Data Services in Academic Libraries

Assessing Needs and Promoting Services

Academic libraries play an important role in making numeric data collections available to their researchers and providing assistance in identifying and accessing appropriate resources. The University of Tennessee Libraries have been working to strengthen the numeric data component of their reference services and have expanded promotion and outreach activities to make this specialized service more visible. In fall 2003, the Data Services Awareness and Use Survey was conducted to learn more about the university's users of research data and assess their awareness of the service and the effectiveness of promotional activities. Results of the survey are being used to plan, promote, and provide data services. The survey portion of this manuscript was presented, in part, at the International Association for Social Science Information Service and Technology Conference on May 26, 2004 in Madison, Wisconsin.

Academic researchers use numeric data resources to study a vast array of research questions. Applications range from using summary statistics to describe a population, the economic condition of a country, or the opinions of a group of people, to analyzing data sets in order to study relationships between different variables. In the academic environment, libraries may play a pivotal role in providing services and collections to help researchers identify and access appropriate data resources.[1] To support these important research activities at the University of Tennessee (UT), the university libraries have, in recent years, committed to expanding reference services and collection development activities to include a numeric data component. This service is called Data Services.

With any relatively new library service, it is important to promote it and raise awareness among potential users. It is just as important to determine the needs of those users to facilitate planning and provide appropriate services. The Data Services Awareness and Use Survey of 2003 was the libraries' attempt to get feedback from current and potential data users about their use of and needs for data, their awareness of Data Services, and the effectiveness of various promotional and outreach activities. This article provides a brief overview of data services in academic libraries and discusses how the survey conducted at UT was used to learn more about its data users in order to develop new services and outreach programs that meet users' needs.

Eleanor J. Read

***Eleanor J. Read** is Associate Professor and Social Science Data Services Librarian, University of Tennessee, Knoxville. Submitted for review June 15, 2005; accepted for publication September 20, 2005.*

DATA NEEDS

A numeric data product may be defined as any information resource with considerable numeric content.[2] These sources may be found in a variety of formats, including print, microform, or electronic formats such as CD-ROMs, online databases, and Web sites. The types of data may range from aggregated statistical data to microdata in large computer files. It should be noted that, while the focus of this article is on numeric data, "data" subjected to secondary analysis by researchers may also come in digital forms such as cartographic data, text, images, and audio and video files.[3]

Numeric data may be used to answer a variety of research questions. A common need of students, for example, is to find information in the form of a few quick facts or statistics that can be used in a report or speech. Students may need to put together a sociodemographic profile of a community, find crime rates for several cities over the last ten years, or look at poll results to assess current public opinion on a given topic. The emphasis, in such cases, is usually on reporting or description, rather than manipulation or analysis.[4] Some popular sources of statistical information include the *Statistical Abstract of the United States* (print and online), the U.S. Census Bureau's American FactFinder, and FedStats (a gateway to statistics from more than one hundred U.S. federal agencies).[5]

At the other end of the spectrum are the researchers who wish to analyze data to examine relationships, trends, phenomena, or policy implications that were not addressed by the original researchers.[6] This level of research usually involves a larger number of variables and observations than does a statistical information question. These researchers will prefer to obtain the data in a machine-readable data file or some other electronic format that facilitates use with statistical computing applications such as SAS or SPSS. A few of the excellent sources of social science data that are often used in academic research are the Inter-university Consortium for Political and Social Research (the world's largest archive of social science data), the Roper Center for Public Opinion Research (the world's largest archive of survey data), the U.S. Census Bureau (including American FactFinder), and the statistical agencies of the U.S. government.[7] A major benefit of using data collected by someone else is that it may eliminate the need for researchers to do their own survey or data collection project, thus saving time, effort, and expense.

DATA SERVICES

Data services in academic libraries may be called many things. Numeric data services, electronic data services, research data services, data library, data center, and data archive are but a few examples. Data librarian and data services librarian are common job titles for a person who provides data services.[8] Unfortunately, the name data services does not necessarily adequately convey the purpose and function of the unit. Even among librarians, there is confusion about what this service is all about.[9]

Whatever the terminology, the core purpose of data services is to help faculty, staff, and students successfully navigate the vast array of available numeric information to find appropriate data for their particular research questions. In short, data services is reference work, and it draws on skills and knowledge that are used regularly in traditional reference services.[10] Often it is more complex and time consuming, though, given the in-depth assistance and "nontraditional strategies" that may be needed to complete the data discovery process.[11]

When researchers wish to analyze existing data for purposes other than those for which the data were originally collected (i.e., to perform secondary analysis), they must discern whether a given study, survey, or program collected data that will be useful. After identifying potentially useful studies based on a study or program description, researchers must follow up by looking at the study documentation (e.g., codebook, user guide, questionnaire, interview script) for important details. For example, for surveys, which are an abundant source of research data, this means reviewing the questions asked of the respondents and learning about the methods used to collect the data. This process is usually a bit more involved than searching for relevant scholarly articles in bibliographic databases. For a complex study, a considerable amount of time and effort may be required before a researcher knows enough about the study to be sure that it is suitable and to be able to prepare, analyze, and interpret the data properly.[12]

Data Service Levels

Researchers need not be alone in their quest for research data. Data librarians provide reference assistance in the data discovery process, helping to identify appropriate sources, studies, or data collections and then locating the data in the library's

collection or through external sources. They advise users on how to access data files from CD-ROMs, the Internet, or library servers, and they may purchase or subscribe to data resources when appropriate. In some cases, they also assist with data file preparation, descriptive or statistical data analysis, or Geographic Information Systems (GIS) applications, among other things.

