THE ORGANIZATION AND MANAGEMENT OF THE NORWEGIAN SOCIAL SCIENCE DATA SERVICES (NSD)  73

Jarle Brosveet, Bjørn Henrichsen, Lars W. Holm, Terje Sande

ON-LINE OR ON TAPE  79

Judith S. Rowe

NEWS AND NOTES

Inventory of Data Sets on Non-Western Nations  85

New Periodical: Software Review  86

Book Notices  87
The IASSIST Newsletter represents an international cooperative effort on the part of individuals managing, operating, or using machine-readable data archives, data libraries, and data services. The Newsletter reports on activities related to the production, acquisition, preservation, processing, distribution, and use of machine-readable data carried out by its members and others in the international social science community. Your contributions and suggestions for topics of interest are welcomed. The views set forth by authors of articles contained in this publication are not necessarily those of IASSIST.

Information for Authors

The Newsletter is published four times yearly. Articles and other information should be typewritten and double-spaced. Each page of the manuscript should be numbered. The first page should contain the article title, author's name, affiliation, address to which correspondence may be sent, and telephone number. Footnotes and bibliographic citations should be consistent in style, preferably following a standard authority such as the University of Chicago Press Manual of Style or Kate L. Turabian's Manual for Writers. If the contribution is an announcement of a conference, training session, or the like, the text should include a mailing address and a telephone number for the director of the event or for the organization sponsoring the event. Book notices and reviews should not exceed two double-spaced pages. Manuscripts should be sent in duplicate to the Editor:

Howard D. White  
School of Library and Information Science  
Drexel University  
Philadelphia PA 19104 USA  
(215) 895-2481 or (215) 895-2475

Book reviews should be submitted in duplicate to the Book Review Editor:

Kathleen M. Heim  
Graduate School of Library Science  
University of Illinois at Urbana-Champaign  
329 Main Library  
Urbana IL 61801 USA

Key Title: Newsletter - International Association for Social Science Information Service and Technology
ISSN - United States: 0145-238X  
Copyright © 1981 by IASSIST. All rights reserved.
THE ORGANIZATION AND MANAGEMENT OF

THE NORWEGIAN SOCIAL SCIENCE DATA SERVICES (NSD)

JARLE BROSVEET, BJØRN HENRICHSEN, LARS W. HOLM, AND TERJE SANDE
NORWEGIAN SOCIAL SCIENCE DATA SERVICES
BERGEN, NORWAY

(This paper was delivered at the 1981 IFDO/IASSIST Conference, Grenoble.)

THE EARLY YEARS

The origins of the Norwegian Social Science Data Services (NSD) go back to 1967 when the Norwegian Research Council could no longer ignore various complaints about long delays and frequent errors in the processing of social science data. As a result the Council decided to set up a committee to study the reasons why bottlenecks occurred in computing services at the universities. The committee was also instructed to make proposals for improvement of social science computer facilities and services.

The committee made a number of recommendations and took the first steps to implement these as part of its work. A series of intensive courses in computing for social scientists was organized, a variety of statistical packages and specialized programs were acquired and installed on the university computers, and steps were taken to build up a first set of data banks easily accessible to researchers at all universities.

In 1970, the committee proposed to institutionalize these activities by establishing a nation-wide organization supplying various types of services to all social scientists. The Research Council approved the proposals and established the NSD for a trial period through 1974. Later a new program of activities was approved for the next three-year period. As the activities proved a success, the NSD was given permanent status from January 1st, 1978.

THE ORGANIZATION

The NSD differs from most other organizations of this type in three ways:

- it is a federally structured organization with offices at four universities and with close working arrangements with the regional colleges;

- it has built up a wide variety of data resources across all fields of the social sciences: the data holdings comprise surveys as well as large databanks for communes and census tracts, an archive of information about organizations, and a series of computerized files of data on the recruitment and careers of various elite groups;

- it has established close contacts with a wide range of users in the social sciences and in administrative bodies, partly through its regular newsletter (Brukermelding), partly through the annual meeting of representatives of a great majority of research institutes and teaching departments active in the social sciences throughout Norway.

The NSD has its headquarters at the University of Bergen. In addition there are external data secretariats at each of the other universities: Oslo, Trondheim and Tromsø.

THE DATA

The NSD has so far built up seven basic types of data holdings:

I  The first and largest is the database for communes. This database covers statistics for all local units of administration since 1800 and is linked up with a facility for computer cartography (POLYVRT, CALFORM, SYMAP, FIGUR). This is our most widely used facility and is constantly expanded and improved. A great amount of work has been expended on finding effective solutions to the problems posed by changes in boundaries and units. The database offers detailed documentation of these changes and provides the user with a variety of coefficients for recalculating data values whenever changes occur.

