Investing in Open Primary Sources

eDesiderata Forum 2018
Forum Agenda

Session I: Existing Open Access and Cooperative Investment Models
12:00-1:25 p.m. CT

Session II: Emerging and Potential New Models for Open Access
1:30-2:55 p.m. CT

Session III: Possibilities for Collective Action (Panel Discussion)
3:00-3:30 p.m. CT
Session I:
Existing Open Access and Cooperative Investment Models
Infrastructure for Social Science Research: the ICPSR Model and its Future

Libby Hemphill
Director, Resource Center for Minority Data and Associate Professor
University of Michigan
ICPSR advances and expands social and behavioral research, acting as a global leader in data stewardship and providing rich data resources and responsive educational opportunities for present and future generations.

ICPSR’s Mission Statement
Data Stewardship: What We Do

- Identify data with analytic utility and/or significance
- Assist data producers in preparing data for reuse
- Process, describe, and preserve valuable data
- Make data FAIR:
  - Findable
  - Accessible
  - Interoperable
  - Re-usable
Data Stewardship: Our Catalog

- 8,381 studies (70,434 datasets) available “on demand” to membership
- 3,090 studies available “on demand” to the public
- 1,671 studies (8,152 datasets) available as “restricted use”
- 1,329 studies have online analysis capabilities
- 280,642 data and documentation files
- 76,310 data-related publications
Data Stewardship: Open Data: Curated

Behavioral Risk Factor Surveillance System (BRFSS), 2003 (ICPSR 34085)

Published: Aug 5, 2013  Cite this Study | Share this study

Principal Investigator(s): United States Department of Health and Human Services, Centers for Disease Control and Prevention

https://doi.org/10.3886/ICPSR34085.v1

Version V1

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Preview</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1 Behavioral Risk Factor Surveillance System (BRFSS), 2003</td>
<td>928 MB</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes

- Data in this collection are available only to users at ICPSR member institutions.

This study is provided by Resource Center for Minority Data (RCMD).
Data Stewardship: Open Data: Crowdsourced

SOI Tax Stats - U.S. Population State and County Migration Data (1990-2016)

Principal Investigator(s): Internal Revenue Service (IRS)

Version: V3

<table>
<thead>
<tr>
<th>Name</th>
<th>File Type</th>
<th>Size</th>
<th>Last Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990-2011-County-Level-Data</td>
<td></td>
<td></td>
<td>03/02/2018 02:47:PM</td>
</tr>
<tr>
<td>1990-2011-State-Level-Data</td>
<td></td>
<td></td>
<td>03/02/2018 02:48:PM</td>
</tr>
<tr>
<td>2011-to-2012</td>
<td></td>
<td></td>
<td>03/02/2018 02:36:PM</td>
</tr>
<tr>
<td>2012-to-2013</td>
<td></td>
<td></td>
<td>03/02/2018 02:39:PM</td>
</tr>
<tr>
<td>2013-to-2014</td>
<td></td>
<td></td>
<td>03/02/2018 02:40:PM</td>
</tr>
<tr>
<td>2014-to-2015</td>
<td></td>
<td></td>
<td>03/02/2018 02:42:PM</td>
</tr>
<tr>
<td>2015-to-2016</td>
<td></td>
<td></td>
<td>03/02/2018 02:45:PM</td>
</tr>
<tr>
<td>SOI-Migration-Data-A-New-Approach.pdf</td>
<td>application/pdf</td>
<td>206 KB</td>
<td>03/02/2018 10:08:AM</td>
</tr>
<tr>
<td>U.S.-Population-Migration-Data-Strengths-and-Limitations.pdf</td>
<td>application/pdf</td>
<td>191.9 KB</td>
<td>03/02/2018 10:12:AM</td>
</tr>
</tbody>
</table>

Published Versions
V3 (2018-03-02)
V2 (2018-03-02)
V1 (2018-03-02)

Export
OAIFPMH
DD2.5
DD3.1

Project Citation:

