ICPSR COUNCIL MINUTES
June 11-12, 2015

Council Members: Christopher H. Achen (Chair), Marilyn Andrews, Tony N. Brown, Robert S. Chen, Colin Elman, Philip N. Jefferson, Carl Lagoze, Chandra L. Muller, Ronald Nakao and William Vega

ICPSR Staff: JD Alford, George Alter, Rita Bantom, Shuming Bao, Dieter Burrell, Alina Conn, Edward Czilli, Linda Detterman, John Garcia, Peter Granda, Lynette Hoelter, Susan Jekielek, Dory Knight-Ingram, John Marcotte, James McNally, Tom Murphy, Asmat Noori, Michelle Overholser, Amy Pienta, Saundra Schneider, David Thomas, Mary Vardigan and Diane Winter

Staff Visitors: David Metcalf, Archive Manager, Substance Abuse and Mental Health Data Archive (SAMHDA), and Justin Noble, Archive Manager, Data Acquisitions

Director's Report

ICPSR Director George Alter welcomed everyone to the June 2015 Council meeting. March Council meeting minutes were reviewed, voted on, and approved.

Alter introduced Dr. Colin Elman, Professor of Political Science at Syracuse University, and indicated that Elman would be continuing the Council term of Janet Box-Steffensmeier, who stepped down due to taking on additional responsibilities. Two visitors, both long-time ICPSR staff members, were in attendance at the meeting: David Metcalf (SAMHDA) and Justin Noble (Acquisitions).

Alter announced that two new staff members would be joining ICPSR at the end of the summer. Harold (Woody) Neighbors, Professor of Health Behavior and Health Education in the U-M School of Public Health and Research Professor in ISR, would be serving as the Interim Director of the Resource Center for Minority Data (RCMD); ICPSR was planning to do a search for a permanent RCMD Director in the fall. Jukka Savolainen, former Associate Professor of Criminal Justice at the University of Nebraska in Omaha, was welcomed as the new Director for the National Archive of Criminal Justice Data (NACJD).

Alter reported that membership numbers were healthy with a total of 757 member institutions; ICPSR added 26 members for the fiscal year with only 13 drops. Collection of membership dues to date stood at $3.8M.

With respect to the budget, the forecast for closing out the fiscal year was more positive than originally projected. While the budget that Council adopted last June included a deficit of $1M, the deficit was down to $580K. The current budget projection for FY2016 was a $683K deficit, but this was a worst-case scenario. A big part of this deficit was a result of the SAMHDA project not being renewed.

Alter presented an in-depth explanation of the underlying reasons for the recent and projected deficits. He noted that ICPSR has four main sources of funding:

(1) Membership – Over 750 members pay a subscription fee
(2) Summer Program – Course fees paid by participants
(3) U-M Provost – Supports some of the ICPSR leadership and minority faculty
(4) Sponsored projects – Provide both direct costs (expenses for doing the work: labor and computers) and indirect costs recovery (facilities and administration costs)

Alter noted that ICPSR, along with the other centers in ISR, had been experiencing rising costs. ICPSR pays ISR for services related to facilities and administration, which are considered indirect costs. The "I/O split" pays for the ISR Director's office and Central Institute Services (transaction processing and human resources) and is divided among the ISR centers by amount of sponsored funding. ICPSR also pays into the Building Fund, which covers rent, construction, maintenance and repair.

Recently, new costs had also been incurred:

1. The Provost Tax: In FY 2014 a new budget model was rolled out to all University of Michigan units, and ISR was no long exempt from the budget model imposed on the rest of the U-M. The Provost Tax is based on expenditures and the rate increases each year from 2014 to 2018, when the tax rate will be 20.4%. ICPSR will be "held harmless" for expenditures at the 2013 level; new expenses over that amount are taxed at 20.4%. Services that ICPSR receives in return include Office of General Counsel, Office of Sponsored Programs & Research – Proposals, Unfunded Agreements, University Human Resources, University Payroll, University Accounting, and University Procurement.

2. Shared Services: Effective May 2015 a new fee for use of the U-M Shared Services Center, regardless of the use of the services, was instituted. The methodology for allocating costs was being disputed by ISR leadership, but the current estimate for ICPSR was $200,000 / year.

ICPSR maintains both committed and uncommitted reserves. Uncommitted reserves are saved funds without designated obligations. Committed reserves are funds that offset obligations to research faculty for bridging, research accounts, and special duty assignment (SDA).

Alter reported that in response to the budget situation, ICPSR had reduced staff through attrition and also made six Reductions in Force (RIFs) since the last meeting. The SAMHDA staff was currently being transitioned and funding would end at the end of August.

Other ways that ICPSR was addressing the budget deficit included:

- Grant applications were continuing to be written and submitted
- Summer Program fees would increase
- An IT services plan was being formulated
- A Membership Budget Group was constituted to look at cost savings
- An Internal Situational Review Group was formed to supplement the external review

In terms of download and website activity, there were 419,828 visits to the ICPSR website since the last meeting with 267,529 unique visitors. The top number of visits to the site originated from the U.S., China, Canada, the U.K., and South Korea in that order.

