Steven M. Frank
Dept. of Surveying, New Mexico State University

Michael F. Goodchild
NCGIA / Dept. of Geography, University of California, Santa Barbara

Harlan J. Onsrud
NCGIA / Dept. of Surveying Engineering, University of Maine

Jeffrey K. Pinto
Dept. of Business, Pennsylvania State University

A Survey on User Requirements for Framework GIS Data

• **Abstract**: Common sets of framework geospatial data -- data usable across many applications -- have been proposed as a method to promote GIS data sharing. While there is much anecdotal information about what the characteristics of such data should be, studies and surveys on GIS data needs have been limited primarily to hardware/software issues or confined to a particular state or region. The National Center for Geographic Information and Analysis (NCGIA) conducted a nation-wide mail questionnaire survey to gather information concerning the technical requirements for geospatial data. The questionnaire targeted existing users of GIS or GIS products (i.e., maps, reports, etc. generated from GIS). These users were asked for responses regarding their data needs for six possible framework data sets. 1360 questionnaires were mailed out. 595 usable questionnaires were completed and returned. The returned information was analyzed across sectors of government, private industry, and academia by geographic region and by professional area of application, showing the technical preferences for framework data sets across each sector profile. A summary of those responses is given in this paper.

Background

There is a general agreement among the users of geographic and land information systems that common sets of geographic data on which users could build to their own particular data needs would promote greater data sharing among the various players in the geographic information system (GIS) community (NRC 1994). Such common data sets -- often referred to as "core" or "framework" data sets -- could be collected by designated government agencies or by participating members of the private sector and added to the public domain, via the National Spatial Data Infrastructure (NSDI), for the use of any and all interested parties.

Within the US geographic information community there has been protracted debate in the past two years over the value of various data sets, and their importance in the National Spatial Data Infrastructure. A range of criteria for prioritization have been suggested, and strong arguments have been presented for the importance of the various types of data, in addition to the framework function described earlier. This paper is not concerned with resolving these arguments, or even with establishing the relative value of various types of data to the user community. Rather, our objective is to help to clarify the debate by identifying the specific aspects of data sets that are perceived to be of most value to specific communities of users.

To help discover which framework data sets should be given priority for incorporation into the NSDI, the Federal Geographic Data Committee (FGDC) has funded a mail questionnaire survey that was conducted by the National Center for Geographic Information and Analysis (NCGIA) with oversight by a peer focus group. The results of this survey are intended to provide a guideline for the selection and/or prioritization of possible framework data sets into the NSDI. A brief summary of results are given in this paper. Full results are available from NCGIA.

Goal

The goal of this survey was to identify the technical criteria that may be used to identify and prioritize the framework data sets for the National Spatial Data Infrastructure. These framework data sets may be identified by defining the technical specifications required by the users of these data sets, including content, tasks for which the data are used, format, geocoding scheme, positional accuracy, vertical accuracy (if needed), updating interval, needs for historical data, and the sources for data currently being used. A peer focus group reviewed the questionnaire, the survey sampling strategy, and the survey results.

It should be recognized that geographic information users may be required to use certain data not because of their technical merit, but because of political or organizational expedience. Organizations may have agreements to share or purchase geographic data from predetermined sources, thus forcing the use of certain types and qualities of data. This questionnaire does not investigate possible political or organizational factors.

It is expected that no single set of criteria will be appropriate for all GIS users, but that identifiable sets of criteria will emerge that can form the basis of selecting current digital geographic data sets that best meet such criteria as NSDI framework data sets and for improving such data sets to meet the full needs specified by the different subsets of criteria.

Survey Data Collection

Data collection was by a structured questionnaire sent and returned by mail. The questionnaire allowed data collection from a large group of people in a relatively short time at a relatively inexpensive cost. The questionnaire was designed to capture background information useful to correlate and aggregate data into meaningful units of comparison and to capture information on the technical details of digital geographic data used or needed by the respondents.

The unit of analysis for the questionnaire was an individual using geographic information. The term "geographic information" as used here is intended to include information about the spatial locations of objects on or near the Earth's surface. Collection of questionnaire data at the individual level allows aggregation of data by many different means, such as discipline, organization, experience, etc.