II The NSD has set up a cartographic service useful to a number of users. Coordinate matrices for all commune boundaries in Norway since 1800 have been computerized. The boundary segments are identified by a time key to allow the production of maps for every year since 1800. Initially, maps were mostly conformant using cross-hatchings to represent the data values. Later programs have been developed to display information at the centerpoint of every commune. This development makes it possible to display not only multivariate statistics, but also discrete data about the presence or absence of particular infrastructure elements such as airports, harbours, hospitals, schools, factories etc.

III To allow analyses at a lower level of aggregation the NSD has taken steps to organize a set of data at the lowest level of official
enumeration: the census tract. This data bank contains data from the censuses of 1960 and 1970 and has proved of great interest to city planners, geographers, anthropologists and sociologists.

IV In dealing with survey data it has been our policy to link up response data for comparable questions over time, so we have deliberately declined to serve as a depository for individual studies. The largest file so far developed covers the surveys of the Norwegian Gallup Institute (roughly monthly ones) from 1964 through 1976. Currently, this archive is being updated with data from the Norsk Opinionsinstitutt A.S. (Norwegian Opinion Institute). A similar linkage is being completed for the electoral surveys carried out by Professor Henry Valen and his colleagues since 1957. Some of the most thorough surveys carried out by the Central Bureau of Statistics since 1967 are now at the disposal of academic users under an agreement with the NSD. The NSD has prepared detailed documentation for some of these surveys and reformatted them for use with SPSS to increase their availability.

V A major effort has been made to develop a systematic databank for elite groups. The first file to be completed contains biographic information on Members of Parliament since 1814. This archive gives information on father's occupation, education, early career, positions in legislative committees etc. The file has recently been linked with one for roll calls since the establishment of party fronts in the 1870's to allow analyses of interrelations between background factors and legislative behaviour. Similar biographical files are being organized for members of the central administration as well as for graduates from the universities since 1811.

VI Another file taken over by the NSD is the archive of information about voluntary associations in Norway 1971/1972.

VII The NSD has also got a file of test data for the recruits to the Armed Forces from 1951 to 1968.

In organizing these as well as other data the NSD has emphasized the paramount importance of documentation and data linkage. The NSD does not wish to serve as a pure depository of data. Consequently, data files are rarely incorporated into our holdings unless they can be fully documented and linked up into a system either topically or across time periods.

To make sure that our facilities are used extensively, the NSD has given high priority to educational activities, courses, lectures and teaching packages. Active support has been offered within the programme of the International Social Science Council as one of the cross-national workbooks in the ISSC series has been produced in close cooperation with the NSD. Steps have also been taken to compile a series of specifically Norwegian teaching packages. One of these is based on fiscal-administrative data for communes, another focuses on time-series data for intermediary levels of regional aggregation (fylker). Further packages are based on data from electoral surveys and the file of biographical information on Members of Parliament.
MATTERS OF POLICY AND ECONOMY

Up until 1977 NSD was almost exclusively financed by grants from the Social Science division of the Research Council. At that time it was realized that NSD holdings were used extensively by groups other than social scientists and for purposes other than research and teaching. Even more important, some governmental agencies as well as the NSD realized that our data archives carried a huge potential for several kinds of governmental planning and research. This potential was likely to grow not least as a result of increased co-operation with the Central Bureau of Statistics. Consequently, our financing had to be altered in one way or another. Several models were discussed, but in the end the NSD is still being financed on a project basis. A major advantage of this arrangement is that it allows the Research Council and the NSD to set the priorities. At the same time we can also choose the governmental projects that appear to give the greatest benefits to the social sciences in general.

From 1978 to 1981 the basic grant to the NSD increased from about 1 mill. to 1.4 mill. Nkr., each year totalling about 8% of the Research Council allocation to social science projects. If this were to be our total budget, it would restrict our expansion considerably. However, we have been able to obtain supplementary funding elsewhere, including grants from other divisions of the Research Council, the Nordic research councils and various governmental agencies.

The current funding of the NSD allows us to keep a reasonably large group of professionals permanently employed. As of today all full-time staff except the secretaries have a university degree. Our recruitment policy has always been to emphasize social science background with experience in data analysis and programming. Such experience is achieved through the training of recruits as assistants either at the NSD or on large research projects at the universities. The academic background of the staff covers almost the full range of the social sciences: economics, sociology, political science, geography, public administration, history and information science. Not only does this provide a good interdisciplinary working group, it also makes it easy to communicate with various research interests. The recruitment policy definitely reflects the tasks that have been assigned priority by the Board of the NSD.

