DOCUMENT RESUME

ED 057 856   LI 003 369

AUTHOR   Dennis, Jack
TITLE   The Relation of Social Science Data Archives to
                Libraries and Wider Information Networks.
SPONS AGENCY   American Library Association, Chicago, Ill.; Office
                of Education (DHEW), Washington, D.C.
PUB DATE   70
NOTE   17p.; (12 References); Working Group B-4
AVAILABLE FROM In "Proceedings of the Conference on Interlibrary
                Communications and Information Networks," edited by
                Joseph Becker. American Library Association, 50 E.
                Huron St., Chicago, Ill. 60611 ($15.00)

EDRS PRICE   MF-$0.65 HC-$3.29
DESCRIPTORS   Archives; Conferences; *Data Bases; *Information
                Networks; *Library Cooperation; *Library Networks;
                Organization; *Social Sciences
IDENTIFIERS   *Interlibrary Communications

ABSTRACT
        Some of the organizational requirements for the
establishment of a network of interlinked information services are
discussed in this paper. In particular, the analysis concerns the
relation of social science data banks to libraries and related
organizations. These problems are approached from the standpoint of a
director of a social science data archive at a university and as one
who has participated in the activities of the national coordinating
organization of social data archives. The paper is divided into three
parts: (1) a short description of organizational developments in the
archiving of quantitative social science information, (2) a
discussion of how local archives could be most efficiently organized
and (3) the relation of social data archives to libraries and thus to
any wider network of information services. (Other papers from this
conference are available as LI 003360 - 003368 and LI 003370 through
LI 003390) (Author/NH)

The Relation of Social Science Data Archives
to Libraries and Wider Information
Networks

Jack Dennis

University of Wisconsin

This paper concerns some of the organizational requirements
for the establishment of a network of interlinked information
services. In particular, the analysis concerns the relation of
social science data banks to libraries and related organizations.
I approach these problems from the standpoint of a director of a
social science data archive at a university and as one who has par-
ticipated in the activities of the national coordinating organiza-
tion of social data archives. The presentation will be brief and
hopefully, to the point. It is divided into three parts: (1) a
short description of organizational developments in the archiving
of quantitative social science information, (2) a discussion of
how local archives could be most efficiently organized, and (3)
the relation of social data archives to libraries and thus to any
wider network of information services.

The Growth of Social Data Archives

The past decade in the social sciences has witnessed an
exponential growth in the number of active researchers, projects
and quantitative data produced. The new store of social data thus
provided has expanded our knowledge of society greatly; but it has
also brought into being new demands for increased capabilities on
the part of institutions that formerly were equipped only for the
more traditional forms of scholarly activity. A number of social
science data archives have been established in response to these
demands, so that social information of a machine-readable type
can be preserved for use by the wider social science community.[1]

The founders of these new archives have tried to devise appropriate means for coping with the social information explosion, from the standpoint of maintaining, retrieving and redistributing valuable project data for further analysis.

As a corollary feature of the creation of these social data archives, an embryonic national and world-wide network of social science information has also begun to appear. Not only such significant service organizations as the Roper Public Opinion Research Center and the Inter-University Consortium for Political Research have been established, but also coordinating organizations have appeared -- such as the International Social Science Council's Standing Committee on Data Archives and the Council of Social Science Data Archives. The Council of Social Science Data Archives, for example, has attempted to coordinate activities within the emerging American archive network, particularly in areas of standardization of policies relating to data format and documentation, as well as policies regulating the costs of data distribution and technical developments. Together, the Standing Committee on Data Archives, CSSDA, and the few major archives such as Roper and ICPR, have formed the nucleus for whatever future progress is to be made in developing a broader system of interlinked social science data banks.

What we shall argue, however, is that important as these steps have been, the system as a whole is at a crucial turning point in its development -- one at which the major attention now needs to be turned to the problem of developing more fully the locally-oriented, general purpose archives. Indeed, only by a greater effort to establish adequate local facilities will the nascent

national and international systems reach maturity; and without a
more pervasive base of local archives, much of the progress made
to this point in developing the broader archival networks could
well be lost.

