A SURVEY OF ICPSR MEMBER INSTITUTIONS

Ann Janda
Vogelback Computing Center
Northwestern University

INTRODUCTION

At the meeting of the ICPSR Official Representatives (ORs) held at Ann Arbor on November 11-14, 1983, a brief survey of member institutions was handed out to attending members. The survey was drawn up at Northwestern University by Lorraine Borman and Ann Janda (Vogelback Computing Center) in an effort to find out what other universities were doing in planning for a distributed computing environment. The questionnaire asked for information about what services are provided to access ICPSR data, mainframe and software usage, current extent of micro usage, and a final open-ended "future plans" question. Out of approximately 150 meeting participants, 38 responded to the questionnaire.

This report discusses the results of the survey and provides two appendices: (A) an index to universities and their mainframes, and (B) a listing by university describing hardware, software, services, and future plans. (Appendix B available from Ann Janda, Vogelback Computing Center, Northwestern University, Evanston, Illinois 60201.)

COMMENTARY ON SURVEY RESULTS BY QUESTION

1. Availability of ICPSR hardcopy codebooks:

18 at departmental site
15 at university library
9 at central computer site
6 at some other location

The above results show that most institutions housed their codebooks at the department with the university library running a close second. The actual totals to this question ran higher because from the 38 responding institutions, 7 located their codebooks at more than one site, generally with a department/library combination. Research institutes largely fell into the "other" category, accounting for 4 of these 6 responses; one research institute identified with a department.
2. Computer programs such as database management systems to identify and access datasets? If yes, which?

   13 Yes
   25 No

This was probably an ambiguous question. For those who answered "yes", the responses fell into two categories: (1) software that provided tape access to the data or (2) software that identified the studies by title or abstract. No two programs mentioned were alike. The software programs are listed in Appendix B.

3. Personal assistance in identifying and accessing datasets?

   23 at departmental site
   12 at central computer site
   8 at some other location
   7 at university library

According to the above results, by far the greatest assistance in identifying and accessing datasets is given in the department. As in question 1, more than one answer was checked, and indeed, 7 departments shared this function with the library and computing center. Most of the responses in the "other" category pertained to social science research laboratories or institutes.

4. Computer programs for graphic analysis of data? If yes, which?

   24 Yes
   11 No

Of those who responded "yes", most used SAS/GRAF. SPSS PLOT was another popular package as was TEL-A-GRAF Display (spelled 3 different ways on the questionnaires). The graphics programs are also enumerated in Appendix B.

5. Consulting on statistical analysis of ICPSR data:

   25 at central computer site
   24 at departmental site
   9 at some other location
   1 at university library

The above figures splitting the statistical consulting load between the department and the computing center accurately reflect the number of multiple responses falling into the combined department/computing center pattern. Slightly more than
half of the institutions provided this combination of statistical consulting. Most of the research labs ("other" location) shared this function with both the computing center and the department.

6. How are your ICPSR data stored?

37 magnetic tape
20 system files, e.g., SPSS or SAS files
15 raw data files
14 disk packs

Almost all of the member institutions store their data on magnetic tape—probably as they are originally received from the ICPSR. In addition to storage on magnetic tapes, more than half convert at least some of the data to SPSS or SAS system files. Disk packs also seem to enjoy considerable use.

7. Do you offer subsetting services for large data files?

20 Yes
17 No

The responses indicate that a little more than half of the institutions provide some form of subsetting for their users. Although the responses were straightforward, it is not clear whether the services provided only consulting and programs enabling the user to subset the large dataset—or actually "doing the subsetting" for the user and delivering a smaller dataset based on user specifications. The latter option was intended.

8. Is documentation available online?

17 Yes
19 No

Although the responses to this question showed a fairly even split among ICPSR members in providing online documentation, some of the additional comments to this question indicated that this may have also been ambiguously phrased. While a few comments referred to availability of "abstracts", a few others referred to the use of machine-readable codebooks and dictionaries on tape. Obviously, online documentation could be construed either as study descriptions/abstracts—or as actual data documentation, e.g., codebooks and data dictionaries. While the online study descriptions were the target of this probe, the online codebooks have more interesting and powerful applications.
9. **What mainframe computer(s) are used for storing and processing ICPSR data?**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>18</td>
<td>DEC</td>
<td>13</td>
<td>CDC</td>
</tr>
<tr>
<td>PRIME</td>
<td>3</td>
<td>Others</td>
<td>3</td>
<td>UNIVAC</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The above figures indicate that leaders in hardware are IBM, DEC, and CDC—in that order. Within the "other" category are two AMDAHLs and a Hewlett-Packard 3000. A more detailed enumeration of specific models by university is indexed in Appendix A.

