BOOK NOTICES / kathleen m. heim

Machine-Readable Social Science Data - Special issue of the Drexel Library Quarterly. Volume 13 (January 1977). Edited by Howard D. White. Issue available for $5.00 by writing: Drexel Library Quarterly, Graduate School of Library Science, Drexel University, Philadelphia, Pennsylvania 19104.

Contents: "Numeric Data Files: An Introduction." Howard D. White.
"The Pre-Acquisition Process: A Strategy for Locating and Acquiring Machine-Readable Data." Alice Robbin.
"Stalking the Wild Data Set: The Acquisition of Machine-Readable Social Science Data at Home and Abroad." David Nasatir.
"Cataloging Machine-Readable Data Files - A First Step?" Sue A. Dodd.
"Social Science Data Files, the Research Library and the Computing Center." Douglas Ferguson.
"The Information Transfer Process on the University Campus: The Case for Public Use Files." Lorraine Borman and Richard Hay, Jr.
"Training the Professional Data Librarian." Judith S. Rowe and Carolyn L. Geda.

"This is a contribution to a growing discussion between and within the computing and library worlds." - Douglas Ferguson

Though Ferguson is referring to his own article in this special number of the Drexel Library Quarterly, his statement quite accurately describes the entire issue, for this marks the first time that the library press has devoted a substantial amount of space to treatment of the problems that confront the data archivist. The contributors are familiar names to those active in the data archive movement--all have published extensively and have been involved in the conceptual definitions of the principles of data archiving.

Because the intended audience of this issue is the broad range of library and information professionals, the treatment of matters familiar to the working data archivist may seem elementary. Howard D. White's introduction to numeric data files is a careful building of concepts and definitions that provides the newcomer with vocabulary and a general framework on which to base an understanding of the technical articles that follow. For the novice, Douglas Ferguson's article follows White's most logically for he assesses the role of the library in data archive development and provides the rationale for cooperation. His use of the specific example of the Stanford University experience provides a cogent model of library-archive interaction. Lorraine Borman and Richard Hay, Jr. extend the concepts defined by Ferguson by presenting the present-day situation in universities for the use of public files. They attack the problem of low use by focusing upon the kind of public use files in a social science file collection; illustrate how these files can be accessed and manipulated by both novice and sophisticated users; and show how this type of use is related to the utilization of public information files and information systems, and assists in the total information transfer process.

From these articles the more specialized essays by Alice Robbin, David Nasatir, and Sue A. Dodd may be approached. Robbin's scrutinization of the pre-acquisition process is a detailed, logical analysis of the complex factors that enter into the decision to obtain a particular data set. After a general introduction which places the data archive squarely as a type of special library, Robbin explores the five factors which influence the pre-acquisition process and the search strategy for locating data. A diagrammatic presentation of the search strategy is presented with each part carefully explained in an appendix. Nasatir's brief piece, "Stalking the Wild Data Set," is by far the most dramatic contribution to the entire issue. It is a simple, yet vivid rendering of the sometimes Machiavellian, sometimes mendicant contortions which a data archivist must endure to procure material. Working data archivists will empathize, and those unfamiliar with the field will marvel, at Nasatir's travails.

Dodd's article on cataloging Machine-Readable Data Files (MRDF) may have been the most difficult piece to compose. It has fallen to Dodd to unravel the lengthy history of both the library and data archive profession's attempts to establish conformity in the cataloging of MRDF. The article is loaded with acronyms, committees and subcommittees of the ALA, ASIS, IASSIST and AECT. Though difficult at the outset her essay provides a solid introduction to the monumental problems involved in hammering out a cataloging format agreeable to the various factions. She examines progress to the present focusing on the IASSIST Classification Group project designed to test the feasibility of cataloging social science data files according to the ALA subcommittee's recommended procedures. An appendix of sample catalog cards from those participating in the project is included.

Finally, Judith S. Rowe and Carolyn L. Geda discuss the training of the professional data librarian. The profession is still in its formative stages and current practioners are drawn from a diverse group: programmers, social scientists or librarians. The only formal training available is via the Inter-University Consortium for Political and Social Research summer program held at Ann Arbor, Michigan, which is described in detail.

As stated above, the intended audience of this special issue is the working librarian or information professional. This audience requires definitions that may be elementary to the data archivist. However, Howard D. White's editing has produced a diverse publication that operates on several levels. His contributors have at once provided a sourcebook of articles comprehensible to one unfamiliar with data archiving and yet valuable to the experienced archivist. Dodd's patient unraveling of the machinations of cataloging MRDF and Robbin's precise articulation of pre-acquisition procedures are the only explanations of these processes in the literature. The essays devoted to the reference and information component of data archiving extend current ideas about service and begin to outline the emergent user-models.

