

ICPSR Council Minutes June 7-8, 2012

Attendees

Council Members: Rogelio Saenz (Chair), Christopher H. Achen, Marilyn Andrews, Tony N. Brown, G. Sayeed Choudhury, Catherine A. Fitch, John Fox, Gregory N. Price, and Linda J. Waite

ICPSR Staff: JD Alford, George Alter, Rita Bantom, Shuming Bao, Bryan Beecher, Dieter Burrell, Linda Detterman, John Garcia, Peter Granda, Lynette Hoelter, William Jacoby, John Marcotte, Dan Meisler, Asmat Noori, JoAnne McFarland O'Rourke, Michelle Overholser, Amy Pienta, David Thomas, Mary Vardigan, and Diane Winter

Visitors: Hans Jørgen Marker, Director, Swedish National Data Service and President of the Council of European Social Science Data Archives, and Matthew Woollard, Director, UK Data Archive

Director's Report

Minutes from the March 2012 Council meeting were reviewed, voted on, and passed without changes.

ICPSR Director George Alter introduced the Council meeting visitors as well as new ICPSR staff member Shuming Bao, Director of the China Data Center.

Alter indicated that ICPSR membership currently stood at 717 with 29 members added and 5 dropped in FY2012. Membership revenue was projected to be \$3.7 million.

ICPSR's Web site activity continued to be strong. According to Google Analytics, from July 1, 2011, through April 30, 2012, there were 365,374 visitors and 579,538 visits to the site. Visitors averaged over five minutes per visit and viewed eight pages. Sixty-one percent of the visitors were new visits. Most visits were from the United States, then China, followed by Canada.

The FY2013 budget projection was for a total budget of 19.1M, based on University accounting. The actual number is closer to 17M. While ICPSR was showing an expected deficit of \$653K, this was a conservative estimate taking into account only those sources of revenue that the organization was certain of receiving. Alter explained that because most ICPSR funding is grant-based and several grants had recently been submitted, it was likely that over the course of the year the organization would receive additional sources of funding.

A large portion of the projected deficit (\$331K) was due to funding for the Gates Foundation Measures of Effective Teaching (MET) project, which was received in FY2012 but would be expended in FY2013.

The FY2012 budget forecast continued to improve. While the original budget projected a \$743K deficit, the latest forecast showed a \$678K surplus (including \$331K early payment from the MET project).

Alter reviewed the proposals that had been submitted since March 2012 and highlighted the following:

- BCC-SBE: Building Communities for Research with Administrative Data (NSF)
- BCC-SBE: Sharing Geo-referenced Research Data with a Virtual Data Enclave (Association of American Geographers (Prime) (NSF)
- Building Community Engagement for Open Access to Data (Alfred P. Sloan Foundation)

The top five studies (measured in number of downloads) released in the previous six months (as of May 28, 2012) were:

- National Survey on Drug Use and Health, 2010 [SAMHDA]
- National Survey of Substance Abuse Treatment Services (N-SSATS), 2010 [SAMHDA]
- Drug Abuse Warning Network (DAWN), 2009 [SAMHDA]
- Uniform Crime Reporting Program Data: County-Level Detailed Arrest and Offense Data, 2009 [NACJD]
- Health Behavior in School-Aged Children, 2005-2006 [SAMHDA]

The top five studies with the highest number of downloads in the previous six months were:

- National Survey on Drug Use and Health, 2010 [SAMHDA]
- National Longitudinal Study of Adolescent Health (Add Health), 1994-2008 [DSDR]
- General Social Survey, 1972-2010 [Cumulative File] [ICPSR]
- Chinese Household Income Project, 2002 [DSDR]
- India Human Development Survey (IHDS), 2005 [DSDR]

After the Official Representatives (OR) meeting last October, Membership and Marketing conducted a survey of ICPSR ORs/DRs and got an excellent response rate of 40 percent. Overall, the ORs and DRs indicated that were very satisfied with ICPSR. The modal OR works in the library and holds a library science degree.

Alter reported that the 2012 Summer Program was on track for another record year with 79 courses compared to 65 in 2011 and 1083 registered participants as of May 25, which was 125 more than in 2011. The Summer Program was holding courses in four locations. A new workshop on health disparities research and minority populations, sponsored by the Robert Wood Johnson Foundation, was a popular offering, with 70 applications for 25 places.