Alter also presented information on ICPSR data usage:

Top 10 Downloaded Studies Released in the Previous 6 Months as of June 7, 2015

Title	Archive
General Social Survey, 2012 Merged Data, Including a Cultural Module [United States]	NADAC
Annual Survey of Jails, 2013	NACJD
National Prisoner Statistics, 1978-2013	NACJD
Health Reform Monitoring Survey, First Quarter 2014	HMCA
Voice of the People End of Year Survey, 2012	ICPSR
United States Agriculture Data, 1840 - 2010	ICPSR
Relationship Dynamics and Social Life (RDSL) Study [Michigan], 2008-2012	DSDR
General Social Survey, 1993, 1998, 2000, 2002 [United States]	NADAC
Survey of Public Participation in the Arts (SPPA), 2012 [United States]	NADAC
Second International Self-Reported Delinquency Study, 2005-2007	NACJD

Top 10 Data Downloads in the Previous Six Months as of June 7, 2015

Title	Archive	Number of downloads
National Survey on Drug Use and Health, 2013	SAMHDA	2481
National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2008 [Public Use]	DSDR	2063
National Survey on Drug Use and Health, 2012	SAMHDA	857
India Human Development Survey (IHDS), 2005	DSDR	704
Chinese Household Income Project, 2002	DSDR	630
Gender, Mental Illness, and Crime in the United States, 2004	NACJD	596
National Mental Health Services Survey (N-MHSS), 2010	SAMHDA	543
National Crime Victimization Survey, 2013	NACJD	477
Drug Abuse Warning Network (DAWN), 2011	SAMHDA	472
Monitoring the Future: A Continuing Study of American Youth (12th-Grade Survey), 2013	NAHDAP	453

Alter pointed out that the Health and Medical Care Archive (HMCA), sponsored by the Robert Wood Johnson Foundation, had already surpassed its processing goals for the current grant and had two collections of interest recently released -- the Health Reform Monitoring Survey, Second Quarter 2014, and the Massachusetts Health Reform Survey, 2013. The General Archive had archived the first two surveys completed under the Army STARRS (Army Study to Assess Risk and Resilience in Service members) Project -- the All Army Study and the New Soldier Study. These data would be available in the Virtual Data Enclave by the end of June 2015.

Acquisitions was awarded a grant from the Mellon Foundation to create data visualization tools around an arts education database that ICPSR will archive.

openICPSR had received over 71 deposits and had made its first (paid) curated openICPSR deposit available to the public: funding was received to support curation of the Connecticut Health Survey (ICPSR 35475), which would become open access. openICPSR also had its first journal subscriber, the *Journal of Economic History*.

Alter reported that the Summer Program was offering 40 courses in the two four-week sessions and 45 two- to five-day workshops and had 1,110 registered participants as of June 8.

Alter noted that this was 11th year of the ICPSR Internship Program and there were currently 56 intern alumni across the world. Two alumni received their PhDs this year, and 7 have returned to ICPSR to work as data processors. ICPSR acquired a no-cost extension of the NSF Research Experience for Undergraduates (REU) funds to support the interns this year. The 2015 interns and the ICPSR units hosting them were: Kimberly Gannon, Michigan State University (NACJD), Christian Nuno, University of Illinois at Chicago (RCMD), Ivies Paniagua, Cornell University (CCEERC), Rebeca Willis-Conger, Reed College (NAHDAP), and Polina Zvavitch, Ohio State University (General Archive).

Alter reviewed the LaCour and Green controversy involving possible falsification of data and ICPSR's role. The data had been shared via openICPSR and thus were openly available for scrutiny. He noted that the responsibility for investigating the case lay with UCLA, and once the review was completed, ICPSR would contact UCLA for direction.

Alter also discussed the Metadata Portal Project, a partnership among ICPSR, ANES, and GSS to enhance documentation for the two data series. He presented two tools built as part of the project -- the Integrated Crosswalk and the Tagging Tool. ICPSR staff involved in the project were George Alter, Sanda Ionescu, Daphne Lin, Asmat Noori, Matthew Richardson and Mary Vardigan.

Alter reminded Council that the 2015 ICPSR Biennial OR meeting was scheduled for September 30 – October 2, 2015, with the theme of *Delivering Effective Data Stewardship*. The meeting would be preceded by a Software Carpentry Workshop hosted by ICPSR (Jared Lyle, organizer) on September 28-29, 2015. Over 17 workshops and sessions were planned for the OR meeting.

Plenary

ICPSR's New Collection Development Policy Fueled by New User Search Data

Amy Pienta, ICPSR Director of Acquisitions, began the plenary with a short report on the revised Collection Development Policy. She indicated that the main purpose of the new Policy was to achieve a balance in determining when to accept new deposits as the spectrum of social science data continues to evolve. The Policy will advocate an approach based on two curation levels: no curation and member-funded curation. The former would include deposits not directly tied to the social and behavioral sciences broadly speaking, that might have very high costs to curate, that could not be disseminated widely, that were available elsewhere, or that had issues

involving direct identifiers or copyright protection. Deposits eligible for member-funded curation would include those that were part of ICPSR data series, were expected to be popular, had methodological rigor and a significant scientific reputation, and were generally considered of high quality.

Justin Noble then proceeded to discuss the results of a study he completed on how users interacted with the search tools on the ICPSR website. In calendar year 2014 there were almost 540,000 site searches with nearly half of them seeking results from a unique search term. A detailed investigation was made of the top 500 search terms where the most frequent type of search focused on specific keywords. There were considerably fewer searches based on the name of an ICPSR data series, specific study, or principal investigator. A number of popular keyword searches resulted in comparatively few hits because ICPSR only had a small number of studies based on these keywords. Specific examples of such keyword searches were: social media, NCAA, LGBT, restorative justice, human trafficking, and second generation immigrant. Examples of popular keyword searches where ICPSR has many studies include: health care, public opinion, voting behavior, criminal justice, law enforcement, and mental health. Since diversity is a major strategic goal of the organization, a special effort was made by the Data Diversity Work Group to code the top keyword searches for diversity. The results of this effort indicate that almost half of the top keyword searches done in 2014 were related to diversity data with geography, disability, and age the most frequent categories of interest.

It is also the case that users are searching for substantive data content-oriented terms much more frequently than they are for curation terms such as data types. This suggests that we should align our metadata fields more closely to what terms our users are actually searching for. Staff intend to repeat and enhance this investigation for calendar year 2015 to see how the current results hold up over time. Staff consider the tool an effective way to better understand ICPSR's users and to identify potential gaps in the holdings as well as to regularly inform the Collection Development Policy and establish high priority topics for possible future data acquisitions. Other goals are to better align our controlled vocabulary lists with user search behavior, enhance the organization of the ICPSR website, increase downloads, and decrease search exits and bounce rates. Next steps include investigating additional Google Analytics features and other text analytics programs and developing membership surveys. The group also wants to compare site search findings against feedback from the ICPSR Council, Official Representatives, and users; results from a review of the research landscape including scholarly publications, grant award databases, and trending research topics in the news; and download/usage statistics.

