BIBLIOGRAPHIC REFERENCES FOR COMPUTER FILES IN THE SOCIAL SCIENCES: A DISCUSSION PAPER

by Sue A. Dodd[1]

BACKGROUND

A recent discussion among the participants of the E-Mail "Informal List for Official Representatives of ICPSR" centered around citing computer files in references, footnotes, and bibliographies; whether to cite a codebook or file (providing you have both); and a discussion on citing primary or secondary sources. With respect to the last two concerns, there appeared to be adequate response indicating that it is better to cite the file as opposed to the codebook, and that one generally cites primary data sources. However, the first concern required more information and the ICPSR OR meeting was targeted as the next opportunity for such a discussion. Note: this paper was first presented at the ICPSR OR Annual Meeting in November 1989, but has been revised for the May-June 1990 IASSIST meeting in Poughkeepsie, N.Y.

IDENTIFYING THE PROBLEM

There is good and bad news. The good news is that researchers are beginning to cite computer files in the reference sections of social science journals. The bad news is that for every person who does cite his or her data source, another twenty to thirty continue to provide no citations. This means that valuable data sources will not be indexed by bibliographical services such a Social Science Citation Index; and more importantly, the next researcher who would like to analyze these data may not have sufficient information to acquire them. Despite efforts to provide researchers with examples and information on how to compile a data file citation, the overwhelming majority continue to describe data sources within the text of their articles, but do NOT follow-up with a citation in the reference section.

The March 1981 issue of Social Forces was the first time that a major social science journal had provided instructions (in the "authors' guide" section) on how to cite a machine-readable data file (MRDF) — currently referred to as a "computer file."

To see if this effort had any impact on the number of computer file citations that could be visibly detected in the reference sections of Social Forces, I took a two year eye-readable sample for 1988 and 1989. Out of approximately 90 articles describing some form of secondary analysis, there were only 12 computer file citations. One of the citations had its own unique style (see below), but

it nevertheless included enough information to gain access to the data. The point being that it is better to err in-the-effort than not to give any information.

> Inter-university Consortium for Political and Social Research. 1979. ICPSR study 7708 Data Description. Police Departments, Arrests and Crime in the U.S., 1860-1920, Principal Investigator: Erick Monkkonen. Ann Arbor.

There were several citations for codebooks, which I assumed to be an indirect way to cite the actual data; and with one or two exceptions, most references to census information came in the form of the GPO printed documents. What this means is that we must renew our eforts to educate and encourage researchers to cite actual data sources. We must also encourage editors and review boards within the various social science disciplines to do likewise. How can we do this and what role can IASSIST members play? Here are some suggestions:

— IASSIST should undertake the task of publishing a small pamphlet or work that would provide sufficient instructions and examples of bibliographic citations for computers files. Such a publication would offer a researcher the luxury of a personal/desk reference sourceeasily retrieved when needed. Apparently, the information provided as part of the data acknowledgement form and that given in certain codebooks is not getting the proper attention. This work should reflect the editorial styles of different social science journals including the American Journal of Sociology, the American Sociological Review, Social Science Information, Government Publications Review, Demography, and the American Political Science Review.

— IASSIST might also want to sponsor an announcement reminding researchers to cite their data sources. This public announcement might read: DON'T FORGET TO CITE YOUR COMPUTER FILES ... It should be sent to the various social science journals, and space permitting, it is likely that they would run it. Another announcement might be jointly sponsored by several editorial boards e.g., THE FOLLOWING EDITITORIAL BOARDS ENDORSE THE PRACTICE OF CITING SOCIAL SCIENCE DATA SOURCES ... The various Associations' Newsletters might also be a vehicle for this type of

announcement.

— Individual IASSIST members should contact editors and discuss the importance of citing computer data sources in references. Point out that researchers are OBLIGED to cite machine-readable sources as well as the printed ones. In addition, this practice should be encouraged so that no data source is described within the text of the article without it also appearing in the reference section.

— Individual members should contact review boards and authors that prepare or oversee "style manuals" — including the Chicago A Manual of Style and Kate L. Turabian's A Manual for Writers. Note: As I was preparing this paper, I discovered a new manual put out by the American Political Science Review entitled Style Manual for Political Science.