The level of reference service provided by a library is influenced by the number of staff and their skills and knowledge, funds for data resources, computing hardware and software, user needs, and availability of related services elsewhere on campus. Jacobs has outlined some examples of service levels for reference data services in academic libraries: (1) data file identification services, (2) basic data file recommendation services, (3) advanced data file recommendation services, (4) data file use advisory services, and (5) data extraction services. Each level requires successively more involvement, knowledge, and skill, and not all libraries may choose or be able to provide all of these levels of assistance. Ideally there would be one or more other units on campus that could partner with the library to provide a broader array of services.[13] For example, at UT, the Office of Information Technology's Statistical Consulting Center (SCC) is a close partner with the libraries' Data Services. Data users are regularly referred to SCC when their work with data files exceeds the capabilities of Data Services and, on occasion, SCC refers faculty and students to Data Services.

Skills and Knowledge

Data librarians must possess a variety of skills, some of which go beyond the normal skill set of a reference librarian. With data services, having knowledge of a broad array of data products is extremely helpful in being able to provide effective assistance at advanced levels. Regardless of the background or training of the librarian, this knowledge takes years of experience to acquire.[14] With each new data resource encountered, the librarian not only learns about its availability, but also the specific topics it covers, the limitations of or problems with its use, the sampling methodology, and other information that facilitates advising users.[15]

Computing skills are essential to a data librarian.[16] In addition to having a basic understanding of and ability to use computer hardware, software, and operating systems, a data librarian should be comfortable working with data files in a variety of formats. Useful skills include downloading, copying, and transferring files; importing and exporting files into or from applications; and unzipping compressed files. Data librarians who are proficient with statistical computing software may help users prepare their data files for analysis by converting a plain text data file into a formatted file, subsetting out selected variables and observations, or merging two or more files together, among other things. Librarians who have GIS skills may help users prepare files for use in mapping applications, or assist with analysis. If these more specialized types of computing assistance are not available in the library, the user should be referred to a campus statistical consulting center, computing center, lab, or department that provides the appropriate programming assistance or software support. Before doing so, however, the data librarian would ensure that the user has and understands the pertinent documentation about the structure of the data in the file and other characteristics that will be necessary to complete the data file preparation process.

Some knowledge of statistical concepts, research methods, and practices is helpful when advising users on selecting and processing data sets. Without having to be an expert in statistics or social science methodology, the librarian should at least be conversant in the relevant terminology. Understanding, for example, the differences between independent and dependent variables, a cross-sectional study and a longitudinal study, and aggregated data and microdata, will facilitate effective communication with researchers.[17]

Service orientation is what brings all of the skills, knowledge, and experience together to produce successful interactions with data users.[18] Users vary widely in their preparedness for conducting secondary data research. New data users in particular may need to be nurtured as they go through the data discovery process and beyond. For many, secondary data research is a totally new and foreign activity. They have few or no statistical skills, limited computing skills, and very little understanding of what will be involved. Working with these users takes more time, explanation, and demonstration to help them understand what they need, and what they need to do. Patience and enthusiasm are important qualities of service orientation in the provision of data services.

Training for Data Services

How does one learn how to be a data librarian? Most data librarians have had little or no formal training for data services. They tend to take their traditional library skills and knowledge, such as reference and collection development, and apply them to their work with data collections.[19] Some

may have taken courses in library schools that introduce students to numeric data resources and issues related to data reference services.[20] Data librarians also may draw upon their educational background and experience in fields such as political science, economics, geography, history, statistics, and computer science when developing and providing services.[21]

Fortunately, there are a number of avenues that data librarians can take to learn about their profession. Participation in data-related associations offers many benefits. Organizations such as the International Association for Social Science Information Service and Technology (IASSIST), the Association of Public Data Users (APDU), and the Inter-university Consortium for Political and Social Research (ICPSR), have annual or biennial meetings with concurrent sessions, workshops, and many opportunities to network with colleagues from other institutions.[22] An invaluable benefit of these organizations is being able to communicate with distant colleagues via electronic mailing lists. The IASSIST mailing list, for example, is a vibrant online community that allows members to tap into the extensive knowledge base that exists among others who are involved in data services. This knowledge is also shared in a more formal manner through the organizations' publications, *IASSIST Quarterly, APDU Newsletter,* and *ICPSR Bulletin,* respectively.

Another excellent educational opportunity is the Summer Program in Quantitative Methods of Social Research offered by ICPSR every year.[23] In addition to courses on statistical and mathematical topics, workshops are offered on selected data products (e.g., U.S. Census, World Values Surveys) and social research areas (e.g., crime and criminal justice, aging and health). Every other year, the program includes a weeklong workshop called Providing Social Science Data Services, which is geared specifically toward individuals who provide data services in academic institutions.

The training and learning opportunities described here are by no means a comprehensive inventory. Some countries have additional resources and training programs for their data librarians. For instance, the Canadian data community has developed a peer-to-peer training program for personnel who provide assistance to data users in academic libraries.[24] Other scholarly organizations, such as library and social science associations, may also contribute to continued learning. Most data librarians learn on the job, cobbling together skills and knowledge from self-study, attendance at professional meetings, participation in workshops, interacting with colleagues from other institutions, and quite simply, just "doing" data services.[25]

DATA SERVICES AWARENESS AND USE SURVEY

As Lamont so aptly said, "Library technology centers must be grown and developed, none will come fully realized out of a box."[26] Data services, being one type of technology center, is no different. The service will evolve over time as resources such as staffing levels and competencies, funds for data products, and computing capabilities expand and are enhanced. Critical to developing the service is identifying current and potential users of the service, understanding their needs, and making sure that they know the service is available to them. Even though most faculty and students know that the library is the place to go for information, they may not realize that it collects data products, particularly machine-readable data files, or that it has staff who can help with the data discovery process and additional tasks such as data file preparation.[27]

UT libraries have provided data services for many years in some form or fashion. This primarily involved providing access to and assistance with the ICPSR data archive and various government-issued data products. These services were generally provided by the libraries' social science coordinator and government documents personnel.