Council members appreciated the depth of information provided during the plenary and indicated that this analysis would benefit the organization as it continues to refine its search processes and consider new sources of data for the membership.

Budget and Policy Committee

Council: Christopher H. Achen (Chair), Tony N. Brown, Robert S. Chen, Phillip N. Jefferson, Chandra L. Muller

Staff: JD Alford, George Alter, Rita Bantom, Diane Winter

I.   Financial and Administrative Issues

A. Financial Updates

Staff reported that the forecast for ICPSR's finances continued to look more favorable than the original budget. The current projection was that ICPSR would end the year with a deficit of around $600K versus the budgeted deficit of $1.036 M. Key changes included increased membership revenue coinciding with a decrease in expenses in membership-funded accounts as well as indirect cost revenue (IDCR) funded accounts.

Staff presented the FY 2016 budget proposal, which showed the organization ending FY 2016 with a deficit of $683 K. This deficit included the unanticipated loss of the SAMHDA project. Prior to learning that funding for the SAMHDA project would not continue, ICPSR had prepared a draft budget that was nearly balanced. Staff emphasized that the proposed budget was a work in progress and changes not included in the budget would have to be made to offset the loss of funding. Staff also pointed out that the deficit included one-time expenses for the upcoming ICPSR Director's search and payments to staff placed on RIF (reduction in force) status.

Staff and Council agreed that ICPSR needed to be aggressive with sponsored funding to balance increasing overhead and tax expenses. Council pointed out that this expectation should be considered during the recruitment of the next director.

Staff finished the presentation of the budget with a new version of the draft Executive Financial Summary. Council found this display helpful in understanding the status of ICPSR's uncommitted fund balance and considered this crucial for addressing the organization's fiscal health.

The Council unanimously approved the budget during the Committee report.

B. Personnel Update

Staff presented the Recruiting and Staffing Report found in the Council binder. During this discussion, Council had questions regarding the decrease in minority temporary staff. It was agreed that this was a concern especially since it was common for temporary staff to progress to permanent positions. It was also noted that the reports did not include staff members placed on RIF status in May nor any SAMHDA staff placed on RIF status.

As part of the Diversity Update, staff briefly reported on the offering of a special onsite professional development opportunity in April called Change it Up! This program educates

participants about bystander intervention. The program was positively reviewed by ICPSR staff and it is hoped that similar opportunities can be offered in the future.

II.　Governance Issues

A. Council Election Process

A list of Council nominees for the four-year term to begin in 2016 was presented to Council. The election will take place this fall.

B. Guidelines for Transparency and Openness Promotion (TOP)

George Alter presented the TOP Guidelines to the full Council during the Committee report. The TOP Guidelines (Guidelines for Transparency and Openness Promotion in Journal Policies and Practices) are intended to improve the reproducibility of research. Council unanimously endorsed the guidelines.

Collection Development Committee

Council: Marilyn Andrews (Chair), Colin Elman, Carl Lagoze, Ronald Nakao, William Vega

Staff: Peter Granda, David Metcalf, Justin Noble, Amy Pienta

I. Strategic Plan Update

The committee discussed recent activities related to Strategy 1 of the first direction of the ICPSR Strategic Plan: Develop national and international partnerships, which states that "ICPSR will enhance its role as a global leader in data stewardship, engaging the global community as a partner, convener, advocate, and supporter."

ICPSR was invited to participate in the **2nd European Data Access Forum**, which took place in March 2015 at the European Commission in Luxembourg. The *European Data Access Forum* provides a unique opportunity for data archives, national statistical offices, other members of the European statistical system and the research community to work together and find acceptable solutions to overcome the barriers to research using government data, both national and European, in the European Research Area. The ambition of the Forum is to gather the different communities in equal proportion, also involving discussions with other important stakeholders such as the national data protection commissioners and research funding councils.

ICPSR also participated in the **Service Providers Forum meeting of the Council of European Social Science Data Archives (CESSDA)** in Bergen in April. This meeting brought together representatives of the European data archives to discuss common issues in the context of building a broad infrastructure for the social sciences across the entire European research area. Both current and prospective members of the new CESSDA organization discussed work plan tasks and other initiatives that are important to ICPSR, including training and the importance of

the Data Seal of Approval. The group also discussed the continuation of "expert seminars," which for many years have brought together technical staff from various archives to discuss common issues and challenges. ICPSR has frequently sent staff members to this meeting since the issues discussed are extremely relevant to the work we do.

III.   Collection Development Policy

Amy Pienta presented and discussed the new policy. She emphasized the importance of the ICPSR Retreat, held in December 2014, in developing new ideas for the revised policy. She asserted that the inclusion of a description about curation levels was one of the most important changes. She also indicated that the web team would suggest more innovative ways of having users access the policy after it is approved. The version presented to the Committee was as follows:

Introduction
ICPSR seeks to acquire, archive, and disseminate data of interest to researchers in the social and behavioral sciences. Our definition of the domain is necessarily broad and encompassing as we recognize that traditional boundaries between disciplines are blurring, research is becoming more integrative, and our designated community is expanding. We focus our collecting efforts on data that can address key questions about the human experience in all of its diversity and richness. To guide the growth and management of ICPSR's collection, we have created a policy that offers clear goals but also a degree of flexibility to enable ICPSR to respond to the rapid pace of change in the research environment. We present the policy here to inform ICPSR membership, users, core partners, prospective depositors, and potential funders of the principles governing our collection development activities.