In the discussion that follows, we will attempt to spell out
this argument in more detail -- including what we mean more precisely
by a locally-oriented, general purpose archive. We will draw espe-
cially upon the experience gained in establishing the Social Science
Data and Program Library at the University of Wisconsin for our re-
commendations. One of the most important aspects of the latter ex-
perience has been the set of relationships that our organization has
had with the emerging national and international archive systems.
Let us first focus upon the nature of these archive systems there-
fore, and subsequently upon the principles that we would recommend
for application in organizing local archives with respect to the
broader networks.

An International System of Social Data Archives?

The founders and early participants both in the International
Social Science Council's Standing Committee and in the Council of
Social Science Data Archives no doubt envisioned a future world
in which there would be an archive network operating at several
levels -- international, national, regional, and local. This net-
work would include both the present, national data banks such as
ICPR, Roper, Steinmetz, Essex, and Zentralarchiv and the more limited
organizations such as those at Oslo, Pittsburgh, North Carolina,
Iowa, Wisconsin, Northwestern, and Yale. When fully established --

this system would have some continuing headquarters and staff, and would help to coordinate policies pertaining to common problems and objectives. It would also serve perhaps in future as the central point in a telecommunications network in which a data user at one archive could gain access to the holdings of other arbhives directly by means of a general, machine-readable index of available data. The user would also be able to analyse the data directly by means of an interactive computing capability.

In the decade of the 1960's some limited progress toward this vision of the future was indeed made. As those who have participated in C.S.S.D.A. activities know, however, many unsolved practical problems of this ideal system have been carried over into the present period. The maintenance of even a small central staff for the C.S.S.D.A. has proven difficult, given the vagaries of available. funding. Nor have even the simpler aspects of inter-archival coordination been more than minimally achieved. For example, the inventory of data holdings is not fully operational; nor is it generally available to member organizations. Provision of common data formatting standards and the keeping of an up-to-date inventory of holdings also remain basic, but unaccomplished tasks at present.

In part, this failure to develop fully has been caused by difficulties of coordination growing out of disparate experience, organization, subject matter specialization and hardware facilities among the various archives. But it is also due, we would guess, to the fact that the various archives have not developed at the same rate. In particular, while the "wholesaler" archives such as ICPR and the Roper Center have experienced marked development in the

past decade, their "retailer" counterparts -- the local archives -- have failed to grow apace. We would strongly emphasize that without substantial growth of these local facilities to make proper use of the resources available from the national archives, the social science information system will probably remain weak and inchoate in the decade of the '70's; and the progress made by the more general archives may very well be eroded by rising costs of ICPR or Roper Center services coupled with inadequate use of data at the local university or other local level.

There are, however, some exceptions to this general pattern, so far as local archives are concerned. The university social data archives at Pittsburgh, Northwestern, and Wisconsin, for example have attempted in recent years to provide facilities which embody, in our view, some essential principles of organization for the coming decade. Let us turn then to what these principles might be.

Organizing Local Archives

How should local archives be organized in order to assure that demands for services by users can be met and also that these archives can more effectively become the foundation stones for a general network of quantitative social science information? Our short answer to this question is that new local archives should self-consciously define their functions along the lines provided by the usual university library. What we are saying is that data archives are most usefully thought of as extensions of the traditional information services of a library -- a library which serves both faculty and students for purposes of research and of instruction.

One of the most important features of this organization-

al perspective is that, unlike the existing "wholesaler" archives, local archives should be fairly general in the scope of their subject matter. The Inter-University Consortium for Political Research specializes in political data, the Roper Center deals with surveys, the Zentralarchiv concerns itself essentially with German-related studies, the Essex Social Science Research Council Data Bank focuses upon British materials, etc. All of these "general" archives specialize more than we would suggest should be the case at the local archive level. Rather, the local data bank should include political, historical, sociological, psychological, anthropological, and economic data, at the very least; and these data should be both individual and aggregate in character -- that is, they should be based upon a variety of data-collection techniques. One might argue further that not simply machine-readable <u>quantitative</u> data be included but -- as at Northwestern University's Inter-societal Information Center -- machine-readable <u>qualitative</u> information ought to be built into the archival holdings and services from the beginning as well. Thus, we would regard the form of local archival organization as being preferably one that is unlike any of the national archives, but rather it would be complementary to all of these higher level archives together -- in the sense that the local archive is designed to receive data from many different sources. Thus it is able to service many kinds of users.

