

Mar 29th, 11:15 AM - 12:00 PM

Government and Online Data: Creation, Access, Preservation

Richard M. Mikulski
Portland State University, mm8@pdx.edu

Follow this and additional works at: <https://pdxscholar.library.pdx.edu/onlinenorthwest>

Let us know how access to this document benefits you.

Mikulski, Richard M., "Government and Online Data: Creation, Access, Preservation" (2019). *Online Northwest*. 10.

<https://pdxscholar.library.pdx.edu/onlinenorthwest/2019/schedule/10>

This Presentation is brought to you for free and open access. It has been accepted for inclusion in Online Northwest by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Government and Online Data: Creation, Access, Preservation



Search

BROWSE BY TOPIC

EXPLORE DATA

LIBRARY

SURVEYS/ PROGRAMS

INFORMATION FOR...

FIND A CODE

ABOUT

// [Census.gov](#) / [Data](#) / [Data Tools](#) / [Census Business Builder \(CBB\)](#)

NOTICE: Due to a lapse in federal funding portions of this website are not being updated. [Learn more.](#)

Data

Rick Mikulski, Portland State University
Online Northwest 3/29/2019
rmm8@pdx.edu

Abstract

From the 1790 Census to the present, the US government has been a major user, producer, and distributor of data. Through its agencies and departments, it creates data; through funding, grants, and data sharing mandates, it makes research data accessible; through the FDLP and various agency platforms, it circulates and stores data online. This presentation discusses the implications of Government's role in data creation, online access, and preservation. What are the potential strengths and weaknesses of this relationship, and how can librarians prepare for the future of data creation, preservation, and access? Special attention given to: Federal Data Mandates; Online Sources of Government Data; Challenges to Preservation and Access.

Overview: Government & Data

- History of relationship between government and Data creation/curation.
- What this relationship means for researchers and librarians.
- Challenges this relationship poses for preservation.
- The role of librarians in this relationship



SCIENCE

Topics, centers,
missions

PRODUCTS

Maps, data,
publications

NEWS

Releases,
I'm a reporter

CONNECT

Contact, chat,
social media

ABOUT

Organization,
jobs, budget

Search

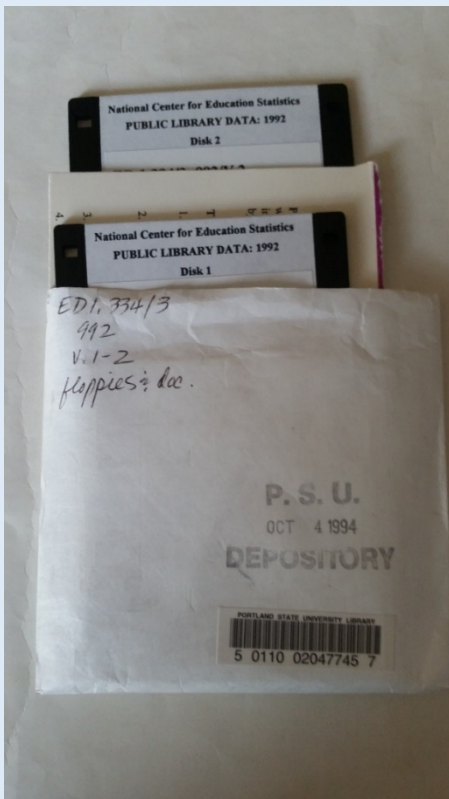


Due to a lapse in appropriations, the majority of USGS websites may not be up to date and may not reflect current conditions. Websites displaying real-time data, such as Earthquake and Water and information needed for public health and safety will be updated with limited support. Additionally, USGS will not be able to respond to inquiries until appropriations are enacted. For more information, please see www.doi.gov/shutdown

Publications



Thinking about the relationship



NCES National Center for Education Statistics MENU

Title: Data File (Public-Use): Public Libraries Survey, Fiscal Year 1992

Description: The Public Libraries Survey is conducted annually by the National Center for Education Statistics as part of the Federal-State Cooperative System for Public Library Data. The data are collected by coordinators appointed by the chief officers of state library agencies in the 50 States. Data are collected on population of legal service area, service outlets, public service circulation, circulation of children's materials, reference transactions, library visits, staff, operating income, operating expenditures, capital outlay, and more.

Online Availability:

- [Download, view, and print the documentation as a pdf file](#) (152KB)
- [Download the full zip file in MS-Access format](#) (2,077KB)
- [Download the full zip file in ASCII format](#) (1,837KB)

[Need Help Viewing PDF files?](#)

[Need Help Viewing Zip files?](#)

Cover Date: November 2002

Web Release: November 5, 2002

Print Release: This Publication will only be available online.

Publication #: NCES 2003308
[General Ordering Information](#)

Center/Program: NCES

Authors: Elaine Kroe

Type of Product: [Data File](#)

Keywords: Libraries
• [public](#)

Questions: For questions about the content of this Data File, please contact: [NCES WebMaster](#).

[Tell a colleague about this product!](#)

IES Institute of Education Sciences Search Go

Contents of the Data Management Plan

The DMP must provide a comprehensive overview of how the final research data will be shared, and should not exceed five pages. DMPs are expected to differ, depending on the nature of the project and the data collected. However, applicants should address the following in the DMP:

- Type of data to be shared;
- Procedures for managing and for maintaining the confidentiality of the data to be shared;
- Roles and responsibilities of project or institutional staff in the management and retention of research data;
- Expected schedule for data sharing, (no later than publication of findings in peer-reviewed publication and available for at least 10 years);
- Format of the final electronic dataset;
- Documentation to be provided;
- Method of data sharing (e.g., provided by the Principal Investigator, through a data archive)
- Whether or not a data sharing agreement that specifies conditions under which the data will be shared is required; and
- Any circumstances that prevent all or some of the data from being shared. This includes data that may fall under multiple statutes and hence must meet the confidentiality requirements for each applicable statute (e.g., data covered by Common Rule for Protection of Human Subjects, FERPA and HIPAA).

As part of the discussion of the roles and responsibilities, applicants should also consider changes that will occur should the Principal Investigator and/or Co-Principal Investigators leave the project or institution. In addition, since access to data is expected for 10 years, the DMP should describe how access to data will be supported after the grant ends.