History
The Inter-university Consortium for Social and Political Research began to build a collection of data to be shared across its member institutions in 1962. The early archive included the American National Election Study and other sample survey data. By the 1970s, several large-scale, social science surveys, including some that were conducted by the various centers of the Institute for Social Research at the University of Michigan, were added. ICPSR's collection of survey data strengthened in the decades following, but the archive also expanded in new ways, partly due to Federal grants and contracts awarded to ICPSR to archive special collections, such as data on criminal justice and aging. In later years, additional topical archives were added: substance/drug use, HIV, education (including early education and childcare), health and medical care, demography, racial and ethnic minorities and several others (see http://www.icpsr.umich.edu/icpsrweb/content/membership/partners/archives.html for a full listing).

Over time, ICPSR began to add new data types resulting from a wide range of quantitative and qualitative methods to its collection. As ICPSR's capacity for curating new kinds of data about the human experience has grown, there has been simultaneous, albeit sometimes slow, growth in the culture of data sharing throughout the scientific research community. ICPSR has positioned

itself to invite data from a full range of social, behavioral, and health science disciplines in recognition that social relationships and status are closely intertwined with biological, cognitive, and clinical processes and experiences, to name but a few.

Who are ICPSR Users?
Data from ICPSR are used primarily by the academic research community, which includes researchers and students around the world. ICPSR data and data related products and services are also used by policymakers, consultants, service providers, journalists and other professionals. As ICPSR continues to increase its supply of free, open-access data, the broader public will increasingly be able to take advantage of ICPSR data.

What Data are In Scope?
Building on broad, inclusive collection development policies from the past and also acknowledging the increasing importance of research infrastructure that supports cross disciplinary research, ICPSR seeks data from many disciplines, in support of many methods, and about wide ranging population groups as described below. These lists are not intended to be exhaustive of what ICPSR is interested in collecting, but rather show the wide range of data that are considered in scope for ICPSR. See also Out of Scope and High Priority Areas below:

Disciplines
- Sociology
- Political Science
- Government
- Economics
- Public Policy
- Law
- Demography
- Education
- Environmental Studies
- Criminology
- Geography
- Anthropology
- Rural Studies/Urban Studies
- Psychology
- Human Development
- Family Studies
- Gender/Women's Studies
- Gerontology
- Epidemiology
- Public Health
- Nursing
- Health Care/Medicine
- Social Work
- Arts & Humanities

Research Techniques

- Survey Techniques (including online) Interviewing Techniques
- Experiments
- Content Analysis
- Field Research and Observational Techniques
- Historical Methods
- Clinical Trials
- Interventions
- Policy Analysis
- Administrative Databases
- Web Scraping
- Data Mining

Population Groups Covered

- U.S.
- International/Cross-National/Comparative
- State/Regional/Local
- Criminal Justice Populations
- Children & Adolescents
- Adults and the Elderly
- Various Race & Ethnic Groups
- Families, Couples, & Households

What Data are Out of Scope?

ICPSR is frequently asked to define what types of data it will not accept. The following list outlines some of the criteria used to define data that are not in scope for ICPSR.

Non-Social and Behavioral Research Data - Data that cannot be connected with or used to expand upon the scientific investigation of the social dimensions of human lives (both antecedents and consequences) will not be acquired. For example, most data in the physical sciences are out of scope for ICPSR.

Cost of Data - ICPSR generally does not purchase data or pass along the costs of access to proprietary data to the user community. Therefore, data with associated fees may be considered out of scope for ICPSR.

Limited Access Data - ICPSR generally does not accept data requiring limitations on use, with the exception of data with access conditions intended to protect the privacy and identity of the human subjects. For example, ICPSR generally does not accept data in cases where access would be governed by an outside entity (e.g., external approval for use, publication review, authorship requirements).

Availability Elsewhere - ICPSR prefers to be the archive of record for a data collection. Data that are permanently available from another trusted repository may instead be linked to from

the ICPSR catalog. Directly Identifying - ICPSR, generally, does not accept data with direct identifiers.

Copyright - ICPSR only accepts data when the data contributor grants ICPSR rights to curate, disseminate and preserve a copy of the data.

Government Data
Historically, ICPSR has acquired and processed government data collections either with support of the ICPSR membership or through topical archives at ICPSR, which make data freely available to the public. As the U.S. government has increasingly become an authoritative and reliable source for the data it collects, ICPSR will acquire government data only when it believes it can add significant value for its users and/or (2) ensure long-term preservation of the data. For other important government collections, ICPSR will provide links to the original data sources in its catalog. ICPSR will continue to accept requests from users about government data that are difficult to locate or use, or of such high interest that their acquisition by ICPSR is justified.

High Priority Data
ICPSR identifies high priority data (data collections ICPSR is especially interested in) through review and analysis of user demand (user behavior, recommendations of ICPSR Council, Official Representatives, and the membership) and scanning the research landscape (review of scholarly publications, grant award databases, and trending research topics in the news). The high priority list evolves as topics are added and others are dropped (updated annually). The purpose is to encourage identification, nomination, and deposit of high value data deemed important to users and found to be limited in ICPSR's current holdings. The current high priority areas are:

Sexual Orientation - As the United States is on the brink of major policy decisions regarding the legality of same sex marriage, ICPSR has observed a large interest in data about sexual orientation.

School Bullying - One of the trending topics in education research is bullying in schools. Research seeks to understand who in schools bully and why, the consequences of bullying on others, and what kinds of school based interventions are effective at reducing bullying and the effects of bullying on children. ICPSR would like to identify and acquire research data about bullying in schools, cyberbullying and so on.

Social Media - ICPSR users frequently search for social media data as this method of research is becoming an increasingly important source for information in the social and behavioral sciences. ICPSR is particularly interested in acquiring studies on the frequency and type of participation in social media and the impact of social media, as well as internet use and behavior. Adding data collections built from social media (e.g. blogs, posts, profiles, search behavior) data would both support secondary analysis of these data and potentially spur collection of new social media data.