There was great difficulty in defining the population of GIS users to be surveyed. We argue that there exists no effective definition of such a population. Instead, we have allowed accessibility and expedience to define the sampling frame for this study, rather than any coherent notion of a population. A list of potential respondents was obtained from sources including: GIS in Business '93 Conference Proceedings (GIS World, 1993), Proceedings of AM/FM International (AM/FM, 1993, 1992, 1991), the Proceedings of GIS/LIS (ACSM, et al., 1993, 1992, 1991), URISA Membership Directory (URISA, 1993). the State Geographic Information Activities Compendium (Warnecke, 1992) and the International GIS Sourcebook (Parker, 1991). Each respondent was identified by state, and by a best estimate of occupation and category. Systematic samplings were then made within each state, occupation, and category to obtain a sample. The final sample consisted of 1360 potential respondents.

We used a "balanced" rather than a "stratified" sampling, reflecting the impossibility of defining a population, and have not attempted to weight responses. We believe it is easier to define the sampled population within each category, occupation, and state, but that severe differences between these sub-populations exist, particularly across occupations. All results from this study are stated as counts, proportions, or averages for the sample and its sub-samples, and while comparisons are made between proportions, no inferential statements are made regarding any larger population.

Although the initial attempt was to collect and analyze responses across each of the fifty states and the District of Columbia, small and possibly heavily biased samples from some states made this impractical. Instead, the states were aggregated into geographic regions corresponding to the US Census use of national regions (Census, 1993).

Sampling Method

Persons currently using geographic information systems or products generated from GIS were the target of the questionnaire. We focused on people using end products which they knew were produced by GIS as well as "hands on" GIS technicians and managers. We assumed that the breadth of future users of GIS is already reflected in the current users of GIS, although the proportion of future users in different categories may vary drastically (e.g., there may be a far greater number of business users in proportion to government users in the future, but these business users will be reflected in the current geographic/land information user community).

Sampling was stratified by three methods. First, it was stratified across the government, private, and academic sectors. Second, it was stratified across occupational sectors. Third, it was stratified by geographic area.

Priority was given to stratification across occupational sectors. Occupational sectors were stratified using US Bureau of Labor Statistics bulletin information statistics giving the numbers of people employed in different sectors of the US work place (BLS, 1992). Thirty application areas were aggregated from this document and listed in the questionnaire (see Table 1). These occupational lists of potential respondents were then pared (if needed) to the best balance possible across government, private, and academic sectors. Finally, the sample population was balanced as best as possible across geographic regions on a state by state basis.

Table 1. Thirty occupational application areas for GIS.

Administration	Engineering (Civil)	Oceanography / Marine
Agriculture	Engineering (Other)	Property / Real estate
Architecture	Forestry	Public relations
Banking / Finance	Geology / Geophysics	Shipping
Biology	Insurance	Social science / Social services
Communications	Law (except police srvcs)	Surveying
Construction	Legislative	Transportation
Economics	Marketing / Advertising	Urban and regional planning
Education	Medical / Health	Utility operations
Emergency srvcs (police, fire, etc.)	Meteorology / Air quality	Wildlife

Each copy of the questionnaire was numbered and cross referenced to the name and address to which it was sent. Follow up letters were sent to those not responding within the first three weeks. Another copy of the questionnaire was mailed to those who did not respond to the second mailing after two weeks.

Questionnaire Content

The questionnaire was divided into two portions. The first portion gathered personal information from each respondent, including discipline, GIS applications area(s), employment level, organization background, and GIS experience. The occupational areas portion of the questionnaire was designed to be correlated to the U. S. Bureau of Labor Statistics bulletin information statistics giving the numbers of people employed in different sectors of the U. S. work place

The members of the sample population were assumed to be well versed in the technical aspects of GIS. Since many users might use GIS in areas other than their occupation or profession, respondents were then asked to check the frequency with which they used GIS for any or all of the thirty occupational areas listed. A space was provided to fill in additional occupational areas if needed by the respondent. Respondents were next asked about their level of employment, the category of their organization in government, private business, or academia, their experience with geographic data in any form, their experience with digital geographic data, and the amount of work time they spent using digital geographic data.