To provide enhanced access to the vast array of secondary data resources available to researchers, the university libraries hired its first social science data services librarian in June 2000. With this addition to the reference and instructional services team, the libraries had a formal, named, and funded service to support UT's secondary data researchers. A key function of the position was to begin to proactively promote Data Services and data resources and extend outreach in a variety of ways.

After several years of working with data users and learning more about data resources, it was time to formally assess the level of awareness of Data Services on campus and the success and effectiveness of various promotional and outreach activities. Trying to identify, to some extent, the level and type of use of secondary data at UT was also an important goal. The resulting information would be invaluable for planning Data Services, including purchasing or becoming familiar with certain resources, targeting particular groups of students or faculty, and engaging in outreach activities that would give this new service more visibility. The forum for this assessment was the Data Services Awareness and Use Survey.

Scope of the Survey

Twenty academic units were selected based on discipline (primarily the social sciences), and a history of and potential for use of secondary data for research. The colleges and departments selected for the survey are listed in table 1. (Note that nursing and social work are presented at the college level because they do not have departments.)

While all faculty, staff, and students at UT are potential secondary data users, the survey focused only on faculty and graduate students because past experience indicated that they were the primary users. Undergraduate students are increasingly using data too, but to keep the project manageable, they were not included. The faculty roster included only regular, research, and clinical faculty, not adjunct or emeritus. The graduate students included both doctoral and master's levels. Through e-mail, all 1,514 faculty and graduate students in the selected departments were invited to participate in the survey, but some of them never received the questionnaire because messages sent to their e-mail addresses were undeliverable. Ultimately, the survey population was 1,429 (304 faculty and 1,125 graduate students).

Methods

The questionnaire began with two important definitions. Knowing that many survey recipients would not know what Data Services or secondary data were, it was important to define the terms to ensure that respondents understood what they were answering questions about. The data definition included an example of how someone might use secondary data.

The questionnaire, shown in the appendix, had five sections: Data Use, Data Services, Related Services, Additional Comments, and Demographic Information. The first question in the Data Use section was used to determine which of the respondents should be classified as data users. Data Services consisted

of the all-important awareness questions. Related Services focused on our important partner, SCC. Additional Comments solicited feedback about secondary data, Data Services, or the experience of conducting secondary data research at UT. The Demographic Information section recorded the status and college/department of the respondent.

Table 1. Participating Academic Units

College/Department	Number of Recipients	Number of Respondents	Response Rate (%)
Agricultural Sciences and Natural Resources			
Biosystems Engineering and Environmental Science	51	9	17.6
Forestry, Wildlife, and Fisheries	83	16	19.3
Family and Consumer Sciences	7	0	0.0
Unspecified Department	–	3	–
Arts and Sciences			
Anthropology	75	15	20.0
Geography	61	10	16.4
Political Science	80	18	22.5
Psychology	159	26	16.4
Sociology	52	16	30.8
Urban Planning	25	6	24.0
Unspecified Department	–	15	–
Business Administration			
Economics	47	11	23.4
Marketing, Logistics, and Transportation	32	4	12.5
Unspecified Department	–	5	–
Communication and Information			
Advertising and Public Relations	38	13	34.2
Journalism and Electronic Media	56	7	12.5
Unspecified Department	–	7	–
Education, Health, and Human Sciences			
Child and Family Studies	84	19	22.6
Consumer Services Management	27	5	18.5
Educational Administration and Policy Studies	90	21	23.3
Health and Exercise Science	82	13	15.9
Nutrition	33	9	27.3
Unspecified Department	–	23	–
Nursing	165	39	23.6
Social Work	182	46	25.3
Unspecified College/Department		8	
Total	1,429	364	25.5

Though staff were not targeted in the survey, the status was included as a choice because numerous graduate students also have staff appointments. This option was provided mainly to avoid confusion and discontent that might arise from its absence. (All respondents who marked staff also marked another status, which was used to classify the respondent.)

The survey was conducted via the Internet. The questionnaire was converted to an online form using SPSS Data Entry Builder and was hosted by SCC. Using a Web-based survey made distribution, completion of the survey, and data collection simpler, and also ensured the respondents' anonymity.

The cover letter opened with a paragraph introducing the Data Services librarian and briefly explaining her role in Data Services. This introduction was intended to orient the recipients if they were not familiar with secondary data or Data Services. The letter, sent to the faculty and graduate students via e-mail, contained links to the online form as well as to a PDF file with the print questionnaire in case a respondent had a preference for print or had trouble with the online form.

When the survey was sent in October 2003, the plan was to allow two full weeks for participation. Near the deadline, people were still actively responding, so the survey was left up for another two weeks or so. E-mail reminder messages were sent a few days before the original deadline and again about a week later.

All responses submitted via the online form were collected into an SPSS data file on the survey server. Responses from three people submitted on print forms were added to the SPSS file. Twelve cases in the file with blanks for all variables were assumed to be false starts, test runs, or problem surveys, and were deleted from the file before analysis began.

Results

No attempt was made to do statistical comparisons among groups or questions. Descriptive analysis was sufficient for the purpose of getting a sense of who UT's data users are and what their needs are, and gathering information about other aspects of the service. For simplicity, in the discussion that follows, "departments" refers to both the departments and the Colleges of Nursing and Social Work. There were no respondents from the department of family and consumer sciences, so it was not included in any of the results.

Response Rates

The Data Services Awareness and Use Survey was completed by 364 respondents (sixty-seven faculty, 293 graduate students, and four unspecified), which yielded a response rate of 25.5 percent. This rate is based on the number of cover letters that were deliverable. Rates for the faculty and graduate students were 22 percent and 26 percent, respectively.