ICPSR recently welcomed a new cohort of Summer Interns. Six interns were selected out of 270 applicants. They are Jane Oliphant, University of Portland; Jeffrey Berg, Carleton College; Joshua Goode, University of Colorado, Denver; Rosa Rincon, California State University Northridge; Thomas Hegland, St. Olaf College; and Tori Thomas, George State University.

Alter reported on the recent work of the Computing and Network Services group. They enhanced search capabilities in the Summer Program portal, harvested variable-level metadata (including variable groups) from existing SDA files for re-use in other applications, and deployed the first streaming video portal for use in the MET project.

ICPSR has been adding persistent identifiers in the form of Digital Object Identifiers (DOIs) to its datasets since 2008 and has been providing data citations since 1990. Alter reported that we are starting to see the impact of researchers using ICPSR data citations and DOIs. When data are cited properly in the references section of publications, automated harvesters like Google Scholar are able to make associations between publications and data used. This will make it easier to add to the ICPSR Bibliography of Data-Related Literature, which currently includes over 60,000 citations. Finding publications related to ICPSR data has been very difficult because journals have been inconsistent in data citation practices. The American Sociological Review (ASR) recently revised its submission criteria to require the use of data citations and persistent identifiers, which is a welcome development.

Alter previewed the new ICPSR Web site for Council members along with the new logo. The launch date is September 4, 2012.

Alter introduced Shuming Bao, Director of the China Data Center, a new archive within ICPSR slated to move into the Perry Building in August 2012. Bao gave some background on the Center and its mission. The Center integrates historical, social, and natural science data into GIS to support research on the human and natural components of local, regional, and global change. An annual subscription fee provides access to several China Data Center data resources and tools, including China Geo-Explorer.

Alter reviewed the 50th anniversary celebrations for the evening. The Open House was to take place at the Perry Building from 3:00 – 5:00 p.m., followed by the Banquet at 6:00 p.m. at the Michigan League.

Budget and Policy Committee

Council: Rogelio Saenz (Chair), Marilyn Andrews, Tony N. Brown, Gregory N. Price, Ann Wolpert
Staff: JD Alford, George Alter, Rita Bantom, Diane Winter

I. Strategic Plan Update

The highlights of the Strategic Plan were presented in the Director's Report earlier in the morning. There was no further discussion by the committee.

II. Financial and Administrative Issues

A. FY2012 Year-End Financial Projection

Staff reported that ICPSR anticipated ending FY2012 with a surplus of \$678K, which was a \$1.4M improvement over the original budgeted deficit. It was pointed out that over half of this difference was due to two major timing issues on sponsored project payments. ICPSR received the full, two-year funding for the MET project in FY2012. The difference between this revenue and the project's FY2012 expenses resulted in a contribution of \$500K to ICPSR's bottom line. Another project contributed \$266K to the surplus as a result of payments received in FY2012 covering FY2011 expenses.

FY 2012 membership revenue (annual dues from member organizations) was up \$200K versus budget while expenses were down \$580K. A big part of the decrease in expenses was a result of the recent Retirement Incentive Opportunity offered to retirement-eligible staff: expenses shifted from salaries on the membership budget to indirect cost funded buyouts. ICPSR also had a non-budgeted vacancy in a Digital Preservation senior staff position. Finally, the Summer Program was budgeted to have a deficit of \$146K but instead ended the last program year with a deficit under \$10K.

Indirect cost revenue was down \$181K while expenses were up \$287K, resulting in a deficit of \$550K versus the budgeted \$80K deficit. The increase in expenses was driven by the Retirement Incentive Opportunity buyouts, spring office moves, and the non-budgeted FLAME project, a project to implement file-level management of ICPSR's digital assets.

B. FY 2013 Proposed Budget

The proposed FY2013 budget was presented to the committee by ICPSR staff. The budget included a deficit of \$653K but the committee was reminded that half of this deficit was being caused by the MET project pre-payment made in FY2012.