To put it simply, data and documentation are given priority over software development and programming. However, the NSD is seldom involved in primary data collection. In effect, we place great store by establishing good working relations with data collecting agencies in administrative bodies, private firms and research institutes.

Obviously, the most important of these is the Central Bureau of Statistics. We now cooperate rather extensively with the Bureau and also exchange services to some extent. This co-operation means that we are able to provide social scientists with virtually all data collected by the Bureau at a very low cost that often amounts to no more than the copying of the magnetic tape. The NSD has also been assigned the responsibility for giving social scientists access to Bureau data that are not completely anonymized, provided that the Bureau receives feedback as to who is given access. Furthermore, the NSD is represented on committees to review census questionnaires and Bureau tables to be publicized using census data.
To a certain extent we are developing similar relationships with other governmental agencies supporting data collection tasks.

Currently, we are cooperating with the Ministry of Consumer Affairs on the computerization of a register of representatives in official boards and councils. We do part of the coding and punching as well as the running of simple tables to be used in the official report. Later, these data will be established as an archive and included in the data holdings available to researchers free of charge.

There are several reasons why such excellent and confident relationships can be developed vis-a-vis the governmental agencies. One of the most important reasons seems to be that there has never been any abuse of data by students or researchers. The NSD has also shown the ability to store and document data in a way that permits simple computer runs, map drawings etc. much faster, cheaper and often with better quality compared with other data distributors. Since the holdings of the NSD contain data collected by means of public funds, there is also a strong argument for maximum public usage of these data.

Before leaving our relationship with governmental agencies, we must also mention the effect of privacy legislation on social research. As soon as the privacy law was proposed, the NSD took the initiative to help fulfill the intentions of the law as well as to counteract some of the negative effects experienced in other countries. After lengthy negotiations we are now in the process of implementing a strategy that we believe is advantageous for all parties concerned. First, the NSD has been given a general concession by the Data Inspectorate to archive and store relevant social science data. Second, we have obtained a grant for acting as a liaison between the research community and the Data Inspectorate. It also enables us to set up a secretariat for the Research Council on matters concerning privacy and data collection.

INTERNATIONAL COOPERATION

On the international side, the NSD has always been cooperating actively with similar organizations in other countries. It was one of the founder members of the International Federation of Data Organizations in 1977 and is active in a Working Group under its auspices for the coordination of local - regional data bases and computer cartography.

Since 1971, information about European archiving efforts has been published regularly in the European Political Data Newsletter set up by Stein Rokkan under the European Consortium for Political Research and later co-sponsored by the Norwegian Social Science Data Services. The EPD Newsletter has developed a classification scheme for the comparison of the contents of European time-series datasets and has repeatedly advocated actions to co-ordinate developments in this field.

First, NSD and the Norwegian Research Council succeeded in securing the sponsorship of the European Science Foundation for a meeting on "Databases for Regional Analysis" in 1977.
Second, NSD took steps to develop a proposal for the establishment of a Joint Nordic Database for Regional Time Series during 1977-1979. Funding for a three-year period was granted by the four Nordic Social Science Research Councils, i.e. the Danish, Finnish, Norwegian and Swedish. The project started in January 1979.

Also, at the IFDO meeting on regional data held in Turin in March 1980, the task of coordinating an effort to summarize and structure information for a joint European database was entrusted to NSD. NSD was asked to complete the matrix or map of available data on the basis of the work started in Turin, to disseminate a report and organize a workshop in which further action can be discussed.

The effort implies the following:

a. to identify organizations and scholars in each nation who are willing to cooperate, and to look for additional researchers and scholars in cases where regional or thematic gaps in the databases have been identified;

b. to establish a common standard for types of variables in these databases;

c. to establish a common standard for documentation of these data;

d. to carry out an effort as far as possible to establish the comparability of these data. This task does not necessarily imply the establishment of formal statistical comparability, but a more substantive kind of theoretical comparability, such as the possibility of using analogous measures for common research projects of analysis of social structures and processes.

NSD concludes from this information-gathering effort that it will be feasible to start the building up of a first version of a Joint European Data Base for Regional Time Series. At this time we have established a network of contacts and compiled information about data matrices and variable lists. A report on the progress of the project has been submitted to the Board of IFDO for further discussion.
ON-LINE OR ON TAPE

JUDITH S. ROWE
PRINCETON UNIVERSITY COMPUTER CENTER
PRINCETON, NEW JERSEY

(This paper was delivered at the 1981 IFDO/IASSIST Conference, Grenoble.)