The departmental response rates ranged from 0 percent to 34 percent (table 1). Departments with at least 20 percent response were, in descending order, advertising and public relations, sociology, nutrition, social work, urban planning, nursing, economics, educational administration and policy studies, child and family studies, political science, and anthropology. Sixty-one respondents did not specify their department, so these rates may reflect lower participation than actually occurred for some departments. This also may be true of the percentages for the questions that follow.

Data Use

Of the 364 respondents, 321 (88 percent) considered themselves data users, having already used secondary data in their research, or anticipating using it in the future. It is not surprising that this percentage was so high because those who were involved in quantitative research were likely to be more interested in participating in the survey than nondata users. With this inherent bias in the respondent group, it is important to note that this result does not mean that 88 percent of the faculty and graduate students in the surveyed units were data users.

Among the faculty, doctoral students, and master's students, the percentages of respondents who said they were data users were similar at 88 percent, 87 percent, and 91 percent, respectively. With the exception of psychology (54 percent), the levels of data use at the department level were relatively high, ranging from 78 percent to 100 percent (see table 2), indicating that all of these departments did have some people using secondary data.

Among the data users, a class project was the most cited reason for using secondary data. This is reasonable because it is the activity that both faculty and students are likely to have been involved in at some point. The results for each project type are shown in table 3. Not surprisingly, there are noticeable differences among the status groups

Table 2. Data Use, ICPSR Awareness, Data Services Awareness, and Data Services Use by Department

College/Department (no. of respondents)	Data Users		ICPSR Awareness		Data Services Awareness		Data Services Use	
	No.	%	No.	%	No.	%	No.	%
Agricultural Sciences and Natural Resources								
Biosystems Engineering and Environmental Science (9)	8	89	1	11	3	33	0	0
Forestry, Wildlife, and Fisheries (16)	13	81	1	6	2	13	1	6
Arts and Sciences								
Anthropology (15)	14	93	3	20	2	13	2	13
Geography (10)	9	90	3	30	8	80	5	50
Political Science (18)	18	100	9	50	7	39	5	28
Psychology (26)	14	54	5	19	8	31	5	19
Sociology (16)	14	88	10	63	11	69	8	50
Urban Planning (6)	5	83	0	0	2	33	1	17
Business Administration								
Economics (11)	10	91	2	18	6	55	2	18
Marketing, Logistics, and Transportation (4)	4	100	0	0	2	50	1	25
Communication and Information								
Advertising and Public Relations (13)	13	100	3	23	7	54	7	54
Journalism and Electronic Media (7)	7	100	2	29	3	43	4	57
Education, Health, and Human Sciences								
Child and Family Studies (19)	17	90	10	53	13	68	10	53
Consumer Services Management (5)	5	100	0	0	1	20	0	0
Educational Administration and Policy Studies (21)	20	95	0	0	8	38	5	24
Health and Exercise Science (13)	12	92	0	0	8	62	1	8
Nutrition (9)	7	78	2	22	6	67	4	44
Nursing (39)	36	92	1	3	13	33	10	26
Social Work (46)	42	91	6	13	15	33	8	17
All Respondents (364)	321	88	69	19	157	43	95	26

*Family and Consumer Sciences is not listed because it had no respondents.

for the different project types. The graduate students obviously have the highest participation for their respective dissertations and theses, and faculty are most involved in faculty research. Note too, however, that graduate students often help faculty with their research and, of course, faculty had to complete a master's thesis or doctoral dissertation at some point before becoming faculty.

Many respondents wrote comments about the topics they studied and types of data resources they used

Table 3. Type of Secondary Data Project by Status

Type of Project	Faculty N=67		Doctoral Students N=135		Master's Students N=158		Overall N=364	
	No.	%	No.	%	No.	%	No.	%
Class project	20	30	65	48	121	77	206	57
Doctoral dissertation	12	18	88	65	10	6	110	30
Master's thesis	7	10	13	10	80	51	100	28
Faculty research	53	79	27	20	5	3	86	24
Institutional or program research	12	18	22	16	28	18	62	17

(305 and 263 respondents, respectively). Research topics included gender issues, lifestyles, marriage and family, aging and gerontology, domestic violence and child abuse, adolescence, juvenile delinquency, health care, public health, drug and alcohol use, obesity, nutrition, mental health, education, school violence, voting and elections, affirmative action, government finances, community and economic development, homelessness, poverty, welfare, environmental issues, forest fire management, soil science, water quality, international trade, transportation statistics, consumer behavior, media effects, and public opinion.

Respondents were interested in data resources from a variety of data producers. Sources and surveys mentioned included the following:

■ U.S. government departments and statistical agencies such as the Census Bureau, Bureau of Labor Statistics, Department of Agriculture, Department of the Interior (National Survey on Recreation and the Environment), Centers for Disease Control and Prevention/National Center for Health Statistics (Behavioral Risk Factor Surveillance Program, National Health and Nutrition Examination Survey, National Hospital Discharge Survey), and Bureau of Justice Statistics (National Crime Victimization Survey)
■ International organizations such as the World Bank, International Monetary Fund, United Nations, and European Commission (Eurobarometer attitude surveys)
■ Various research centers and institutes (General Social Survey, National Longitudinal Study of Adolescent Health, National Survey of Families and Households, National Youth Survey, and Panel Study of Income Dynamics)

Only sixty-nine respondents (19 percent) had heard of the ICPSR data archive. The faculty had a higher level of awareness of ICPSR (31 percent) than did the doctoral students (23 percent) or the master's students (11 percent). Length of time at UT and amount and type of research experience may explain the differences between the three groups. ICPSR has been a major data resource at UT for many years, so the faculty would have had more time to become familiar with it. The levels of awareness of ICPSR within departments ranged from 0 percent to 63 percent (see table 2). Most aware of ICPSR were the faculty and students in sociology, child and family studies, political science, geography, journalism and electronic media, advertising and public relations, nutrition, anthropology, psychology, and economics. These results

are not unexpected, for the most part, because the biggest users of ICPSR at UT in the last few years have been sociology, communication (includes advertising and public relations and journalism and electronic media), political science, child and family studies, social work, education, geography, anthropology, and psychology. Given their actual use of ICPSR, social work and education had conspicuously low levels of awareness on the survey (13 percent and 0 percent, respectively).

