Archival Development Committee Minutes  
March 5, 1999

Notes from meeting held on March 5, 1999 - Los Angeles, UCLA Campus

**Attendees:** Charles Betsey, Chris Dunn, Peter Granda, Ed Nelson, and Libbie Stephenson (co-chair)

Several areas for further investigation were identified at the October 1998 meeting, and committee members agreed to bring ideas and materials to the March meeting:

1. Evaluation of substantive data quality as a selection criteria.
2. Scope of ICPSR's acquisitions in relation to emerging research areas.
3. ICPSR's role in national (U.S.) archival policy.
4. Privacy, security and confidentiality and effect on selection criteria.

Data Quality

ICPSR has accepted virtually any data any researcher has wished to deposit. Evaluations of deposited data have focused on the completeness of documentation, data cleaning (a series of procedures to clear up missing data, blanks, multi-punch, identifiers and the like) and data formats (raw, system, etc.). An assessment of the scholarly quality of the data is not made. The Committee was asked to make recommendations on data quality issues in the acquisition/collection policy. The following points resulted:

1. Determining the research value of data is best accomplished through peer review and publications.
2. Data may have methodological problems and still be useful for certain types of analysis.
3. Use of clear and complete documentation is of great value in determining the validity/value of data.
4. Documentation of research design is as important as documentation of data file structure.
5. ICPSR staff should be given the option of refusing poorly documented data.
6. A set of criteria and conditions under which data will not be accepted will be provided in the acquisitions policy and these will also be added to the Data Preparation Manual for depositors. (See attached draft prepared by Ed Nelson)

Emerging Research Areas

The Committee believes ICPSR acquisitions should depend on the requirements of data users in the social science community. The ICPSR holdings should, first and foremost, support core subject areas in the social sciences. At the same time new research areas and new research techniques are evolving, and ICPSR should acquire these materials as well. The Committee has been able to review current ICPSR collection habits in terms of what has been acquired and used in the past. However, these details may not be useful in determining future research areas for which ICPSR should acquire data.

The Committee supports a proactive rather than reactive approach to acquisitions. The Committee recommends that procedures be put in place and administrative support be provided so that Senior staff can establish sustained links to the research community, in order to keep track of new research areas and methodologies. Other avenues for data discovery are the Summer Program, the OR meeting, attendance at key conferences, and subscriptions to major social science journals. The proposed conference on archiving could also heighten awareness of new types of data being used and created. The Committee considered the idea of holding such a conference at regular intervals (every 2-3 years, for example), perhaps in conjunction with the biennial OR meeting.

National Archival Policy

At the October 1998 full Council meeting, Hal Winsborough (Chair, ICPSR Council) proposed the idea of holding a conference addressing ICPSR's role (as well as the place of social science data) in the national archival arena. In discussion that followed, suggestions included:

1. Conference planning should involve the professional archival community.
2. ICPSR member institutions and their archival/data service situation would merit discussion.
3. Representatives from the National Archives, Center for Electronic Records, as well as the professional archival staff at ICPSR should be part of the program.
4. Conference program should include key historians and legal experts in the field of electronic information.

Names of suggested participants were forwarded to Winsborough. (Note: A draft program for the proposed conference has been developed. See [below](https://www.icpsr.umich.edu/web/pages/about/governance/minutes/1999-03.html#agenda).)

In the March meeting, discussion of the proposed conference focused on the following:

1. Need for involvement by ICPSR professional archival staff.
2. Need for participation by representatives from NARA, SAA, APDU, and IASSIST.
3. A session on the archival setting at academic (ICPSR member) institutions would be useful. Consideration should be given to the international nature of membership.
4. It is not clear from the title of the conference what is the intended outcome. (i.e., ICPSR already has a working archival policy)
5. How is Council involved? How should Council proceed as a result of the Conference?
6. Should the Archival Development Committee make a presentation regarding ICPSR's archival and acquisitions policies?
7. How does the conference affect ICPSR current and developing archival and acquisitions policies?

Further discussion of the conference will take place at the June Council meeting.

In the working document for a policy on acquisitions the Committee continues to support a substantial role for ICPSR in the social science research community. The Committee has focused on a desire to build and maintain a "comprehensive" rather than "complete" collection of the highest use and value to social science research.

The Committee also supports coordinated relationships with archives internationally. Informal relationships between ICPSR and European archives exist even now. One manifestation of a comprehensive archive could be accomplished through joint acquisition/collection relationships with other archives.

The Committee discussed the financial aspects of a national role for ICPSR. If ICPSR takes on a role of national archival proportions, then the cost to accomplish this would need to be considered. The Committee believes the current level of funding acquired through member dues and funded archives would not be adequate. Further, the concept of a national archive suggests that ICPSR's holdings would be freely available to anyone. This could imperil the membership structure of the Consortium and could result in significant loss of revenue.