Persistent URL: http://doi.org/10.3886/E101745V3

Report a Problem
Found a serious problem with the data, such as disclosure risk or copyrighted content? Let us know.
Data Stewardship: Open Data: OpenICPSR

City of Hamilton 2018 Municipal Election Issues Poll
Principal Investigator(s): Anthony Piscitelli; Michael Borrelli; Beasley Neighbourhood Association
Version: V2

<table>
<thead>
<tr>
<th>Name</th>
<th>File Type</th>
<th>Size</th>
<th>Last Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton-Beasley-Survey-Script.docx</td>
<td>application/vnd.openxmlformats-officedocument.wordprocessingml.document</td>
<td>15.4 KB</td>
<td>09/14/2018 04:16:PM</td>
</tr>
<tr>
<td>Hamilton-Beasley-Survey-Report-September-10-to-11-2018.doc</td>
<td>application/msword</td>
<td>610.5 KB</td>
<td>09/14/2018 04:29:PM</td>
</tr>
<tr>
<td>HamiltonBeasleyFullCampaignSurveyAnswers.csv</td>
<td>text/csv</td>
<td>16.1 KB</td>
<td>10/02/2018 03:37:AM</td>
</tr>
<tr>
<td>HamiltonBeasleySurveySept12to112018.sav</td>
<td>application/x-spss-sav</td>
<td>16.5 KB</td>
<td>09/14/2018 02:09:PM</td>
</tr>
<tr>
<td>HamiltonBeasleySyntax.sps</td>
<td>text/x-spss-syntax</td>
<td>5.1 KB</td>
<td>09/14/2018 02:09:PM</td>
</tr>
</tbody>
</table>

Published Versions
V2 (2018-10-03)
V1 (2018-10-03)

Export
DAI-PMH
DDI 2.5
DDI 3.1

Project Citation:

Persistent URL: http://doi.org/10.3886/E106542V2

Project Description
Research

- Digital archiving and preservation
- Data management practices
- Data re-use
- Metadata structures and standards
- Human subjects protection
- Disclosure risk
# Educational Opportunities

<table>
<thead>
<tr>
<th>Opportunity Type</th>
<th>Recent Offerings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1 Consulting</td>
<td></td>
</tr>
<tr>
<td>Hosted Series</td>
<td></td>
</tr>
<tr>
<td>Data Fair</td>
<td>2</td>
</tr>
<tr>
<td>Official Representative meeting</td>
<td>1</td>
</tr>
<tr>
<td>Present at Conference</td>
<td>21</td>
</tr>
<tr>
<td>Sponsored Workshop</td>
<td>16</td>
</tr>
<tr>
<td>Static or Persistent Resources</td>
<td>9</td>
</tr>
<tr>
<td>Summer Program</td>
<td>84</td>
</tr>
<tr>
<td>Webinar</td>
<td>23</td>
</tr>
</tbody>
</table>
ICPSR Governance and Structure

• Council
  – 12 elected members
  – 3 meetings/year
• Standing Committees
  – Budget and Policy
  – Collection Development
  – Membership and Marketing
  – Preservation and Access
  – Training and Instruction
• Advisory Committees
ICPSR Funding

<table>
<thead>
<tr>
<th>Source</th>
<th>Amount</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membership Fees</td>
<td>$4M</td>
<td>• General Archive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Infrastructure (including technology)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Strategic and capacity-building activities</td>
</tr>
<tr>
<td>Sponsored Projects</td>
<td>$8M</td>
<td>• Topical/specialized archives</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Research</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data tools and infrastructure</td>
</tr>
<tr>
<td>Summer Program Fees/Tuition</td>
<td>$2M</td>
<td>• Instructor salaries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Staff</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Facilities (including technology)</td>
</tr>
</tbody>
</table>
ICPSR Funding: Sponsored Projects

• Data
  • NACJD (BJS, NIJ, OJJDP)
  • Data Lumos (Casey)
  • Individual studies (Gates)