Sixty-five respondents (18 percent) had used or thought they would use GIS technology in their research. Given the nature of the subjects, it was not surprising that urban planning (83 percent); geography (80 percent); biosystems engineering and environmental science (44 percent); forestry, wildlife, and fisheries (44 percent); and anthropology (40 percent) were among the most common users of GIS applications. Less obvious GIS users were sociology (38 percent) and consumer services management (60 percent).

Data Services

Of the respondents, 157 (43 percent) were aware that the university libraries had a Data Services librarian. Ten of these respondents said they were not data users. All three status groups had greater awareness of Data Services than of ICPSR: faculty (36 percent versus 31 percent for ICPSR), doctoral students (44 percent versus 23 percent), and master's students (47 percent versus 11 percent). The faculty were more aware of ICPSR than were the students, but the reverse was true for awareness of Data Services. While graduate students are at UT for less time than faculty, they may actually have more chances to learn about library services through library instruction sessions and workshops, doing research in the library, and word of mouth from their classmates, among other things. Within departments, awareness of Data Services ranged from 13 percent to 80 percent (see table 2), and the departments with notable levels of awareness were geography; sociology; child and family studies; nutrition; health and exercise science; economics; advertising and public relations; marketing, logistics, and transportation; and journalism and electronic media.

Among the various ways that people had learned about Data Services, being referred by a professor, instructor, or advisor was by far the most prevalent (see table 4). As would be expected with this response, these referrals were mostly to the students. Other avenues of exposure to Data Services were a class or workshop on secondary data resources, library guides or handouts, a col-

Table 4. Exposure to Data Services by Status

Avenue of Exposure	Faculty N=67		Doctoral Students N=135		Master's Students N=158		Overall N=364	
	No.	%	No.	%	No.	%	No.	%
Professor/instructor/advisor	1	2	33	24	48	30	82	23
Data class or workshop	8	12	19	14	27	17	54	15
Library guides or handouts	10	15	12	9	21	13	43	12
Colleague or classmate	9	13	20	15	12	8	41	11
University Libraries Web site	7	10	16	12	12	8	35	10
Another librarian	6	9	9	7	14	9	29	8
Data Services Web site	2	3	8	6	3	2	13	4

league or classmate, the university libraries or Data Services Web sites, and a librarian other than the Data Services librarian. A few people also learned about Data Services from the libraries' social science newsletter, annual data and GIS users' forum, ICPSR Summer Program promotion, and the university's SCC.

Ninety-five respondents (26 percent) said they had used Data Services in some way since the service was launched in 2000. The level of use was about the same for the faculty and graduate students (25–28 percent). The use of Data Services by department ranged from 0 percent to 57 percent (see table 2). Departments reporting the most use were journalism and electronic media, advertising and public relations, child and family studies, sociology, geography, and nutrition. It was interesting to see nutrition reporting above-average use of Data Services, because at the time of the survey, their faculty and graduate students had made little contact with the Data Services librarian. Social work, on the other hand, was on the lower end of both awareness and use of Data Services, which was inconsistent with being one of the top three user groups since 2000.

Of those who said they had used Data Services, 17 percent had contact with the Data Services librarian in person or by phone or e-mail, 11 percent had attended a class or workshop conducted by the Data Services librarian, 11 percent had visited the Data Services Web site, and 6 percent had used Data Services guides or handouts. Little difference was seen between the faculty and students with respect to how they used the service.

Respondents who consulted with Data Services most often did so for help with locating data and documentation, identifying appropriate studies or data resources, and accessing, downloading, or purchasing data files or resources. As shown in

table 5, the graduate students tended to consult Data Services more than the faculty for these aspects of their research, perhaps because they were less familiar with the available resources and the process used to find and access them.

Related Services

Of the respondents, 218 (60 percent) were aware of SCC. The faculty had the highest level of awareness (85 percent), as might be expected given their longer terms of service at the university. Doctoral students and master's students had levels of 72 percent and 41 percent, respectively. Sixteen percent of the respondents had contacted SCC for assistance with a secondary data project at some time; of these, three had been referred to SCC by Data Services.

Comments

Eighty-four respondents (23 percent) provided feedback in the additional comments section at the end of the survey. Several themes emerged. Some comments were included in more than one thematic category if they touched on more than one issue, and a few were not categorized at all if the intent was not clear. The most prominent message conveyed by the comments was the importance of promotion and awareness of Data Services.

■ Awareness of Data Services (twenty-nine respondents): These comments were mostly about being or becoming aware of Data Services. Many indicated that they became aware of Data Services when they got the survey and were glad to know about it. This illustrated that there is interest in using secondary data resources at UT, and that researchers may be

Table 5. Data Services Assistance Requested by Status

Type of Assistance	Faculty N=67		Doctoral Students N=135		Master's Students N=158		Overall N=364	
	No.	%	No.	%	No.	%	No.	%
Locate data and documentation	13	19	30	22	40	25	83	23
Identify studies or resources	9	13	26	19	29	18	64	18
Access/download/purchase data	7	10	21	16	20	13	48	13
Review or utilize documentation	5	8	10	7	8	5	23	6
Prepare SAS or SPSS data files	1	2	7	5	4	3	12	3
Data analysis	2	3	3	2	5	3	10	3

more likely to seek help from the university libraries now that they know about Data Services. It also showed that the survey itself had a promotional effect.