At Wisconsin, where the data archive has been organized along these lines since 1966, members of over thirty different departments of the university have used our services; and these subject matter specialities range from political science, sociology and economics --

the major users -- to mass communications, social work, home eco-
nomics, demography, urban and regional planning, and physical edu-
cation. In this sense, the Data and Program Library Service at
Wisconsin acts as a general acquisition, storage, and distribution
facility, just as does the university library. As a result of this
broad subject matter emphasis, usage of the data library grew, for
example, from 355 visits in academic year 1966-67 to over
1300 requests in each of the academic years 1967-68, and 1968-69.
Unless there is this broadening of scope of subject matter from earl;
in the life of the archive a large and diverse group of potential
users will necessarily be excluded.

At the beginning of course, there are likely to be only a
few core disciplines represented -- that is, the disciplines in
which quantitative social data needs are already paramount and for
which the national data archives - such as ICPR, the Roper Center
or the Bureau of Labor Statistics already provide ready sources of
data supply. But, in the second period of development, we would
suggest that these services be extended as we have extended ours
at Wisconsin.

Now a second major feature of this form of organization --
modelled on the idea of a central, library-like information center
at a university -- is its considerable emphasis upon servicing users.
Were archives at the local levels conceived only as repositories
of machine-readable data, then such emphatic user-orientation might
not be necessary. But we conceive the main point of a local archive
not to rest upon its "archival" functions per se, that is, in ac-
quiring, indexing and preserving these data in some antiquarian

8

spirit for posterity. Rather, we see the local archive as the primary means for keeping these data alive -- to put them continually, without delay, and at minimal cost into the hands of potential users beyond their originators. By making data easily available, new tests of hypotheses can be made; and the considerable effort and expense that goes into producing reliable social information is matched by the increased marginal benefits of use in the wider academic, industrial or governmental communities.

To be user-oriented means also that various other kinds of services will need to be created -- in order that data-requestors, however technically unsophisticated, can make proper use of the evidence provided. In our archive, we have found it necessary to provide not only the usual library functions of acquisition, cataloging, indexing, referencing, storing, maintaining, reproducing, and distributing information. We have also discovered that we must edit ("clean") the data, redocument them in some instances (by updating and editing codebooks), provide computer programs for use in file handling and statistical analysis, give consulting on the proper use of the programs, and furnish more general education both to students and to faculty about the existence and proper use of these resources in their scholarly activities. In addition, we maintain a variety of hardware facilities -- for example, a remote terminal for the university's central computer, and an array of unit record equipment. We thus go very far in making a full range of services available in a single location, so that the user has a convenient mode of utilizing the data resources of the archive. Each of these functions is important, moreover, given the fact that they are incompletely available elsewhere in the university. Indeed, our program

library, for example, is so valuable a part of our services that it has become a coordinate activity with the data library. It provides, furthermore, the core of a current attempt to extend progra library abstracting and indexing services to a range of users beyond Wisconsin.[2]

We thus suggest that a local archive be general in types of information and user-oriented in the sense of providing a range of auxiliary services necessary for making efficient use of archival data files. Together, these two features provide important means of attracting usage and thus a means for improving the quality of social research. Scholarship is advanced when local users are able to draw upon data not otherwise efficiently available and when there is an improvement in the cost performance of information-generating activities for the intellectual community as a whole.

Another important and perhaps more basic product of increasing data usage is that a popular base is created for improving the services already established. Once university or foundation administrators begin to see that the archive is a going concern in the sense that many and various types of people are using the facilities not simply that the archive has acquired a good collection of data files or that it has relatively efficient modes of indexing and processing the data -- then and only then are they likely to upgrade the facility significantly and put its financing on a more permanent basis. Administrators, after all, are a pragmatic lot; and pragmatists typically want to see what the organization is doing: what services it is in fact providing for users, rather than simply how

comprehensively, or efficiently, organized it is.

As potential data users become acquainted with these services through actual experience, moreover, they are also likely to begin to understand the necessity of sharing their own project data with the wider community -- once their own project needs have been met. Short of requiring as a condition of foundation support that researchers deposit their project data in archives, this is perhaps the most effective way of informing members of the social science community of their wider responsibilities. They are also able to see, by this method, the necessity for proper data documentation, cleaning, formatting and maintenance. Once their own data have been released to an archive, their errors become a matter of visible public record! In this sense, encouraging contributions to archives is a way of maintaining standards of social science evidence, at least at the level of data preparation and documentation.