ICPSR was budgeting \$3.51M for membership dues revenue, a small increase over what was expected in FY2012. Staff pointed out that \$3.75M had already been invoiced so there was some hope that FY2013 revenue would exceed expectations. Even with increased expenses from senior staff hires (Associate Director and Associate Archivist) built into the budget, the membership portion of the ICPSR budget was balanced.

Staff reminded the committee that the indirect cost revenue number included only awarded or continuing projects. ICPSR had submitted several proposals in the spring that, if funded, would help to increase the revenue amount. The proposed budget included a deficit of \$415K in this area due to the conservative revenue projection, increased expenses from the FLAME infrastructure project, and ICPSR's contribution to a University of Michigan "Administrative Services Transformation" project.

The proposed FY2013 budget was recommended for approval during the Budget and Policy Committee report and passed unanimously by Council.

C. Personnel Update

During the personnel update, staff provided an update on the recruitment committees for the three senior staff open positions.

Staff agreed to investigate the possibility of including race and ethnicity characteristics of job applicants in recruitment reports. It was discussed that while these data are available in the University of Michigan system, applicants voluntarily supply the information.

D. Diversity Goals and Strategies

Staff presented the report found in the Council briefing material. This report was the result of an October 2011 action item. Much of the brief discussion focused on diversification of ICPSR's workforce and how to build a pipeline of diverse applicants.

III. Governance Issues

A. China Data Center

This item was discussed earlier during the Director's Report.

B. Plan for Updating the Strategic Plan

The report in the Council Briefing book was the result of an action item from the March 2012 Council Meeting. Staff started the discussion by pointing out the long list of accomplishments that were generated by the current Strategic Plan. The committee agreed that a considerable amount of work went into creating the existing Strategic Plan and it was now the right time to tweak this Plan rather than create a new one. The objective is to keep the mission and strategic directions and focus on updating the goals. The committee discussed the details of the consultant's proposal and the full Council agreed that ICPSR should move forward with hiring the consultant and updating the Strategic Plan.

Collection Development Committee

Council: Christopher H. Achen (Chair), G. Sayeed Choudhury, Catherine A. Fitch, John Fox,
Linda J. Waite

Staff: Peter Granda, Jared Lyle, John Marcotte, Amy Pienta

Visitors: Hans Jørgen Marker, Matthew Woollard

I. Strategic Plan Update

Staff updated the Committee on Strategic Plan activities regarding Direction I (Leadership).

ICPSR holds a seat on the Board of the International Federation of Data Organizations (IFDO) and in this capacity participated in several conference calls in the fall and winter which focused on completing an online survey to gather information about data sharing policies and practices from as many countries as possible. These policies would then appear in conjunction with the completion of IFDO's new Web site. The Board met on June 5 in Washington, DC, to discuss the results of the survey.

As an observer, ICPSR also attended the General Assembly of the Council for European Social Science Data Archives (CESSDA) held at the Swedish National Data Service (Svensk Nationell Datatjänst) in Gothenburg on April 20. The group discussed current developments in building a formal social science data infrastructure within the European context and various cross-national projects that were already under way. ICPSR provided a brief overview of the environment for data infrastructures in the United States.

With regard to Direction III (Broadened Collections), staff has submitted an R03 application to the National Institute on Aging to complete conversion of paper records associated with the Cornell Study of Occupational Retirement, conducted in the early 1950s, to electronic form to facilitate research use of these important data.

II. Self-Archiving Option for Depositing Data at ICPSR

Following the practices of some of its sister archives in other countries, ICPSR is now studying the feasibility of providing users with a self-archiving option for depositing their data and metadata. Factors involved in considering this option include:

- 1) A desire to work more closely with journals that are increasingly requiring authors to obtain persistent identifiers and deposit the data that they use in their research articles
- 2) Providing an additional archiving mechanism for grantees who must now supply data management plans as part of their proposals to the National Science Foundation and provide open access to their data
- 3) A defined need, after further internal discussion, to perform some kind of confidentiality review of any deposited materials using this option

Since data management plans are now required from NSF applicants, ICPSR has received numerous requests for letters of support or cost estimates for archiving and processing data that will result from grant awards.

NSF has determined that ICPSR's current organizational structure (where data collections processed using member dues are only available free of charge to member institutions) does not satisfy their requirement for free and open access. Self-archiving with a limited time commitment for preserving data files produced from NSF awards is one option under discussion within ICPSR to provide archiving services for these data.