■ Promotion of Data Services (fourteen respondents): This is related to awareness, but is separated here to make an important point. These respondents, in some cases rather emphatically, indicated that we need to do more to get the word out about Data Services, particularly to graduate students. Several respondents were frustrated that they did not know about the service when they were beginning their thesis or dissertation research and felt that it might have saved them some time and effort if they had known about it earlier.

■ Kudos for Data Services (sixteen respondents): These comments were from respondents who had a positive experience with Data Services and were satisfied with the result of the interaction.

■ General Library Issues (seventeen respondents): This group includes anything library-related that was not specifically directed toward Data Services. These comments ran the gamut from positive comments about services and resources, to dismay over canceling UT's subscription to a popular full-text database, and problems with online or remote access to electronic resources.

■ SCC (eleven respondents): There were a variety of comments, mostly positive, about past or future use of SCC.

DISCUSSION

With slightly more than a quarter of the recipients responding to the survey, a definitive profile of UT's secondary data users cannot be established

from its results. However, the feedback from 364 faculty and graduate students, the majority of whom were data users, provides some sense of who the most likely user groups are and where UT libraries should put its energies and funds for promotion, outreach, instruction, and resource allocation. Information collected in-house on Data Services consultations and ICPSR usage helps to give some context when interpreting the survey results.

In the years leading up to the survey, more than three-quarters of the interactions between the Data Services librarian and data users were with sociology, social work, child and family studies, advertising and public relations, journalism and electronic media, and political science. Those departments, plus education, geography, anthropology, and psychology, were the heaviest users of ICPSR. For some groups, survey results indicating high use of services and resources were in line with this past experience. For other groups, levels of use or awareness indicated on the survey were not always consistent with each other or with actual Data Services or ICPSR use. In some cases, departments that had little or no contact with Data Services reported high or moderate levels of use or awareness (e.g., nutrition). The reverse was also true, with some departments having numerous past Data Services interactions but relatively low levels of use and awareness indicated on the survey (e.g., social work). These inconsistencies may have much to do with the nonresponse rate of about 75 percent, the fairly small group sizes for some departments, and the obvious differences between actual-use data and self-reported survey responses.

Given actual Data Services contacts and ICPSR use, and relatively high awareness and use percentages from the survey, the departments that unquestionably are the primary user groups at

this time are sociology, advertising and public relations, journalism and electronic media, social work, child and family studies, and political science. The survey indicates that the remaining departments also have secondary data users. In some of these disciplines, secondary data research is a well-established practice (e.g., economics, education, nutrition, and urban planning). These departments' lack of contact with Data Services and use of ICPSR may reflect that their faculty and graduate students have been using resources other than ICPSR and have not needed assistance finding data. Since there is potential for secondary data use in all of the departments involved in this survey, it would be wise to continue to reach out to them in order to increase their chances of knowing about Data Services if and when the time comes that they need assistance.

It would have been helpful to get some direction from the survey in terms of topics studied and types of data used to proactively focus energies and funds on appropriate data resources. The write-in comments on data and topics varied so widely, though, that it was not really possible to do that. While the variety may preclude specific planning activities, it certainly illustrates the rich array of research interests that abound at UT.

Awareness was a key issue in this survey. One hundred fifty-seven respondents were aware of Data Services, but that was slightly less than half of the 321 respondents who said that they did or would use data in their research. Clearly, the libraries would like more than half of UT's data users to know they can get assistance with secondary data resources from the Data Services librarian. Interestingly, in contrast to their knowledge of ICPSR, which was relatively low, the graduate students were more aware of Data Services than were the faculty. Perhaps this is because graduate students have more opportunities to be exposed to data-related promotion and outreach activities through graduate student orientations, library instruction sessions, data workshops, and consultations with the Data Services librarian and other subject librarians.

ICPSR is the world's largest archive of social science research data and is one of UT's oldest and most widely promoted secondary data resources. Consequently, the rather low level of awareness among the respondents (19 percent) was both disheartening and unexpected. It is obvious that the university libraries need to become more effective at promoting ICPSR, particularly to the graduate students who could benefit greatly from its resources.

The results of several survey questions and the comments provided by the respondents resoundingly indicated that there is room to improve when it comes to promotion and outreach activities. Graduate students need to be hit "early and often" with information about data services and sources because they are not at UT for very long, and they probably have the most to gain by using secondary data (it can be a tremendous time saver). It is also important to pay attention to the faculty, to help them as data users as well as conduits of information for their graduate students. They play an important advisory role and have referred their students to Data Services, so keeping them well-informed about data services and resources is a necessity.

In addition to interactions with the Data Services librarian in classes or workshops, respondents learned about Data Services from colleagues, classmates, and other librarians, indicating that word of mouth and library referrals help raise awareness. Faculty and graduate students apparently do look at and use the various library handouts, guides, and Web sites related to Data Services and data resources. It is difficult to assess the effectiveness of these kinds of noncontact promotion and outreach, so the survey has affirmed that they are worth maintaining and should be improved and updated whenever possible.

As a result of the survey, the UT libraries implemented several new or expanded outreach initiatives. Targeting graduate students, introductory and announcement e-mail messages have been sent through a network of departmental secretaries, more data-related workshops have been offered, and subject librarians have been enlisted to help promote Data Services in graduate student orientations and library instruction sessions. These initiatives have already paid off, generating several contacts from students in response to the e-mail messages or after attendance in workshops. A series of e-mail messages sent to faculty have also generated inquiries or requests for assistance.