— Individual members should assist ICPSR in their efforts to provide quality control over bibliographic descriptions of data files and accompanying documentation. Better control over bibliographic elements facilitates the citation process. The ICPSR staff is making valiant efforts in this regard, but need more guidance and feedback. For example, when IASSIST members discover any discrepancies between the bibliographic elements on a title page and those presented in a citation on the verso of the title page, then this should be pointed out so that it can be corrected.

— IASSIST members should get more involved with the national and international groups dealing with standards associated with computer publishing, production and access. Social science data producers are in the minority compared with computer software producers. Without more active involvement and visibility, decisions are made that exclude the needs of social science data users. For example, there is a National Information Standards Organization (Z39) committee — known as the NISO Committee FF: Computer Software Description — that includes information on bibliographic citationsfor computer software. However, there is no similar effort for text or data files.

It is not possible in this discussion paper to provide anything but brief examples, but more detailed instructions on the components of a bibliographic citation are provided in the IASIS article (Dodd, 1979) and in part three, chapter 9 of Cataloging Machine-Readable Data Files (Dodd, 1982).

Using different editorial styles, the following examples of bibliographic citations are given below. In most cases, the computer file in question is considered a "published work" or book equivalent — even though some computer files in the social sciences do not have "sewn or glued bindings" nor are they always boxed and sealed in packages. The composition of the citations are according to the style of the respective publication and are adhered to by the author with the exception of the bracketed information designating the computer-readable format.

SOCIAL FORCES, AMERICAN POLITICAL SCIENCE ASSOCIATION, and DEMOGRAPHY

U.S. Bureau of the Census. 1989. American Housing Survey, 1985: National Core File [computer file]. ICPSR ed. Ann Arbor: Inter-university Consortium for Political and Social Research.

U.S. Bureau of the Census. 1979. 1979 Census of Popululation and Housing, Fourth Count Population Summary Tape [computer file]. Washington: U.S. Bureau of the Census [producer] Arlington, Va.: DUALabs [distributor].

AMERICAN SOCIOLOGICAL REVIEW

Gentemann, Karen M.
1978   Survey of North Carolina Women [computer file]. Chapel Hill: Institute for Research in Social Science, University of North Carolina.

Verba, Sidney and Norman H. Nie. Political
1967   Participation in America [computer file]. ICPSR ed. Ann Arbor: Inter-university Consortium for Political and Social Research.

University of North Carolina. North Carolina
1969   Information System [computer file]. Chapel Hill: Institute for Research in Social Science [producer].

AMERICAN ECONOMIC REVIEW

Elkins, David J., Blake, Donald E. and Johnston, Richard, British Columbia Election Study: 1979-1980 [computer file], Vancouver, B.C.: Department of Political Science, University of British Columbia, 1980.

SOCIAL SCIENCE INFORMATION STUDIES and POLITICAL PSYCHOLOGY

Davis, J.A., Smith, T.W. and Stephenson, C.B. (1978). General social surveys 1972-1978: Cumulative data [computer file]. Chicago: National Opinion Research Center

Louis Harris and Associates (1973). Harris 1973 Confidence in Government, Study no. 2343 [computer file]. New York: L. Harris and

Associates (producer); Chapel Hill, N.C.: L. Harris Data Center (distributor).

ELECTRONIC JOURNAL MATERIAL

A new phenomenon brought about by computer technology is the so called "electronic journal." Computers have changed the way that scholarly articles are created. For example, articles may be prepared using computers and word processing programs, then sent to colleagues for review via EMail and computer networks, and later returned to the author. Once completed, they are submitted to a discipline-related electronic journal and/or computer list-server for storage and access on demand.

An example of an electronic journal is the Public Access Computer System Review (PACS Review). This journal focuses on "public access" computer systems that libraries make available for patron use. Articles are stored as files on the PACS Forum list-server. The table of contents section of the Review is sent to all PACS Forum users, who can retrieve articles of interest from the list-server by following the instructions contained in that section. PACS Review is published three times a year, has an editorial board and is copyrighted. It also features special departments and reviews of others works. The first volume and issue appeared in January 1990.