The DMP will be included in an appendix in the grant application, and will not be counted towards the Project Narrative's limit of 25 pages. IES Program Officers will be responsible for reviewing the proposed DMP to ensure that it addresses all of the components listed under "Contents of the Data Management Plan." Once awards are made, IES Program Officers will be responsible for monitoring the DMP over the course of the grant period in regular monitoring activities (i.e., update calls and annual reports). The grantee will also describe plans for executing the DMP in the final report.

Timing of Providing Access to Data

Timely data sharing is important to the scientific process. IES thus expects that data will be shared no later than when the main findings

There is no going back.

Ticket closed by Ed Hill.

On Mar 04, 2019 @ 01:41 pm, Ed Hill wrote:

Ticket closed: Set up a Linux VM, installed DOSBox. Mounted the USB disc drive in the Linux VM (Virtualbox).

Made disc images of the floppy discs with ``dd bs=512 count=2880 if=/dev/fd0 out=/home/ed/disc1.img status=progress``

Made empty directory and mounted disc 1 with ``mount -o loop disc1.img /home/ed/install``.

Made empty directory for dosbox data.

In Dosbox mounted data directory to f: with ``mount f /home/ed/dosboxdata``

Mounted disc image to a: with ``mount a /home/ed/install -floppy``

When it asks, for disc 2, unmount disc 1 with ``umount ~/install`` and mount disc2 with same command as above.

Resume install.

Once install is complete, Excel should be able to open database files and Word should be able to open WordPerfect files.

For a full-on version of this, you would have something like:

- A (preferably **Linux**, to my mind) station with a floppy drive to make disk images of the floppies. Each disk only takes a couple minutes to image. Ideally this would also involve checking the files in the image for bit-level accuracy. These images would be duplicated to an archive and a copy for processing. Ideally, if you were doing this for reals real, you would have a forensic hardware block in place so that you couldn't accidentally overwrite the data on the disks, but that's not *strictly* necessary.

- A (again, preferably **Linux**) station with Dosbox. The workflow for multi-disk installs would be:

- 1 - Create a directory in **Linux** for dosbox data
- 2 - Create a directory in **Linux** to mount disk images. Mount image 1.
- 3 - In Dosbox, mount the data directory; mount the disk directory as -t floppy
- 4 - Change to the disk directory and initiate install
- 5 - For each requested disk image change, in **Linux** unmount the current disk and mount the next disk to the same **Linux** directory.
- 6 - Repeat 5 as needed until fully installed.

A [brief] history of Government & Data

The Return for SOUTH CAROLINA having been made since the foregoing Schedule was originally printed, the whole Enumeration is here given complete, except for the N. Western Territory, of which no Return has yet been published.

DISTRICTS	Free white Males of 16 years and upwards, including heads of families.	Free white Males under sixteen years.	Free white Females, including heads of families.	All other Free persons.	Slaves.	Total.
Vermont	22435	22328	40505	255	16	85539
N. Hampshire	36086	34857	70160	630	158	141885
Maine	24384	24748	46870	538	NONE	96540
Massachusetts	95453	87289	190582	5403	NONE	378787
Rhode Island	16019	15799	32652	3407	948	68825
Connecticut	60523	54403	117448	2808	2764	237946
New York	83700	78122	152320	4654	21324	340120
New Jersey	45251	41416	83287	2762	11423	184139
Pennsylvania	110788	106948	206363	6537	3737	434373
Delaware	11783	12143	22384	3899	8887	59094
Maryland	55915	51339	101395	8043	103036	19728
Virginia	110936	116135	215046	12866	292627	747610
Kentucky	15154	17057	28922	114	12430	73677
N. Carolina	69988	77506	140710	4975	100572	393751
S. Carolina	35576	37722	66880	1801	107094	249073
Georgia	13103	14044	25739	398	29264	82548
	807094	791850	1541263	59150	694280	3893635
Total number of Inhabitants of the United States exclusive of S. Western and N. Territory.	Free white Males of 21 years and upwards.	Free Males under 21 years of age.	Free white Females.	All other Free persons.	Slaves.	Total
S. W. territory	6271	10277	15365	361	3417	35691
N. Ditto	—	—	—	—	—	—

1790 Census



1960s IRS "Brain Machines"

A [brief] history of Government & Data

INDIANA UNIVERSITY BLOOMINGTON

Virtual CD-ROM / Floppy Disk Library


Home | List All | 1843 items | Next >>

About | Search | List All

1. "Dichte, Dichter, tage nicht!" die Europäische Schriftsteller-Vereinigung in Weimar 1941-1948
2. "Lepers outside the gate" excavations at the cemetery of the Hospital of St. James and St. Mary Magdalene, Chichester, 1986-87 and ...
3. "Viewing the violent universe" the Compton Gamma Ray Observatory
Contains computer-viewable version of a 5-minute video on the launch and early achievements of the Compton Gamma Ray Observatory.
NAS 1.86:UN 3/3/CD
4. 100-percent summary file Census 2000 dress rehearsal
Summary data files, profiles, and geographic products for Sacramento, California; 11 counties in South Carolina, including the city of Columbia
C 3.275:SU 6/CD
5. 108th Congressional district summary files census of population and housing, 2000
Provides census data for U.S. congressional districts. Summary tape file 1D provides complete population data by age, race, and sex and housing
Summary tape file 3D contains sample data weighted to represent the total population.
C 3.282/4:108
6. 1961-1990 global climate normals
"... a collection of DOS ASCII files which describe the 1961-1990 global standard climate normals for over 4000 stations worldwide computed by
C 55.281/2-2:G 51/2/CD
7. 1987 census of agriculture Volume 1 Geographic area series. State data files, county data files (including U.S. totals)
Contains summary statistics. Subject matter description based on data for all farms and for all farms with sales of \$10,000 or more; geographic
C 3.277:Ag 8/987/CD/v.1/rev.2
8. 1987 economic censuses
Files contain statistical data on retail trade, wholesale trade, service industries, transportation, manufactures, mineral industries, and constructi
C 3.277:EC 7/987/CD/V.1-C 3.277:EC 7/987/DOC/V.1-C 3.277:EC 7/987/FLOPPY 3.277:EC 7/987/FLOPPY 2-3C 3.2:Ec 7/11/989
9. 1987 full panel longitudinal studies
Provides statistical data related to income and government entitlement programs (AFDC, social security, veterans' benefits, etc.).
C 3.275:SU 7/CD
10. 1988 dress rehearsal census of St. Louis City, east central Missouri and eastern Washington
Contains three sets of products which include P.L. 94-171 data files, Summary tape file 1A data files, and the TIGER/line files for Boone County,
C 3.275:D 81
11. 1988 full panel longitudinal files
Provides statistical data related to income and government entitlement programs (AFDC, social security, veterans' benefits, etc.).
C 3.275:SU 7/2/994/CD
12. 1990 census of population and housing Ancestry of the population in the United States Subject summary tape file (SSTF) 2
C 3.286:CD 90 SSTF-02
13. 1990 census of population and housing characteristics of adults with work disabilities, mobility limitations, or self-care limitations Subject summ
Provides sample census data weighted to represent the total population, as well as 100-percent counts and unweighted sample counts for total
transportation.
C 3.286:CD 90 SSTF 04