• Research
  • Common Data Elements (NIH)
  • Crowdsourcing Metadata Enhancements to Improve the Discoverability and Reusability of Scientific Data: Experimental Evaluations (NSF)
  • Researcher Credentialing (Sloan) and Open Badge Researcher Credentials for Secure Access to Restricted and Sensitive Data (NSF)

• Education
  • Summer Program in Quantitative Methods of Social Research
  • CyberTraining: Preparing the Public Sector Research Workforce to Impact Communities through Data Science (NSF)
ICPSR Funding: Opportunities

- Bibliography
- Research Impact Measures
- Social Media Archive (SOMAR)
- Cloud-based Virtual Data Enclave
- Researcher Passport
ICPSR Funding: Opportunities: Bibliography
ICPSR Funding: Opportunities: SOMAR
ICPSR Funding: Opportunities: Research Impact Measures
Discussion and Q/A

Contact information
ICPSR: help@icpsr.umich.edu
RCMD Director: Libby Hemphill  libbyh@umich.edu
Moving from Grant to Membership Models: the Biodiversity Heritage Library and the Medical Heritage Library

**Constance Rinaldo**
Librarian, Ernst Mayr Library of the Museum of Comparative Zoology, Harvard

**Melissa Grafe**
John R. Bumstead Librarian for Medical History, Cushing/Whitney Medical Library, Yale
Free & Open Access to Biodiversity Literature

Constance Rinaldo
Librarian of the Ernst Mayr Library of the MCZ, Harvard University, & Biodiversity Heritage Library Chair
“The cultivation of natural science cannot be efficiently carried on without reference to an extensive library.”

Charles Darwin, et al (1847)
MISSION

The Biodiversity Heritage Library improves research methodology by collaboratively making biodiversity literature openly available to the world as part of a global biodiversity community.
Biodiversity Literature/Archives = SCIENTIFIC DATA

- Names
- Specimen numbers
- Species descriptions
- Distribution records
- Climate records
- History of discovery
- Observations
- Illustrations/images
TAXONOMIC IMPEDIMENT

“Science is all about disseminating knowledge and building upon what has come before, yet so much of our knowledge of plants and animals has remained inaccessible to those who could make use of it. This has been a big part of the ‘taxonomic impediment.’”

Dr. John Sullivan
Evolutionary Biologist
BHL is a Global Consortium with a Membership model

20 MEMBERS

21 AFFILIATES

AS OF AUGUST 2018

80+ WORLDWIDE PARTNERS
MEMBERS

- American Museum of Natural History Library
- BHL Australia
- BHL México
- Cornell University Library
- Field Museum of Natural History Library
- Harvard University Botany Libraries
- Harvard University, Museum of Comparative Zoology, Ernst Mayr Library
- Library of Congress
- The LuEsther T. Mertz Library, The New York Botanical Garden
- Missouri Botanical Garden, Peter H. Raven Library
- Museum für Naturkunde Berlin
- Muséum National d’Histoire Naturelle
- National Library Board, Singapore
- Natural History Museum Library, London
- Royal Botanic Gardens, Kew, Library, Art & Archives
- Smithsonian Libraries
- United States Department of Agriculture, National Agricultural Library
- University Library, University of Illinois Urbana-Champaign
- University of Toronto Libraries
- Yale University
AFFILIATES

- Academy of Natural Sciences of Drexel University, Library and Archives
- BHL Africa
- BHL China
- BHL Egypt
- BHL SciELO (Brazil)
- Bibliothèque Cantonale et Universitaire - Lausanne
- California Academy of Sciences Library
- Canadian Museum of Nature
- Chicago Botanic Garden, Lenhardt Library
- Internet Archive
- Lloyd Library & Museum
- Los Angeles County Arboretum & Botanic Garden
- Marine Biological Laboratory/Woods Hole Oceanographic Institution Library (MBLWHOI Library)
- Mendel Museum
- Narodni Museum (National Museum, Prague)
- Natural History Museum Los Angeles County
- Naturalis Biodiversity Center
- Oak Spring Garden Foundation
- Smithsonian Institution Archives
- United States Geological Survey Libraries Program
- University Library Johann Christian Senckenberg
FUNDING AND SUSTAINABILITY