Privacy, Security, and Confidentiality

Currently ICPSR Bylaws provide that data are acquired to be freely available to any members. Some data producers are now requiring substantial user registration and legal documents from users of data deposited at ICPSR. Much of the time these measures are thought to provide respondents with greater privacy and confidentiality and still provide unique resources for research. The Committee discussed the effects increases in producer-related distribution restrictions will have on data access, and on the liability ICPSR or Official Representatives might face if data are misused.

In terms of an acquisitions policy, the Committee has addressed the potential for developing the ICPSR collection through acquisition of sensitive materials. The goal is toward public data and equitable access, with an acceptance for reasonable levels of restriction. This approach should actually result in access to many kinds of data not now available. The Committee hopes to develop an acquisition policy that will accommodate data producers but not place undue responsibility or work load on the ICPSR archival staff. The Committee should review and discuss acquisition of specialized data of great value to research, but which have significant access limitations.

Current and evolving legislation on data access and information policy should be monitored. The acquisition policy should be routinely evaluated as a result of any new laws.

The Committee briefly discussed the issue of donor rights. For example, does a donor have the right to have data returned? If so, how should ICPSR monitor and retrieve copies that have been acquired by member institutions? How can ICPSR ensure that deposited data are accurately cited? How does ICPSR maintain and guarantee version control for data files supplied to users. Are there copyright issues to be addressed for instructional materials acquired by ICPSR? These issues will be further addressed at the June Council meeting.

Archiving Social Science Data:  
Tentative Agenda for Fall Meeting

1. Existing Archives and Distribution Nodes: Status Reports (use of Websites and background documents prior to meeting) and Self Appraisals (facing up to challenges and limitations as viewed by the unit representatives)
   1. US and European Archives and Networks/Consortia
      * ICPSR: Richard Rockwell (Michigan)
      * International/European: Mochmann (Cologne) or Essex Rep
   2. Restricted Data Centers/Data Enclaves
      * Census-inspired: David Card/Haltiwanger/Joe Hotz (Chicago)
      * Other, re "sensitive" data:
   3. Large Database Distributors Using Servers/Web
      * PSID or HRS/AHEAD: Frank Stafford/Bob Willis (Michigan) or Lux Income Study: Tim Smeeding (Syracuse)
      * Historical censuses: Steve Ruggles
      * Landstat/GIS: Shiller/Roberta Miller (Columbia)
   4. Federal Statistical Suppliers
      * Statistics Canada: Ernie Boyko
      * Bureau of Labor Statistics: Katherine Dippo
   5. Nonacademic Suppliers/Archives
      * Roper: Everett Ladd
      * Commercial: J. J. Card
2. Critical Appraisal of Existing Infrastructure by Users/Sponsors
   1. Funders/Sponsors
      * NSF: Bill Butz
      * NIA: Richard Suzman
      * Foundations: Richard Quandt (Andrew Mellon Fdn)/ Michael Teitelbaum (Sloan)
      * European Science Foundation or Brussels: John Smith (ESF Secretariat)
   2. Users
      * Large policy "think tanks": Jim Smith (RAND)
      * Federal policy: Rebecca Blank (Council of Econ Advisors/Michigan) Katherine Wellman (OMB)
      * Undergrad teaching/instruction: Henry Farber (Princeton)
      * Major research universities: Robert Fogel (Chicago)/Larry Bartells (Princeton)
3. New Scientific Opportunities/Challenges and Implications for the Nature of Data, Archiving, and Distribution
   1. Functional Magnetic Resonance Imagery for brain/behavior interactions: Bill Eddy (CMU) [Should we try for a European?]
   2. Sequential Analysis of Classroom Interaction (TIMSS): Jim Stigler (UCLA)/Bob Baruch (Penn)
   3. Streetcorner Sounds: Mitch Danier (Wisconsin)
   4. Keystroke Assessments in Cognitive/Education Research: Ken Koedinger (CMU)
4. Next Steps Toward Meeting the Challenges of New Opportunities for the 21st Century
   1. Metadata and delivery: Gary King (Harvard)/Merrill Shanks
   2. Oversight and governance: David Featherman (Michigan)/Barbara Torrey (CBASSE)
   3. New analytical tools: Finis Welch (UCLA)/Steve Fienberg (CMU)
   4. Confidentiality: Fritz Scheuren/Norman Bradburn
   5. System design: Dan Atkins/Jose Marie Griffiths (Michigan)
   6. Research and digital libraries: Sid Verba (Harvard)/Wendy Lougee (Michigan) or Dan Waters/Clifford Lynch (Coalition for Networked Information)