Immigration - Policies around immigration reform and understanding the experiences of legal and illegal immigrants remain front and center on the American stage and stem from our roots as

country built by immigrants. ICPSR seeks to archive data that yield a new understanding of immigrant populations and help further our knowledge about the impact of immigration. Individual Well-Being - Understanding how men and women achieve and maintain psychological and social well-being in contemporary settings remains an important research goal, yet data sharing in this area has been somewhat limited. ICPSR seeks data on a wide range of topics about mental or psychological well-being, social well-being, happiness, depression, and demoralization.

Longitudinal Data - Longitudinal data are often seen as a gold standard in human development research and many social and behavioral science disciplines. Also, longitudinal data help to establish causality. Because of the strengths of this approach, ICPSR remains very interested in archiving longitudinal data.

International Data - ICPSR seeks data originating from one or more non-U.S. countries. We are especially interested in data from Asian countries as well as countries and regions of the world that do not have a national structure for archiving, disseminating, and preserving research data. Also, we seek comparative data that could be used to support cross-national, comparative research.

Data Formats Accepted
1. ICPSR prefers data in a readily useable format (see the Library of Congress' Recommended Format Specifications), accessible in a variety of computing and technological settings.
2. ICPSR prefers data formats that promote easy access and use without compromising research value.
3. ICPSR prefers that data files deposited in a raw format be transformable or convertible into formats usable by a variety of statistical or analytical software.
4. ICPSR prefers data files unaccompanied by value-added software.
5. Data in obsolete, proprietary, or hard-to-use formats may still be accepted by ICPSR, although these characteristics may compromise any future use of the data other than as-is, bit-level access.

Security, Privacy, and Confidentiality Considerations
1. ICPSR requires that data deposited in the archive meet recognized standards for privacy and confidentiality of subjects studied. (For information on these standards, see the University of Michigan's Human Subjects Protection Web page, specifically the section titled "Use of Human Subjects in Research.").
2. ICPSR prefers to acquire data that can reside in the public domain.
3. ICPSR requires that data intended for public use be formatted so that identifiers inadvertently included in the data can be removed using standard practices without reducing the research value of the original data.
4. Any access limitations that ICPSR might apply to specific data collections (e.g., a requirement that restricted-use agreements must be signed) should be legally justified and manageable given ICPSR's resources, goals, and mission.

Curation Levels

ICPSR maintains a broad policy of inclusion because it has developed two levels of curation services. ICPSR offers (1) no curation where data are made available to the user community in the condition deposited and (2) member funded curation where data are reviewed, enhanced, and quality checked to ensure usability and findability. The selection criteria employed for the two levels of curation services are:

No Curation - The least restrictive stream of data entering ICPSR is data that receive no curation. Any depositor with data meeting the terms of ICPSR's broad Collection Development Policy may deposit and publish data in openICPSR, an open-access repository. Fees may be collected from non-member institutions for this service.

Confidential data in openICPSR requiring restricted access will be supported with fees collected from users. Curation services, paid for by the contributor, are available.

Member Funded Curation - ICPSR also accepts and curates data that are to be valuable (either in the present or future) to the membership of ICPSR. There are additional selection criteria placed on data that are curated for ICPSR membership. Potential value is determined by the following criteria:

> Popularity - Data in openICPSR that are used but could be improved through additional curation will undergo review for possible member funded curation. Series - ICPSR maintains longstanding series and will continue to curate new data that are part of these series to maximize the historic investment in the data by funding agencies, data producers, researchers, and ICPSR itself.

> Methodological Rigor - Data that are methodologically sound, including but not limited to nationally representative sampling designs, will be identified and acquired for member funded curation. Data stemming from an ineffective or flawed research design will not be curated for the membership and deferred to openICPSR.

> Scientific Reputation - Highly cited data collections, data collected by frequently cited scientists, and data resulting in high quality citations (impact) will be identified and acquired for member funded curation.

> Data and Documentation Quality - Quality of the data and documentation are considered when reviewing incoming data. Due to higher cost of curation and prospects for more limited use, data will not be curated for the membership if there is inadequate documentation and/or data are of poor quality. Data not meeting these criteria will be deferred to openICPSR.

> High Priority - Data that are in demand and/or that represent known gaps in ICPSR holdings will be acquired for curation for the ICPSR membership (see High Priority Data above). ICPSR understands that new areas of research may be more experimental and as

such the data might not otherwise meet the criteria for curation. ICPSR considers these high risk - high reward data as being worthy of curation for the membership at lower levels of quality, methodological rigor, and reputation.

ICPSR also has several grants and contracts to provide data archiving services to various research communities by creating topical archives, but each of these projects has developed its own set of selection criteria that fit within the broad Collection Development Policy of ICPSR.

Policy Review
This policy is subject to a five-year review and re-issuance of policy by Collection Development Committee (approved last on July 7, 2015). This policy is open for review and comment by membership at any time.

Council pointed out the need for an Executive Summary of the Policy that should appear on the ICPSR web site with a special emphasis on the need to explain the value of curation for depositors. Other suggestions included adding language in the Policy about the importance of data for teaching, biometric data, the need to continually keep current on overall social science research trends, and the advantages of using the network of Official Representatives to contribute to setting priorities for acquisitions and curation decisions. Additional information about what role administrative and geospatial data will play in the new Policy as well as the overall challenges posed by qualitative data were also suggested.

Council asked the staff to make modifications to the Policy based on the discussions in the Committee and during the Committee report to Council and send it all Council members by June 30 for a vote.

III. Report on Data Sharing for Demographic Research (DSDR) Archive

Staff provided an overview of this project, which began a five-year renewal in funding from NICHD beginning in September 2014. The project aims to strengthen its capacity to provide advice to researchers seeking to manage and archive their data, especially in the biomedical area. It seeks to build networks focusing on the topic of genomic data and is planning a summer workshop in this area in 2016. There are also plans to collaborate with the University of Michigan's School of Public Health to build partnerships with researchers in the school, particularly in the area of disclosure analysis with the goal of producing new data files for secondary analysts.

