Introduction

KATHLEEN M. HEIM

AUGUST COMTE'S GRAND VISION of the final synthesis of the natural sciences and social physics focused on sociology as the universal link "completing the upward flight of our contemplation of reality."[1] The end result would be an applied social science that would ameliorate the state of humanity. Comte's call for a general social theory that would consist of a set of general, testable, explanatory propositions applicable to the total area of collective human behavior was shared by later nineteenth-century statisticians, sociologists and anthropologists. These were the individuals who endeavored through the comparative method to establish an internationally and interculturally valid body of knowledge about variations and regularities in the functioning and development of human societies.

Such an aim proved difficult to reconcile with other compelling objectives as the social sciences developed. The need to establish strict canons of evidence and inference in order to achieve a high level of analytical precision, as well as the need to gain academic recognition for these emerging disciplines, tended to force early social scientists to abandon universalistic theorizing and focus their inquiries on the local, the concrete and the specific. The social sciences achieved academic status by increasing attention to methodological rigor and deliberate concentration on well-delimited local or national inquiries. Though the gain in precision has been great, the original aim of the social sciences has tended to be sacrificed and the idea of an encompassing or

Kathleen M. Heim is Assistant Professor, Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign.

grand theory that synthesizes universal concepts is not a prominent one in the current panoply of social science philosophy. Nevertheless, the idea of the eventual achievement of a universal theory holds a persistent grasp on the imagination and presents to the layperson the most obvious rationale for the importance of the social sciences.

Before social scientists could begin to reconsider postulating theories at a highly general level, the raw information required to develop such ideas has had to undergo four revolutions. Karl Deutsch has described these as: (1) the collection of largely disjointed facts and figures chiefly for administrative, tax and military purposes in the seventeenth and eighteenth centuries; (2) the use of historical data dealing with successions of types of societies by such social scientists as Herbert Spencer or Karl Marx; (3) the rise of new methods in the mid-1930s for gathering partial and sectoral data, along with new quantitative techniques for organizing and interpreting them in order to put discrete, disjointed data in relationship (helped along by advances in survey research and sampling theory); and (4) the rise of multiple methods and complex data bases with the eventual aim of all-to-all comparison.[2]

The fourth data revolution, with the possibilities of all-to-all comparisons, has come into being because of technological advances in the computer which allow ever more complicated statistical analysis of data and expanded capacity for storage of these data in facilities such as data archives or data libraries. The data library contains machine-readable collections of survey, census, polling, or legislative voting information and provides a laboratory for the social scientist to analyze data in order to make sometimes narrow and sometimes highly general observations about the nature of society.

Data archive and data library development was the focus of intense national and international debate in the late 1960s and early 1970s. The massive programs of international economic and political integration undertaken after World War II demanded much comparative research which was unavailable because the theoretical underpinnings of any efforts at cross-national comparison were poor and fragmentary. Unesco support to forward the state-of-the-art in the comparative social sciences often focused on the data library as a primary component in the information system required to enable researchers to understand and provide solutions which might alleviate disparities in the development of various nations.[3] In spite of the extensive international and national efforts to develop not only individual data libraries throughout the world but networks of such information as well, the data library move-

ment has persisted but not flourished. Although researchers at every university in the world could make use of data libraries, the high cost and poor understanding of their role in the social science information system has mitigated against full-scale development on the level of more traditional information facilities, such as libraries. Put most simply, the material in data libraries is a vital source of information for social research, yet because of the nontraditional format of these data, has been ignored by the information community at large. The failure of most models of the social science information system to include the data library in schemata and diagrams indicates that this resource is poorly understood. If the highly motivated librarian attempts to provide information resources to scholars wishing to study voting behavior and can only direct them to books, periodicals and government reports, major lacunae in the information infrastructure are evident, for machine-readable collections on the topic are probably more pertinent than anything published.