Few of us here remember the early days in which data were transported from one installation to another using metal tapes or punched cards. Needless to say, not many data sets were transported. Metal tapes are now historical artifacts, although most of our computers can still read punched cards. But for almost twenty years data archives and data libraries have used magnetic tapes as the medium of choice for data exchange. With the advent of microcomputers some exchange uses floppy disks, but since no standard has been established for these and since their capacity is quite limited, they are not widely used for this purpose.

To a large extent, access to data by local users is also via magnetic tapes, although under certain circumstances data can be made available permanently or temporarily on disks. The decision to provide local access and certainly to provide permanent storage by one device or another is essentially one of cost and is based largely on the local charging algorithms for the storage and use of the various media. The two key variables are the size of the data set and the frequency of its use.

Although two computer centers may both be attempting to recover costs, they may not necessarily be doing so in the same way. Charges are normally set to encourage particular types of user behavior. Most centers, for example, offer cheaper off-hour rates in order to even out the flow of work. However, depending on the size of the machine and/or storage area, and the number and the type of input-output devices, charging may encourage the use of tape storage or disk storage, high density or low density tapes, permanent or mountable disks.

At Princeton it clearly makes sense -- and will increasingly in the future as we eliminate mountable disks -- to store small, frequently accessed data sets on permanently mounted disks and large, infrequently accessed data sets on high density tapes. However, it should be noted that in order to keep the number of tapes in the machine room manageable, there is a small penalty for archival tapes which are seldom or ever mounted. As the computer center's largest tape owner, the data library owns its own tape rack and hence does not pay the "no use tax" except for rented tapes in the public use area.

With the present array of storage options now available to center users, decisions concerning the middle-sized data set may be less clear. Princeton provides charts to help users pick the most economical storage medium for a given data set. These have now been augmented with a small, publicly available program which allows the user to insert data set size and anticipated use variables in order to obtain a storage medium recommendation. The charts, however, are useful if one is not in a gray area.
Until recently our concerns with these issues were purely local ones, but with the increasing amount of data now available from on-line services, a new dimension has been added. It is no longer enough to think in terms of appropriate media for local storage and use. Many of us are already faced -- and all of us will be increasingly in the future -- with new decisions on the form in which data should be acquired. These decisions are not easy ones, and they will not get easier in the future. There is no single answer which applies to every data archive or data library, nor is there a single answer which applies to every data collection.

We cannot say, e.g., that Princeton should buy tapes and the University of Grenoble should use on-line services, although depending on the local environment some institutions may lean in one direction or the other. Nor can we say that everyone should acquire data set A on tape and data set B through an on-line service, although there, too, we may all veer in one direction or another.

In making these decisions there are three major components for consideration: our local environments, which although essentially fixed are different for each of us; our user communities, which may vary both among us and for each of us over time; and the characteristics of the various data products available to us, essentially the same for each but differing from product to product. It's a little like running a restaurant. We know what our local resources are -- the size of the ovens, the number of tables -- and generally speaking, we can find out what products are on the market and the characteristics and costs associated with acquiring or using them -- fresh, canned and frozen meat, fish and vegetables. The big question is how many people will turn up for dinner.

If we have lots of group reservations made well in advance, we can plan more easily; but if we have only a walk-in clientele, then we have only our experience and the complaints of the customers to rely on. Nonetheless, in spite of these local and market variabilities, it is likely that in the future we will each find ourselves using a mix of "on tape" and "on-line" resources. The remainder of this presentation focuses on the elements which will determine your particular mix and mine.

Let us first look further at local computing environments. Although even within a given institution some of these elements may vary for different classes of users, by and large there is a pattern which characterizes each environment.

The first and perhaps the most critical element in this pattern is the type and the amount of money available. In most computing environments, we find two types of money. The first is real money, money which can be spent as easily outside the institution as inside, money to which there are no strings attached.

The second type of money is internal -- general funds or "funny money" -- money which may be available only for local computing (or, in some cases, for other local services) but which cannot be spent outside. In each institution both the data archive or data library and the individual user have some combination of these moneys. The total amount of each and the amount of computing and ancillary services which the internal component can purchase are key in the on-line or on tape decision. If there is little real money, but adequate
internal money and cheap computing, the inclination will be to keep it local, to buy tapes.

Assuming, however, that the money issue is not so elastic, what are the other elements in the local environment which will affect our decisions? Certainly the quality of computing is one. Although the reliability of the system, the cost and availability of storage devices, and turnaround or response time are important, the quality and variety of available software and programming support are key. If it is not possible for the user to do locally the kinds of analyses required for his work, and on-line services to do these things are available, the pressure to use them will be great.