The second portion of the questionnaire asked the respondents to detail their information needs for six classes of geographic information: 1) transportation feature data, 2) water feature data, 3) other well-defined cultural feature data, 4) elevation data, 5) land parcel data, and 6) boundary data. A glossary of terms used to define the technical features of framework data was included in the questionnaire. Use of the glossary was encouraged to resolve any ambiguities about the meanings of the terms described. Respondents were asked to reply only to questions in categories of data which they used or might use in the future. They were asked for information regarding their data needs for that data category, including content, tasks for which the data are used, format, geocoding scheme, positional accuracy, vertical accuracy (if needed), updating interval, needs for historical data, and the sources for data currently being used.

Questionnaire Results

The survey responses were tabulated and analyzed using Statistical Program for Social Scientists (SPSS) software and custom developed software. Preliminary results reflect the data needs and uses of the aggregated respondents. Results have also been compiled for various user occupations, for various levels of government and private business, and for various geographic sections of the United States for certain survey questions. Data content, data tasks, data formats, data geocoding schemes, data positional accuracy, and data vertical accuracy were broken down into tables showing the response frequencies for each geographic region, for each application area, and for each employment sector.

The application area responses were further broken down with separate tables showing the response frequencies across all application area users (those giving a response between 2 and 5 for using a certain application area) and across heavy application area users (those responding with a 4 or 5 for a certain application area use). Responses in which persons replied that they did not use GIS or products generated from GIS in the performance of their jobs were not processed for technical contents, since such users were asked to not complete the questionnaire.

The first wave of questionnaires were returned by mail during the first week of June, 1994, and questionnaires were collected until mid-August, 1994. Thirty five questionnaires were returned because of incorrect address information. A total of 595 usable responses (44.9%) were received and keyed into a raw data. The responses were cross-checked and validated against the range of response values for each question by use of custom software. The numbers of usable responses were tabulated by geographic region, by the employment sector of the respondent, and by the application areas in which he or she is using GIS or products generated from GIS.

Respondent's Background

Respondent's background was obtained through the questions in Section 1 of the questionnaire. Respondents were asked questions regarding the area of application for GIS, their level within their organization, the type of organization they were with, their total experience using geospatial data, their experience using digital geospatial data, and the amount of work time that they used such data on their job.

Of the 30 application areas listed, an average of 10.6 application area uses per respondent was noted. An average of 4.3 application areas were answered with a value of 4 or 5, indicating frequent or heavy use, while an average of 6.3 application areas were marked as being used less frequently (value of 2 or 3). Five respondents did not select any of the 30 listed application areas, instead listing their individual primary uses of GIS or products derived from GIS as demographic analysis, emergency management, cartographic mapping, planning, and groundwater / drinking water supply. Responses are shown in Table 2.

Table 2. Respondent's background information (% of total useable responses)

Employment level

Clerical	Technical	Prof.	Mid-mgmt	Up-mgmt
0	8.7	46.1	29.2	15.8

Employment Sector

Fed. Gov't	State Gov't	Local Gov't	Private	Academic
6.2	35.8	36.3	17	4.7

Total number of years using geospatial data in any form

< 1 yr	1-5 yrs	6-10 yrs	11-20 yrs	> 20 yrs
2.2	19.5	23.9	27.9	26.6

Number of years using GIS

< 1 yr	1-5 yrs	6-10 yrs	11-20 yrs	> 20 yrs
6.9	45.2	30.1	14.1	3.5

Percent of worktime spent using GIS or GIS products

0-20%	21-40%	41-60%	61-80%	81-100%
27.1	17.1	16.1	17.3	22

Transportation Feature Data Responses

Transportation feature data was used or needed by 91.9% of the respondents. More than half (58.0%) responded that they needed better transportation data while about one-fourth (25.2%) responded that they did not need better transportation data. Responses by heavy users -- those responding with a frequency value of 4 or 5 -- for transportation feature data are given in Table 3.