• Federal Funding
• Member and Affiliate Dues
• Institutional Endowments
• Donations
• Product Development
• Institutional Subventions
• In-Kind Contributions
• Grants**

GOVERNANCE

• Members’ Council
• Executive Committee
• Secretariat
• Technical Advisory Group
Users, Content and Outreach

• 7+ million users to date
• Visits from 243 countries/territories
• 54+ million pages, 231,000+ volumes
• 184 million taxonomic name instances
• 780+ in-copyright titles with permissions
• 108,000+ social media followers
  • 48,000+ Facebook
  • 19,500+ Twitter
  • 15,500+ Flickr with 124,000+ images
• Tools: DOI assignment, taxonomic name searching, PDF downloads, Article indexing, APIs/Data export
• Citizen science participation
Biodiversity Heritage Library

Hortus canalius
21 photos • 718 views

Botanique (Atlas)
66 photos • 822 views

An atlas of the fossil conchology of Great Britain and Ireland with descriptions of all the species
115 photos • 1,027 views

Herbier General de l'Amateur. Tome 7 (1824)
72 photos • 2,921 views
PARTNER ENGAGEMENT

• Annual Meeting: All partners
• Governance participation: All partners
• Committees/calls: all partners
  • members, staff, collections
• Outreach: All partners
• Grant opportunities: All partners
• Monthly statistics view: members only
• Training: all partners
Next Generation BHL: Data Extraction and Connections

• Full text search launched
• Completed transcriptions underway
• More flexible metadata model
• Image search
• Better name search
• Better taxonomic name backbone
• Connection to specimens
• Opportunities for crowdsourcing
CHALLENGES

• Continued **SUSTAINABILITY**
• **FUNDING** new initiatives
• Keeping partners **ENGAGED**
• Moving partners to **MEMBERSHIP**
• **CAPACITY** for supporting new partners
Melissa Grafe
President, Medical Heritage Library, Inc.
Head, Medical Historical Library, Yale University
Medical Heritage Library, Inc.

www.medicalheritage.org/

Under construction -- please pardon our dust!

Search the MHL now!

Collaborate

Create

Content

The Medical Heritage Library, Inc. is an online research, education, and digitization collaborative that provides the means by which readers and scholars across a multitude of disciplines can examine the interrelated nature of medicine and society, both to inform contemporary medicine and to strengthen understanding of the world in which we live.
Mission: Digitization and Discovery

• Founded in 2010.

• Content-centered curated digital collaborative and community

• Supporting research, education, and dialogue
  • Examine the interrelated nature of medicine and society, both to inform contemporary medicine and to strengthen understanding of the world in which we live.
Membership Model

• Incorporated in May 2018, 501(c)3 status

• 9 Principal members with governance and voting rights

• 30+ additional contributors with MHL tagged collections in Internet Archive.

• Scholarly advisory committee
Who are we?

Principal Members:
- Augustus C. Long Health Sciences Library at Columbia University
- Bibliothèque interuniversitaire de Santé (BIU Santé)
- The College of Physicians of Philadelphia
- The Cushing/Whitney Medical Library at Yale University
- The Francis A. Countway Library of Medicine at Harvard University
- The New York Academy of Medicine
- UCSF Library and Center for Knowledge Management
- U.S. National Library of Medicine and NIH Library, Otis Historical Archives, etc.
- Wellcome Library and UK-MHL partners

Contributors:
- Health Sciences and Human Services Library, University of Maryland, the Founding Campus
- United States Navy’s Bureau of Medicine and Surgery’s Office of Medical History
- University of Toronto, Gerstein Science Information Centre
- University of Toronto, Dentistry Library (Harry R. Abbott)
- University of California, San Francisco, Legacy Tobacco Documents Library
- University of Illinois at Chicago, Library of the Health Sciences, Special Collections
- United Kingdom Medical Heritage Library
- The Welch Medical Library, Library of the Institute of the History of Medicine, and the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions
- More at: http://www.medicalheritage.org/about/
Funding Sources

Membership:
- Principal Membership dues of $4000/year
- In-Kind Contributions

Membership dues fund part-time administrative support and a new fellowship program, as well as annual expenses.

