

## Content Identification

The Data-PASS partnership is continuing the work of identifying at-risk social science data. These efforts are taking place both at the partnership level as well as the individual organization level. One of the biggest tasks has been undertaken by ICPSR in our search of the National Science Foundation (NSF) Award database for funded social-science projects. This ongoing search is producing leads for ICPSR and our partner organizations.

### ICPSR

#### NSF Grants Database Search

- Review research grants funded by NSF
- Using the NSF Award Search at [www.nsf.gov/awardsearch/](http://www.nsf.gov/awardsearch/)
- NSF Database is divided into
  - > Historic awards (pre-1976)
  - > Current awards (1976 to the present)
  - > Expired awards (1976 to the present)

#### NSF Results: Pre-1976

- 103,674 records reviewed
- Selected 1,476 “promising” titles
- Historic records are incomplete

#### NSF Results: 1976 to the Present

- Initial selection criteria — NSF program codes
  - > Social, Behavioral, and Economic Sciences (SBE)
  - > Division of Behavioral & Cognitive Sciences (SBE/BCS)
  - > Division of Social and Economic Sciences (SBE/SES)
  - > Division of Science Resources Statistics (SBE/SRS)
- 17,195 records downloaded from the NSF Award Search database
- Secondary screening criteria (ongoing)
  - > Data collection or new data resource created

#### Examples of Studies Identified:

- Beck, Elwood. *Racial Violence in the American South in the Decade of the 1980s* (NSF 9308169, 1993)
- Bose, Christine. *Reframing Women's Work in 1900* (NSF 9320663, 1994)
- Hurlbert, Jeanne. *Social Resources Activated in Response to a Natural Disaster* (NSF 9224444, 1992)
- Huston, Ted. *A Longitudinal Study of the Antecedents of Distress and Divorce* (NSF 9311846, 1993)

### The Roper Center

- National Opinion Research Center
  - > A list of projects has been obtained.
  - > Data for over 80 surveys not in the archive have been identified.
- Traces
  - > Surveys for which the Roper Center has reports, but not the actual data.
  - > Special surveys by the Harris organization have been targeted. The Odum Institute archives the regular monthly surveys while the Roper Center will continue to archive the special surveys.

### The Odum Institute

- State polling data / The National Network of State Polls (NNSP)
  - > Database of state polling organizations has been updated. Began contacting the identified state polling organizations as well as non-NNSP organizations (e.g., Elon University).
- Harris Polls
  - > After identifying gaps in the Odum collection of Harris Polls and making contact with Harris Polls, we have begun to receive non-proprietary missing polling data. Receipt of this data has provided a test-bed for developing and testing processes and procedures.

### The Murray Research Archive

- Qualitative data
  - > Identified approximately 150 individual datasets containing relevant data.
  - > A small number of leads were selected for initial follow-up. Currently in discussion with six leads interested in contributing data in the near term.

### NARA

- NARA recently accessioned the following records:
  - > National Ambulatory Medical Care Survey, 1973–1976
  - > National Hospital Discharge Survey, 1979–2002
  - > Various Census surveys, e.g., Truck Survey and Use, 1977 and 1987
  - > Multiple Birth Data Set, 1995–2000

## Partner Collaboration

Collaborative efforts among the partners to identify relevant data continue. We have also identified additional ways to collaborate, both between partners and with other data resources.

- Data-PASS will prepare guidelines on how the partnership will process fragile material such as punch cards and magnetic tape.
- The Roper Center and the Odum Institute will coordinate efforts regarding their collections of Harris datasets. They will develop a strategy to approach the Harris Organization to obtain Harris polls not contained in either of their collections.
- NARA and the Roper Center have begun to copy and transfer US Information Agency (now part of the Department of State) international surveys to and from NARA and to locate the missing data. As a result, a large block of records from 1993 to 1999 were transferred, including 263 data files on seven CDs and 27 cubic feet of printed technical documentation.

### Other Collaborative efforts

#### The Odum Institute

- RTI International
  - > RTI has agreed to identify contracts that include data collection. Their initial results number about 19,000. Further screening efforts aim to identify a smaller number of important studies.

#### The Roper Center

- National Opinion Research Center
  - > Data for at least 80 surveys that are not part of the Roper Center archive will soon be acquired.
  - > Additional inquiries will investigate surveys on the NORC list that are not archived at either the Roper Center or NORC.

### NARA

- NARA has embarked on a major outreach effort within the federal community to encourage agencies that create and use electronic records of permanent value to transfer those records to NARA as soon as possible.

#### The Murray Research Archive

- Identified 26 research centers with relevant data.
  - > One center has been identified for initial follow-up to determine interest in contributing data in the near term.

### Building an Open Infrastructure for Preservation

As part of our collaborative effort, a shared catalog of collected material will be made available using the Virtual Data Center software. Currently, the holdings of the Murray Research Archive are available through the VDC. The Odum Institute has transferred its catalog to the VDC, and is in the process of moving its VDC node to Chapel Hill, North Carolina. NARA is in the beginning stages of making its records available through the VDC.

Virtual Data Center (VDC) software is a comprehensive, open-source, digital library in a box, designed to help curators and researchers disseminate, cite, and preserve quantitative data.

Individual archives provide:

- Specialized study preparation
- Quality assurance
- Local, secure, redundant storage
- Administration
- Localized dissemination

VDC tools support:

- Study ingest and management
- Interoperability
- Dissemination and online analysis
- Distribution and federation
- Citation management

Figure 1. Sample of Virtual Data Center output.