Last but not least is the quality of the data archive or data library staff and the systems available for acquiring, storing, locating and accessing data. If these are all well-organized, ceteris paribus, users may be less interested in paying real money and learning new systems to use on-line services. Other considerations peculiar to each local environment may also be at issue, but those I have mentioned are always relevant.

Now what of the individual data products? The most fundamental question is: Is there a choice? Some data are available only on-line and other data only on tape. One may characterize data files as consumable, durable, or permanent. The consumable files are like food. One must replace them daily, weekly or monthly. These are files, e.g., which provide input to economic models where the presence of timely data is essential. Timeliness is generally far more essential in business than in academia. These data are generally available only on-line, which accounts in part for the fact that 90 percent of the use of on-line services for accessing numeric data is by business and only 10 percent by academia. For the rest of the explanation we must work back to differences in local environments, to the differences in how commercial and academic users measure costs.

Durable files are like cars, refrigerators, or even husbands. They may be inert files or dynamic files in which the currency of the models is valuable but not essential. The availability of these files on-line is usually a function of size and generality. Large, small-area aggregate files or special-purpose microdata files are less likely to be available on-line than files of national, state, or provincial annual time series data.

Permanent files are like fine paintings or Oriental rugs. They have long lives and small groups of devoted admirers, but seldom enough admirers to justify their availability on-line, whereas their long lives make them good investments for purchase on tape.

Assuming one has a choice between on-line or tape access, what other characteristics of the data product should be explored? Size, cost, and subject matter are certainly factors, as are the relative ease and appropriateness of access and use -- i.e., the source of the data, the form in which they are provided, and the availability of software.

If the file is small and inexpensive, can be purchased and delivered within the appropriate time, can be analyzed locally, is of a permanent nature and relatively general interest, buy tapes.

This decision, however, will also be affected by expected use. For exam-
ple, during the life of the file will there be many users or few? Will they require access to the entire file or only to a small subset? If the latter, will they each require a different subset? And what will be the nature of their use? Do they require only a one-time display of data or will they be subjecting the data to complex and repeated analyses? Can one purchase only a subset of the data on tape or is it necessary to purchase the whole collection? Is there a subscription fee for on-line access, or does one pay per use?

If the cost of on-line access is moderate and the user needs to display only a few numbers, then purchasing the tape may not be justified unless there are clear indications of high future use.

I have identified many variables in the decision equation which we may someday construct. However, our experience is as yet too limited to assign values to those variables. At this point we can only make subjective judgments, based on our awareness of the relevant factors. Each of us behaving rationally may make different decisions, because our computing environments and our user communities are different. Moreover, the advent of cheaper, larger on-line storage makes these decisions dynamic ones.

Some data we may all acquire on tape. Other data we may all access on-line. But for an increasing amount of data our decisions will be individual ones. There is no single answer.

A detailed look at a few specific examples of available data which some of us have acquired or might consider acquiring may serve to illustrate the interaction of local environment, user community, and data product characteristics in determining data access choices. I have tried to choose an international group for these illustrations, but you will forgive me if I speak more about the data I know best.

The United States Bureau of Labor Statistics (BLS) maintains a collection of 100,000 time series covering such subjects as national labor turnover; industry, producer, and consumer price indices; and national, state, and area employment. This collection is known as LABSTAT, and in BLS parlance LABSTAT includes both the data and the software for its analysis. Many of us over the years have purchased individual time series or groups of time series from the Bureau for the cost of approximately $100 per reel, and have used locally installed time series packages such as TROLL, TSP or SAS/ETS for data analysis. A few of us have used one or more of the selected series which a number of the on-line services make available with accompanying analytic software.

Recently Lockheed, which specializes in bibliographic data files, has made available to its users almost half of the LABSTAT data collection. The primary omission is the series relating to unemployment. Lockheed acquires each new update as soon as it is released and provides access to it, using the software which is familiar to its bibliographic data file users as DIALOG. There is no provision for any manipulation of the data, but selected displays are simple and inexpensive.

As we discovered at Princeton in providing access to the United States decennial census data, there is a large group of individuals who use machine-readable data products to find numbers which are not available in print or microform or which are more easily accessible in machine-readable form.
Lockheed has made the same discovery, and the "Labor Statistics (LABSTAT)" file is already being heavily used. A United States Bureau of the Census file called "U.S. Exports" was added at the same time; and it is likely that Lockheed and its competitors will soon add similar data files to their available holdings, and that they will concentrate on the display market rather than on the analysis market. For those of us who wish access to the full database and to analytic capabilities, it is still necessary to purchase tapes.