topics in text

U.S. Census Bureau *the Official Statistics*



Your Source for Social, Demographic, and Economic Information

News

Search

About the Bureau

User Manual

Access Tools

CenStats CenStore

New on the Site

Subjects A-Z

Just For Fun

Current U.S. Population Count

Current Economic Indicators

[Other Official Statistics](#)

Last Revised: December 14, 1996

Between 1995 and 2002, GPO distributed 4,890 disks.

-Cruse, Patricia, Charles Eckman, John Kunze, and Heather Christenson. 2003. Web-Based Government Information: Evaluating Solutions for Capture, Curation, and Preservation. An Andrew W. Mellon Funded Initiative. Berkeley, CA :: California Digital Library. http://www.cdlib.org/services/uc3/docs/Web-based_archiving_mellon_Final.pdf.

2009 Open Government Directive & Data.gov

The screenshot shows the White House Open Government Initiative website. The header includes the White House logo and navigation links for Briefing Room, Issues, The Administration, and 1600 Penn. Below the header, there's a search bar and a breadcrumb trail: Home > The Administration > Open Government Initiative. The main content area is titled "Open Government Initiative" with the tagline "TRANSPARENCY + PARTICIPATION + COLLABORATION" and an "OPEN GOV" logo. A navigation bar lists "About Open Government", "Open Gov Blog", "Open Government Partnership", "Participation", and "Collaboration". On the left, a sidebar titled "ABOUT OPEN GOVERNMENT" lists "The Directive", "Policy", and "Working Group". The main content area features the "Open Government Directive" document, dated December 8, 2009, with ID M10-06. It is a memorandum for the heads of executive departments and agencies, issued by Peter R. Orszag, Director. The subject is "Open Government Directive". The text of the memorandum explains the President's instruction to the Director of the Office of Management and Budget (OMB) to issue an Open Government Directive, responding to the principles of transparency, participation, and collaboration. It notes that the Directive was informed by recommendations from the Federal Chief Technology Officer and public comment through the White House Open Government Initiative. The three principles of transparency, participation, and collaboration are highlighted as the cornerstone of an open government, promoting accountability, public contribution, and effectiveness of government.

Open Government Directive

This document is also available as pdf, txt, doc or view on Slideshare

December 8, 2009

M10-06

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Peter R. Orszag, Director

SUBJECT: Open Government Directive

In the Memorandum on Transparency and Open Government, issued on January 21, 2009, the President instructed the Director of the Office of Management and Budget (OMB) to issue an Open Government Directive. Responding to that instruction, this memorandum is intended to direct executive departments and agencies to take specific actions to implement the principles of transparency, participation, and collaboration set forth in the President's Memorandum. This Directive was informed by recommendations from the Federal Chief Technology Officer, who solicited public comment through the White House Open Government Initiative.

The three principles of transparency, participation, and collaboration form the cornerstone of an open government. Transparency promotes accountability by providing the public with information about what the Government is doing. Participation allows members of the public to contribute ideas and expertise so that their government can make policies with the benefit of information that is widely dispersed in society. Collaboration improves the effectiveness of Government by encouraging partnerships and cooperation within the Federal Government, across levels of government, and between the Government and private institutions.

This Open Government Directive establishes deadlines for action. But because of the presumption of openness that

The screenshot shows the Data.gov website. The header includes the "DATA.GOV" logo and navigation links for Home, Catalog, About, FAQ, Contact Us, and Suggest Other Datasets. The main content area is titled "DISCOVER. PARTICIPATE. ENGAGE." and features a search bar with the text "Search Data.gov catalog by category, agency, or both". Below the search bar are two dropdown menus: "All Categories" and "All Agencies". A "SEARCH" button is present, along with a link to "Click here for advanced search". To the right, there is a "FEATURED TOOL" section for the "BUREAU OF TRANSPORTATION STATISTICS" tool, "Airline On-Time Performance and Causes of Flight Delays". The description of the tool states: "The Airline On-Time Performance and Causes of Flight Delays table contains on-time arrival data for non-stop domestic flights by major air carriers, and provides such additional items as departure and arrival delays, origin and destination airports, flight numbers, scheduled and actual departure and arrival times, cancelled or diverted flights, taxi-out and taxi-in times, air time, and non-stop distance." Below the description is a "VIEW THIS TOOL" link. At the bottom, there is a "VIEW THE ENTIRE DATA CATALOG" link and a small graphic of a data storage device.

DATA.GOV

HOME | CATALOG | ABOUT | FAQ | CONTACT US | SUGGEST OTHER DATASETS

DISCOVER. PARTICIPATE. ENGAGE.

Search Data.gov catalog by category, agency, or both

All Categories

All Agencies

SEARCH (Click here for advanced search)

VIEW THE ENTIRE DATA CATALOG

FEATURED TOOL:

BUREAU OF TRANSPORTATION STATISTICS

Airline On-Time Performance and Causes of Flight Delays

The Airline On-Time Performance and Causes of Flight Delays table contains on-time arrival data for non-stop domestic flights by major air carriers, and provides such additional items as departure and arrival delays, origin and destination airports, flight numbers, scheduled and actual departure and arrival times, cancelled or diverted flights, taxi-out and taxi-in times, air time, and non-stop distance.

VIEW THIS TOOL


Data.gov is launched in 2009 as an informal partnership with various agencies. Beings with 47 datasets. As of 3/28/19 there are 236,345.

