

r(E)volution within





Margaret Levenstein

I celebrated my one-year anniversary as director of ICPSR in July 2017. It has been a busy and rewarding year. I have learned so much from the dedicated ICPSR staff, which has tackled challenges we face with enthusiasm and energy. In looking over the past year, we have done so much! Reorganizing our data curation team is a big and ongoing step to prepare us to do better what has long made ICPSR so special: our data is understandable, findable, and easy to use now and for posterity. In the midst of all the organizational transition, the curation team has continued to process and release data at a rapid pace. In the past fiscal year, we added or updated 677 studies. In addition, openICPSR published 175 studies. We launched a newly designed website, and our web and social media team has been sharing our work with our 10,000 friends and followers. Have you looked at the [ICPSR YouTube channel](#) recently? Be sure to watch “What is Data Curation?” and What are Metadata?”

Left: ICPSR Director Margaret Levenstein gives remarks during a reception welcoming her to ICPSR in July 2016. (Photo by David Thomas)

“The launch of Archonnex, ICPSR’s new data asset management system ... will provide us a platform to host new types of data and interact with our user community in new ways.”

The launch of Archonnex, ICPSR’s new data asset management system, is also a big and ongoing step for the organization that will provide us a platform to host new types of data and interact with our user community in new ways.

During the past year, we launched [DataLumos](#) and built the [Open Data Flint](#) collaborative. We received new awards from the National Science Foundation, the National Institutes of Health, and the Sloan, Arnold, and Gates Foundations. I made 29 out-of-town trips in the last year, one third of which were to our nation’s capital. While I could do without all the time in the air, these trips have given me a chance to learn more about how the broader community views ICPSR. I never make a trip without someone telling me how much they enjoyed some long ago summer in Ann Arbor at the ICPSR Summer Program. These interactions also regularly remind me of the important role ICPSR plays in articulating the values of data sharing and access to both scientific progress and a well-functioning democracy.

So, let’s celebrate the contributions that everyone in the ICPSR community makes to building and sustaining those values as ICPSR continues to revolutionize what it means to be an international data steward.



Chandra Muller


It has been an exciting year at ICPSR. Long-run investments are bearing fruit. Our new director, Maggie Levenstein, has introduced new and already successful initiatives while updating the organization to better serve member needs. Several staff leadership positions have been filled with highly accomplished people. Yet, even positive developments can bring demands to adjust, fill in, and adapt in ways that can be taxing.

The dedication of the staff during this period of transition has been inspiring. They are a highly professional group of individuals with valuable skills, flexibility and positive attitudes. They've kept ICPSR running the core activities—a vibrant summer program, data acquisition, curation, and archiving, and various other member services and outreach—that are so valuable to students and researchers at our member institutions. Council members have had a front row seat to observe and applaud these developments. Maggie Levenstein spent considerable effort listening to input from stakeholders in the ICPSR community about what is working, what is not, and ideas for development. Her efforts are paying off. She has completed her first year as director with outstanding successes, and ICPSR's future is bright.

Left: The ICPSR Council meets in Ann Arbor in October 2016. *(Photo by Dory Knight-Ingram)*



Council members were enthusiastic about the expansion of ICPSR's branded archives. DataLumos was developed to preserve and archive government data. Open Data Flint is an innovative development to provide a space for archiving all available data on the water crisis in Flint, Michigan. Branded archives for journals, such as those published by the American Educational Research Association and Journal of Economic History, provide authors with an excellent repository as journals and professional organizations encourage or even require the archiving of data used in articles. When other scholars, including students, can access the data it enhances the impact of the published work and builds an enduring infrastructure of knowledge. These activities are coupled with ICPSR's leadership to ensure that scientists get credit for producing data resources in addition to the credit for published articles. ICPSR offers scholars at member institutions tailored services for data management and sharing, such as those required by many federal grants. ICPSR will heighten the impact of their research by improving quality and efficiency of data archiving.



"Council members were enthusiastic about the expansion of ICPSR's branded archives."