The International Monetary Fund produces several major data collections, including the World Financial Series (WFS) and the Direction of Trade. The former contains over 40,000 time series of financial statistics on over 160 countries, plus aggregate data for the world and over 50 regions. These data go back to 1948 and have been updated monthly since 1965.

The Direction of Trade contains approximately 100,000 time series of import and export statistics for 230 nations and their trading partners. These data also go back to 1948 but have been updated monthly only since 1977.

Monthly tapes for IFS may be purchased directly from IMF for an annual subscription of $400. ICPSR members may obtain both collections on a less timely basis without cost. Regardless of the source, when the time series package arrives it is necessary to use a special COBOL program to unpack it and write it on disk, an assembly program to reformat it to tape, and a FORTRAN program to retrieve it for analysis before using the data. Data Resources Inc. (DRI), possibly the largest of the commercial on-line sources of economic and other statistical data, ADP Network Services Inc., FRI Information Services Limited, Rapidata Inc., and Telesystemes-Eurodial, among others, provide subscribers with access to these data and to software for analyzing them.

At Princeton the International Finance Section of the Economics Department has adequate general funds money for computing and a graduate assistant to preprocess the tapes and to analyze each month's data. As a result, on-line offerings have not proved attractive.

DRI also provides United States Current Population Survey annual time series data from 1968 to the present. These time series are based on the aggregate data published by the Bureau of the Census. For individuals interested in CPS data in this form, there is no competing product. However, for individuals interested in cross-sectional microdata and specifically in the monthly supplement questions, the Bureau sells each month's file for under $300. The quality of the files themselves and of their accompanying documentation has improved markedly in recent years, and normally the data are easily analyzed if software for processing hierarchical files is locally available. The popular March file, known as the Annual Demographic File, is available from ICPSR.

A somewhat unique case is the 1970 United States Census of Agriculture. A preliminary tape containing a subset of the preliminary published aggregate data is sold by the Bureau at its customary $110 per reel for 5 reels. At the request of some members of the user community, final data were produced on tape as a special tabulation and sold for $1000. In the most recent Directory of Online Data Bases, published by Cuadra Associates, an organization called On Line Research was reported to be making Census of Agriculture data available on-line. Further investigation, however, uncovered the fact that this organization is no longer in business. In this instance it would appear that the 'display-only' customer will do well using the printed reports, and that
the researcher whose data needs are not excessive may find it more economical to reenter the necessary data himself -- particularly if student or clerical time is freely available.

This latter alternative is not one I would normally encourage, but it should not be totally overlooked. It may sometimes prove to be the most practical alternative.

In discussing on-line access I have not distinguished between commercial vendors and academic sources. The issues are essentially the same. When should data be acquired and used locally, and when should it be accessed on-line from a remote site? We must each make our own decisions, and perhaps when we gather together next time, we will have a larger body of experiences to share.

I realize that many of you had hoped for more specific recommendations on these issues. However, as with so many issues concerning the management of data archives and data libraries, we are each bound by our unique institutional settings and by the peculiar nature of our user communities. In the case of our local environments we must consider the amount and the type of money available for computing, the cost and the quality of local computing, the software available and supported, the amount of cheap labor, and the nature of our own operations. In terms of the user community we must consider the number of anticipated users and uses, the type of use, and the amount of data required. And finally, in looking at each data collection we must be aware of the options from which we may choose: the size of the file, its durability, its content, its completeness, its currency, and the software required and available for on-line or on tape use.
ANNOUNCING: AN INVENTORY OF QUANTITATIVE DATA SETS RELATING TO NON-WESTERN NATIONS

Machine readable data files relating to non-Western European, non-Anglo-American nations are scattered among local and general service archives throughout this country and western Europe. The need for a full enumeration of these data files has long been recognized; just such a comprehensive listing is being undertaken by the Civilizations Course Materials Project of the University of Chicago.

The Civilizations Course Materials Project (CCMP) is a curriculum development project sponsored by the National Science Foundation through its program for Comprehensive Assistance to Undergraduate Science Education (CAUSE). Its basic aim is to produce primary documents for undergraduate courses in the non-Western civilizations. The six non-Western areas included are Latin America, the Far East, South Asia, sub-Saharan Africa, the Middle East and other predominantly Islamic areas, and eastern Europe. CCMP's main work has centered on the selection, translation and annotation of documents. By way of providing guidelines for the classroom use of these documents Professor Ralph A. Austen, of the University of Chicago History Department, has undertaken the composition of a CCMP Handbook.