2013 Memorandum

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren 
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

1. Policy Principles

The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

Scientific research supported by the Federal Government catalyzes innovative breakthroughs that drive our economy. The results of that research become the grist for new insights and are assets for progress in areas such as health, energy, the environment, agriculture, and national security.

Access to digital data sets resulting from federally funded research allows companies to focus resources and efforts on understanding and exploiting discoveries. For example, open weather data underpins the forecasting industry, and making genome sequences publicly available has

“The Office of Science and Technology Policy (OSTP) hereby 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.”

Department of Agriculture,
Department of Commerce,
Department of Defense,
Department of Education,
Department of Energy,
Department of Health & Human Services,
Department of Transportation,
Department of Veterans Affairs,
Environmental Protection Agency,
National Aeronautics & Space Administration,
National Science Foundation,
Office of the Director of National Intelligence (ODNI),
Smithsonian Institution,
US Agency for International Development,
US Geological Survey

Thorny Issues

Public Data vs. Open Data

- “Public” can be requested or obtained.
- “Open” is readily accessible and should be machine readable.

Federal Data Vs Federally-funded Data

- “Federal Data” is produced/collected by the government.
- “Federally-funded Data” is *final research data* produced in a funded study.

Terminology

- “Final Research Data,” “timely,” “reasonable cost” not defined.

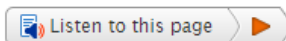
Data Management Plans

- b) Ensure that all extramural researchers receiving Federal grants and contracts for scientific research and intramural researchers develop data management plans, as appropriate, describing how they will provide for long-term preservation of, and access to, scientific data in digital formats resulting from federally funded research, or explaining why long-term preservation and access cannot be justified;
- c) Allow the inclusion of appropriate costs for data management and access in proposals for Federal funding for scientific research;
- d) Ensure appropriate evaluation of the merits of submitted data management plans;
- e) Include mechanisms to ensure that intramural and extramural researchers comply with data management plans and policies;
- f) Promote the deposit of data in publicly accessible databases, where appropriate and available;
- g) Encourage cooperation with the private sector to improve data access and compatibility, including through the formation of public-private partnerships with foundations and other research funding organizations;
- h) Develop approaches for identifying and providing appropriate attribution to scientific data sets that are made available under the plan;

H.R.4174 - Foundations for Evidence-Based Policymaking Act of 2017 (signed 1/14/19)

Summary: H.R.4174 — 115th Congress (2017-2018)

[All Information](#) (Except Text)



There is one summary for H.R.4174. [Bill summaries](#) are authored by [CRS](#).

Shown Here:

Introduced in House (10/31/2017)

Foundations for Evidence-Based Policymaking Act of 2017

This bill requires departments and agencies identified in the Chief Financial Officers Act to submit annually to the Office of Management and Budget (OMB) and Congress a plan for identifying and addressing policy questions relevant to the programs, policies, and regulations of such departments and agencies. The plan must include: (1) a list of policy-relevant questions for developing evidence to support policymaking, and (2) a list of data for facilitating the use of evidence in policymaking.

The OMB shall consolidate such plans into a unified evidence building plan.

The bill establishes an Interagency Council on Evaluation Policy to assist the OMB in supporting government-wide evaluation activities and policies. The bill defines "evaluation" to mean an assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency.

Each department or agency shall designate a Chief Evaluation Officer to coordinate evidence-building activities and an official with statistical expertise to advise on statistical policy, techniques, and procedures.

The OMB shall establish an Advisory Committee on Data for Evidence Building to advise on expanding access to and use of federal data for evidence building.

Open, Public, Electronic, and Necessary Government Data Act or the OPEN Government Data Act

This bill requires open government data assets to be published as machine-readable data.

Each agency shall: (1) develop and maintain a comprehensive data inventory for all data assets created by or collected by the agency, and (2) designate a Chief Data Officer who shall be responsible for lifecycle data management and other specified functions.

The bill establishes in the OMB a Chief Data Officer Council for establishing government-wide best practices for the use, protection, dissemination, and generation of data and for promoting data sharing agreements among agencies.

Overall Historical Trend

- Government has produced and collected data for centuries.
- Government has Computerized their data since the mid 20th century.
- Electronic government data circulated physically after 1980s via FDLP.
- Government data widely accessible online by late 90s
- Government data slowly centralized beginning in 2009.
- Government began setting data standards in 2013 through funding.
- Government requiring agency data to be machine readable in 2019.

```
VIEW/UPDATE DATA ENTRY SCREEN - ADMINISTRATIVE ENTITY Pg 1
IDENTIFICATION
01 LIB. ID#:003          1A FSCS. ID#:WY0002
02 NAME: KAMPBELL COUNTY PUBLIC LIBRARY
03 ADDRESS:5101 4-J ROAD          04 CITY:BILLETTE
4A County:                05 ZIP1:92716    06 ZIP2:0000
07 PHONE: (307) 682-3223
```

Part II

What does this relationship mean for librarians?

- Government agencies produce incredible amounts of data, which our users want/need to access.
- This centralization has resulted in an irreversible dependence.
- Government funding requirements is slowly setting standards.
- These standards are increasingly making librarians key to grant funding, as we assist with DMPs and archiving.

It is within this environment that Librarians must...

- 1) Provide access to data.
- 2) Work to preserve government data.
- 3) Assist researchers in preserving their own data.