Table 3. Transportation feature data heavy users (4 or 5) responses (% of all usable responses)

Fwys	57.7
Hwys	66.8
Local	60.8
Rural	56.5
Streets	61.6
Trails	26.2
RRs	44.2
F. rail	17
Water	31.7
Airports	27.4
Pipes	28.4
Bus	17.9

Data contents

Accident	10.6
Address	40
Env. mon.	27.6
Fac. des.	22.7
Fac. main.	18.2
Inventory	30.4
Res. mgt.	32.4
Site anal.	39.2
Veh. rout.	19.7

Data tasks

Vector	73.9
Photo	37
Imagery	19
Raster	15.1

Data format

42.7	Lat. & long.
33.7	UTM
54.3	State plane
49.8	St. Addr.
14.7	Mile posts

Data geocode

0.01 M	11.6
0.1 M	15.5
1 M	29.4
10 M	38.5
20 M	30.1
50 M	21.1
100 M	13.1
250 M	8.6
500 M	7.9
1000 M	6.9

Data positional accuracy

0.01 M	9.1
0.1 M	14.3
1 M	18.3
10 M	15
20 M	7.9
50 M	4.9
100 M	3.7
250 M	2.8
500 M	2.7

Data vertical accuracy

No maint.	Ad hoc	Any chng.	Major chng	Weekly	Monthly	Yearly	Other
9.6	14.5	22.7	22.4	1	3.4	12.8	3.7

Data updating

Federal	State	Local	Priv-repack	In-house	Other
57.1	25.9	28.4	10.8	23.5	1.8

Data sources

Few, if any, notable differences were found between occupations, regions, or employment sector for transportation feature data contents. Differences in transportation data tasks seemed primarily across employment sectors. For example, environmental monitoring was cited much more frequently by those working for the federal or state government that by those working in local government, the private sector, or academia. Data format choices were predominantly for vector data across all occupations, regions, and employment sectors. Differences in geocode preferences were found along both occupational lines and employment sector lines. Those in natural science occupations (biology, forestry, geology, etc.) had far less use for street address geocodes than those in other occupations. Those in federal or state government had greater need for data in latitude and longitude coordinates.

Many respondents seemed a little confused about their use or need for positional and vertical accuracy. Many respondents did not answer the questions on accuracy at all. The responses were nearly even for and against the need for elevations for transportation feature data. Fewer than 20% wanted vertical accuracy of 1 meter, the most popular heavy use response for transportation data. There were some differences in accuracy needs among occupational lines. Those in engineering oriented occupations required greater accuracy than those in most other occupations.

Transportation feature data updating was needed primarily as changes were made. The major source for transportation feature data was the federal government.

Water Feature Data Responses

More than 85% of the respondents replied that they used of needed data about water features with more that 48% replying that they need better data about water features (27.1% said that they did not need better water feature data). Water feature data responses are given in Table 4.

As expected, those respondents living in a region near coastal waters did have a higher need for inshore and offshore ocean data than other regions. Also, as expected, those listing oceanography / marine as a GIS application area had significantly higher needs for ocean data. However, these responses seem to indicate a need for inshore, rather than offshore, ocean data. No other deviations in responses seemed significant.

Environmental monitoring appeared to be more popular among users of natural science applications. Agriculture, biology, forestry, geology, meteorology / air quality, oceanography / marine, and wildlife application users cited this use much more frequently than other users. Users in these applications areas also responded more frequently to water management as a needed task. No other significant variations were noted.

The other data category response differences were the same as noted for transportation feature data.

employment
data tasks
onitoring was
ient that by
mat choices
oyment
onal lines and
stry,
occupations.
longitude

nal and
cy at all. The
ortation
popular
ccuracy
equired

nade. The

it water
eatures
data

a higher
d, those
r needs for
her than

al science
,
e frequently
ently to
ed.

portation

Table 4. Water feature data heavy users (4 or 5) responses (% of all usable responses)

Inshore	21.4
Offshore	8.1
Rivers	64.2
Lakes	60
Const.	42.7
Wetlands	46.6
W. sheds	44.5

Data contents

Env. mon.	29.3
Fac. des.	21.9
Fac. main.	14.8
Inventory	32.8
Site anal.	39.7
Veh. rout.	10.4
Wat. mgt.	21.7

