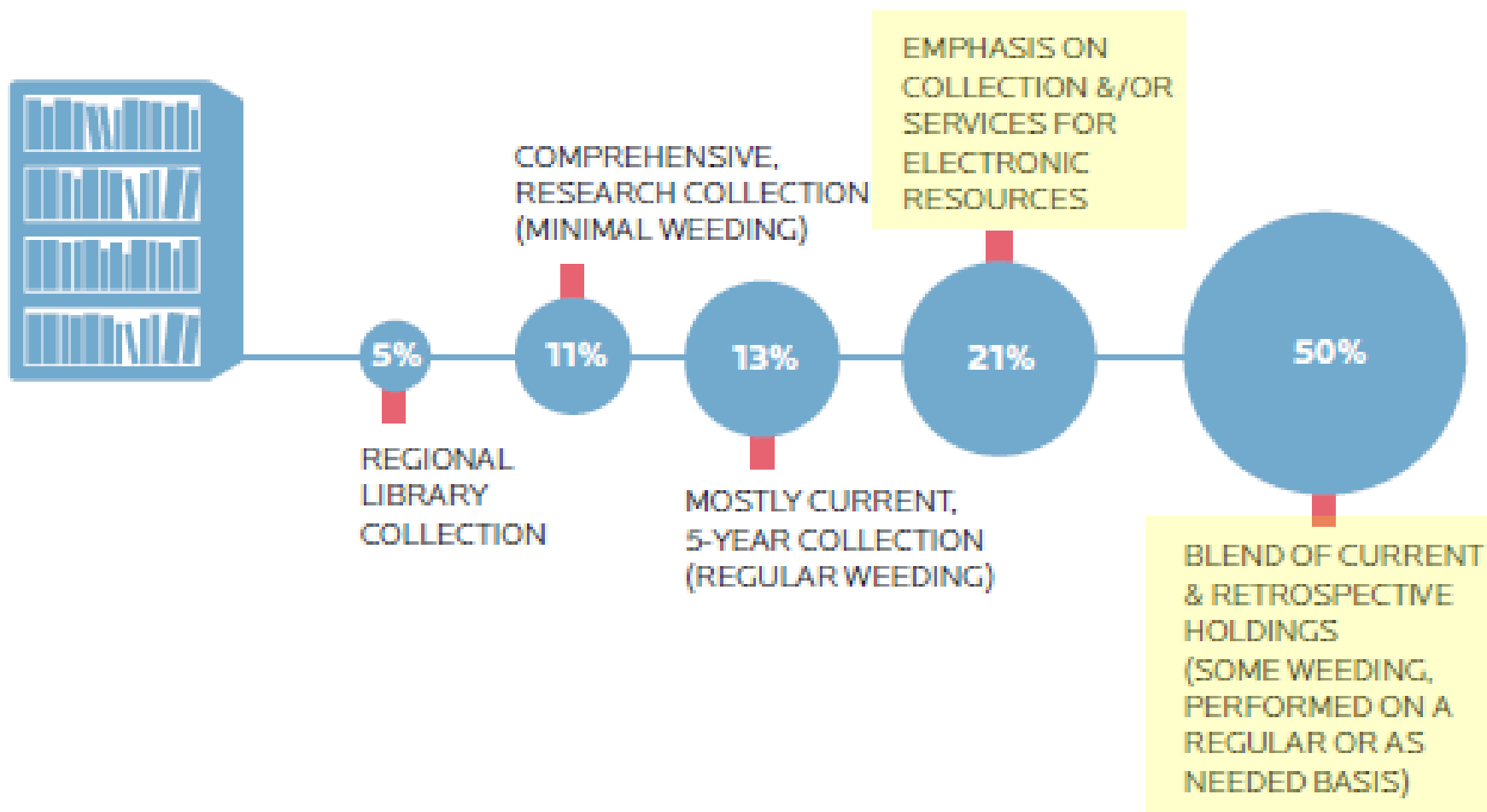
Where we're at: Ithaka

- Libraries are moving from the “just in case” model of collecting tangible materials locally.
- Instead, relying more on online materials
 - often, moving large portions of the collections into offsite storage or even deaccessioning them.



Most libraries do not host their own digital content, but instead license collections of journals and books from centralized service providers.

The typical depository collection



User behaviors and expectations

- **Users don't really care where their info comes from, just that they get it.**
 - 84% of people begin their search with a search engine
 - 55% of FDLP users list “Google or other search engine” as a frequent source for US government information.
- **24/7 self-service access to information resources**
 - The top most-desired improvement to government information access is more materials online (50%).
- **Expect data they can interact with and reuse for their own purposes**



http://www.oclc.org/content/dam/oclc/reports/pdfs/studentperceptions_part1.pdf

<http://www.fdlp.gov/component/content/article/19-general/1011-depositoryusersurveyreport>

More than just documents

“Government Information” includes so much more than it used to.