Perhaps the most important product of all of building local usage is acquaintance that clients of these archives gain with the wider social information system. The local archive serves as an essential conduit for building recognition and support for the national archives, for CSSDA and for other aspects of the emerging international social science information system. As social scientists -- particularly those among the current student generation -- find these services available and valuable at their present institutions, when they move on to other university settings they are likely to want to build or improve similar resources at these locations. Thus, they will add to the total potential base for establishing the

world-wide information system. The local archive -- even if it
currently exists, for example, only as a means of acquiring and
storing political surveys from ICPR -- is likely to serve as a
stimulus for the development of expanded services of this type --
in a sense similar to what has been true in the initial period of
development in which the existence of ICPR, the Roper Center, and
other major collections have served to stimulate the creation of
whatever relatively limited local archives have sprung up in various
universities during the past decade.

Further Developments: Articulation of Social Data Archive and Library

If, during the coming decade, there is indeed the kind of
full-scale effort to develop local social science data archives
along the lines that we suggest, what then should be the next
steps? Two developments are partially explicit in our discussion
thus far. One is the need for more effective growth of inter-
archival coordination at regional, national and international
levels. The other is the necessity for closer integration of the
local archive into existing local university information services --
particularly those typically provided by the traditional university
library. In that we have already said a good deal about the possi-
bilities for the growth of the national and world archival networks
based upon a more comprehensive system of local archives, let us
turn our main attention at this point to the relation of the social
science data archive to the university library.

When we suggested that data archives be organized or re-
organized to represent the informational interests of a wide variety
of students of society and that these services be user-oriented --

especially in that they be provided at minimal cost to the user --
we were saying that the functions and policies of the data archive
should be brought into line with the usual operating doctrine of a
university library. In an important sense, a library is a university-
wide facility servicing all comers on the basis of minimal direct
cost to the user. Furthermore, we see the data archive functions of
acquiring, documenting, indexing, maintaining, reproducing and dis-
tributing quantitative information as precisely those of the tradi-
tional library. Thus, the extension of these library functions to
the handling of primary quantitative data is but a very short step.
Libraries already go quite beyond simply collecting and making avail-
able information that exists in printed form. They store and dis-
tribute microfilm and microfishe, for example, and even maps, paint-
ings, and phonograph records. "Libraries," as Ralph Bisco perhaps
overoptimistically observed, "have been exceptionally responsive to
the information needs of their users."[3] While one can argue that
data archives have almost in every case originated with social
scientist users or data-generating agencies, rather than with uni-
versity libraries, this does not mean that librarians are necessar-
ily unresponsive to these new information needs.

Our own experience -- and this has been matched in other
instances that we know of, e.g., at U.C.L.A. and Northwestern Uni-
versity -- is that the members of the staff of the library are
keenly aware of their potential responsibilities in cooperating with
social scientists to expand the scope of these services and in
coordinating such services with standard bibliographic and other
services of the library. Our suggestion would be therefore that

organizers of archives bring the library into organizational planning as soon as possible, in much the same spirit that the computing center would be a point of natural coordination from the beginning of archive operations.

Furthermore, after an initial period of cooperation, we would think it feasible to begin building the archive into the library system itself. Obviously, this further development depends upon how interested the library staff is in moving with the changing informational needs of its social science users. Certainly not every university library will be able, by reasons of shortage of competent staff, appropriate facilities or financial resources, to enter immediately into this new area. But many libraries will be willing to do so; and in our opinion, this movement should be encouraged by those concerned with archival development. As more and more aspects of library operations become computerized in other respects library administrators will undoubtedly begin to see the necessity for advancing the credentials in computing of their staff and in acquiring appropriate hardware. Once these developments have been made then it is mainly a matter of organizational transfer to incorporate social data archives into libraries. The library ought to be given every opportunity to maintain its traditional role as the center of university informational services. This means that library administrators should be encouraged to play a leading role in developing the local archive and the national and international social science information systems.