Data tasks

Vector	64
Photo	36
Imagery	17.6
Raster	10.9

Data format

Lat. & long.	34.5
UTM	27.9
State plane	49.4
Mile posts	10.3
EPA ID	13.5

Data geocode

0.01 M	12.4
0.1 M	12.9
1 M	29.6
10 M	36.5
20 M	26.7
50 M	18.4
100 M	10
250 M	5.4
500 M	4.9
1000 M	4.5

Data positional accuracy

0.01 M	7.2
0.1 M	14.9
1 M	22.9
10 M	18.8
20 M	10.3
50 M	7.7
100 M	3.3
250 M	2.9
500 M	2.9

Data vertical accuracy

No maint.	Ad hoc	Any chng.	Major chng	Weekly	Monthly	Yearly	Other
14.3	14.5	13.9	22.5	0.5	2	10.8	3.2

Data updating

Federal	State	Local	Priv - repa	In-house	Other
47.2	19.2	19	5.2	17.2	1.7

Data sources

Other Well-defined Cultural Feature Data Responses

Although 83.3% of the respondents replied that they used or needed data about other well-defined cultural features, fewer than 40% responded with a heavy use or need for any one particular feature content item. However, 48.6% responded that they needed better cultural feature data while only 17.3% responded that they did not need better data. Responses for all of the cultural feature questions are shown in Table 5.

Cultural feature content differences were mostly along occupational lines. For example, more in the medical profession needed hospital locations while more in education needed school locations. Differences in responses across occupational, regional, and employment sectors were primarily as found in transportation feature and water feature data.

Table 5. Other well-defined cultural feature data heavy users (4 or 5) responses (% of all usable responses)

Comm.	24.2
Power	32.8
Gas/oil	32.9
Water	35.5
Waste	35.8
Hydrants	24
Schools	31.2
Hospitals	27.6
Libraries	20.6

Data contents

Address	33.8
Env. mon.	21
Fac. des.	21.2
Fac. main.	16.5
Inventory	27.7
Res. all.	18.8
Res. mgt.	24.7
Site anal.	35.6
Veh. rout.	13.6

Data tasks

0.01 M	10.6
0.1 M	15.6
1 M	31.1
10 M	33.6
20 M	23
50 M	17.3
100 M	9.6
250 M	5.6
500 M	4.9
1000 M	4.5

Data positional accuracy

Vector	61.7
Photo	34.3
Imagery	14.9
Raster	9.3

Data format

Lat. & long.	35.2
UTM	28.4
State plane	51.5
USPLSS	23.7
Mile posts	9.4
St. addr.	35.4

Data geocode

0.01 M	7.8
0.1 M	14.2
1 M	20.2
10 M	13
20 M	7.2
50 M	4.3
100 M	2.1
250 M	1.7
500 M	1.5

Data vertical accuracy

No maint.	Ad hoc	Any chng.	Major chng	Weekly.	Monthly	Yearly	Other
10.6	13.3	18.7	20.5	1.2	2.4	8.4	2.5

Data updating

Federal	State	Local	Priv-repack	In-house	Other
32.3	12.8	21.7	4.2	20.5	1.7

Data sources

Elevation Data Responses

Elevation data was used or needed by 70.6% of the respondents with 41.3% asking for better elevation data (11.3% were satisfied with current elevation data). Elevation data responses are shown in Table 6.

Digital contours were the most popular content for elevation data with 39.7% claiming a heavy use or need for digital contours. Federal government employees were more frequent users of digital elevation models (DEMs) than the average user, but no other significant deviations are noted.

Table 6. Elevation data heavy users (4 or 5) responses (% of all usable responses)

DEMs	27.1
DTMs	22
Contours	38.7
Bathym.	6.1

Data contents

Env. mon.	22.5
Fac. des.	22.3
Fac. main.	10.1
Res. all.	16.1
Res. mgt.	22.9
Site anal.	41.2
Veh. rout.	6

Data tasks

Grids	25.8
TINs	26.4
Contours	41.7

Data format

Lat. & long.	27.8
UTM	24.4
State plane	39.2

Data geocode

0.01 M	10.3
0.1 M	16.5
1 M	28.1
10 M	23.9
20 M	16.8
50 M	11.1
100 M	6.5
250 M	3.5
500 M	2.6
1000 M	2.7

Data positional accuracy

0.01 M	11.1
0.1 M	18.5
1 M	29.3
10 M	22.2
20 M	12.8
50 M	7.3
100 M	3.9
250 M	2.6
500 M	2.7

Data vertical accuracy

No maint.	Ad hoc	Any chng.	Major chng	Weekly	Monthly	Yearly	Other
18.3	10.4	9.1	17.1	0.2	1	5.7	2.2

Data updating

Federal	State	Local	Priv-repack	In-house	Other
28.1	7.6	10.6	1.2	11.8	0.3

Data sources

Land Parcel Data Responses

Land parcel data was used or needed by 80.9% of the respondents. More than half (51.3%) replied that they needed better land parcel data while 12.6% were satisfied with their current data. Table 7 gives the responses for land parcel data.