- **Traditional documents, reports, and publications**
- **Statistics**
- **Microdata**
- **Geographic data**
- **Data and publications created by government contractors**
- **Data and publications funded by government grants**



Open, Open, Open

- **Open Access**
 - Finished articles.
 - About sharing findings and knowledge for the greater good (especially funded by tax payers) .
- **Open Data**
 - Data for replication (scholarly accountability)
 - and re-purposing (new research, new tools).
- **Open Government**
 - Transparency and government accountability

A vertical neon sign with the word "OPEN" in a stylized, glowing red font. The letters are outlined and have a bright red glow. The sign is set against a dark background and is partially obscured by a vertical black line.

Is it gov info now?

- **National Institutes of Health (NIH)**
 - “It requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication... accessible to the public... no later than 12 months after publication.”
- **National Science Foundation (NSF)**
 - “Investigators are expected to share... the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants.”
- **NEH Office of Digital Humanities**
 - The plan will need to address two main topics:
 - What data are generated by your research?
 - What is your plan for managing the data?
- **Office of Science and Technology Policy (OSTP) memo**
 - “directs each Federal agency with over \$100 million in annual conduct of research and development expenditures to develop a plan to support increased public access to the results of research funded by the Federal Government.”

<http://publicaccess.nih.gov/>

<http://www.nsf.gov/bfa/dias/policy/dmp.jsp>

http://www.neh.gov/files/grants/data_management_plans_2013.pdf

http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Increased access to government-produced data

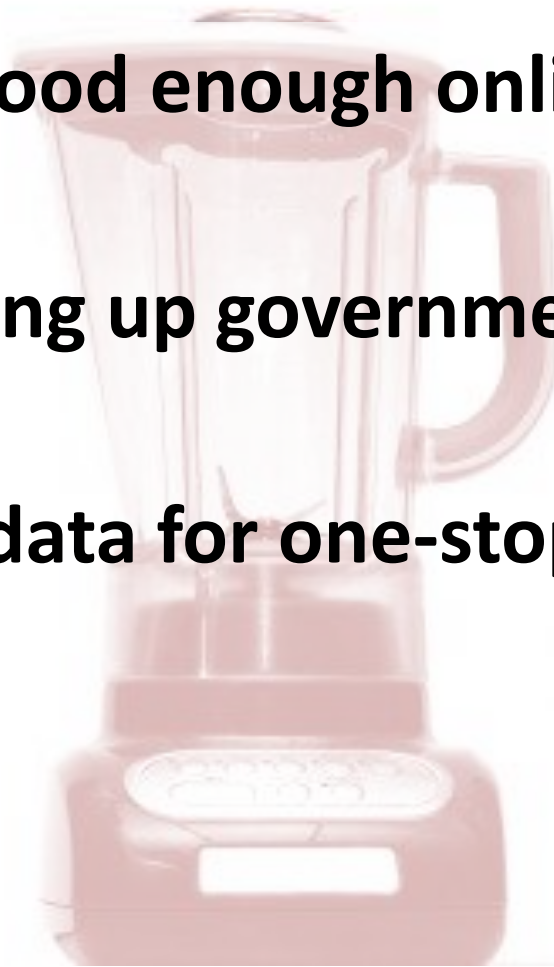
For example...

- **Data.gov**
- **Census, American Factfinder**
- **Bureau of Labor Statistics**
- **FRED (Federal Reserve Economic Data)**
- **EPA Data Finder**
- **DOE Data Explorer**



Remixing the government

- The pdf isn't good enough online access any more
- APIs for mashing up government information into new tools
- Reformatting data for one-stop-shopping



Projects done by

- **Researchers**
 - IPUMS
 - A New Nation Votes (Tufts)
 - American Presidency Project
- **Nonprofits**
 - Sunlight Foundation
 - National Institute on Money in State Politics:
followthemoney.org
- **By open government nuts who see a need**
 - Govtrack.us
 - Code for America Brigade
 - Open San Diego



Costs of distributing Government Information

- In the print model,
 - Depository libraries bear the costs of managing tangible materials, staff, and physical plant needs, providing public access, etc.
- In the electronic model,
 - Costs of providing digital materials fall on GPO for FDSys and the other agencies that provide online content.