The Handbook provides a statement of the pedagogical principles guiding the project and suggests specific methods for analyzing the kinds of documentary materials it has been producing. One section is being devoted to a consideration of the use of quantitative data in civilizations courses. While it is recognized that local circumstances may not always permit the actual manipulation of aggregate and survey data by large numbers of undergraduate students, it is hoped that exposure to these materials, and an explicit consideration of the thinking behind them, will help students develop a more critical understanding of how professional social scientists work.

Professor Austen decided to include an "Inventory of Machine Readable Data Files Relating to Non-Western Nations" as an appendix to the Handbook; he delegated responsibility for this project to Joseph Hanc. Though it began life as an afterthought, the Inventory has generated far more interest than expected. This note is therefore aimed at colleagues and institutions who may want to acquire the Inventory for purposes other than civilization course instruction and/or can supply the CCMP with additional archival data.

Information on non-Western European, non-Anglo-American nations is found in studies of varying geographic scope. "National" studies, which contain information on individual countries, should be distinguished from "cross-national" studies, which examine large numbers of countries. A further discrimination may be made between cross-national studies with "regional" and "world-wide" samples. National and regional studies are noted in the Inventory, as are "local" or "community" studies. World studies are excluded.

The Inventory contains information on the holdings of over twenty general and local service archives. The data files themselves have not been examined. Rather, the Inventory reproduces titles, periods, dates of assembly, and file numbers provided by the disseminating archives themselves. Individ-
ual files are arranged first by world area, next by topic. The world areas represented are Latin America, the Far East, South Asia, Sub-Saharan Africa, the Middle East, and Eastern Europe; the topic headings are drawn from the Guide to Resources and Services published annually by the Inter-university Consortium for Political and Social Research.

A simple listing cannot hope to convey the character and variety of the quantitative social science data available on non-Western civilizations, but it is hoped that this inventory will stimulate more frequent recourse to this little known and underutilized source of information. The inventory will be distributed to a select group of archives for comment. Following revisions, the finished inventory will be deposited in both the library of the National Opinion Research Center (6030 S. Ellis Ave., Chicago, Illinois, 60637), and the CCMP Office (5845 S. Ellis Ave., University of Chicago, Chicago, Illinois, 60637); copies are available for the cost of reproduction and handling. Any archives which hold unique data files on non-Western nations and have not yet been contacted by Professor Austen or Mr. Hanc are encouraged to contact either of these offices so that their holdings can be incorporated into the inventory.

NEW PERIODICAL: SOFTWARE REVIEW

Microform Review Inc. has announced a new publication Software Review, to be published twice yearly (starting October 1981) at $38 (U.S.) or $43 (Canadian and foreign). It is aimed at the library and educational market and provides information about pre-written software. Each issue features:

- Articles on software concepts, evaluation, and selection;
- Reports on available software products suitable for educational and library applications;
- Reviews of books and other recent publications pertaining to computer software; and
- Advertisements of available products.

Software Review currently seeks contributions in the above areas. Send them to the editors, William Saffady and Rhoda Garoogian (both of Pratt Institute), c/o Microform Review Inc., 520 Riverside Ave., Wesport, CT 06880.

Companion publications, also new, are Computer Equipment Review and Videodisc/Teletext.

Meckler Books, a division of Microform Review Inc., has announced for publication in 1982 a Bibliography of Statistical Bulletins and Yearbooks, covering all areas of the world; and a Guide to Statistical Materials Published by Governments and Associations in the United States. Both may be useful in reference work in data libraries and archives.

This volume is an alphabetical index to those free black heads of households that were listed in the Federal censuses for New York state in the five decennial censuses from 1790 to 1830. The compilers intend it to be used with Black Genesis, their previous work published by Gale (Vol. 1 in the series named above).

Tracing black families is complicated greatly by the fact that many lived in white households and are not listed in the censuses, while others are listed only by first name. Obviously an index of this kind cannot cover the first group, but the second group is indexed by first name, interfiled with the surnames.

Each listing includes the census year, the New York county and township, and the page number. For the 1790 census, page number refers to the printed volume compiled by the Census Bureau in 1908: Heads of Families at the First Census of the United States Taken in the Year 1790. Page numbers for subsequent censuses are from the microfilms available from the National Archives and Record Center.

This is a well organized and clearly printed volume that should greatly simplify research in the area.