On the other hand, we do not mean to suggest that social scientists and other organizers of data archives should drop from

14

the scene as the archive becomes more integrated with the library.
Indeed, to do so would destroy one of the great virtues of present
archival activity, namely, the felicitous involvement of social
scientists in providing themselves with adequate information
storage and retrieval facilities. One needs to work out a new set
of relationships in which the archive obtains both library inte-
gration and support, but also continues to have active inputs of
administrative energy, fund-raising and policy participation by
social scientists.

The experience we have had in recent years at Wisconsin suggests,
therefore, that there is a considerable potential market for the
kinds of data repository services that we have begun to provide.
Ready access to machine-readable data collected from a variety of
sources improves the quality of research and teaching, and extends
the informational services traditionally provided by university li-
braries. We would urge that these local archives be more extensively
organized on a wide subject matter, user-oriented basis, to provide
the natural complement of the already established wholesaler archives;
and to provide a firmer base of organization for social data archives
nationally and internationally. Our own experience provides a strong
impetus toward a library-like facility with a variety of auxiliary
services such as a remote access computer terminal, a computer pro-
gram library and computational-statistical consulting.

In the future, one needs to expand these services in volume
and scope at a local level by building in most cases upon the or-
ganizational base of the traditional university library, thus moving
the library into new but highly relevant areas of informational ser-
vice. In this way in the decade of the 1970's, the benefits of a more

comprehensive social information system will be enhanced and the aggregate costs to the community of scholars of social science data collection will be greatly lowered.

I would also suggest that until these organizational developments have been made, more advanced techniques such as interarchival telecommunications links, while technically feasible, will lack the other resources necessary for their maintenance. It would be most uneconomical to install telecommunications equipment in libraries where there is no organization to provide proper data storage and retrieval for local use, where there is also therefore few users who have experience and understanding of archival information usage, and where one lacks an administrative staff able to understand the hardware and software requirements of users such as social scientists and   have an appreciation of the limitations and possibilities of an inter-archival network. My argument therefore is that we need to expend our greatest immediate effort in expanding the archival capabilities of local institutions and to forge a set of archive-library linkages which can provide the anchor points for interarchival networks, both technical and organizational.

Author's note:

I am indebted to Margaret Adams who assisted me in an earlier version of this paper, published in <u>Social Science Information</u> and to Robert Alford, Michael Aiken and <u>David Elesh of the Wisconsin</u> Sociology Department faculty -- all of whom have been former directors of the Social Science Data and Program Library Service -- and to Alice Robbin, present Data Librarian, Richard Lochrie, Program Librarian, and James Taylor, Project Administrator of the recently formed National Program Library Service.

[1]For discussion of this emerging network of data archives see, for example: R.L. Bisco, "Social Science Data Archives: Progress and Prospects", <u>Social Science Information</u> 6 (1), 1967, pp. 39-74; S. Rokkan (ed.), <u>Data Archives for the Social Sciences</u>, Paris, Mouton, 1966; D.K. Stewart, "Social Implications of Social Science Data Archives", Technical Memorandum 3729/000/00, Santa Monica, Calif. System Development Corporation, 7 November 1967; W.A. Glaser and R.L. Bisco, "Plans of the Council of Social Science Data Archives", <u>Social Science Information</u> 5 (4), 1966, pp. 71-96; D. Nasatir, "Social science data libraries", Publication A-89, Berkeley, Calif., Survey Research Center, University of California, November 1967; and R.L. Bisco, "Social science data archives: Tehnical Considerations", <u>Social Science Information</u> 4 (3), 1965, pp. 129-150.

[2]<u>Social Science Information</u> has published descriptions of some of these organizations. See R.L. Bisco, "Social science data archives ...", <u>op. cit.</u>; R.L. Merritt and R.E. Lane, "The Training Functions of a Data Library", <u>ibid.</u>4 (3), 1965, pp. 118-126 mainly concerned with the Political Science Research Library of Yale University; and S. Rokkan and F. Aarebrot, "The Norwegian Archive of Historical Ecological Data: Progress Report. August 1969", <u>ibid.</u> 8 (1), 1969, pp. 77 - 84.

[3]Under a recent grant from the National Science Foundation, the Social Science Data and Program Library Service at the University of Wisconsin has begun a pilot project on establishing a National Program Library Service - beginning in the initial two-year period to compile a machine readable inventory of available computer programs of interest to social scientists. The goals of this project will be outlined in a subsequent report in <u>Social Science Information.</u>