10. **If more than one mainframe is used, how are data transferred?**

In all, there were 12 responses to this question, ranging from the use of DECNET, ETHERNET, tapes, to locally developed software and utilities. Specifics are found in Appendix B.

11. **What mainframe analytical packages are used with ICPSR data?**

<table>
<thead>
<tr>
<th></th>
<th>Mostly</th>
<th>Sometimes</th>
<th>Seldom</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPSS</td>
<td>35</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SAS</td>
<td>10</td>
<td>11</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>BMDP</td>
<td>1</td>
<td>13</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>OSIRIS</td>
<td>0</td>
<td>8</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>SIR</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>27</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

It is no surprise that SPSS and SAS are the most commonly used statistical packages for the mainframe. BMDP comes in as the next contender with high usage in the used "sometimes" and "seldom" categories. The comparatively low frequency of usage for SIR could most likely be explained by its relative newness in the market. Among packages named by respondents in the "other" category were MINITAB, ABC, TSP, SCSS, MIDAS, DATATEXT, and TROLL. From this group SCSS was noted as most commonly used by one institution; the remaining packages fell largely into the "sometimes/seldom" category with MINITAB usage noted at 5 institutions.

12. **What use is made of microcomputers in analyzing ICPSR data?**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a great deal</td>
<td>0</td>
<td>some</td>
<td>8</td>
<td>none or virtually none</td>
</tr>
</tbody>
</table>

Micros have not made much headway as consistently used tools for statistical analysis—yet. Although the number of responses in the "none or virtually none" category seems high, a few were
qualified by remarks referring to the early arrival of additional micros, pending changes to increase the use of micros, and current use by individual faculty and user groups. Also as more statistical packages become available, more use is anticipated.

13. If microcomputers are used for analyzing ICPSR data on your campus, please describe how the datasets are downloaded for use:

There were a total of 9 responses to this question. Two institutions used KERMIT; the remainder variously responded with COMTTY, SMARTTERM PC, locally developed software, public domain software, an interactive mainframe program via SUPERWYLBUR—on through "the problem is currently being addressed."

14. If microcomputers are used, what are the most common machines and analysis programs?

From 11 responses, the most frequently mentioned micro was the IBM PC followed by the APPLE IIE. IBM-XTs appeared among the sprinkling of Victors, Zeniths, TERAKS, Radio Shacks (probably TRS-80s), and TI Professionals. No clear pattern of software use emerged. Comments referred to uses other than statistical, e.g., as spreadsheets and as terminals hooked to mainframe statistical programs. It appears that presently micro stat packages are largely under consideration for use/purchase, e.g., one of the institutions is reviewing Statgraphics, Statpak, and other programs.

15. Are you planning to change your existing method of access and distribution in the next two-three years? If so, how?

Twenty three respondents answered "yes" to this question. Their elaborations mainly pointed to the increased use of micros. For many this increase is linked with networking, workstations, shared storage, and subsetting datasets for micros. A few mentioned a move toward creating a social science laboratory. The complete references to future plans are printed in Appendix B.

SUMMARY: WHAT HAVE WE LEARNED?

We have reviewed information on hardware, software, and services provided to users of ICPSR data at 38 member institutions. For most institutions, the department is the most important site providing services to data users, e.g., codebooks and personal assistance in identifying datasets. The central computing site is next most important, and the library's role
is limited largely to storing codebooks. SPSS is by far the most common program for analyzing ICPSR data, and about half the institutions store data prepared for analysis in SPSS or SAS system files. ICPSR data are still analyzed primarily through computers at the central site, despite the popularity of microcomputers. Indeed, no respondent reported extensive use of microcomputers, and 30 said that they made little or no use of microcomputers at present. Although most respondents reported future plans for analyzing ICPSR data with microcomputers, the revolution has not yet occurred.