--Barbara Roberts, Archivist Urbana IL Free Library


The U.S. government gave official consideration in 1966 to developing social indicators comparable to its economic and labor statistics. Generally social rather than economic, urban indicators are statistics that provide data on the quality of life in our cities in such areas as health, education, and social services. Murphy's Guide balances the theory, methodology, technology, and application of urban indicators to achieve an interdisciplinary representation of readily available materials.

An introduction and six chapters provide a comprehensive bibliography of urban indicators: their development and future; types; problems and applications; and program evaluation. Three appendixes are devoted to 1) relevant periodicals, 2) abstracts and indexes, and 3) public interest and professional associations. There are author and title indexes, and a detailed subject index.
Built in 1888, the Hotel del Coronado is considered to be one of the last extravagantly conceived seaside resort hotels and is a fine blend of Victorian charm and modern-day convenience. Attractions include white sand beaches, spas, tennis courts, heated pools and boat rentals. Restaurants and meeting rooms open onto decks and boardwalks overlooking the ocean front.

The Program Committee recommends participants and attendees consider taking advantage of the Memorial Day weekend and the Monday holiday by taking vacation and extending hotel reservations to include Sunday night, May 30.

Single or double occupancy will cost the same amount, $69.00 per night for the Main building. Rooms may be shared to reduce conference costs or a spouse may be included at no extra charge. For those wishing to stay Sunday, the hotel has agreed to extend the conference rate.

Sea World, the San Diego Zoo, and the Wild Animal Park are within easy reach. Trips to Mexico can also be scheduled since the border is only 15 minutes from the hotel.

Call Don Trees (213) 393-0411, ext. 489, or Ilona Einowski (415) 642-6571 for additional information.

Plan now to join us!

Program Committee
Ilona Einowski, UC Berkeley, Co-chair
Jackie McGee, Rand
Elizabeth Stephenson, UCLA
Don Trees, Rand, Co-chair

Site Arrangements
Jackie McGee

Workshops
Elizabeth Stephenson
ADMINISTRATIVE COMMITTEE

President: Alice Robbin, Data and Program Library Service, University of Wisconsin, 4452 Social Science Building, Madison, Wisconsin 53706, U.S.A.

Vice President: Per Nielsen, Danish Data Archives, Niels Bohrs Alle 25, DK-5230 Odense M, Denmark

REGIONAL SECRETARIES

Asia: Naresh Nijhawan, Indian Council of Social Science Research, Data Archive, 35 Ferozshah Road, New Delhi 11001, India

Canada: Sue Gavrel, Machine Readable Data Archives, Public Archives of Canada, 395 Wellington Street, Ottawa, Ontario, K1A ON4, Canada

East Europe: Krzysztof Zagorski, Instytut Folosofii i Socjologii, Polskiej Academii Nauk, Nowy Swiat 72, Palac Staszica, 00-330 Warszawa, Poland

West Europe: Henk Schrik, Steinmetz Archives, Herengracht 410-412, 1017 BX Amsterdam, The Netherlands

United States: Judith S. Rowe, Computer Center, 87 Prospect Street, Princeton University, Princeton, New Jersey 08544, U.S.A.

MEMBERS-AT-LARGE

Patrick Bova, National Opinion Research Center, 6030 South Ellis Avenue, Chicago Illinois 60637, U.S.A.

Nancy Carmichael McManus, Social Science Research Council, 1755 Massachusetts Avenue, N.W., Washington, D.C. 20036, U.S.A.

John de Vries, Department of Sociology, Carleton University, Ottawa, Ontario, K1S 5B6, Canada

Sue A. Dodd, Social Science Data Library, Manning Hall, University of North Carolina, Chapel Hill, North Carolina 27514, U.S.A.

Carolyn Geda, Inter-university Consortium for Political and Social Research, P.O. Box 1248, University of Michigan, Ann Arbor, Michigan 48106, U.S.A.

Guido Martinotti, Archivio dati e programmi per le scienze sociali, via G. Cantoni 4, 20144 Milan, Italy

Harold Naugler, Machine Readable Archives, Public Archives of Canada, 395 Wellington Street, Ottawa, Ontario K1A ON4, Canada

EX OFFICIO

Treasurer: Ed Hanis, Social Science Computing Laboratory, University of Western Ontario, London, Ontario, Canada

Editor: Howard D. White, School of Library and Information Science, Drexel University, Philadelphia, Pennsylvania 19104, U.S.A.
IASSIST

JUDITH S. ROWE, U.S.A SECRETARIAT
PRINCETON UNIVERSITY COMPUTER CENTER
87 PROSPECT STREET
PRINCETON, NEW JERSEY

08544, U.S.A.