



The Kalmanovitz Library and
The Center for Knowledge Management

A Data Sets Service in the Library: Overview and Evaluation

Peggy Tahir, Interim Manager, Access Services;
Min-Lin Fang, Information Services Librarian;
University of California, San Francisco
May 23, 2004





Data Sets Access Problem

- Researchers at UCSF need access to data sets
- Prior to project, no coordinated program available to assist them with finding and accessing data sets



Initial Project

- Investigate the use of data sets on campus
- Identify resources and availability of data sets
- Determine what the library could do to promote access to data sets resources and services
- Initial project began in mid-2000



Needs Assessment

- Availability of data sets and tools in the library's collections and on campus
- Determine needs of UCSF researchers for health data sets resources and services
- Used Facilitate.com for online survey, a web-based conferencing software
- Multi-disciplinary collaboration



ICPSR Service

- Library Administration approved additional funding for ICPSR membership
- Data archive contains preprocessed data for longitudinal studies in social sciences with documentation, related reports and SAS and SPSS
- Allows UCSF affiliates to download raw data sets and documentation from ICPSR
- Coordinated by Data Management Services
- Started in September 2001



Outcomes

- Reference staff developed web page with links to data sets resources
- Collaboration between reference staff and data manager to assist patrons
- Reference staff initial contact for ICPSR assistance



Phase II: Evaluation of ICPSR Service

- End of Year 1: before ICPSR Direct
 - Increasing use of service
 - Well-attended orientations
 - Objection to usage fee for selected studies
 - Desire for free and direct access



New ICPSR Direct Service

- Library changed terms of ICPSR subscription to ICPSR Direct
- Started in July 2002
- Provides direct access to the ICPSR data holdings for all students, faculty, and staff at UCSF
- Users with authorized UCSF IP address may download data from ICPSR, without going through a local campus representative



ICPSR Direct Usage

- Significant increase in data sets usage since library moved to ICPSR Direct
- Year 1: before ICPSR Direct (2001-2002)
 - 79 data sets downloaded
- Year 2: ICPSR Direct (2002-2003)
 - 493 data sets downloaded
- Data sets downloads in this period was more than six times the number of downloads vs. the period before ICPSR Direct



ICPSR Direct: Mean Usage

Comparison with similar sized institutions

- Year 1: before ICPSR Direct
 - Average use = 176
 - UCSF use = 23
- Year 2: ICPSR Direct
 - Average use = 384
 - UCSF use = 493



2004 Survey

- Email survey
- Very short - 8 questions
- Sent it out three times
- Spring 2004

May 23, 2004

The Kalmanovitz Library and the
Center for Knowledge Management

UCSF



Target Groups

- Datausers listserv subscribers
- Researchers who downloaded data sets from ICPSR
- Researchers who borrowed OSHPD data sets



Survey Questions

- Recently used data sets
- How they obtained data sets
- If they obtained data sets from ICPSR / satisfaction with service
- If they obtained data sets from OSHPD



Response Rate

- Survey sent to 125 individuals
- Received 26 usable responses



Responder Profile

- 18 faculty
- 7 academic staff
- 1 post-doctoral fellow

May 23, 2004

The Kalmanovitz Library and the
Center for Knowledge Management

UCSF



Responses - ICPSR

- Nearly half of respondents had downloaded data from ICPSR
- Were able to find all they required
- Found the service very useful
- Two respondents commented specifically on the ability to obtain historical data sets via ICPSR



Commonly Used Data Sets

- National Health Interview Survey
- California Health Interview Survey
- Medical Expenditure Panel Survey
- AHEAD (Asset and Health Dynamics Among the Oldest Old)
- National household survey on drug abuse



Responses - OSHPD

- Nearly half of respondents had retrieved data from OSHPD
- 9 researchers used patient discharge data
- 8 researchers used hospital financial data
- 1 used long term care financial data
- Respondents indicated more than one type of data used



Conclusions

- Patrons want to be able to access data sets on their own
- Even small costs are barriers to access
- Project provided excellent venue for outreach to researchers on campus
- Successful collaborative project between library units
- Increased knowledge of data sets and patron access issues
- Difficulty in obtaining statistically significant sample using both online and email survey methodologies



Additional Resources

- California Office of Statewide Health Planning and Development's (OSHPD) Data Sets
- Bureau of Health Professions' Area Resources File
- Counting California
- RAND California