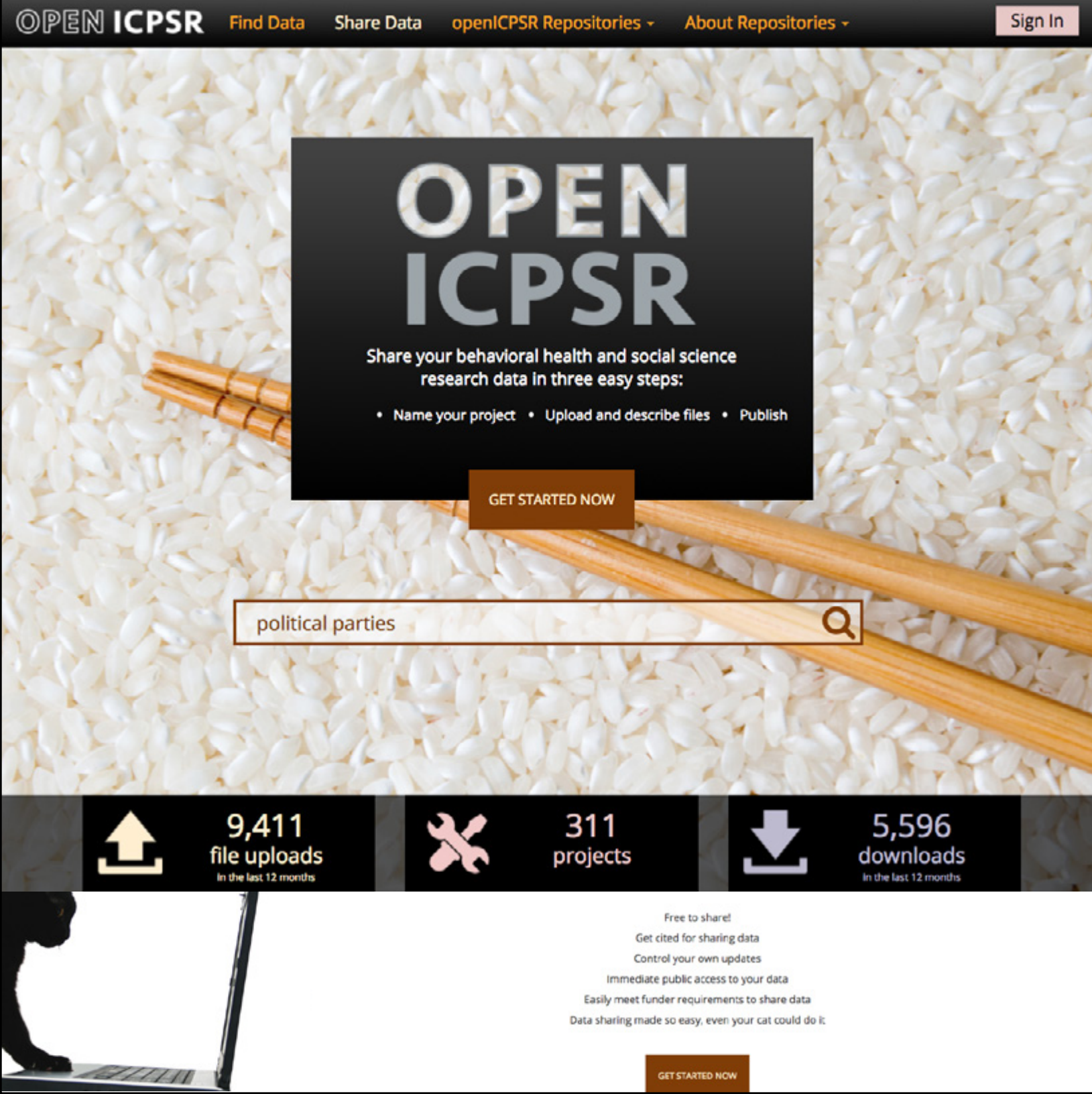
With all the development and success, it is easy to lose track of the challenges ICPSR staff members have faced this past year. They have had to compete for two renewals to fund major archives, have written many additional grants to continue or develop new initiatives, and they must innovate to serve members. The good news is that they have been highly successful. As we all know, data science is an expanding and rapidly changing area. Increasingly, people are learning data and computation sciences at earlier ages and in more diverse settings, and ICPSR's data and teaching resources are in high demand. This comes with both opportunity to expand to new user groups, but also challenges in adapting to the changing environment. We discuss these successes, challenges, and ideas for future initiatives at Council meetings. Please feel free to contact Council members with your own ideas and insights. Our goal is to represent member organizations.



Left: ICPSR staff and Council members gather at the Council meeting in June 2016 in Ann Arbor. Pictured from left are Susan Leonard, Phillip Jefferson, Michael Shove, Tom Murphy, and Robert Stine. (Photo by Dory Knight-Ingram)



The coming year is a transition for half of the Council. We owe many thanks to the valuable insights and dedication of Robert Chen (Columbia University), Lisa Cook (Michigan State University), Colin Elman (Syracuse University), Philip Jefferson (Swarthmore College), and Ronald Nakao (Stanford University) who will be ending their terms on Council. These members have been vital in leading the Council during a transitional period that included a change in ICPSR leadership and other challenges that required reflection and understanding of the short-term considerations and long-run vision for the organization. I will also be ending my term as Chair. As I reflect on the privilege of serving in this position, I am not only grateful for the dedication of my fellow Council members, but also the staff and leadership of ICPSR. Observing Maggie Levenstein's leadership has been personally inspiring. She is authentic and genuine, original, creative, and dedicated to excellence. In only one year she has a record of many accomplishments. But most of all, and speaking for the entire Council, what is most valuable are the member organizations. The communities of data enthusiasts that each member organization brings to ICPSR produces a vibrant resource for us all.