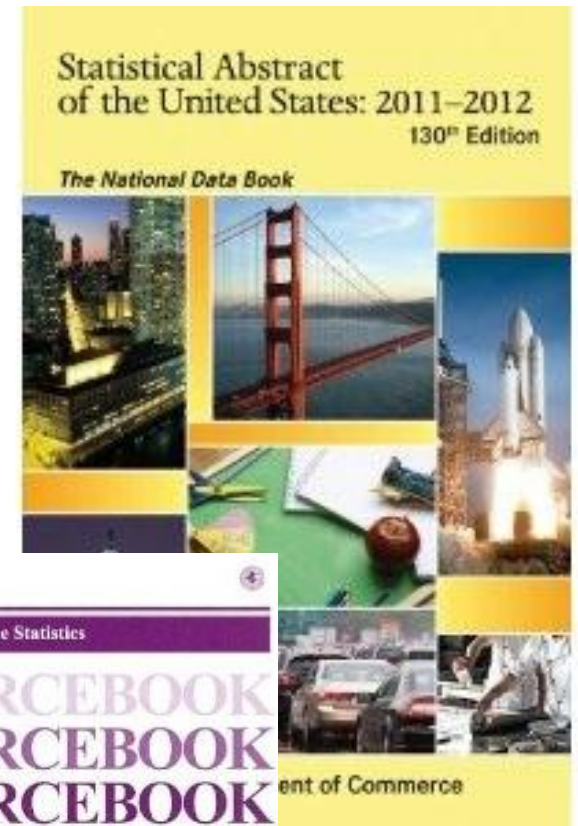


Congressional
Research
Service

Smaller Government

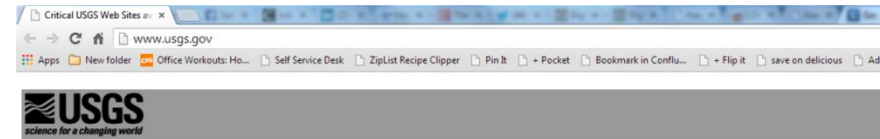
Good-bye

- Statistical Abstract
- Sourcebook of Criminal Justice Statistics



Government Shut-Down

- Bureau of Economic Analysis: unavailable
- Census Bureau: unavailable
- Centers for Disease Control: available but not being updated (with exceptions for websites supporting critical functions)
- data.gov: unavailable
- Department of Agriculture: unavailable
- Department of Education: available but not being updated; however, the National Center for Education Statistics website and the ERIC database are unavailable.
- Department of Energy: available but not being updated
- Department of Health and Human Services: available but not being updated (with exceptions for websites supporting critical functions)
- Department of Labor: available but not being updated
- FDSys is being updated only with Congressional materials, as well as Federal Register documents that protect life and property.
- Federal Communications Commission: unavailable
- Federal Trade Commission: unavailable
- Institute for Education Sciences: unavailable
- International Trade Administration: unavailable
- Library of Congress: access to entire website restored 10/4
- NASA: unavailable
- National Archives: available but not being updated (except for daily Regi)
- National Institutes of Health: available but not being updated
- National Institute of Standards and Technology (NIST): unavailable
- National Library of Medicine (including PubMed database): available but updated
- National Park Service: unavailable
- National Science Foundation: unavailable
- NOAA-associated: unavailable, except for weather.gov
- USGS: mostly unavailable



Due to the lapse in federal government funding, this website is not available.

After funding has been restored, please allow some time for this website to become available again.

For information about available government services, visit usa.gov

To view U.S. Department of Agriculture Agency Contingency plans, visit: <http://www.whitehouse.gov/omb/contingency-plans>

[Message from the President to U.S. Government Employees](#)

It was safe in the fiche...



NASA & ERIC review projects

- **Full text documents taken off-line pending review**
- **ERIC: “privacy concerns about information contained in some of the collection”**
- **NASA: “to ensure that it does not contain technical information that is subject to U.S. export control laws and regulations”**

So what's a library to do?

Whatever's best for our users,
of course!



Buy it

Readex
A Division of NewsBank

 **HEINONLINE**

 **GEOGRAPHIC
RESEARCH, INC.™**

 **DATA-PLANET™**
A UNIVERSE OF DATA

LexisNexis®

 **Social Explorer®**
Professional

 **Bernan**
Since 1952

ProQuest

WestlawNext™

LLMC DIGITAL
Law Library Microform Consortium

east view
information services

Pros and Cons

- Not really a new concept
- Sometimes it's worth the money to have a better interface or more (perceived) stability
- Vendors have the resources to carry out these big digitization projects, update and maintain interfaces, provide value-added context and/or metadata
- But they also have to rely on the government to produce that data in the first place
- And we have to justify spending the money on something that is "free."