Value

Assistant Librarian - Data Management
Assistant/Associate Librarian - Social Sciences and Data Services
Associate Librarian, Data Management Services
Asst/Assoc/Full Professor of Data Science
BUSINESS & DATA ANALYSIS LIBRARIAN
Business and Data Reference Librarian
Business, Economics, and Data Access Librarian
CLIR POSTDOCTORAL FELLOWSHIP IN DATA CURATION FOR THE SCIENCES AND SOCIAL SCIENCES
Data & Evaluation Coordinator
Data Acquisitions Librarian (Knowledge Analyst) - 21074
Data Analyst/Curator, Senior
Data Analytics & Insights Director
Data and Digital Scholarship Librarian
Data and Research Support Specialist, Strategy Unit,
Data Architect, Senior
Data Curation Librarian
Data Curation Librarian
Data Curation Network Project Coordinator (UL392)
Data Hunter
Data Librarian
Data Librarian
Data Librarian and Assistant/Associate Professor
Data Literacy Librarian
Data Management Analyst 2
Data Science Capability Lead
Data Science Specialization
Data Services and Digital Scholarship Librarian
Data Services Librarian
Data Services Librarian
Data Services Librarian
Data Services Librarian
Data Services Librarian / R98111
Data Services Specialist
Data Services Specialist
Data Visualization/Lab Manager
Data Wrangler
Director of Princeton Research Data Service
Enterprise Data Management Project Manager
Geospatial Data Services Postdoctoral Fellowship
Geospatial/Data Visualization Librarian
Head Librarian, Data & Assessment Services

IT Systems/Data Analyst
Librarian - Data Management (S01401P)
Librarian 2/Head of Cataloging and Metadata
Librarian for Finance, Accounting, & Business Data
Librarian for Research Data Management & Reproducibility
Librarian for Research Data Management and Reproducibility
Library Data Analyst
Library Data and Repository Analyst
Map and Geospatial Data Librarian
Metadata & Linked Data Strategist
Metadata & Linked Data Strategist
Open Source Data Scientist
Reference/Database Librarian
Research Data Librarian
Research Data Management Librarian
Research Data Management Librarian
Research Data Management Librarian
Research Data Specialist (RDS)
Research Data Specialist (RDS)
Science & Engineering Data & Information Specialist
Sciences Data Librarian
Senior Manager for Data Analytics - Public Library
Social Science Data Curator and GIS Librarian
Social Sciences and Data Librarian
Social Sciences Data and Sociology Librarian
Spatial Data Science Librarian
Third Party Data Asset Manager
WD171753 - Data Curator - Unstructured Data
WD177351 - Metadata Engineer, Big Data Solutions

Postings on ALA Joblist with “data” in the title
March 2018 – March 2019
(Collected by rmm for future project)

Providing Access to Government Data

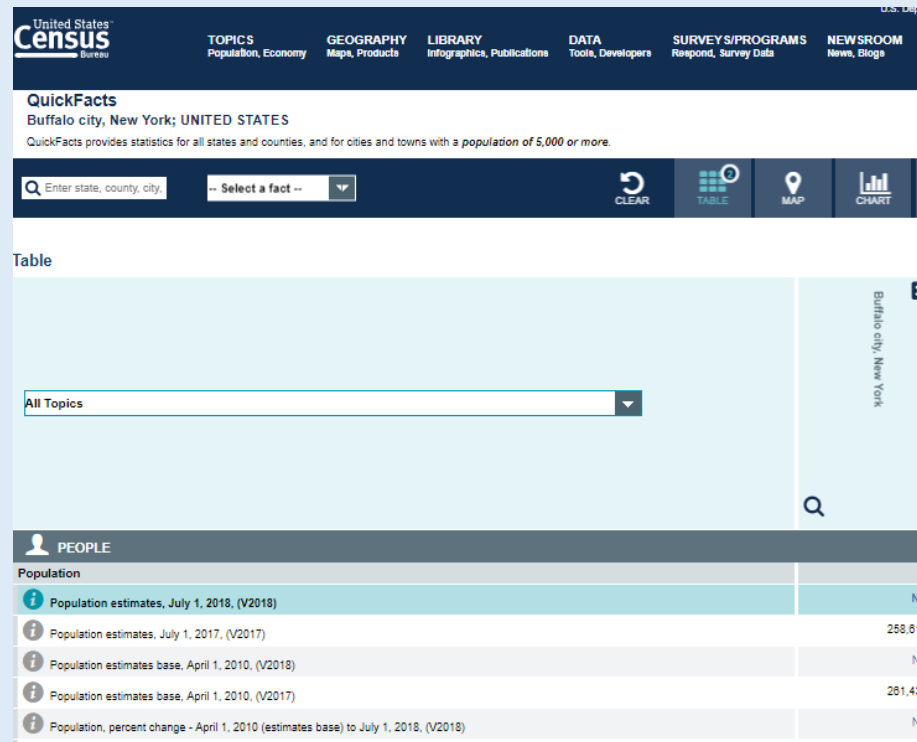


DATA.GOV DATA TOPICS IMPACT APPLICATIONS DEVELOPERS CONTACT

The home of the U.S. Government's open data

Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and [more](#).

GET STARTED
SEARCH OVER 236,345 DATASETS



United States Census Bureau

TOPICS: Population, Economy; GEOGRAPHY: Maps, Products; LIBRARY: Infographics, Publications; DATA: Tools, Developers; SURVEYS/PROGRAMS: Respond, Survey Data; NEWSROOM: News, Blogs

QuickFacts
Buffalo city, New York; UNITED STATES

QuickFacts provides statistics for all states and counties, and for cities and towns with a population of 5,000 or more.

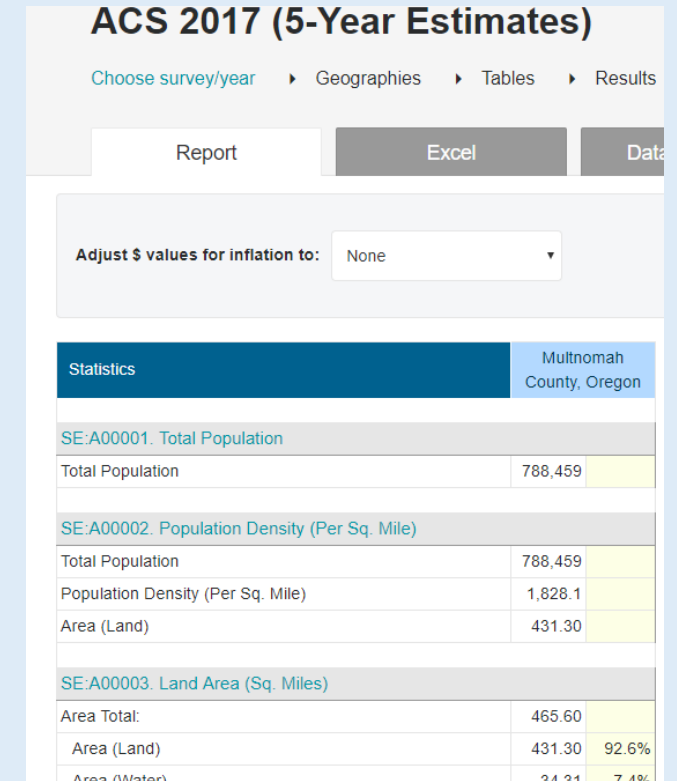
Enter state, county, city. Select a fact