NEW INITIATIVES

openICPSR 2.0 pilots new data management platform, Archonnex

In August 2016, we launched openICPSR 2.0, ICPSR's data preservation option for researchers who want to self-publish research data, and for institutions desiring to build and brand their own data repositories. openICPSR 2.0 was built entirely on the new data science management platform, Archonnex. Learn more about using or sharing data via openICPSR, or how to start an institutional repository at openicpsr.org.



ICPSR celebrates launch of new data deposit system

Our new Deposit Manager, a pivotal step in the development of ICPSR's new digital asset system, Archonnex, launched in the summer of 2016. Among the benefits to data depositors:

- Customizable workspaces
- More control over the presentation of data and documentation files
- More project status communication
- More enhancements will roll out over the summer

On June 20, an ICPSR team gave a live demonstration of ICPSR's new Deposit Manager. A [recording of the webinar](#) can be found on our YouTube channel.

New agreement with DPN further enhances ICPSR's commitment to long-term data preservation

A new agreement with the [Digital Preservation Network \(DPN\)](#) further enhanced ICPSR's commitment to long-term data preservation. ICPSR is a longtime member of the Data Preservation Alliance for the Social Sciences ([Data-PASS](#)), which recently joined the Digital Preservation Network (DPN). Data-PASS members who can now deposit data into DPN, a membership organization dedicated to developing the services, tools and infrastructure needed by research and cultural heritage institutions for the long-term preservation of digital content. DPN has also developed the legal agreements that enable succession and protect against institutional failure.

ICPSR Summer Program announces Heitowit Scholarship for past Summer Program participants

Just in time to benefit 2017 Summer Program participants, a [scholarship fund](#) was established in honor of Henry “Hank” Heitowit, who served as the Director of the ICPSR Summer Program from 1981 to 2007. The Heitowit Scholarship will enable a former participant to return to the ICPSR Summer Program. Hank Heitowit played a major role in making the Summer Program the leading resource for methodological training across the social, behavioral, and health sciences. During his tenure as Director, Hank Heitowit worked tirelessly to provide an optimal experience for participants, maintain the scholarly excellence and high-quality of instruction, and enhance diversity across the Summer Program.

Left: Participants take notes during the 2017 workshop “Regression Discontinuity Designs,” taught by Rocio Titiunik. (Photo by Stephanie Carpenter)