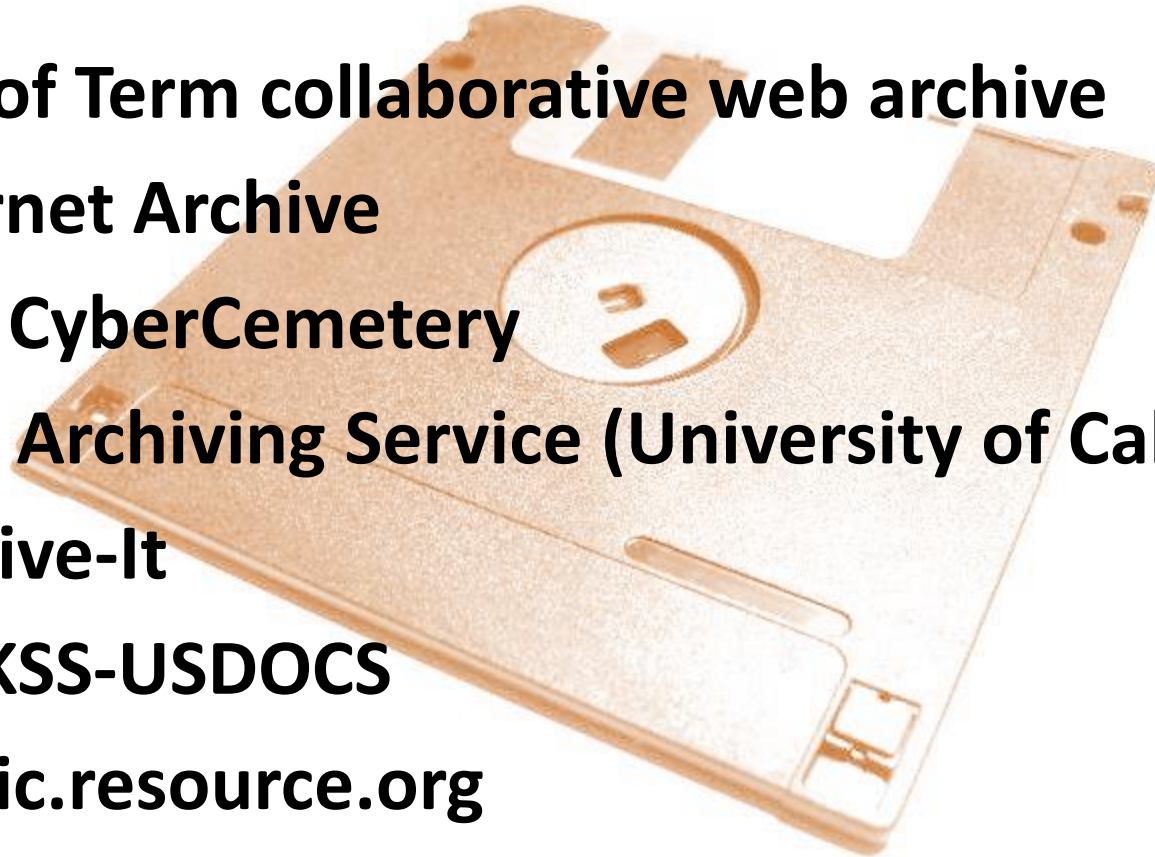


Digitally preserve it

- There's a lot of government information out there, waiting to be digitized
- Copyright is not an issue for digitizing FDLP materials
 - But you can't guillotine a Regional's collection!
- GPO (FDSys) and LOC (American Memory Project)
- HathiTrust
- LLMC
- Internet Archive
- UNT A-Z digitization project, CRS archive
- Indiana Virtual CD-ROM Library

Back it up

- End of Term collaborative web archive
- Internet Archive
- UNT CyberCemetery
- Web Archiving Service (University of California)
- Archive-It
- LOCKSS-USDOCS
- public.resource.org
- GODORT SLDTF Inventory of Projects



But...

- **Digitization has a relatively short history.**
 - Less is known about the long-term, archival retention of digitized or born digital materials
 - (We're figuring it out as we go along.)
- **Born digital materials—such as databases, websites, and publications—may also be dynamic**



Congressional
Research
Service

Still to come

Our biggest challenge is figuring out how to preserve not just the data but the ability to interact with the data

- **Old CD-ROMs**
- **Databases that go offline**
 - **USGS National Biological Information Infrastructure (NBII)**
- **Online maps**
- **American Factfinder**





**So how do
we make this
work?**

I really wish I had the
answers