To specifically address increasing awareness of ICPSR, an ICPSR workshop has been provided at least once every term instead of just once a year. In several semesters, an additional session was added to meet the needs of faculty and students who were not able to attend the workshop at its scheduled time. Additionally, the new e-mail networks have been used to promote the ICPSR Summer Program in Quantitative Methods directly to all graduate students and faculty in selected departments, instead of relying on department heads to disseminate the information. In response, more

inquiries about the program have been received than in previous years.

FUTURE ASSESSMENT

This survey provides an important baseline for future evaluation of awareness and use of Data Services. It has provided a snapshot of data use at UT and awareness of ICPSR and Data Services, among other things. In order to measure the effectiveness of promotional and outreach activities, and track levels of Data Services awareness and secondary data use, it will be beneficial to occasionally repeat the survey, perhaps every four years or so. The effort is too intensive to do it more often than that, and there is no desire to bother the faculty and longer-term graduate students by surveying them too often.

CONCLUSION

Providing data services is a natural fit for the academic library's core mission of helping users find information in a variety of formats. Data services is challenging in its breadth of subject matter; the required skills in computing and methodology; the vast array of possible sources of data; and the in-depth assistance that is often necessary to identify, locate, acquire, and prepare an appropriate data resource. With those challenges come many rewards, including

- learning about topics and research methods in a variety of subject areas;
- supporting a methodological approach to research that is important in many academic disciplines; and
- helping new data users sort out terminology and concepts, acquire new computing skills, and become comfortable dealing with quantitative research.

Because of the unique characteristics of data services, it is necessary for a library to assess needs and determine the likely audience for these services. At UT libraries, user needs assessment is an important ongoing activity. Consultations, ICPSR use, and workshop attendance will continue to be tracked on a regular basis to learn about the use of data services and resources. Additionally, interactions with data users will demonstrate what they need in terms of both data resources and guidance with the process of doing secondary data research. An occasional survey will help the libraries mea-

sure against new benchmarks and gather information that cannot routinely be tracked through logs and personal interactions. The university libraries will continue to look for innovative and effective ways to reach out to current and potential data users to keep them informed about the resources and services that are available to facilitate their quantitative research. Training and learning opportunities will be pursued to expand the knowledge and skills necessary to provide high-quality services in a timely manner. UT wants to make the most of this specialized service called Data Services, and all of these things will be important factors in making that happen.

References and Notes

1. Judith Rowe, "Electronic Research Data for Scholars and Students," *Serials Review* 22 (Summer 1996): 5–6, 8; Michael N. Cook, John J. Hernandez, and Shawn Nicholson, *Numeric Data Products and Services,* SPEC Kit 263 (Washington, D.C.: ARL, Office of Leadership and Management Services, 2001), 9.
2. Cook, Hernandez, and Nicholson, *Numeric Data Products and Services,* 9.
3. Diane Geraci and Linda Langschied, "Mainstreaming Data: Challenges to Libraries," *Information Technology and Libraries* 11 (Mar. 1992): 10–11.
4. Ann S. Gray and Diane Geraci, "Complex Reference Services: Data Files for Social Research," in *Social Science Reference Services,* ed. Pam M. Baxter (New York: Haworth, 1995), 136.
5. Statistical Abstract of the United States: www.census.gov/prod/www/abs/statab.html; American FactFinder: http://factfinder.census.gov/; FedStats: www.fedstats.gov/ (accessed Jan. 16, 2007).
6. Geraci and Langschied, "Mainstreaming Data," 11.
7. Inter-university Consortium for Political and Social Research: www.icpsr.umich.edu (accessed Jan. 16, 2007); Roper Center for Public Opinion Research: www.ropercenter.uconn.edu (accessed Jan. 16, 2007); U.S. Census Bureau: www.census.gov (accessed Jan. 16, 2007); the U.S. statistical agencies may be found through FedStats or an Internet search of the agency name.
8. Cook, Hernandez, Nicholson, *Numeric Data Products and Services,* 18.
9. Terrence B. Bennett and Shawn W. Nicholson, "Interactions Between the Academic Business Library and Research Data Services," *portal: Libraries and the Academy* 4, no. 1 (2004): 108–09.
10. Geraci and Langschied, "Mainstreaming Data," 15.
11. Gray and Geraci, "Complex Reference Services," 137.
12. Ibid., 146–47.
13. Jim Jacobs, "Providing Data Services for Machine-Readable Information in an Academic Library: Some Levels of Service," *Public-Access Computer Systems Review* 2, no. 1 (1991): 145, 154–57.
14. Gray and Geraci, "Complex Reference Services," 139.
15. Jacobs, "Providing Data Services," 155–57.
16. Diane Geraci, Chuck Humphrey, and Jim Jacobs, "Data Basics: An Introductory Text" (course material, Providing Social Science Data Services, Inter-univer-

sity Consortium for Political and Social Research, 2000), 4.2–3.

17. Ibid., 4.1–2.

18. Ibid., 4.3.

19. Judith S. Rowe and Carolyn L. Geda, "Training the Professional Data Librarian," *Drexel Library Quarterly* 13 (Jan. 1977): 101.

20. Edmond Mignon, "Numeric Data Bases in the Professional Librarianship Curriculum: Implications for Behavioral and Social Sciences Librarians," *Behavioral and Social Sciences Librarian* 1 (Spring 1980): 183; Laine G. M. Ruus, "Training of Data Services Professionals: Past, Present, and Future," *Library Trends* 30 (Winter 1982): 458.

21. Cook, Hernandez, and Nicholson, *Numeric Data Products and Services,* 19.

22. IASSIST: www.iassistdata.org/; APDU: www.apdu.org/; ICPSR: www.icpsr.umich.edu (accessed Jan. 16, 2007).

23. ICPSR Summer Program in Quantitative Methods of Social Research: www.icpsr.umich.edu/sumprog/ (accessed Jan. 16, 2007).