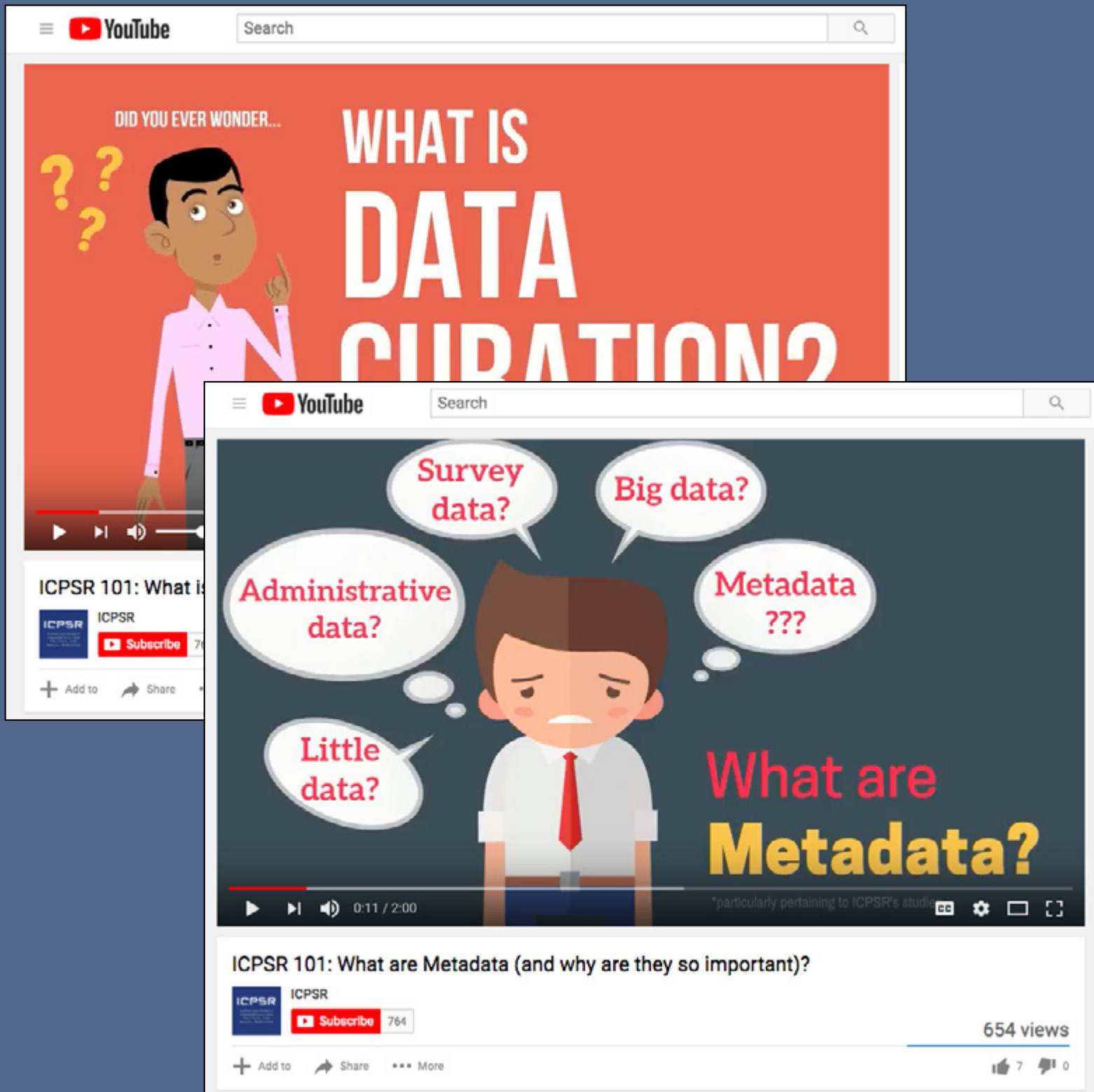
DataLumos launches

We officially announced the launch of [DataLumos](#) during [Love Your Data Week](#) with a [webinar](#) that gave an overview and answered viewers' questions. DataLumos is an ICPSR archive for valuable government data resources. It accepts deposits of public data resources from the community and recommendations of public data resources that ICPSR itself might add to DataLumos.



ICPSR 101 video series begins

We launched a fun and informative video series called ICPSR 101, with the first videos being “What is Data Curation,” and “What are Metadata (and why are they so important)?” More videos are planned, with topics including restricted-use data and informed consent.



New look for ICPSR website

Our redesigned website went live in August 2016. The revamped home page was designed by our web team to make it simpler for our millions of site users to find or share research data. Updates include full site navigation from each page of the website, hover menus, and the capability for website users to conduct ICPSR website and data searches in one convenient place.



Working together to make Greater Flint Region data available



Healthy Flint Research Coordinating Center
OPEN DATA FLINT

Data • Publications • Share Data • About Us • ODF for -



Open Data Flint serves as a central location for housing Flint-based data gathered from academic institutions, local organizations, and federal agencies to encourage a healthier and informed community. The data are made available to researchers, policy makers, general public, and others interested in Flint.



Healthy Flint Research
Coordinating Center

A collaborative partnership with CBOP/CERB, ICPSR, University of Michigan-Flint, University of Michigan-Ann Arbor, and Michigan State University

INITIATIVES cont'd

Open Data Flint launches

[Open Data Flint](#) made a big splash during its debut in the fall of 2016. Our Data Acquisitions Director Amy Pienta traveled to Washington to showcase the new archive at the first annual White House Open Data Innovation Summit. Open Data Flint is a community and multi-university partnership to provide open access to data on the Flint community. Open Data Flint is a project of the Healthy Flint Research Coordinating Center, which is a collaboration between community leaders, community-based organization partners, and researchers at the University of Michigan-Ann Arbor and -Flint, and Michigan State University. Open Data Flint is a public data repository that aims to assist the community of Flint, MI, bring together data to help build the evidence base to achieve a healthier Flint community. The archive also aims to help the community gain a deeper understanding of the far-reaching impact of the water crisis on the Flint population.