Table

All Topics

PEOPLE

Population	
Population estimates, July 1, 2018, (V2018)	NA
Population estimates, July 1, 2017, (V2017)	258,612
Population estimates base, April 1, 2010, (V2018)	NA
Population estimates base, April 1, 2010, (V2017)	261,433
Population, percent change - April 1, 2010 (estimates base) to July 1, 2018, (V2018)	NA



ACS 2017 (5-Year Estimates)

Choose survey/year Geographies Tables Results


Report Excel Data

Adjust \$ values for inflation to: None

Statistics	Multnomah County, Oregon	
SE:A00001. Total Population		
Total Population	788,459	
SE:A00002. Population Density (Per Sq. Mile)		
Total Population	788,459	
Population Density (Per Sq. Mile)	1,828.1	
Area (Land)	431.30	
SE:A00003. Land Area (Sq. Miles)		
Area Total:	465.60	
Area (Land)	431.30	92.6%
Area (Water)	34.31	7.4%

(Social Explorer)

Efforts to Preserve Federal Data



[HOME](#) [OUR STORY](#)


Data Refuge is a community-driven, collaborative project to preserve public climate and environmental data. When we document the many ways diverse communities use data, we can also advocate for future data. We want to hear your data stories!



FREE GOVERNMENT INFORMATION
Because government information needs to be free

freegovinfo.info

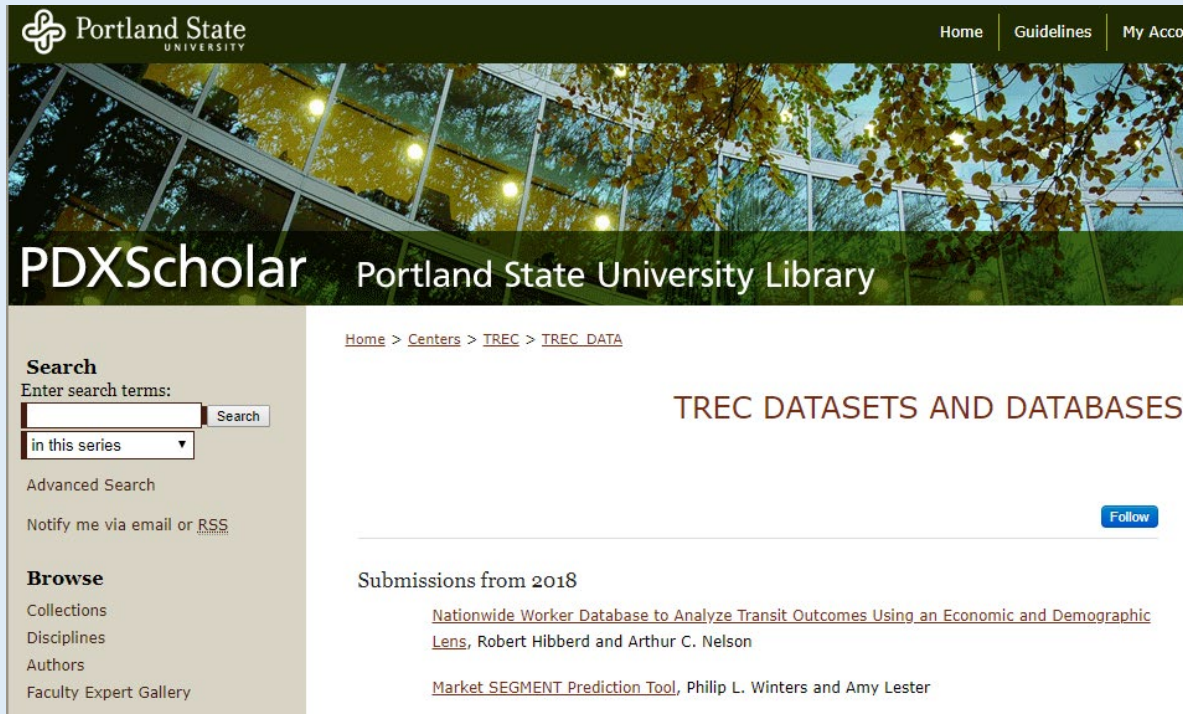
Because Government information needs to be free



Virtual CD-ROM / Floppy Disk Library

Home	Search Virtual CD-ROM/Floppy Disk Library	
About	<input type="text"/>	<input type="button" value="query"/>
Search		

Preserving Research Data



The screenshot shows the PDXScholar website interface. At the top left is the Portland State University logo and name. Navigation links for Home, Guidelines, and My Account are in the top right. A large banner image of a modern building with a glass facade is visible. Below the banner, the text 'PDXScholar Portland State University Library' is displayed. On the left side, there is a search section with a text input field, a 'Search' button, and a dropdown menu labeled 'in this series'. Below the search section are links for 'Advanced Search' and 'Notify me via email or RSS'. A 'Browse' section lists categories: Collections, Disciplines, Authors, and Faculty Expert Gallery. The main content area shows a breadcrumb trail: Home > Centers > TREC > TREC DATA. The title 'TREC DATASETS AND DATABASES' is centered, with a 'Follow' button to its right. Below this, a section titled 'Submissions from 2018' lists two items: 'Nationwide Worker Database to Analyze Transit Outcomes Using an Economic and Demographic Lens, Robert Hibberd and Arthur C. Nelson' and 'Market SEGMENT Prediction Tool, Philip L. Winters and Amy Lester'.



Library

Home → Help & Services → Data Management

Data Management

What is Data Management?

Data management covers all steps in the collection, processing, storing and sharing of research data, including:

- Processing and analyzing your data
- Verifying and documenting your data
- Data formats, naming, and organization
- Data archiving and sharing
- Ethics and attribution of data

The Library can help you with data management planning and archiving of research data; [Research Computing](#) can help with storage and processing needs while you are working on your project.