Members have contributed some money from institutional endowments, in support of projects, in the past.

Grants:
- Alfred P. Sloan Foundation
- Council on Library & Information Resources
- Andrew W. Mellon Foundation
- National Endowment for the Humanities
- National Network of Libraries of Medicine
Membership Engagement

• Annual Meeting: all principal members, opened to contributors and advisors
• Governance participation: All principal members
• Committees: principal members
• Outreach: Principal members and contributors
• Grant opportunities: All principal members, contributors, and we reach out to collections beyond MHL
Scope of the MHL

• Over 266,000 items in Internet Archive, and growing!!:
  https://archive.org/details/medicalheritagelibrary

• Text, audio-visual, audio
  – Journals, monographs, pamphlets, ephemera, video
• Circa 1200-2000s
• Primarily in English, but also French, German, Spanish, Latin, Portuguese...
By the numbers

73,599,986 downloads via Internet Archive

5,620 Twitter followers
1,800 Facebook followers
Searchability: Internet Archive

The Medical Heritage Library

The Medical Heritage Library (MHL), a digital curation collaborative among some of the world's leading medical libraries, promotes free and open access to quality historical medical resources.
Welcome to the Medical Heritage Library Search

Description
The purpose of this tool is to allow you to search the full text of one or more items, all at once. Each page of each book has been digitized and software has been used to extract and index each word that occurs on every page, thus the term "full text". The MHL searches the full text and the descriptive information (metadata) compiled for each item (such as title, author, publication date, and subject terms). As with any digitization project, an OCR error rate of 1% is expected. Factors affecting the error rate include non-Roman typefaces and font size. By comparison, the Internet Archive search allows full text search on one item at a time. It does not offer simultaneous full text search of multiple items within the MHL. Additionally, the Internet Archive does not index every page of every text. For this reason, search results may vary between the two tools. The results from this tool are more comprehensive.

What you will and won’t find
In addition to OCR errors, results are influenced by the completeness of the descriptive information (metadata) provided by the contributing libraries. We know, for example, that for 4,000 items in the MHL, no date has been provided. There is no full text for video items. Serial publications present special problems: volume numbers and volume dates are not always complete or uniformly formatted. For each volume a second date may appear (“Date”). This indicates the first year of publication.

How to Search
Each field is labeled with the behavior it applies:
Reaping images out of books

MHL Inc. Flickr (we also have a sandbox site!)
In the not too distant future...

Millions of images...
Challenges

• Infrastructure and capacity as the MHL continues to scale up.

• Rethinking membership and contributor models

• And like BHL, funding new initiatives and keeping members engaged.
Discussion and Q/A

Contact information

Constance Rinaldo  crinaldo@oeb.Harvard.edu
@BioDivLibrary
Melissa Grafe  Melissa.grafe@yale.edu
medicalheritage@gmail.com
Supporting Open Access at the Grassroots: USAIN and Project Ceres

Michael Cook  
Head of Collections, Mann Library, Cornell University

Amy Paster  
Head, Life Sciences Library, Pennsylvania State University
A NATIONAL PRESERVATION PROGRAM
FOR
AGRICULTURAL LITERATURE

By
Nancy E. Gwinn
Assistant Director, Collections Management
Smithsonian Institution Libraries

with Assistance from the
Advisory Panel on Preservation
U.S. Agricultural Information Network

May 16, 1993

National Preservation Plan
A NATIONAL PRESERVATION PROGRAM FOR AGRICULTURAL LITERATURE

By

Nancy E. Gwinn
Assistant Director, Collections Management
Smithsonian Institution Libraries

with Assistance from the
Advisory Panel on Preservation
U.S. Agricultural Information Network