C²METADATA

Continuous Capture of Metadata

George Alter⁴, Darrell Donakowski², Jack Gager³, Pascal Heus³, Sanda Ionescu¹
 Jeremy Iverson⁴, H.V. Jagadish⁵, Carl Lagoze⁵, Jared Lyle¹, Thomas Murphy¹
 Ornulf Risnes⁶, Dan Smith⁴, Tom Smith⁷, Jie Song⁵

Supported by the Data Infrastructure Building Blocks (DIBBs) program of the National Science Foundation through grant NSF ACI-1640575

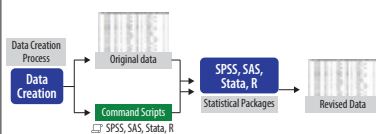
¹ICPSR ²ANES ³MTNA ⁴Colectica ⁵University of Michigan ⁶NSD ⁷NORC

The Problem: Many scientists in a wide variety of disciplines use statistical analysis packages (SPSS, Stata, SAS, R) for data management and data transformations. These packages include limited metadata, and they do not record provenance information about data transformations. Consequently, metadata are created by hand after data transformations have been completed. This is a slow, inefficient process and much valuable information is lost. The C²Metadata Project will create a continuous workflow for metadata creation by automating the capture of data transformations performed by statistical analysis software.

c2metadata.org

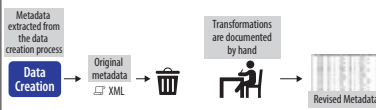
The Data Creation Pipeline

Most data are born digital and pass directly into data transformation and analysis. The leading statistical analysis packages (SPSS, Stata, SAS, R) are often used for data management and data transformation.



The Metadata Data Creation Pipeline before C²Metadata

Metadata can often be extracted directly from the data creation process. However, changes made by statistical analysis packages (SPSS, Stata, SAS, R) make the original metadata obsolete. Data transformations are documented manually and incompletely.



The C²Metadata Solution

Researchers use command scripts to operate statistical analysis packages. We can extract data transformation information from command scripts and use that information to update metadata files.

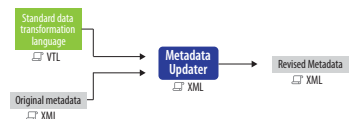
Step 1. Parsing the command script and creating VTL

The Script Parser reads the command script, outputting a description of the data transformations in a standardized form – Validation & Transformation Language (VTL).

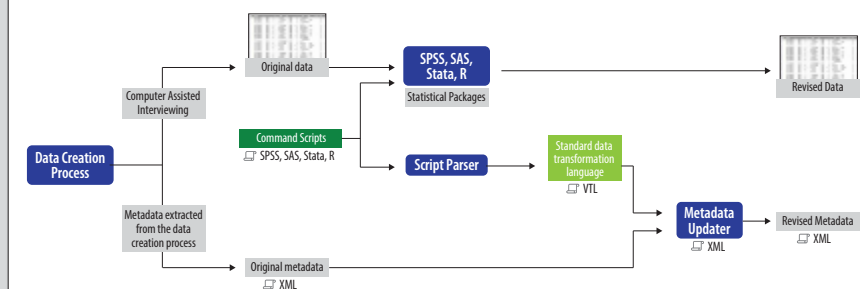


Step 2. VTL is used to update the metadata file

The Metadata Updater uses the VTL to add data transformations into the original metadata file.



The C²Metadata Pipeline: Data and metadata are changed in parallel.



Continuous Capture of Metadata project begins

We announced in August 2016 that ICPSR is among collaborators in a project called “Continuous Capture of Metadata for Statistical Data,” otherwise known as [C2Metadata](http://C2Metadata.org). The project will develop new tools that will work with common statistical packages to automate the capture of metadata at the granularity of individual data transformations. Software-independent data transformation descriptions will be added to metadata in two internationally accepted standards, the Data Documentation Initiative (DDI) and Ecological Metadata Language (EML). These tools will create efficiencies and reduce the costs of data collection, preparation, and reuse.