As expected, land parcel are more popular among those in property / real estate applications. They are also more frequently cited among administrators, architects, construction users, civil engineers, law applications, and surveyors -- and especially among those classified as heavy users in those applications. Predominant use of land parcel data is in local government, followed by private industry.

Regionally, the US Public Lands Survey System was not popular in areas where it is not used, such as New England and the Middle Atlantic states. Positional accuracy needs for land parcel data was higher than for other possible framework data sets. The respondents were not asked about vertical accuracy for land parcel data, although one respondent did cite a need for vertical accuracy for "flood analysis." The aggregate trend in higher positional accuracy for land parcel data seems to be reflected by the larger numbers of engineering / surveying application users than earth science application users who need land parcel data. Also unlike other possible framework data sets, the primary source for land parcel data was from local, not federal, government.

Table 7. Land parcel data heavy users (4 or 5) responses (% of all usable responses)

Land owner	60.6
Consolidat	50.1
Pub. ease.	48.7
Priv. ease.	40
Pub. ROW	52.6

Data contents

Address	40.3
Env. mon.	23.6
Fac. des.	25.2
Fac. main.	18.3
Inventory	36
Res. all.	21
Res. mgt.	26.2
Site anal.	40.3
Veh. rout.	12.8

Data tasks

0.01 M	13.6
0.1 M	21.8
1 M	38.6
10 M	26.9
20 M	16.8
50 M	10.1
100 M	5
250 M	3.9
500 M	2.8
1000 M	3.1

Data positional accuracy

Vector	62.2
Photo	35.3
Imagery	14.6
Raster	10.4

Data format

Lat. & long.	30.1
UTM	24.7
State plane	51.7
USPLSS	27.2
Parcel ID	45.9
St. Addr.	41.9

Data geocode

No maint.	Ad hoc	Any chng.	Major chng	Weekly	Monthly	Yearly	Other
8.2	8.4	25.1	11.1	2.5	5.4	8.9	3.5

Data updating

Federal	State	Local	Priv-repack	In-house	Other
9.9	13.3	34.6	3.4	16.5	1.3

Data sources

Boundary Data Responses

Boundary data was used or needed by 94.3% of respondents. While 20.0% said that they did not need better boundary data, 62.7% replied that they did need better data. Heavy user responses for boundary data are given in Table 8.

Table 8. Boundary data heavy users (4 or 5) responses (% of all usable responses)

Census	47.1
PO Zip+4	20.3
PO Zip	26.9
Voting	23.1
Zoning	37.8
Othr admin	52.8
City/Town	73.1
County	71.1
State	49.8

Data contents

Address	40.4
Demogr.	41.9
Env. mon.	27.2
Fac. des.	19.4
Fac. main.	14.5
Inventory	32.6
Res. all.	27.2
Res. mgt.	30.6
Site anal.	39.5
Veh. rout.	16.3

Data tasks

0.01 M	11
0.1 M	16.8
1 M	37.6
10 M	39.5
20 M	28.1
50 M	16.6
100 M	11.1
250 M	6.9
500 M	5.7
1000 M	6.1

Data accuracy

Vector	74.1
Photo	31
Imagery	13.6
Raster	10.4

Data format

Lat. & long.	42.3
UTM	32.1
State plane	55.7
St. Addr.	42.2

Data geocode

No maint.	Ad hoc	Any chng.	Major chng	Weekly	Monthly	Yearly	Other
11.4	12.4	29.2	21.3	1	2.5	10.1	2.2

Data updating

Federal	State	Local	Priv-repack	In-house	Other
47.6	19.3	27.1	9.9	19.2	2.4

Data sources

City and county boundaries were the most frequently used or needed boundary content items among heavy users. Census tract and ZIP code boundaries are very popular among banking / finance, economics, insurance, legislative, marketing / advertising, medical / health, and social science / social services applications. Demographic analysis was most frequently boundary data task noted among users in architectural banking / finance, economics, education, insurance, legislative, marketing / advertising, public relations, shipping, and social science / social service applications.