APPENDIX A: UNIVERSITIES AND MAINFRAMES INDEX

ARIZONA STATE UNIVERSITY
AUBURN UNIVERSITY
AUSTRALIAN NATIONAL UNIVERSITY
BARDOIN COLLEGE, CUNY
BOWDOIN COLLEGE
BRYN MAWR COLLEGE
CALIFORNIA STATE UNIVERSITY, FRESNO
CARNegie-MELLOn UNIVERSITY
CENTRAL MICHIGAN UNIVERSITY
COLBY COLLEGE
CORNELL UNIVERSITY
FLORIDA STATE UNIVERSITY
HUNTER COLLEGE, CUNY
INDIANA UNIVERSITY
LOYOLA UNIVERSITY, CHICAGO
NORTH TEXAS STATE UNIVERSITY
NORTHWESTERN UNIVERSITY
OHIO WESLEYAN
PRINCETON UNIVERSITY
RUTGERS UNIVERSITY
SSRC DATA ARCHIVE
SwarTHMORE COLLEGE
TEMPLE UNIVERSITY
TEXAS TECH UNIVERSITY
TRINITY COLLEGE, YALE FEDERATION
UNIVERSITY OF ALBERTA
UNIVERSITY OF CALIFORNIA, LOS ANGELES
UNIVERSITY OF FLORIDA
UNIVERSITY OF IOWA
UNIVERSITY OF OREGON
UNIVERSITY OF PITTsburgH
UNIVERSITY OF SOUTHERN CALIFORNIA
UNIVERSITY OF UTAH

IBM, MODEL: 3081
IBM, MODEL: 3033
DEC * UNIVAC
IBM
DEC, MODEL: 10
OTHER: HP 3000
CDC, MODEL: 170 SERIES, 720 LOCALLY---730 AND 760 CENTRALLY IN LOS ANGELES
DEC, MODEL: 20"S AND VAXS
CDC * WE HOLD TAPES AND USE AS NEEDED
DEC, MODEL: VAX
IBM, MODEL: 3081 * DEC, MODEL: 1020
CDC, MODEL: 760, 730
IBM, MODEL: 370 SYSTEM, 3081
DEC, MODEL: VAX & PDP 11/44 * CDC, MODEL: 170/855 (MOST USE) * PRIME
IBM, MODEL: 30335
IBM COMPATIBLE, MODEL: NAS 8040
DEC, MODEL: VAX 11/780 * CDC, MODEL: CYBER 170/730
DEC, MODEL: VAX 750 AND TWO 730"S
IBM, MODEL: 3081
IBM COMPATIBLE, MODEL: NAS/9000-2 * DEC, MODEL: VAX 730, VAX 780
DEC, MODEL: 10
PRIME, MODEL: 750
CDC, MODEL: 172, 174 (SOON, 750)
IBM, MODEL: 3033
IBM (AT YALE)
OTHER: AMDAHL 5860
IBM, MODEL: 3033, AND 4341
IBM
IBM, MODEL: 3033 * PRIME, MODEL: 750
IBM, MODEL: 4341 * DEC, MODEL: 1051
DEC, MODEL: 1099
IBM, MODEL: 360
UNIVAC
BOOKS

POLITICAL TERRORISM:
A RESEARCH GUIDE TO CONCEPTS, THEORIES, DATA BASES AND LITERATURE

by ALEX P. SCHMID, Centre for the Study of Social Conflicts, State University of Leiden, Leiden, The Netherlands

1984 XIV + 586 PAGES
PRICE: US $40.50 (IN USA & CANADA)
DFL. 95.00 (REST OF WORLD)
ISBN 0-444-85602-1 PAPERBACK
APPROX. PUBLICATION DATE: 12/84

This extensive handbook surveys contemporary social science thinking on political terrorism. As a reference work it provides the reader with the largest bibliography on the subject—a computer-based, partly annotated, 4000+ item, multi-disciplinary, multi-lingual literature survey covering aspects of theory and practice. Apart from regional and country entries, it carries subdivisions on such varied aspects as nuclear terrorism, hostage saving measures, state terrorism, etc. Divided into 21 major and 46 minor categories, this author-index bibliography covers legal, psychological, sociological, military and ideological aspects of political terrorism.

The handbook also includes a 130-page "World Directory of 'Terrorist' Organizations and Other Groups, Movements, and Parties involved in Political Violence as Initiators or Targets of Armed Violence," which has been compiled by A.J. Jongman.

An 80-page survey of current thinking on the origins of terrorist violence in various contexts provides insight into the sociological, psychological, conspiratorial theories on the subject, and attention is given also to the theories of regime terrorism and those of the terrorists themselves.

A further section in the volume discusses and evaluates available data bases for the study of terrorism such as those of the CIA, RAND, etc. The accessibility and the reliability of data on terrorism are discussed and data requirements for social science research are indicated.

Finally, the present state of the literature on terrorism is discussed, and research desiderata and strategies pointed out.


For information or to order, write:
Faxon Europe
P.O. Box 197
1000 AD Amsterdam
The Netherlands