Summer Program

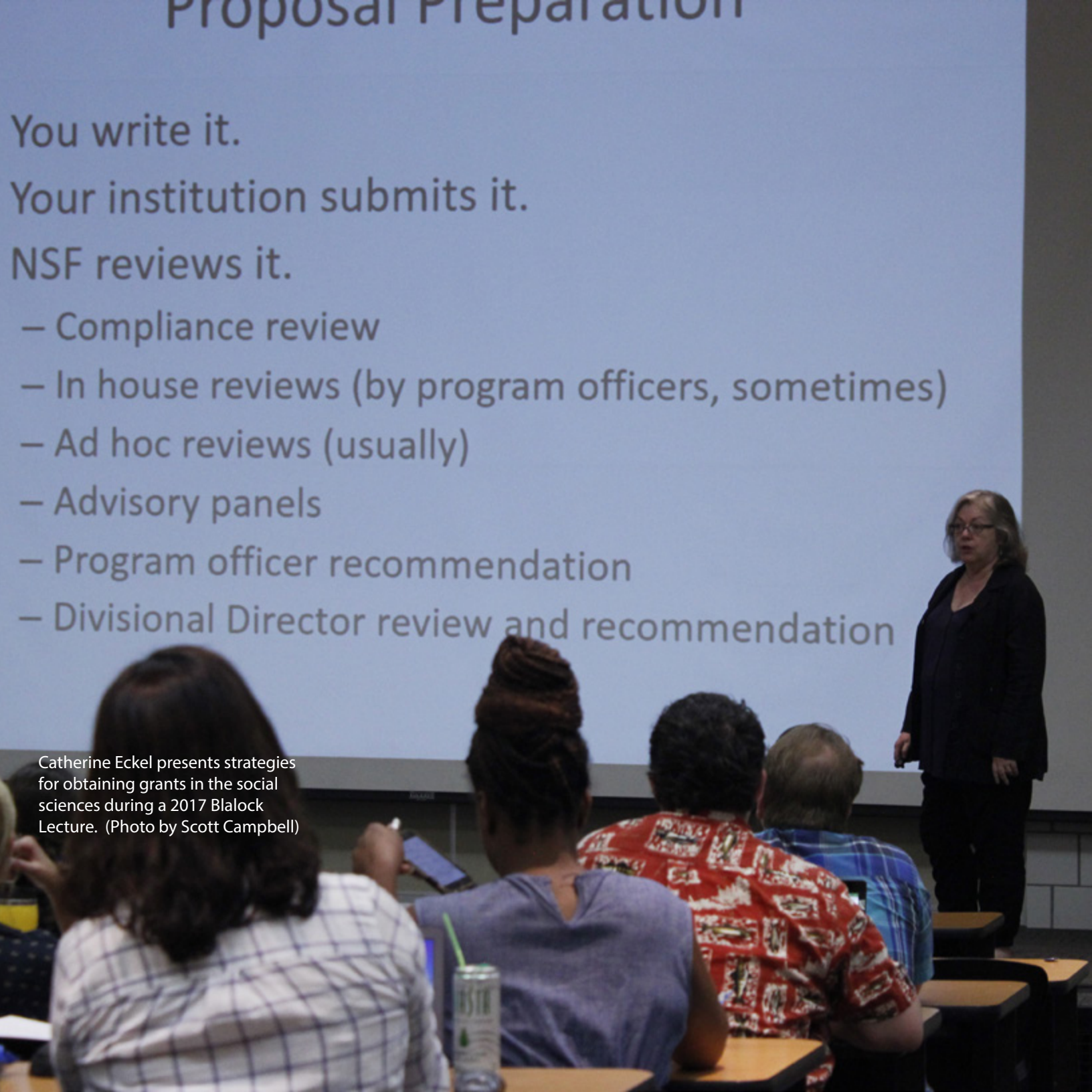
The 2017 ICPSR Summer Program in Quantitative Methods of Social Research drew 974 participants from more than 330 institutions in more than 30 countries and at least 30 disciplines. The Summer Program employed 127 instructors and 45 teaching assistants, who taught a total of 109 courses.

The Summer Program added these courses in 2017:

- Multilevel Models II: Advanced Topics (four-week workshop)
- Applications of Models for Longitudinal and Multilevel Data in R and Stan (short statistical workshop)
- Bayesian Latent Variable Analysis (short statistical workshop)
- Bayesian Multilevel Models (short statistical workshop)
- Panel Data Using Stata (short statistical workshop)
- Machine Learning: Applications and Opportunities in the Social Sciences (short statistical workshop)

Summer Program summary by Stephanie Carpenter/Photo illustration by Jenna Tyson





Catherine Eckel presents strategies for obtaining grants in the social sciences during a 2017 Blalock Lecture. (Photo by Scott Campbell)

The Summer Program also offered 19 sessions in the Hubert M. Blalock Lecture Series, which features presentations by respected researchers and professionals on cutting-edge methodological topics, professional socialization, and diversity issues. Topics included:

- “Text Matching to Understand the Citation Gap” by Molly Roberts (UC-San Diego)
- “How to Measure Legislative District Compactness If You Only Know It When You See It?” by Gary King (Harvard University)
- “How to Compete with Fake News and Cat Videos: Political Persuasion in the Internet Era” by Arthur Lupia (University of Michigan)
- “Grants in the Social Sciences” by Catherine Eckel (Texas A&M University), Rick Wilson (Rice University), and Sara Mitchell (University of Iowa)
- “Academic Publishing” by William G. Jacoby (Michigan State University), Brian Pollins (Ohio State University), and Sandy Marquart-Pyatt (Michigan State University)
- “Fixed, Random, & Mixed Effects: Modern Approaches to Dealing with Nested, Panel, and Longitudinal Data” by John Poe (University of Kentucky)
- “The New White Flight: Democratic Appeals to Latinos and White Support for the Democratic Party” by Mara Ostfeld (University of Michigan)