IV. Update on Status of the National Archive of Criminal Justice Data (NACJD)

Staff provided a brief update of the recent activities NACJD, which is comprised of:

- The Bureau of Justice Statistics (BJS), which collects, analyzes, publishes and disseminates information on crime, criminal offenders, victims of crime, and the operation of the justice systems at all levels of government.

- The National Institute of Justice (NIJ), which is the research and development agency established to prevent and reduce crime and to improve the criminal justice system. The NIJ Data Resources Program (DRP) collects, preserves, and disseminates data from NIJ grantees. DRP supports researchers using archived data in order to reproduce the original findings or test new hypotheses.
- The Office of Juvenile Justice and Delinquency Prevention (OJJDP), which provides national leadership, coordination and resources to prevent and respond to juvenile delinquency and victimization. OJJDP provides formula and block grants to states, territories and localities, as well as funding private organizations, including faith-based institutions through discretionary grants.

Recent developments included the search for a new permanent Archive Director, ongoing discussions with all three program officers from BJS, NIJ, and OJJDP, and a possible roundtable at the American Society of Criminology Annual Meeting in November.

Membership Services Committee

Council: Tony N. Brown (Chair), Marilyn Andrews, Philip N. Jefferson, William Vega

Staff: Linda Detterman, Lynette Hoelter, Dory Knight-Ingram, David Thomas

Visitors: David Metcalf and Justin Noble (staff guests)

I.   Strategic Plan Update

Linda Detterman discussed the two strategic initiatives most closely tied to the Membership Services committee, the Organizational Insights Team and the Innovations Group. The Organizational Insights Team (OIT) was continuing to assemble and evaluate data already collected about ICPSR data users and planned to undertake a survey about use of instructional materials this summer. The Innovations Group submitted nominations for some campus innovation awards and was working internally to recognize and reward "out-of-the-box" thinking, such as that required to launch openICPSR.

II.   Instructional Resources Update

Lynette Hoelter highlighted two instructional initiatives: (1) the work currently being done for the Bureau of Justice Statistics using some of their most common datasets (e.g., National Crime Victimization Survey) and (2) a module using the National Politics Study (NPS). Both of these efforts would result in new material for the ICPSR site. The NPS module was created by R. Khari Brown of Wayne State University and was modeled after the SETUPS resource. Khari would soon be meeting with ICPSR's Web Team to discuss how to bring the data to life on the Web (the data were currently being processed, so the goal was to have the module up by the fall).

Council noted some interesting institutions in the list of "high users" of Instructional Resources, including two South Korean universities. A short discussion followed about whether the usage was mainly a few faculty/classes in a couple of institutions who were heavy users or whether it was more broadly spread. Lynette noted that, according to the status attached to MyData accounts of users, the audience for instructional resources had become even more varied: whereas in the past it appeared to be primarily faculty driving usage, now graduate students comprised 36% of users, faculty 31%, and undergraduates 24%. This makes understanding who is using the resources and how they are using them even more critical.

It was noted that only one staff member in the Instructional Resources remained.

III.   RCMD Update

David Thomas discussed the status of the RCMD Director position: Director John Garcia would be retiring and be replaced by an Interim Director, Harold "Woody" Neighbors. Thomas reminded the group of the Provost's contribution – current and promised – to the funding of the directorship position. He also referred to the plenary presentation by Justin Noble and Amy Pienta and pointed out that two of the topics that people are interested in and searching for -- LGBT and immigrant data – were areas in which RCMD was already seeking to grow its collection.

Most of the time for this update was spent discussing how the interim period for the Director could be most productively used. Council acknowledged that the uncertainty that will accompany the layers of change in the next year (new Directors of both ISR and ICPSR) could make it difficult to define what is needed in RCMD's next Director but encouraged staff to think carefully about structuring the interim period so that momentum was not lost. The archive staff were currently undertaking a review of their short- and longer-term goals to assist in the process as well. Council asked for details about how the search would be handled, who would comprise the committee, and the like, but those details were not yet known.

IV.   Membership Status & Activities

A. State of the Membership

Detterman reported that 26 new members joined ICPSR in FY 2015 and 13 dropped, leaving a net gain of 13 for a total of 757. Revenues had exceeded what was budgeted by about $100K ($3.8M rather than the estimated $3.7M) and $600K of FY 2016 fees had already been collected. Detterman noted that the FY 2016 projected budget used $3.8M as the revenue estimate, leaving very little "wiggle room." A brief discussion about whether targeting corporate entities for membership made sense concluded that it was not worth putting a lot of effort in that direction because that is a group that is high churn, low fee ($2100 each), and for which we do not have much in terms of relevant data.

B. Membership Pricing

A discussion of membership pricing was held over from the March meeting because not enough Council members were present at that time to discuss the issues and vote. The Committee reviewed pricing information and recommended to the full Council a 5% increase in fees across the board to take effect in FY2017 and to be held constant for two years. This would result in a range of increases from $105 per year for Carnegie BA/Specialized institutions to about $825 per year for Extensives. This fee increase was seen as reasonable and the Council passed the motion unanimously in the open session following discussion.

V. Membership Outreach Activities

Detterman reminded Council of the list of activities and sessions for the upcoming OR meeting. She pointed out that there is a growing audience for the kinds of things ICPSR does, as evidenced by very high attendance at recent Webinars and presentations about how we can help researchers meet the federal data sharing requirements.