The Committee discussed in some detail the experience of the UK Data Archive with a self-archiving option. The Archive ingested almost 500 data collections through this mechanism but very little metadata had accompanied these deposits, making it difficult to determine their quality, integrity, or completeness. The Archive now believes that these datasets should first be stored in institutional repositories with some migrating into the UK Data Archive after further evaluation. It is still an open question as to whether institutional repositories, no matter which country they are located in, are ready, willing, and able to deal with the many peculiarities of numeric data files.

The Committee emphasized that self-archiving was not a solution for long-term preservation while recognizing that funding agencies and journals are increasingly asking for standardized identifiers for datasets. If an organization provides different levels of service to data depositors, these policies must be made clear. The Committee also recommended that both the legal and reputational aspects of initiating a true, self-archiving option at ICPSR should be studied carefully before proceeding and that carefully branding such a service was an important consideration.

III. Revision of the Collection Development Policy

The Committee was not able to discuss this item because of time constraints and asked that it be included on its agenda at the October meeting.

IV. Report on Developments in One of the Topical Archives: Substance Abuse and Mental Health Data Archive (SAMHDA)

The SAMHDA Director, John Marcotte, provided an outline of recent developments in the Archive focusing on current initiatives regarding restricted data. Two online systems for restricted data were described. The first, Secure SDA or Restricted Data Analysis System (R-DAS), provides users with access through a Web browser but only produces tables and frequencies. The system has built-in vetting capabilities to suppress cell sizes if they are below minimum thresholds. The second system, the Virtual Data Enclave (VDE) portal, requires completion of a restricted data agreement and two-factor authentication, and it can be accessed only from authorized locations. It does provide full modeling and regression analysis capabilities, but all output must be vetted by SAMHDA staff before release. Since the goal is to provide a set of tools for the research community to obtain access to restricted data, the Committee emphasized the importance of training researchers on the proper use of these data before they actually starting working with the files.

The Committee had an extensive discussion about the growing amount of restricted data which is intended to protect respondent confidentiality when current cultural norms seem to be moving in the opposite direction: individuals more freely providing information about their private lives through various social media.

Membership and Marketing Committee

Council: Catherine A. Fitch (Chair), Rogelio Saenz, Marilyn Andrews, Tony N. Brown
Staff: Linda Detterman, David Thomas
Visitor: Matthew Woollard

I. Strategic Plan Progress

Staff updated Council on the progress of the Strategic Plan for Membership and Marketing. As a means to “sustain membership, improve services, and expand the value” of ICPSR membership, staff noted the following activities dedicated to this objective within the Strategic Plan: 2012

OR/DR Membership Feedback Survey; attendance at several conferences and sponsorship of several receptions as a result of the 50th Anniversary celebration; communications including the Winter DataBytes, the Spring and Special Edition Bulletins, and participation on social media platforms; and several onsite campus visits to member institutions.

Lastly, the objective, “build the data community and advocacy,” was satisfied in the past few months through several webinars supporting quantitative literacy and the planning of the 2012 Data Fair focused on election data.

II. Membership Activity and Utilization

Linda Detterman reported that membership continued to show net positive growth this fiscal year and has reached the 717 member level. Receipt of membership dues was running similar to previous fiscal years. ICPSR met its FY2012 membership revenue forecast in March. Utilization data for FY2012 to date as measured by the unique users count (the count of active downloaders) looked to be on par or slightly below FY2011. Since July 1, 2011, over 365,000 people visited the ICPSR Web site, spending an average of 5.3 minutes on the site. Staff noted that a FY2012 fiscal end revenue, membership, and utilization review would be presented at the October Council meeting.

ICPSR was active in the communications area, releasing both the Spring and Special Edition Bulletins and the Winter DataBytes e-newsletters since Council last met. ICPSR currently had 738 Twitter followers and 564 Facebook fans.