An Introduction to Civic Learning, Engagement, and Action Data Sharing

Webinar

17 November 2016



ICPSR

Webinars

ICPSR broadcast 29 webinars with 1,500 attendees in fiscal year 2016-2017. They included:

- Orientation to DataLumos — ICPSR's Archive to Preserve Valuable Government Data
- An Overview of the 2017 ICPSR Summer Program
- ICPSR's New Deposit System has Launched! Join Us for a Live Demonstration!
- Read Time in Space: The Spatio-temporal Analysis with China and US Geo-Explorers"
- NACJD Archiving Orientation for OJJDP Research Grantees
- Discovering Solutions to Researchers' Common Data Management Problems
- An Introduction to Civic Learning, Engagement, and Action Data Sharing (CivicLEADS)
- Biomedical Data: What is it? Who is Involved? What Data are Available?
- Orientation to ICPSR with a Fresh New Look!

Webinar videos and slides are available on the [ICPSR YouTube channel](#).

ICPSR Web Presence

537,000+ users

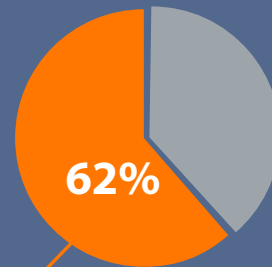
Individuals who have had at least one session during the fiscal year.

> 5 million views

The total number of pages viewed during the fiscal year.

854,794 sessions

A session is the period of time a user is actively engaged with our website.



1st-time visits
to www.icpsr.umich.edu

This year in Social Media

ICPSR

4,415 likes
258 posts

3,176 followers
247 posts

40 videos added



Summer Program

2,412 likes
164 posts

1,204 followers
461 posts

65 videos added

Membership Trends

ICPSR's membership at the end of fiscal year 2017 was at 759 (+4 from FY2017). The biggest increase was in the Master's category (+2). There were slight decreases in the following categories: BAs (-2), and Foreign members (-3).

Web Presence

In fiscal year 2016, the ICPSR website had 537,358 users, with 854,794 sessions; 5,124,541 page views; 531,624 new users, and 323,170 returning users. New sessions (first-time visits) were at 62 percent.

Fall 2016 Data Fair

Over 216 attendees from 131 institutions joined us for 17 webinar sessions September 26-29, 2016, to explore strategies in Navigating the Data (R)Evolution! The online sessions covered ICPSR data services from depositing to curation to accessing data; the benefits and challenges of open data and data curation; and several data projects hosted within ICPSR. Did you miss it? Recordings and presentation slides can be found on ICPSR's YouTube channel.



Left: Fall 2016 Data Fair presenters included 1. Harsha Ummerpillai and Thomas Murphy; 2. Kathryn Lavender; 3. Amy Pienta; 4. Jared Lyle; 5. Linda Detterman, David Thomas, and Lynette Hoelter; 6. Jessi Carrothers, Piper Simmons, and Justin Noble; 7. Linda Detterman and Wendi Fornoff. (Photos by Dory Knight-Ingram)

Staff and researchers from ICPSR and its topical archives exhibited at, participated in, or presented papers at a variety of conferences around the world, including the following:

- The Administration for Children and Families' National Research Conference on Early Childhood, July 2016, Washington, DC
- American Sociological Association, August 2016, Seattle, WA
- American Political Science Association, September 2016, Philadelphia, PA
- The Hispanic Association of Colleges and Universities, October 2016, San Antonio, TX
- Michigan State University Graduate and Law School Fair, October 2016, Lansing, MI
- American Public Health Association, October 2016, Denver, CO
- Gerontological Society of America's 69th Annual Scientific Meeting, November 2016, New Orleans, LA
- American Society of Criminology, November 2016, New Orleans, LA
- Allied Social Sciences Association Annual Meeting, January 2017, Chicago, IL
- Southern Political Science Association, January 2017, New Orleans, LA
- Michigan State University Diversity Career Fair, January 2017, Lansing, MI
- Child Care and Early Education Policy Research Consortium, March 2017, Washington, DC
- Academy of Criminal Justice Sciences, March 2017, Kansas City, MO
- Association of College and Research Libraries, March 2017, Baltimore, MD
- Midwest Political Science Association, April 2017, Chicago, IL
- Society for Research in Child Development, April 2017, Austin, TX
- American Educational Research Association, April 2017, San Antonio, TX
- Population Association of America, April 2017, Chicago, IL
- American Planning Association's National Planning Conference, May 2017, New York City, NY
- American Association for Public Opinion Research, May 2017, New Orleans, LA
- International Association for Social Science Information Services and Technology, May 2017, Lawrence, KS
- Society for Prevention Research, May 2017, Washington, DC