Preservation and Access Committee

Council: Carl Lagoze (Chair), Christopher H. Achen, Robert S. Chen, Ron Nakao

Staff: Sanda Ionescu, Tom Murphy, Asmat Noori, Mary Vardigan

I. Strategic Plan Update

The group discussed recent leadership-oriented activities at ICPSR related to the Strategic Plan. Staff had attended several meetings and conferences, many with a focus on metadata, including:

- A Research Data Alliance (RDA) workshop, held at the National Institute of Standards and Technology (NIST), on the topic of metadata

- An RDA-supported meeting on adding semantics to metadata

- A meeting of the National Data Service (NDS), an emerging national infrastructure for research data

- The third annual North American DDI Users Conference

II. FLAME Update

Tom Murphy led the group through the plans for developing FLAME (File Level Archive Management Engine), a comprehensive OAIS-conformant software system to ensure file-level control and management of data throughout the archival and data processing pipeline. He reported that requirements had been solidified and a three-year plan had been developed.

In terms of architecture, it was decided to start with the recently redesigned and revamped open source product Fedora 4.0 from Duraspace (ICPSR is a contributing member of Duraspace) with a Hydra front-end and to implement a Service Oriented Architecture along with a Message/Queuing component. FLAME would combine openICPSR and the legacy ICPSR for an integrated system going forward.

To arrive at the plan, CNS investigated different repository technologies, including Dataverse, as it is no longer feasible to write customized software. An important requirement was that there be an active open source community around the software. Fedora best met the requirements, and it also fit nicely with SEAD (Sustainable Environment, Actionable Data), an NSF DataNet project that recently moved to ICPSR. Another plus was that ICPSR sits on the Fedora Leadership Council and thus can have an impact on the product going forward. In addition, the University of Michigan Library would be implementing Fedora and Hydra.

It was reported that a Fedora instance was up and running at ICPSR and CNS was starting high-level development, The plan was to phase in new technology as we sunset older technology. The competition is getting better faster and ICPSR needs to be competitive again. The new architecture will permit ICPSR to handle new data types like GIS data, video, biomedical data, etc.

A question was posed about whether Fedora architecture would be robust enough for functions to migrate to cloud services. The response was that one of the newest features of Fedora 4.0 was asynchronous storage so that instances could be hosted in the cloud. A challenge going forward with cloud computing, however, is getting funding agencies to understand that cloud computing is secure.

It was pointed out that ICPSR created a gold standard for data curation through its data pipeline, which was built on the workflows that data curators need. Going forward a new more robust architecture is required and we need to standardize the workflow. Fedora would allow us to extend and build in a sustainable way. Incremental development and evolution are safer and sometimes make more sense, but they often end up costing more in the long run. ICPSR needs to take a bolder step and abandon the incremental approach.

A question was asked about whether ICPSR could fund any of this development out of other related projects. Actually, CNS planned to leverage some of the SEAD work, so there would be synergies there.

It was mentioned that while Fedora was good for broad access, there might be a challenge with respect to ICPSR's domain-specific DDI-oriented approach. The more general Dublin Care is fine for other organizations but is not adequate for ICPSR.

Council summarized that CNS had put together a good plan. Currently, ICPSR has a home-grown repository architecture that is not up to modern standards. We should not build something home-grown again but instead build an implementation that is based in open source. Being OAIS-conformant was an important requirement and the decision was to go with Fedora as the

best alternative. Fedora has been around for 16 years and it has an active user community. To succeed, we need modern architecture. This is a significant budget outlay but it is an investment in the future.

III.   CNS Staffing Update

Tom Murphy gave a verbal staffing update. The recent SAMHDA loss affected staffing, and CNS was holding off on new hires because of budget issues, despite the fact that CNS was down two people (in March it was down four people and had recently hired two). There were 13 current members of CNS and a total of 7 developers. Given the plans for the future in implementing FLAME, there were some mismatches in skillsets for the tasks to be performed. Developers were being retrained to mitigate this problem.

Council pointed out that ICPSR is a technology enterprise and IT is critical to its success. In terms of strategic priorities, it might not be prudent to hold steady or reduce the size of the CNS staff despite budget concerns. Also, there had been a dip in morale, and uncertainty in terms of job security was causing CNS staff to look for new jobs. This was troubling because it is difficult to hire and retain IT professionals. In IT there is a lot of competition for people with these kinds of skills, and losing staff can have a higher impact than in other areas. The team was already below strength, and it might be worthwhile to take steps to reduce the risk of higher impact losses. Also, IT is the source of bigger grants so it is a core function and an increasingly important component of being competitive with respect to big data science and related initiatives.

Another issue raised was that ICPSR had not investigated mobile apps and its current staff did not have those skills. The newer technology would be mobile-friendly in the sense that websites were being reconfigured according to responsive design, but more effort on this front would be required.

Council reiterated that it was its job to look at overall priorities and to make recommendations in the best interest of ICPSR in the near and future term. Given the programmatic objectives and the centrality of IT to those objectives, it was Council's view that IT infrastructure development needed to continue but that CNS also needed to deal with mismatches of skillsets and to stabilize after recent cuts.

IV.   Update on Archival Backup Procedures and Policies

At the previous meeting in March, Council advised that it was not necessary for ICPSR to maintain the large number of replication copies (eight) it was currently keeping and requested that staff look for ways to reduce this number. In response, staff made the decision to cut one incomplete copy and to decentralize the distant copies that previously had all been synchronized through DuraCloud since this was a potential vulnerability.

These actions would result in a small cost savings of about $5K, but the diversification of copies was seen as a more important outcome. Council asked about a worst-case scenario – for example, a disaster occurring at the University of Michigan with the result that the local copies and tapes were destroyed. ICPSR would then need to pull data from the backups.

It is important to address such potential threats. Another vulnerability is restricted data, which are encrypted. If we lost the encryption keys, we would not be able to access the data. A worm or virus that destroys all copies is another potential disaster, but in those cases we could restore from our tape backup.

ICPSR was in contact with the Digital Preservation Network, which advocates for the use of diversified technologies for preservation to guard against loss and there is a possibility to work with them in the future. ICPSR's problems are not unique and it makes sense to collaborate with others.

