MACHINE READABLE ARCHIVES: USER SURVEY

by

SUE GAVREL

Machine Readable Data Archives
Public Archives of Canada

This article is an abridged version of a report written by S. Paula Mitchell and Stavko Manojlovich of the Social Science Computing Laboratory, the University of Western Ontario, London, Ontario, Canada.

Background

In November, 1981, the Machine Readable Archives Division undertook an analysis of the requests and inquiries it had received for its services over the last few years. It became evident during the analysis that although the division provided a great deal of information on the establishment of machine readable programs, procedures and practices used to archive machine readable records, the actual holdings of the division were not being used. There were several possible reasons for the lack of use of the files, but the main conclusion from the analysis was that the MRA did not have sufficient information on who the user population was and therefore was not distributing information about its holdings to the appropriate areas. The users of the traditional archival records were not the same as those who would use machine readable data. It was decided that a survey of our existing users and potential users should be undertaken.

The Social Science Computing Laboratory of the University of Western Ontario was contracted to undertake the survey. Several steps were involved: the drafting of the questionnaire, the collection of the data and the analysis. The project spanned the period February 1, 1982 to October 30, 1982 and the total cost of the project was $12,478.07, which included staff charges, computing services, printing, stationery and postage.

The main objective of the survey was to obtain information on the needs of Canadian social scientists who created or used machine readable data. The results of the survey would be used by the MRA to help plan an expanded public service program and determine priorities for the allocation of resources. A secondary aim has been to identify sources of machine readable data in Canada. Specific objectives of the survey were as follows:

- to provide a measure of the relative usefulness of a number of products and services to the user community:

- to provide an indication of the preferred mode of acquisition for machine readable data files and the preferred type of codebook;

- to identify the location of machine readable data files for the purpose of expanding a union list of machine readable data files in Canada and identifying files of historical research value and national significance falling

within the archival mandate of the MRA;

- to provide an indication of the public's awareness of the holdings and services of the MRA;

- to provide an indication of the need for a national organization to coordinate selected information services for users of machine readable data (including special workshops and/or conferences);

- to provide a mailing list of individuals interested in receiving the MRA's publications or announcements; and

- to identify the need for products or services not covered by the first objective noted above.

Survey Population and Data Collection

The survey population was derived from two social science information systems, the Canadian Register of Research and Researchers in the Social Services (Social Science Register) developed by the Social Science Computing Laboratory, and the Canadian Directory Service of Social Scientists (SocScan), established by the Social Science Federation of Canada between 1976 and 1978. The Register supercedes the SocScan database.

The target population was derived from these two data bases by selecting variables defining methodology and research orientation. Although the target population does not include all social scientists who are current or potential users of machine readable data, the large number (4,065) does provide a good representation of users of social science data in Canada.

The questionnaire, designed to be as simple and short as possible, contained 15 questions. Data were collected using a mail-out/mail-back process. A covering letter explaining the objectives of the survey was sent with the questionnaire. A follow-up reminder was sent to all non-respondents

three weeks (15 working days) after the initial mailing.

FINDINGS

Survey Response

As mentioned earlier, 4,065 questionnaires were sent out. Of the 4,065 questionnaires, 82 were returned as undeliverable mail, thus making the actual survey population 3,983. Response to the survey was very good with an overall response rate of 48.95 percent, or 1,950 completed questionnaires.

Survey Results

The following section provides a brief summary of results. Findings are reported on a question-by-question basis.

Q1. *Prior to receiving this questionnaire, were you aware of the activities and services of the Machine Readable Archives?*

The results of the survey indicate that awareness of the activities and services of the MRA among social scientists is low. Of the 1,930 respondents answering this question, 14.2 percent indicated that they were aware of the Division's activities and services prior to receiving the questionnaire packet. There was no significant difference in the level of awareness between respondents in the academic and non-academic sectors (14.2 percent and 14.1 percent, respectively).

Q2. *Do you use machine readable (computer) data in your research, teaching, or other activities?*

Of the 1,908 social scientists who responded to this question, 1,293 (67.8 percent) reported that they are users of machine readable data, and 615 (32.2 percent) reported that they do not use machine readable

data. Given the quantitative orientation of the survey population, it is somewhat surprising that close to one-third do not use machine readable data. However, the very high interest shown by this "non-user" group in receiving information on the services and holdings of the MRA (76.8 percent wished to be on the MRA mailing list) suggests that the survey was on target in terms of reaching quantitative social scientists who are potential users of the MRA.

It is possible that some respondents in this "non-user" group found the question ambiguous: either limited the inclusive time frame of "use" to the present, or interpreted the question to mean use of secondary data rather than their own primary data.

The distribution of users of machine readable data geographically and by sector of employment is proportional to the distribution of all respondents in these two respects. Except for Ontario and Quebec, the percentage distribution by province of social scientists who use machine readable data does not vary by more than plus or minus 0.6 percent relative to the percentage distribution by province of all respondents. In Ontario, there are 1.9 percent more users relative to the percentage of social scientists from Ontario constituting the total respondent group, and 3 percent fewer in Quebec.

The distribution of users by employment sector (academic/non-academic) varies by only 0.3 percent relative to the sectoral distribution of social scientists in the responding population.

Q3. *Is there a facility within your department or organization which provides information and/or other services for machine readable data?*

1,269 users of machine readable data responded to this question. A very high percentage (93.2 percent) reported that there is a local facility which provides machine readable data services; 86 users (6.8 percent) reported that there is no local facility providing such services.

Q4. *Please identify the facility within your department or organization which provides information and/or other services for machine readable data.*

More than half (53.1 percent) of 1,172 users who are availed of a local facility are serviced by a computing center only. While the range of services provided by any one computing center may approach the full range of information, access, and utilization services provided by specialized local facilities, computer centers typically provide limited support services in these areas. It is a significant finding that such a high proportion of machine readable data users are availed of limited informational support services for finding and using machine readable data. 369 Respondents (31.5 percent) indicated that more than one facility provides local machine readable data services. 25.1 percent of users have a specialized local data support unit (such as a data archives) available to them. For 23.3 percent of users, the local research library integrates some level of service for machine readable data with other services.

Q5. *Please indicate the potential usefulness of each of the following information products or services to your research, teaching, or other activities:*

  a) *Catalogue of Holdings of the Machine Readable Archives;*

4