Revenue	
General Fund Transfers	\$(484,239)
General Income	5,626,057
Gifts and Grants	95,75,878
Indirect Cost Recovered	26,76,193
Investment Transactions	222,936
Transfers	67,529
Internal Rebills	1,152,905
Total Revenue	\$18,837,261

Expenses	
Staff Salaries	\$8,729,734
Fringe Benefits	2,025,830
Consultants	195,432
Stipends	31,861
Student aid	300
General Expenses	1,375,728
Data Processing	1,146,706
Postage	2,260
Telephone	49,338
Subcontract Under 25K	52,786
Subcontract Over 25K	956,746
Travel/Hosting	511,446
Equipment	-
Transfers	(85,275)
Transfers and Distributions	5,016
Indirect Costs	2,676,193
Total Expenses	\$17,674,102

Revenue from Membership and Grants/Contracts		
Fiscal Year	Membership	Grants/Contracts
1989	\$1,654,781	\$1,413,092
1990	1,403,297	1,434,630
1991	1,690,709	1,578,484
1992	1,481,713	1,583,566
1993	1,712,031	1,832,865
1994	1,915,000	2,099,753
1995	1,969,783	1,893,296
1996	2,066,499	2,146,822
1997	2,181,661	2,310,968
1998	2,258,676	2,742,252
1999	2,307,856	2,651,291
2000	2,254,963	2,456,876
2001	2,835,791	3,358,040
2002	2,259,161	3,531,002
2003	2,656,468	4,151,227
2004	2,666,290	5,289,738
2005	2,791,442	5,770,366
2006	2,935,642	6,252,685
2007	3,093,900	6,776,695
2008	3,165,587	7,584,089
2009	3,311,411	8,138,624
2010	3,465,415	9,849,709
2011	3,502,368	10,015,225
2012	3,531,236	9,991,773
2013	3,738,255	9,099,902
2014	3,775,109	8,487,492
2015	3,820,064	7,867,161
2016	3,963,483	9,041,421
2017	4,084,406	9,525,580

ICPSR closed fiscal year 2017 with revenues totaling \$18.8 million; representing a 3 percent increase from fiscal year 2016's total. Membership dues income, totaling \$3.96 million, rose 3 percent since 2016 and contributed 22 percent of the organization's total revenue. Sponsored project awards contributed \$9.5 million and accounted for 51 percent of ICPSR's revenue. Summer Program revenue was \$1.99 million for fiscal year 2017. ICPSR's overall expenses were \$17.7 million, a 2.8 percent decrease from fiscal year 2015. The organization closed fiscal year 2017 with a surplus of \$1,163,159. As of June 30, 2017, ICPSR maintains a total fund balance of \$5,740,893.

677

added or updated
studies in Fiscal
Year 2016-2017

TOP DOWNLOADS



Most popular releases in FY 2016-2017

Rank	Study
1	National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2008 [Public Use]
2	National Health and Nutrition Examination Survey (NHANES), 2007-2008
3	National Health and Nutrition Examination Survey (NHANES), 2005-2006
4	Maternal Lifestyle Study in Four Sites in the United States, 1993-2011
5	National Health and Nutrition Examination Survey (NHANES), 1999-2000
6	Youth Development Study, 1988-2011 [St. Paul, Minnesota]
7	National Health and Nutrition Examination Survey (NHANES), 2001-2002
8	Uniform Crime Reporting Program Data [United States]: 1975-1997
9	National Health and Nutrition Examination Survey (NHANES), 2003-2004
10	American Community Survey (ACS): Public Use Microdata Sample (PUMS), 2005



Lisa D. Cook joins ICPSR Council

ICPSR welcomed Lisa D. Cook of Michigan State University to the ICPSR Council in June 2017. Cook will serve on the Council through February 2018. Cook is an Associate Professor of Economics and International Relations.



Matthew Richardson named 2017 ICPSR Staff Innovator of the Year

Matthew Richardson was the recipient of the [2017 ICPSR Innovator Award](#). Colleagues noted Richardson's knack for problem-solving and his crucial involvement in setting up tools and interfaces that allow staff across ICPSR to do work independently and efficiently.

Left: Matthew Richardson accepts the 2017 ICPSR Innovator of the Year Award.
(Photo by Dory Knight-Ingram)

This report was compiled and edited by Dory Knight-Ingram, and designed by Jenna Tyson, with special thanks to Elizabeth Moss, Shelly Petrisko, Stephanie Carpenter, Peggy Overcashier and JD Alford.