Failure of traditional libraries to consider machine-readable data files as within their purview has caused data library development to take place, in the main, outside of traditional library settings. Worse, most librarians in large research libraries do not even recognize the disservice they do their clientele in omitting information provision in this area. This issue of *Library Trends* is intended to place the data library and its holdings squarely in the forefront of vital information resources for social sciences research. Judith S. Rowe provides a general introduction to the importance of Machine-Readable Data Files (MRDF) in the social sciences reference exchange and demonstrates examples of their use in typical library situations.

Sue A. Dodd describes the arduous struggle to develop bibliographic control over MRDF and thus legitimatize them to the library community. She observes that "communicating the availability of usable data is an inseparable part of research and an integral part of librarianship. In the near future libraries will have no choice but to become more involved."

The enormity of information collected in MRDF is characterized by Joseph W. Duncan, who notes the underutilization of federal social data due to a lack of adequate information. Duncan discusses the data access policies of the federal statistical system and selected source documents which aid researchers in accessing these data.

Margaret O'Neill Adams introduces four models of data library development and notes that "there is no clear administrative structure for services related to social science data files nor for numeric informa-

tion systems that is ideally suited for all institutional settings." The four models which Adams presents provide alternative functional strategies for meeting information needs for MRDF. These models include the integrated University of Florida facility outlined by Ray Jones, the University of British Columbia Data Library operated jointly by the University Library and Computing Centre described by Laine G.M. Ruus, the Data and Program Library Service at the University of Wisconsin characterized by Alice Robbin, and Adams's online numeric Kentucky Economic Information System at the University of Kentucky.

Once the need for data libraries has been made clear, the type of resources they might hold particularized, strategies of bibliographic access delineated, and model facilities described, the question arises: Who will staff and maintain such services? Laine G.M. Ruus explores the training of data services professionals and includes a suggestion that graduate schools of library education consider incorporating such training in their curricula.

The next three papers in this issue examine use of data files, the role of data files in social science teaching, and issues of confidentiality and privacy. Howard D. White explores the current state of affairs vis-à-vis citations to MRDF through an analysis of *Social Sciences Citation Index* detailing the complexities and vagaries of accurate bibliographic control over these files.

Jeff Sobal, a social scientist and a user of MRDF in his own research, discusses the role of secondary data analysis in teaching the social sciences, and suggests ways in which traditional libraries might assist in expanding awareness of MRDF as an information resource.

A pressing concern for data librarians is the confidentiality of the information retained in data files. David F. Linowes, former chairman of the U.S. Privacy Protection Commission, and Michele M. Hoyman of the Institute of Labor and Industrial Relations at the University of Illinois analyze the broad issue of information privacy in general as well as its relationship to the function of the librarian and archivist.

Observations by Barton Clark, head of the Social Sciences Libraries at the University of Illinois conclude this issue. Clark is not optimistic about the wholesale integration of traditional libraries and data libraries, but rather takes a middle stand in recommending that traditional libraries become bibliographic brokers for data libraries.

If long-range goals for the systematization of a social science infrastructure as outlined by the UNISIST International Committee for Social Science Information and Documentation are met, the data library will be fundamental to the aim of creating a world system of social

science information and documentation.[4] I hope that this issue of *Library Trends* contributes to an understanding and greater visibility for this vital social science resource.

References

1. Comte, August. *System of Positive Polity*. Vol. 1. Translated by J.H. Bridges. London: Longmans, Green and Co., 1875-77, p. 12.

2. Deutsch, Karl W. "The Impact of Complex Data Bases on the Social Sciences." In *Data Bases, Computers, and the Social Sciences*, edited by Ralph L. Bisco, pp. 19-41. New York: Wiley-Interscience, 1970.

3. For a detailed account of national and international efforts to develop a network of data archives, *see* Heim, Kathleen M. "Social Science Data Archives: A User Study." Ph.D. diss., University of Wisconsin-Madison, 1980.

4. Unesco. International Committee on Social Science Information and Documentation. Ad Hoc Committee on Social Science Information. "Background Paper on Institutional Aspects and the Organization of Social Science Information and Documentation Activities," (SS-77/Conf. 607/Col. 4, Nov. 14-18). Mimeographed. Paris: Unesco, Social Sciences Documentation Centre, 1977.

This Page Intentionally Left Blank