Data Management Services

Data Management Planning

My Dashboard Create plan

TITLE

Project Details

Plan overview

Write Plan

Share

Download

expand all | collapse all

0/6 answered

+ Data and Materials Produced (0 / 1)

+ Standards, Formats and Metadata (0 / 1)

+ Roles and Responsibilities (0 / 1)

+ Dissemination Methods (0 / 1)

+ Policies for Data Sharing and Public Access (0 / 1)

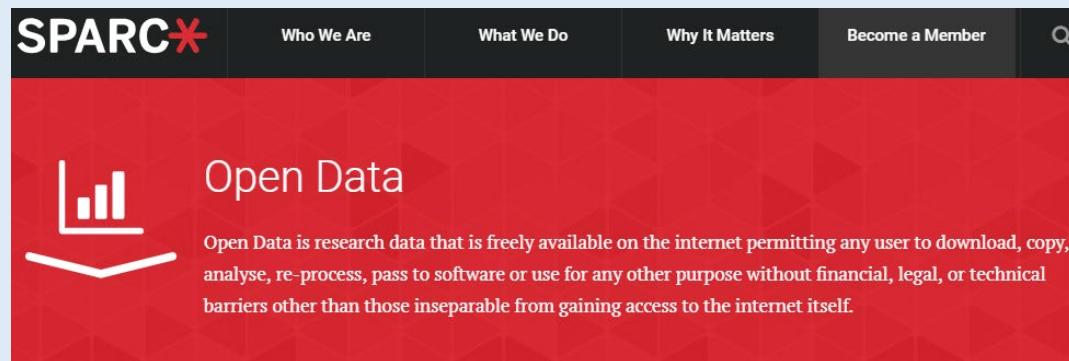
+ Archiving, Storage and Preservation (0 / 1)

<https://dmptool.org/>

Government policy directly shapes data curation, preservation, and access

- Through agencies, it produces data, which is increasingly accessible (only!) online.
- Through funding it is requiring federally-funded data to be freely accessible online.
- By implementing Data policies and setting the criteria, it is setting the standards.
- For better or worse, librarians and researchers are largely dependent upon the online infrastructure the government has set in place for continued access.

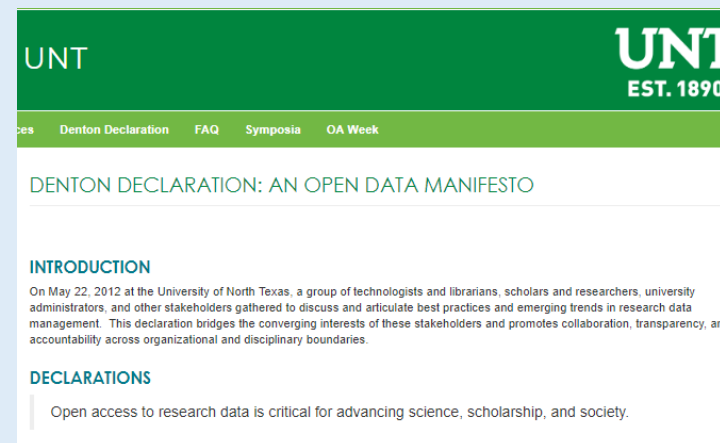
If we keep active, it'll be fine.



SPARC Who We Are What We Do Why It Matters Become a Member

Open Data

Open Data is research data that is freely available on the internet permitting any user to download, copy, analyse, re-process, pass to software or use for any other purpose without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.



UNT EST. 1890

es Denton Declaration FAQ Symposia OA Week

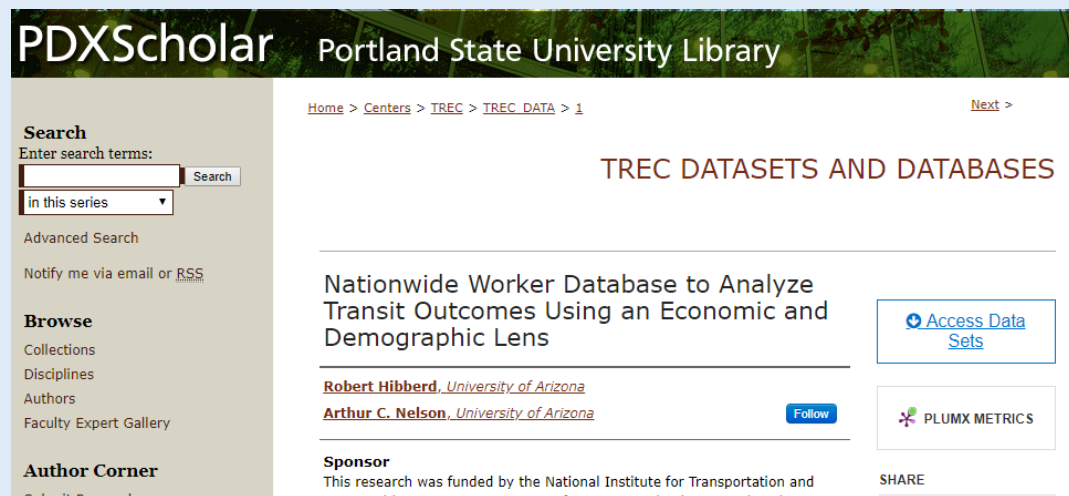
DENTON DECLARATION: AN OPEN DATA MANIFESTO

INTRODUCTION

On May 22, 2012 at the University of North Texas, a group of technologists and librarians, scholars and researchers, university administrators, and other stakeholders gathered to discuss and articulate best practices and emerging trends in research data management. This declaration bridges the converging interests of these stakeholders and promotes collaboration, transparency, and accountability across organizational and disciplinary boundaries.

DECLARATIONS

Open access to research data is critical for advancing science, scholarship, and society.