Preservation and Access Committee

Council: G. Sayeed Choudhury (Chair), Paul N. Courant, Ann Wolpert
Staff: Bryan Beecher, Jared Lyle, Matthew Richardson, Mary Vardigan

I. Strategic Plan Progress

Mary Vardigan presented an update on ICPSR’s execution against the Strategic Plan in the areas of preservation and access. She noted several highlights: (1) a recent grant application to the Sloan Foundation on “Building Community Engagement in Data Citation and Open Access to Data”; (2) the 2012 redesign of the ICPSR Web site; (3) a survey of institutional repositories; (4) a guest blog post by Director George Alter on open access to data; and (5) Bryan Beecher’s regular blog posts on technology at ICPSR.

II. MET Project Update

Beecher presented an overview of two new, complementary Bill and Melinda Gates Foundation grants under way at ICPSR related to education, Measures of Effective Teaching (MET) and MET Extension. These grants will significantly increase the size of ICPSR’s archival holdings and will move ICPSR into the “video business.”

III. New ICPSR Web Site

Matthew Richardson demonstrated the new ICPSR Web site, which was to launch in September 2012. Highlights included new navigation, member institution search, auto-complete while searching, new facets, and new study home pages, among many other new features.

IV. Data Science Course Update

Vardigan reported on a new data science course that was to be offered in July as part of the ICPSR Summer Program. The week-long course, to begin on July 23, 2012, was fully subscribed.

V. Data Management Plans

Vardigan reported progress and difficulties in building a library of data management plans. A major obstacle is that investigators are not keen to share such plans. Council noted that their experience was very similar, and further noted that the goal of building a library may be too hard to achieve.

VI. Data Curation Working Group

Jared Lyle noted a joint effort between ICPSR and Stanford University to develop a Data Curation Working Group, which would meet via Web conference to discuss issues arising from curating data. One area of exploration is the suitability of expanding the use of ICPSR's own curation tools to this broader community. Council asked if this was a revenue generation activity, and Lyle noted that it was not; rather the goal is to spread the knowledge of how to curate data properly.

VII. Strategic Plan Revision

Vardigan noted that the action item (from March) to update the ICPSR Strategic Plan had been handed off to the Budget and Policy Committee.

Training and Instruction Committee

Council: Gregory N. Price (Chair), Christopher H. Achen, John Fox
Staff: Dieter Burrell, John Garcia, Henry Heitowit, Lynette Hoelter, William Jacoby
Visitor: Hans Jørgen Marker

I. Strategic Plan

Staff reported on activities related to Strategic Direction IV, Educational Strength, which were organized under three broad categories. The first category concerned the maintenance and expansion of the Summer Program's curriculum. Course offerings for 2012 increased in both the four-week and three- to five-day formats with the Program holding three new four-week classes

and 13 new short courses. Among the innovations were three workshops on data science, one new short workshop on statistical computing, a new emphasis on course sequences to provide more comprehensive coverage of relevant topics, and several workshops on topics of special interest to diverse audiences.

Second, staff reported on the Program's ongoing diversity initiatives directed at recruiting the Summer Program faculty and a diverse pool of participants. Staff described new diversity-related courses and lectures, including two new five-day workshops and several guest lectures on methodological issues in quantitative research on minority populations. Committee members also discussed ideas about possible scholarships directed at HBCUs or other minority-serving institutions. What institutions outside of ICPSR might be interested in funding such endeavors?

Third, staff reported on efforts to extend the activities of the Summer Program, including the new off-site (non-Ann Arbor) location in Boulder, Colorado. Committee members also discussed other possible off-site locations, including such urban areas as Atlanta or Washington, DC. Finally, the Committee discussed ways to encourage participation in the Summer Program by scholars from less-intensive resources institutions such as HBCUs.

II. 2012 ICPSR Summer Program

Staff reported on the beginning of the 2012 Summer Program, with the four-week sessions scheduled to start on June 18. The Program planned to hold 78 courses this year, and enrollments were already up over previous years.

III. Online Learning Center

Staff reported plans to consolidate several ICPSR instructional resources into one location. The Data-Driven Learning Guides (from the Online Learning Center), ICPSR's in-depth instructional modules, and other instructional resources would soon be located under one "Resources for Instructors" site on the ICPSR Web site. Staff also reported that ICPSR was hiring new staff to help oversee the instructional resources projects.

IV. Community College Report

The Committee continued the discussion from the March meeting about efforts to engage community colleges. Staff reported the intention to prepare a grant proposal related to instruction and community colleges during the 2012-2013 academic year.