Data without Borders

How libraries can promote preservation and use of public and research data

Katie Fearer
Open data portals are becoming widespread
Government support of open data is growing

Passed into Law: OPEN Government Data Act (S. 760 / H.R. 1770)

Passed: The OPEN Government Data Act was passed by the US House and Senate on December 31, 2018 and was signed by the President on January 14, 2019.

Reintroduced in both the House of Representatives and the Senate on March 29, 2017, the Open, Permanent, Electronic, and Necessary (OPEN) Government Data Act (S. 760 / H.R. 1770) is a bill to codify former President Obama’s Executive Order making “open and machine readable” the default for all government data. The sponsors of this legislation are Rep. Derek Kilmer (D-WA), Rep. Blake Farenthold (R-TX), Sen. Brian Schatz (D-HI), and Sen. Ben Sasse (R-NE).

The bill institutionalizes the federal government’s commitment to Open Data, and codify a policy of “open by default” for all government data. The bipartisan legislation would require federal agencies to

SPARC (the Scholarly Publishing and Academic Resources Coalition)
Libraries are getting involved too

Data is a powerful tool for change

When government data are made accessible and re-usable, they enable individuals, organizations and even governments themselves to innovate and collaborate in new ways.

-World Bank
Using data to effect change is hard
Data isn’t always available

EPA Staff Prevented from Updating Climate Change Indicators Website

Date: November 21st, 2018
Agency: EPA, Federal
Explanation: Government Censorship, Self-Censorship
Scientist: Climate

On November 21, 2018, E&E News reported that staff at the Environmental Protection Agency (EPA) had been prevented from updating a website providing information about climate change. The site, titled “Climate Change Indicators in the United States,” is intended to “compile a key set of indicators related to the causes and effects of climate change.” During the Obama Administration, the site was updated every six months to include new climate change indicators, such as temperature increases or changes in sea ice cover. However, according to EPA staff, the site has not been updated since President Trump took office. One staffer told E&E news:

“Every year, we get a new data point. The Arctic sea ice extent for that year, for example. Or a new number from [the Centers for Disease Control and Prevention] on Lyme disease cases. So usually we update that site every six months, but we haven’t...”

Silencing Silence Tracker, Columbia School of Law, Sabin Center for Climate Change Law.
How do we democratize data?

Perhaps no place in any community is so totally democratic as the town library. The only entrance requirement is interest.

–Lady Bird Johnson

Book Shelves and Computers by Alex Proimos, licensed under CC BY-NC 2.0.
Strategy 1: Connect researchers with data

- Anchorage Open Data Portal
- data.gov
- Arctic Data Center
- Directory of Open Access Repositories
- CoreTrustSeal Certified Data Repositories
- Registry of Research Data Repositories
Strategy 2: Create a data user guide

Data User Guides

Data user guides are designed to go beyond metadata and give data users a sense of the purpose for which the data was collected, the business processes, software, and applicable standards involved, suggested applications for the data (and how to use it), and other details that will help data users.

It’s our hope that data users will contribute to the knowledge base through data user guides. We also expect that data user group activities will also lead to the development of more and more user guides.

We’ve been developing and testing a framework for data user guides. Our first dataset specific user guides are now available for Allegheny County Property Assessment Data and City of Pittsburgh 311 data, and we released a broader guide for criminal justice, court, and corrections data in LibGuide format. As always, if you have any questions, suggestions or comments about documenting data or anything else, please let us know.

- Draft user guide framework
- Allegheny County Property Assessment Data User Guide
- City of Pittsburgh 311 Data User Guide
- Guide to Crime, Courts, and Corrections in the City of Pittsburgh
- A Guide for Data Related to Cardiovascular Disease Risk in Allegheny County
Strategy 3: Data literacy workshops

- School of Data’s Gentle Intro. to Extracting Data
- Urban Institute’s Guide
- Data Equity on Main Street Curriculum
- Data storytelling tools from Berkeley

Connect with Open Data
Mondays, November 6th-27th • 5:00 pm • Central Library, 4th floor, Wells Fargo Technology Center

How clean is my tap water? How much traffic is there in my neighborhood? Who contributed to local political campaigns? Learn to use open data to make a difference in your life and community. Join us for a series of four cumulative workshops that will teach you to identify, analyze, manipulate, and share open data. Space is limited. Please register at: http://sandiego.librarymarket.com/connect-open-data-2.

- Nov. 6: What Is Open Data?
- Nov. 13: Finding and Analyzing Open Data
- Nov. 20: Working with Open Data
- Nov. 27: Making Your Case with Open Data

For more information, please contact: pgunderson@sandiego.gov

San Diego Central Library  @ Joan & Irwin Jacobs Common
330 Park Blvd, 92101 • 619-236-5800
Strategy 4: Save data – ICPSR's DataLumos

DataLumos is an ICPSR archive for valuable government data resources. ICPSR has a long commitment to safekeeping and disseminating US government and other social science data. DataLumos accepts deposits of public data resources from the community and recommendations of public data resources that ICPSR itself might add to DataLumos.
Strategy 5: Bring data to the classroom

- ICPSR teaching resources
- Teachingwithdata.org
Strategy 6: Collaborate with your city

Sunlight Foundation’s Roadmap to Informed Communities

https://www.policedatainitiative.org/
Strategy 7: Support data producers

- Data management plans
- Metadata
- Data curation workshop
- Connect data producers with repositories

Digital Curation Center’s Guide for Higher Education Institutions
Strategy 8: Learn

- DataONE Modules
- UC DMPTool
- N.E. Collaborative Data Management Curriculum
- Digital Curation Centre
Strategy 9: Endangered Data Week

• Feb. 25 – March 1, 2019
• 2020 at your library?
• Resources on GitHub
Strategy 10: Advocacy

- Sunlight Foundation
- GODORT
- DLF Records
- Transparency / Accountability Group
- SPARC
But I’m a librarian not a data scientist
What will you do?