Council also raised the issue for future discussion of our incomplete backup of video data. If ICPSR brought the video data into the current replication system with the same number of copies to guarantee archival protection, the cost would be about $45K. This is something we need to think about as the data are only minimally protected right now.

V.   Demonstration of Metadata Portal Tools

Sanda Ionescu presented two new DDI-based tools to assist in data discovery and classification: an integrated ANES-GSS crosswalk and a tagging tool for classification and comparison. These tools were deliverables of the NSF-funded Metadata Portal project. For that project converted all of the ANES and GSS documentation to DDI at the study and variable levels.

The crosswalk incorporates variables from 30 ANES time series and 29 GSS waves, and variables are classified by concepts and year. For input, we used the ANES cumulative file and the ANES core utility as well as the GSS online interactive subject index and the GSS codebook subject index.

To create the integrated crosswalk, staff matched the GSS variables to the ANES variables by topic and harmonized the subject headings into unique ICPSR concepts for the items common to both time series. The result was 144 sets of mapped variables/concepts, which were grouped into eight broad categories for browsing.

For the GSS ICPSR also created extracted variables by year so that each wave was searchable and comparable on its own. Cross-wave comparison within and across series was enabled.

The new tagging tool allows users to create their own lists of keywords and to tag variables and build customized crosswalks. While this functionality was currently only available on the Metadata Portal site for the ANES and GSS, it could be extended to the full ICPSR holdings.

Council expressed its enthusiasm for this project because it showed that ICPSR was leveraging its strength in DDI to enhance discovery of and access to data in the archive. This would continue to have an impact and ICPSR would be seen as a leader in terms of tools development and leveraging DDI for others to use. The functionality would likely be adopted by others because they have the same needs.

Summer Program Committee

Council: Chandra L. Muller (Acting Chair), Colin Elman

Staff: Saundra Schneider, Edward Czilli, Dieter Burrell

I.   Strategic Plan Update: Inclusion and Diversity Efforts

Summer Program Director Sandy Schneider reviewed the Summer Program's past accomplishments in increasing diversity among its instructional staff, within its participants, and across course offerings. The 2014 Summer Program had 37 instructors (33% of the total) from under-represented groups and 422 minority participants (45% of the total). In addition, the 2014 Summer Program curriculum offered a four-week workshop on "Methodological Issues in Quantitative Research on Race and Ethnicity," a short (five-day) workshop on "Designing, Conducting, and Analyzing Multi-Racial and Ethnic Political Surveys," and a series of three presentations in the Hubert M. Blalock Lecture Series on topics pertaining to inclusiveness and diversity.

Schneider reported on the Summer Program's plans to further enhance diversity, particularly in the 2015 Program's course offerings. These included: continuing the four-week workshop on "Race, Ethnicity, and Quantitative Methodology" (previously titled "Methodological Issues in Quantitative Research on Race and Ethnicity"), offering a short workshop in conjunction with the Robert Wood Johnson Foundation New Connections Group on health disparities among immigrant populations, and offering a series of presentations in the Hubert M. Blalock series on "Race, Ethnicity, and Gender: Identification, Measurement, and Analysis Issues." This was followed by a discussion of the Program's outreach efforts to increase diversity among its participant population (through partnerships with other programs within and outside of the University of Michigan).

II.   ICPSR Summer Program Scholarships

Summer Program staff reiterated the decision of the ICPSR Council at the March 2015 Council Meeting to approve a budget of $101,500 to fund Summer Program scholarship awards in 2015. As of June 5, 2015, the Summer Program had received 97 applications for these scholarships.

The status of other Summer Program scholarship awards was also discussed. The Society for Political Methodology had received 18 applications for the Garcia/Box-Steffensmeier scholarships, making 6 awards; 17 University of Michigan graduate students received Rackham Summer Training Awards to attend the Summer Program; and 5 students received EITM scholarships to cover fees for Summer Program participation.

This was followed by a brief discussion of the status of the Under-Resourced Institutions Awards. The ICPSR Council approved the Summer Program's request to put the Under-Resourced Institution Awards for 2015 "on hold" until the eligibility criteria and implementation procedures for these awards could be reviewed more carefully.

III.   Overview of the 2015 ICPSR Summer Program

Schneider gave an overview of the schedule of courses for the 2015 Summer Program. A total of 85 courses were planned: 40 courses in the four-week sessions running from June 22 through August 14, and 45 short (two- to five-day) courses offered from early May through mid-August.

Several changes to the 2015 Summer Program course schedule since the March 2015 Council Meeting were announced: one new short statistical workshop ("Introduction to Regression Analysis") was added, to be offered on the Uniformed Services University of the Health Sciences Campus in Bethesda, Maryland; five short courses that had originally been planned were cancelled due to low enrollments.

The Summer Program staff discussed the advantages and disadvantages of offering ICPSR Summer Program short courses in remote locations. The courses can serve as important "feeders" to promote participation in the regular four-week sessions at Ann Arbor, but can be difficult to organize and monitor.

IV.   Overall Developments in the Summer Program

The Summer Program staff and ICPSR Council members discussed some of the issues confronting the Summer Program and the plans to address these issues. These included: maintaining and increasing the ability to offer a comprehensive set of courses in statistical analysis and research methodology, from introductory-level classes to workshops on advanced, innovative, and "cutting-edge" topics; expanding the reach of the Summer Program into academic disciplines and fields that have not been well-represented in the past through more targeted marketing initiatives; assessing the impact of the ever-growing number of methodological training programs, summer boot camps, intensive learning workshops, and online instructional modules; and re-assessing the Program's off-site course offerings to determine the best locations and the right mix of courses across the sites.

The Committee also discussed the financial and organizational challenges facing the Summer Program. The Summer Program had made concerted efforts to keep the Program's budget in check even through there had been substantial increases in Program expenses and charges. The Program planned to update and reconfigure its administrative operations so that it would be better positioned to conduct promotional activities, compete with other training programs, and attract a broader, more diverse set of participants.