Conclusions

Many current GIS users appear to be using GIS across more than one disciplinary application area. An average number of almost ten applications were cited per user. This could mean that many of the problems being solved through GIS are interdisciplinary in nature. It could also signify that common uses of many data sets are being applied by centrally located GIS and that the respondent at that centrally located site performed tasks for a variety of other departments. It might also reflect user responses to organizational uses, rather than individual uses, of GIS. Since more than one application area was cited by many users, possible correlation in the answers stratified by occupational area use will need to be further investigated.

Many respondents had difficulty in addressing some of these technical issues, especially questions about positional and vertical accuracy. Several possibilities arise that might explain this phenomenon. First, users might not have documentation (metadata) that can readily give them answers to such technical questions. Second, users may depend on secondary positional and vertical accuracy measures, such as map scale and contour interval, to provide accuracy measures. Third, users might depend on the judgment of others to provide them with data of sufficient accuracy for their individual needs. Fourth, users might be using whatever data sets they find available without worrying about the possible technical qualities of that data. Although much has been done to promote concerns about these issues -- such as the FGDC Metadata Content Standard and the recent International Symposium on the Spatial Accuracy of Natural Resource Data Bases held in Williamsburg, Virginia, in May 1994 -- perhaps more education and emphasis on understanding the implications of these issues may be in order.

Differences in user requirements appear to be primarily application oriented. Some regional differences were noted, primarily concerning proximity to oceans and the use of the US Public Lands Survey System. Differences among employment sectors seemed to be primarily in areas related to the amount or scale of data being used. This is particularly apparent in the geocoding scheme responses, where federal government employees, users of typically large spatial areas of analysis, seem to prefer a scheme that will deal effectively with the representation problems of those areas -- namely latitude and longitude coordinates. Street addresses, primarily a geocode useful for relatively small spatial areas of analysis, are not very popular among federal employees. However, street addresses, like state plane coordinates, are much more popular among users who typically deal with relatively small spatial areas of analysis, such as local government and private sector employees.

Future analyses planned for this data include a look at the possible clustering of GIS applications. As noted in the section relating the positional accuracy responses, it appears that certain GIS applications appear to have similar data technical needs with other applications. The question arises whether GIS users could naturally fall into a limited number of user categories.

We have attempted to find the current digital geospatial data technical requirements of users with this study. We do not attempt to predict future needs. While some technical requirements, such as data format and data contents, may remain stable for the foreseeable future, other technical requirements, such as positional and vertical accuracy, may increase in importance with time. This study attempts to define technical guidelines that may be useful for selecting current digital geospatial data sets that may be useful for the greatest numbers of users which may then be prioritized for inclusion in the National Spatial Data Infrastructure.

This study reflects perhaps the first large-scale effort to sample the US population of GIS users in a systematic fashion, regarding their current and future uses of geospatial data and related technologies. The analyses reported here are a small fraction of the results which might be extracted from the data, and we anticipate that others may have specific questions they wish to ask of the database. NCGIA will provide copies of the database on request to any bona fide group or agency wishing to conduct further analyses.

References

ACSM, ASPRS, AAG, URISA, and AM/FM (1993, 1992, 1991). Proceedings of GIS/LIS.

AM/FM (1993, 1992, 1991). Proceedings of AM/FM International. Aurora, CO: AM/FM International.

BLS (1992). Outlook: 1990-2005: Occupational Employment. Washington, DC : The United States Bureau of Labor Statistics (May, 1992).

GIS World (1993). GIS in Business '93 Conference Proceedings.

NRC (1994). Toward a Coordinated Spatial Data Infrastructure for the Nation. Mapping Science Committee, National Research Council. Washington, DC: National Academy Press.

Parker, H. Dennison, ed. (1991). 1991-1992 International GIS SOURCEBOOK.

URISA (1993). URISA Membership Directory 1993-1994. Washington, DC: Urban & Regional Information Systems Association.

Warnecke, Lisa (1992). State Geographic Information Activities Compendium.