May 16, 1993

NATIONAL PRESERVATION PLAN FOR AGRICULTURAL LITERATURE

CORE HISTORICAL LITERATURE
(Popular and Trade Journals)
University-based nationally coordinated

FEDERAL DOCUMENTS
NAL

MANUSCRIPTS AND ARCHIVES
Local Initiatives

UNIQUE COLLECTIONS
- subject-based
- geography-based
- format-based
- etc.
Local Initiatives

NON-PRINT AND AUDIOVISUAL COLLECTIONS
Local Initiatives

PRE-1962 IMPRINTS
NAL-ass project

STATE AND COUNTY DOCUMENTS
State-based, nationally coordinated

LAND-GRANT PUBLICATIONS
Assess status, fill in gaps

NAL areas
State & Local Literature
1993  NPPAL approved at USAIN Auburn Conference.

1994  NPPAL steering committee established.

1995  NEH funds “Preserving the History of United States Agriculture and Rural Life” Phase I.

1996-1998  Phase I participants: AL, CA, CT, FL, NE, NY, PA, TX, WI

1998-2000  Phase II participants: AZ, AR, CA, FL, HI, IA, MN, MT, NE, TX

2000-2002  Phase III participants: AR, IA, KS, MI, MN, NC, ND, NM

2002-2004  Phase IV participants: GA, IL, MI, NC, OH

2004-2006  Phase V participants: CO, IL, MD, OH, OK, WA

2006-2008  Phase VI participants: AZ, CO, KY, MD, SD, WA

2008  NEH funding ends. Phase VII unfunded.
Agricultural Network Information Collaborative
Global Resources Agriculture Partnership—Project Ceres

In 2012, CRL formed a partnership with the United States Agriculture Information Network (USAIN)® and the Agriculture Network Information Collaborative (AgNIC) to support the ongoing preservation and digitization of collections in the field of agriculture. Project Ceres has two primary goals:

1. Sustain consensus-based, cooperative archiving of primary serial collections in the field of agriculture, defined broadly;
2. Expand electronic access to digital and print resources from all world regions to support agricultural research.

The Collection
A list of complete Project CERES Titles is available through CRL's online catalog.

Awards
Project Ceres
THE GUERNSEY! A BUTTER BREEDER.
SOME CHARACTERISTICS OF THE BREED.
Where It Came From; What It Is Good For.

It is pretty well understood by all who know anything about cattle, that the Jersey and Guernsey lighter and more graceful Jerseys ran in their herds. The Guernsey men passed the strictest laws—absolutely prohibiting the importation of breeding animals. The ship that landed such an animal was to be forfeited, with imprisonment of master and crew. The result was that the Guernseys remained as pure sterling qualities have been retained and slowly developed without any loss due to chance or speculative breeding. As between average specimens of the two breeds, the Guernsey is the healthier and hardier animal. She is also the better barn cow, and will give better results on the care usually given the

Project Ceres
A rich trove of once-hidden Cornell agricultural research from journals and other serial publications is now available to the public.

Cornell University Library collaborated with Kathryn J. Boor ’80, the Ronald P. Lynch Dean of the College of Agriculture and Life Sciences (CALS), to release copyright restrictions on more than 1,700 documents published between 1880 and 1996. This material – produced by scholars at CALS, including the Cornell University and New York State Agricultural Experiment Stations, as well as Cornell Cooperative Extension – can now be searched, downloaded and read on HathiTrust, a digital library of more than 15 million volumes, of which Cornell University Library is a partner.
Pomology: Apples and Cider

Pomology is the study of and cultivation of fruit. Cider (sometimes called hard cider) is a fermented beverage made from apple juice. This comprehensive collection highlights the historical importance of apple (Malus ×domestic and other Malus species) and cider production around the world. Many texts contain color illustrations of rare, historic, and heirloom apple cultivars. Topics in this collection include the following:
Discussion and Q/A

Contact information

Michael Cook  mnc2@cornell.edu
Amy Paster    alp4@psu.edu
Stay tuned for:
Session II  Emerging and Potential New Models for Open Access