24. Ernie S. Boyko et al., "Lifting Ourselves by Our Bootstraps: Developing a National Peer-to-Peer Training Program for Data Librarians in Canada," (paper presented at the World Library and Information Congress: 69th IFLA General Conference and Council, Held in Berlin, 1–9 Aug. 2003), www.ifla.org/IV/ifla69/papers/136e-Boyko_Hamilton.pdf (accessed May 11, 2005).

25. Geraci and Langschied, "Mainstreaming Data," 15.

26. Melissa Lamont, "Critical Human Factors in Emerging Library Technology Centers," *Library Hi Tech* 17, no. 4 (1999): 393.

27. Ibid., 391.

APPENDIX

Data Services Awareness and Use Survey

Data Services: A service of the University Libraries that facilitates access to secondary data for analysis by researchers at the University of Tennessee, Knoxville.

Secondary Data: Existing data used for purposes other than those for which the data were originally collected. Data may be presented in print, electronic, or Web-based formats. Using secondary data may eliminate the need to collect one's own data. For example, a sociologist interested in environmental racism might examine the correlation between sociodemographic characteristics of an area (obtained from decennial census data files) and locations of hazardous waste sites.

Directions: Please check the response that most appropriately represents your experience or knowledge. Note that some questions allow more than one response. There is space for written comments on page 4 of the questionnaire.

Data Use

1. Have you ever used, or do you anticipate using in the future, secondary data from electronic, Web-based, or print resources for research at UT? Mark all that apply.
   ___ Have used secondary data
   ___ Anticipate using secondary data in future research

   If yes to either, for what type of project have you used, or do you anticipate using, secondary data? Mark all that apply.
   ___ Class project
   ___ Master's thesis
   ___ Doctoral dissertation
   ___ Faculty research
   ___ Institutional or program research
   ___ Other research (please specify)

   _____

2. If you have used, or anticipate using, secondary data in your research:
   a. What topic(s) did or will you study? (e.g., domestic violence, consumer spending)
   b. What type(s) of data or specific data resources did or might you use? (e.g., public opinion polls, National Crime Victimization Survey, Current Population Survey)

3. Have you ever heard of, or are you familiar with, the Inter-university Consortium for Political and Social Research (ICPSR) data archive?
   ___ Yes ___ No ___ Not Sure

   If yes, have you ever used data or codebooks from the ICPSR data archive?
   ___ Yes ___ No ___ Not Sure

4. Did or will your secondary data research involve the use of GIS (Geographic Information System) technology for mapping and analysis?
   ___ Yes ___ No ___ Not Sure
   ___ Not Applicable

Data Services

5. Before receiving this survey, were you aware that the University Libraries has a Data Services librarian to help faculty, staff, and students identify, locate, and acquire data for secondary research?
   ___ Yes ___ No ___ Not Sure

   If yes, how did you learn or hear about Data Services? Mark all that apply.

   ___ Data Services Web site
   ___ University Libraries Web site
   ___ Attended class or workshop on secondary data resources and services
   ___ Professor/instructor/advisor
   ___ Colleague or classmate
   ___ Another librarian
   ___ Library guides or handouts
   ___ Information for Social Scientists newsletter
   ___ Data/GIS Users Forum
   ___ ICPSR Summer Program
   ___ Statistical Consulting Center
   ___ Other (please specify) _____
   _____

6. During the past three years, have you used Data Services at Hodges Library by interacting with the Data Services librarian, visiting the Data Services Web site, or using Data Services handouts or guides?
   ___ Yes ___ No ___ Not Sure

If yes, in what way? Mark all that apply.

___ Contact with the Data Services librarian in person or via phone or e-mail
___ Attendance at a class or workshop conducted by the Data Services librarian
___ Data Services Web site
___ Data Services guides or handouts
___ Other (please specify) _____
_____

7. If applicable, for which aspects of your secondary data research did you consult the Data Services librarian or make use of Data Services information resources? Mark all that apply.

___ Identifying appropriate studies or data resources
___ Locating the data and documentation
___ Reviewing/utilizing the documentation
___ Accessing/downloading/purchasing data files or resources
___ Preparing SAS or SPSS data files
___ Data analysis
___ Other (please specify) _____
_____

Related Services

8. Were you aware that UT has a Statistical Consulting Center that provides faculty, staff, and students with assistance in planning their research, converting raw data sets into usable files, using statistical computing applications, and analyzing data, among other things?

___ Yes   ___ No   ___ Not Sure

9. Have you ever contacted the Statistical Consulting Center for assistance with any phase of a secondary data research project?

___ Yes   ___ No   ___ Not Sure

If yes, were you referred to the Statistical Consulting Center by Data Services?

___ Yes   ___ No   ___ Not Sure

Additional Comments

Please write any comments you have about secondary data or data services. Feel free to offer observations about your experience conducting secondary data research at UT.

Demographic Information

Please identify your status and field of study. Note that this is for the purpose of comparing groups, not for identifying individuals in any way.

Status:
___ Faculty
___ Staff
___ Doctoral Student
___ Master's Student

College and Department:
Please mark the appropriate college and write your department following the college name.

___ Agricultural Sciences and Natural Resources
_____

___ Arts and Sciences
_____

___ Business Administration
_____

___ Communication and Information
_____

___ Education, Health, and Human Sciences
_____

___ Nursing
_____

___ Social Work
_____

Please return your completed survey to **Data Services Survey, 145 Hodges Library**, by **October 31, 2003**. If you have any questions about this survey or Data Services, please contact the Data Services Librarian at 974-0011 or dataserv@utk.edu.

Thank you for your participation. Your feedback will help the University Libraries provide better service to secondary data researchers at the University of Tennessee.