This issue will stand as a vital resource for data archiving until some more comprehensive publication supersedes it. The only glaring defect is the lack of a simple introduction that provides a framework for the whole. Nevertheless, this issue of the Drexel Library Quarterly is a necessary item for all data archivists. It will provide the profession with a stronger realization of its role in the information transfer process and it will provide traditional librarians with an understanding of the role of data archives in that process.

Richard C. Roistacher and Barbara C. Noble. Computer Network Support of Social Research Communities. University of Illinois. Graduate College. Center for Advanced Computation. Urbana, Illinois 61801.

In March of 1977 following the First IASSIST North American Conference, Richard Roistacher provided participants with publications and technical reports which fell into three major classes: 1) LEAA Data Archive Material; 2) Data Archiving and Analysis Methods; 3) Sociology of Computing. All of the papers advance the conceptualization of the role of data archives in the research community. Mr. Roistacher has agreed to provide copies to interested data archivists who contact him at the Center for Advanced Computation.

In this notice we would like to focus upon one of the papers which falls into the category of the Sociology of Computing: Computer Network Support of Social Research Communities. This paper, a pre-print of an article to be published by Human Organization, sets forth the idea that present computer technology will allow geographically separated social researchers to share data and data analysis tools, collaborate on research and writing, and communicate with their colleagues much as if they were in a single department. A description of the Law Enforcement Assistance Administration (LEAA) computer network based research support center which operates over a commercial communication network is described in detail.

The LEAA data archive and research facility for social scientists and program evaluation groups which has been developed by the Center for Advanced Computation at the University of Illinois to provide research support services to clients scattered around the country is viewed by Roistacher as a prototype. The services offered by a computer network fall into three major categories: data management and archiving; computer and statistical consulting; and communications and documentation.

Roistacher uses the generic term, Research Support Facility (RSF), to describe the multi-purpose service which this prototype provides. Services fall into two modes: those which do not require a computer network connection to the client (data archives; consulting on analysis of archive files; consulting on client-collected data; agency for other archives; bibliographic information retrieval; and software distribution) and those which can only be provided over a direct network connection (shared archival data; shared software; on-line consulting; shared user data; remote technical services; network mail; network conferencing; facilitating the formation of working groups; and a network journal). All services are described in detail. Potential clients for RSF services include academic researchers, independent research organizations, state planning agencies, and program evaluation groups.

Computer and network costs are compared favorably with the costs of using local facilities. Though charges seem high Roistacher demonstrates that they are much lower than travel, postage, typing, duplicating and time considerations under the present system. Management considerations in remote services are treated. These include training users, administrative procedures, accomodation of clients in distant time zones, terminal acquisitions, and clear and self-sufficient documentation.

The social consequences of networking as perceived by Roistacher's observations of the Department of Defense's ARPAnet are discussed and the implications for collaboration among colleagues regardless of geographical considerations explored.

Though data archives are only one component in the larger RSF, they are a vital component and data archivists will want to ponder Roistacher's predictions for computer-based research support. His experiences with a highly developed facility provide a model for future developments in behavioral science communication and augur a new perhaps more democratic era for the invisible college.

Statistical Policy Division. Office of Management and Budget. In Cooperation with the Federal Statistical Agencies with Responsibility for the Collection, Processing, Analysis and Dissemination from Major Federal Statistical Programs. Framework for Planning U.S. Federal Statistics 1978-1989. (Reprinted from the Statistical Reporter, 1974-1976). Copies available from the Statistical Policy Division, Office of Management and Budget, 726 Jackson Place, N.W., Washington, D.C. 20503.

This collection of reprints from the Statistical Reporter details the efforts of the Statistical Policy Division of the Office of Management and Budget to provide for a more systematic review of the needs for improved statistical planning and coordination. A topical outline of the planning framework is presented along with articles on its nature. The goal is improved coordination of the highly decentralized U.S. Federal Statistical System and setting of statistical priorities.

Of particular interest is the article, "User Access-Data Banks," which includes a bibliography of major agency guides and indexes to statistical data. Much concern is expressed over the failure of government agencies to comply with government regulations which provide for improvement of the dissemination of statistical information on a timely basis.

Feedback is requested from users of government generated data. IASSIST members can obtain a copy of the Framework for Planning U.S. Federal Statistics 1978-1989 from the above address and reply to the same on matters concerning shortcomings in the statistical services of the U.S.