Because standards and past traditions fall behind technology and the capability to produce computer-generated works, there are no definitive "guidelines" for citing an article that appears in an electronic journal. However, common sense plus building on what is currently in place, makes the leap from print to computer-readable amanageable feat. The following examples reflect articles that have appeared in PACS Review.

Morgan, James Jay. 1990. "Expansion and Testing of a Meridian CD-ROM Network" [computer file]. Houston, Texas: Public-Access Computer System Review. Electronic journal. 1(1) 34-42. (Access via EMail "GET MORGAN PRVIN1"

LISTSERV@UHUPVM1 or LIB3@UHUPVM1.BITNET]

Stigleman, Sue. 1990. "Text Management Software" [computer file]. Houston, Texas: Public-Access Computer System Review. Electronic journal. 1(1) 5-22. (Access via EMail "GET STIGLEMA PRV1N1" LISTSERV@UHUPVM1 or LIB3@UHUPVM1.BITNET).

ON-LINE DATABASES TREATED AS SERIALS

Many computer works take the form of true serials or ongoing databases (sometimes called "dynamic databases"). They can be cited in bibliographies just as other types of computer files. The only difference is that they have a beginning date and ending date — provided the serial is complete. For those that are continuing, then only the beginning date is given, followed by a hyphen and blank spaces.

University of North Carolina. 1989- IRSS Catalog of Data Holdings [computer file]. 3rd ed. Chapel Hill, N.C.: Institute for Research in Social Scienc On-line database. (uirdss@uncvml.bitnet).

_____. 1989- Questions from the Louis Harris Surveys 1958 to the Present [computer file]. Chapel Hill, N.C.: Institute for Research in Social Science. On-line database. (uirdss@uncvml.bitnet).

_____, 1989- Questions from the USA TODAY Polls 1983 to the Present [computer file]. Chapel Hill, N.C.: Institute for Research in Social Science. on-line database. (uirdss@uncvml.bitnet).

EMAIL AND COMPUTER RELATED ITEMS

Use of electronic mail and networks among social scientist has grown rapidly in recent years, but most are only using a fraction of the power and resources available world-wide. Networking in the future will be the way to access and disseminate data resources — especially if the cost remains so low. Tapping into these data resources and alerting others to their availability becomes the responsibility of all the EMail and network users. Just as with unpublished manuscripts, thesis, dissertations, or letters, there are ways to give credit to authors and provide sufficient information for subsequent access.

Works created using the computer and later circulated via EMail and networks would most likely fall into the category of unpublished material and more specifically "typescripts." In fact, to coin a new phrase, they would more aptly be

called "computerscripts."

When citing unpublished or forthcoming computer works such as an EMail letter, thesis, computer-readable article, etc., be descriptive about the nature of the item and include as much information as is reasonable.

Jones, Paul. 1989. "What is the Internet?" Academic Computing Services, University of North Carolina at Chapel Hill. EMail. (pjones@samba.acs.unc.edu).

Holland, Alecia, ed. 1989. Institute for Research in Social Science Newsletter (Jan.). University of North Carolina at Chapel Hill. Computer-readable mimeo.

Cox-Byrne, Sarah. EMail letter to Laura Ann Guy dated October 5, 1989. EMail correspondence.

Petterson, Lynne M. 1990."The Impact of International Competition, Technological Adoption and Industrial Restructuring on the Evolving Geographic Distribution of the Contemporary American Machine Tool Industry. "[computer file]. Ph.D. diss. University of North Carolina at Chapel Hill.

Updegrove, Daniel A., John A. Muffo, John A. Dunn, Jr. "Electronic Mail and Networks: New Tools for Institutional Research and University Planning." [computer file]. AIR Professional File. Forthcoming.

REFERENCES

Dodd, Sue A.
1979   "Bibliographic References for Numeric Social Science Data Files: Suggested Guidelines." Journal of the American Society for Information Science, 30:77-82.

Dodd, Sue A.
1982   Cataloging Machine-Readable Data Files: An Interpretive Manual. Chicago: American Library Association. Pp.169-172

[1]Presented at the IASSIST 90 Conference held in Poughkeepsie, N.Y. May 30 - June 2, 1990.

Institute for Research in Social ScienceUniversity of North Carolina
Chapel Hill, N.C. 27599
usdodd@uncvml.bitnet
Rev. May 1990