PDXScholar Portland State University Library

Home > Centers > TREC > TREC DATA > 1

TREC DATASETS AND DATABASES

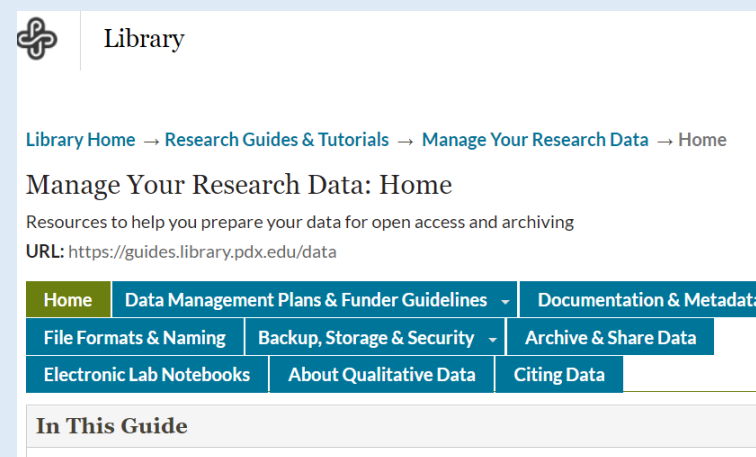
Nationwide Worker Database to Analyze Transit Outcomes Using an Economic and Demographic Lens

[Access Data Sets](#)

[Robert Hibberd, University of Arizona](#)
[Arthur C. Nelson, University of Arizona](#) [Follow](#)

Sponsor
This research was funded by the National Institute for Transportation and Community and Systems Research (NITCS) a program of TREC at Portland State University.

[SHARE](#)



Library

[Library Home](#) → [Research Guides & Tutorials](#) → [Manage Your Research Data](#) → Home

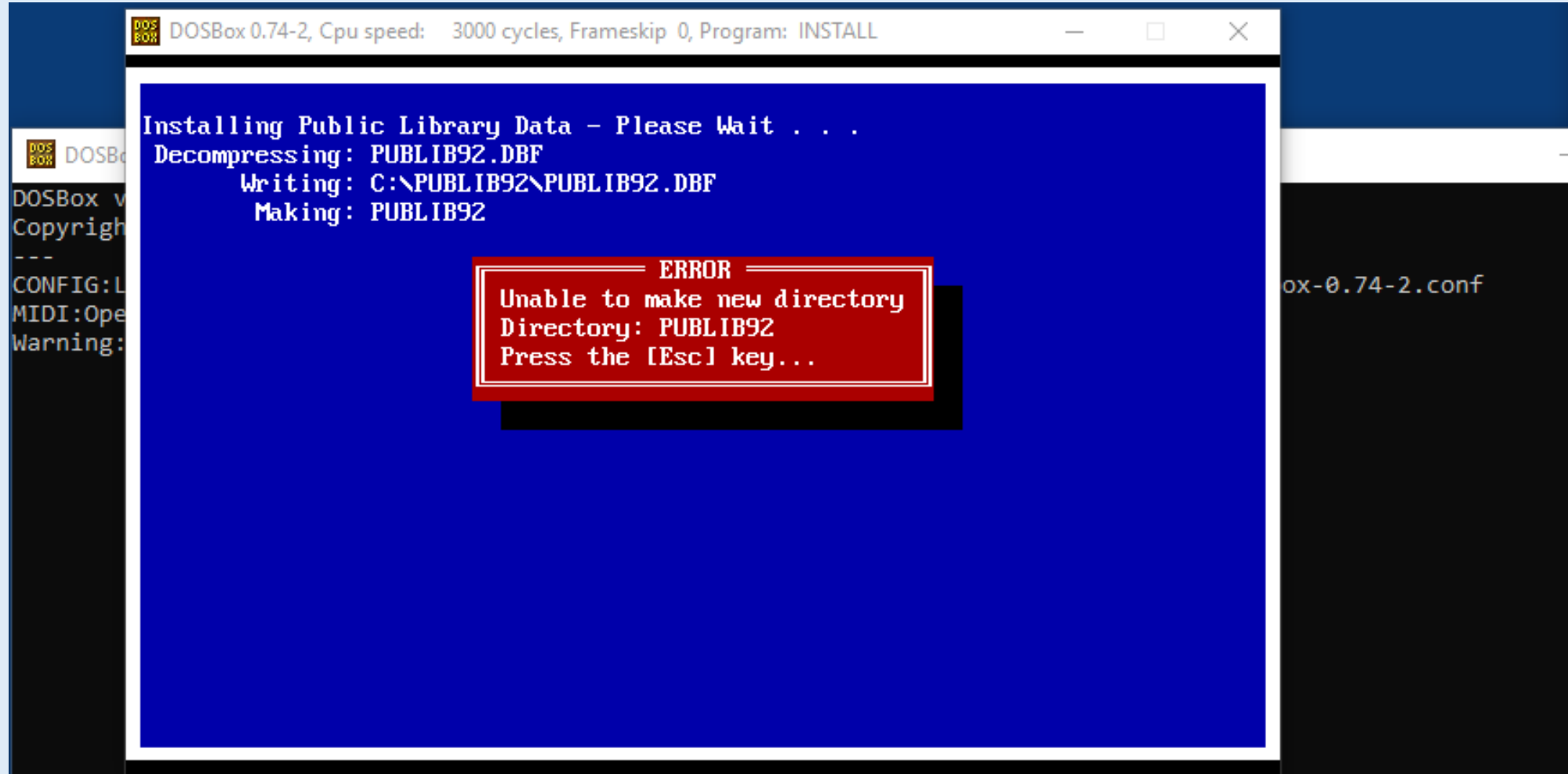
Manage Your Research Data: Home

Resources to help you prepare your data for open access and archiving
URL: <https://guides.library.pdx.edu/data>

Home	Data Management Plans & Funder Guidelines	Documentation & Metadata
File Formats & Naming	Backup, Storage & Security	Archive & Share Data
Electronic Lab Notebooks	About Qualitative Data	Citing Data

In This Guide

Questions?



Acknowledgements

Special thanks to Ed Hill and Sho Ikeda for their work and patience regarding my floppy disk antics. Colleagues on the Govdoc-L list (Susanne Caro, Jim Jacobs, Brooke Johnson, Julia Cater) similarly assisted in gathering material on archiving outdated FDLP data.

Selected References

<https://www.datarefuge.org/>

https://webapp1.dlib.indiana.edu/virtual_disk_library/

<https://shsulibraryguides.org/c.php?g=622180&p=5782944>

http://www.cdlib.org/services/uc3/docs/Web-based_archiving_mellon_Final.pdf

<https://library.pdx.edu/services/data-management/>

<http://guides.library.pdx.edu/data>

<https://freegovinfo.info/>

<https://dmptool.org/>

<https://openaccess.unt.edu/denton-declaration>

<https://sparcopen.org/open-data/>



Government and Online Data: Creation, Access, Preservation by Richard M. Mikulski is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).