



N E W S L E T T E R

Data Quests: Geospatial Data Clearinghouse Links Distant Resources

At one time or another every user of geographic information has searched for data. Numerous techniques are employed, including browsing indexes from Federal agencies, phoning friends and coworkers, checking library references, or querying electronic networks. When a data set is located, the struggle begins to obtain a copy of it in a reasonable time. All geographic data consumers long for a more efficient means to locate and access data.

The Federal Geographic Data Committee (FGDC), concerned that access to spatial data is inadequate for today's information requirements, is creating a network-based clearinghouse for geographic data. The intent is not to centralize all geographic data in one location, but to provide a link to all the distributed sites where data are produced or maintained. The Geospatial Data Clearinghouse uses the rapidly growing Internet network to tie together computer-based information sources. Internet links nodes across the country and the world.

One of the first requirements for creating such a clearinghouse is to use software that supports queries over a network. The FGDC is testing the public domain software tool known as Wide Area Information Servers (WAIS) for conducting searches over the Internet network. WAIS, in its public domain implementation, enables users to discover and browse both data sets and the metadata files that provide

descriptive information about the data sets. If the metadata indicate a particular data set might be appropriate, steps can be taken to access the data. In some cases, if the data are stored online by the production agency, they may be directly accessed over the network. In other cases, the metadata will provide a point of contact for retrieving the data manually.

WAIS supports both general and specific text-based queries across the network. WAIS also uses "relevance feedback" to fine-tune responses to queries to assist a user in identifying the most pertinent data sets. The FGDC is supporting additional enhancements to the software to enable searches for data on the basis of geographic coverage, as well as content.

Complete and consistent data docu-

mentation is another consideration in developing the Geospatial Data Clearinghouse (see related article, p.3). The FGDC is incorporating the recently revised draft of the metadata standard into a test of the prototype clearinghouse. The test of the prototype clearinghouse involves over 50 individuals representing Federal, State, regional, and local governments, as well as members of the academic and private sectors. In its initial phase, the Geospatial Data Clearinghouse is designed to involve those individuals who are using UNIX-based platforms with access to Internet. These individuals create metadata files adhering to the metadata standard to document a small number of their active GIS data sets. They then generate WAIS indexes on these metadata files to make them accessible over the network. In the future, additional platforms other than UNIX will be included.

The aim of the Geospatial Data Clearinghouse is to enable access to all types of geographic data sets, regardless of whether they originate from Federal, State, regional, or local government, or through partnerships between governmental agencies, academia, and the private sector. This is a key part of the development of the national spatial data infrastructure.

For more information about the clearinghouse or to participate in the prototype test, please contact the FGDC Secretariat at the numbers listed in this newsletter.

In This Issue

FGDC Contact Information	2
FGDC Publications Request Form	8
Data Quests: Geospatial Data Clearinghouse Links Distant Resources	1
Traveling a Difficult Road: Talks Begin on Data Content Standards	2
Testing Begins for the Revised Draft Spatial Metadata Standard	3
The National Spatial Data Infrastructure: A Technical Perspective	4
The National Spatial Data Infrastructure: A Policy Perspective	5
Workshop Explores GPS Base Station Control Issues	6
Data Sharing Opportunities: A USGS Example	6
Spotlight on FGDC Subcommittees	7

Traveling a Difficult Road: Talks Begin on Data Content Standards

FGDC Contact Information

The *FGDC Newsletter* describes the activities of the spatial data community and the development of a national spatial data infrastructure. Subscriptions are free of charge. Subscription renewals as of issue no. 1, spring 1993, have been incorporated in the newsletter mailing list. Correspondence or contributions may be directed to the FGDC Secretariat as indicated below:

In writing:

FGDC Secretariat
590 National Center
Reston, VA 22092

By telephone:

703-648-4533

By facsimile:

703-648-5755

By email (via Internet):

gdc@usgs.gov

By anonymous FTP via Internet (for FGDC information and documents):

Address: [isdres.er.usgs.gov](ftp://isdres.er.usgs.gov)
(130.11.48.2)

User name: anonymous

After connecting: `cd gdc`

(File README.DOC has current information on the files in each subdirectory.)

By modem (for FGDC information and documents):

703-648-4200 (up to 2400 Baud)

703-648-7800 (9600 Baud)

At screen prompt (call, display or modify) type: `c isdres`

User name: `gdc`

(File README.DOC has current information on the files in each subdirectory.)

To move to a subdirectory: `set def`

[subdirectory name]

To return to home directory: `[-]`

When is a road a road? What are the criteria that define it? Can you build a transportation network using roads from separate data bases? Questions such as these face all data producers as they struggle to standardize the content of geographic data bases that are increasingly

shared with other data users. The 12 thematic data subcommittees of the Federal Geographic Data Committee (FGDC) are attempting to find common approaches to classifying and defining geographic features for all themes of data: base cartographic, bathymetric, boundaries, cadastral, cultural and demographic, geologic, ground transportation, soils, vegetation, water and wetlands.

To accelerate the content standards process, the FGDC held a Technical Forum on Spatial Features, March 9-10, 1993, at the U.S. Geological Survey in Reston, Virginia. This meeting brought together nearly 100 representatives from Federal agencies that produce or use geographic data sets. Small group sessions held at the Forum discussed key issues:

- The goal of the standards setting process is not to create one universal standard, but to determine the degree of standardization required for each data theme to support data sharing.
- The standards developed for separate data themes overlap with one another and must be integrated.
- An overarching mechanism to manage the process of developing and maintaining data content standards is needed, involving all levels of government and the private and academic sectors.
- Individuals from within thematic data communities must lead the standards setting process, or consensus will never be reached.
- Federal agencies must direct funding and personnel into standards setting activities to accomplish this task.

At the forum, participants spent considerable time discussing the relationship of the newly adopted Spatial Data Transfer Standard (SDTS) to data content standards. The primary intent of the SDTS is to provide a vendor-neutral format for data. The SDTS requires that the data producer document the data dictionary used to define the geographic features and attributes appearing in the transferred data sets. Part 2 of the SDTS offers a preliminary list of generic definitions for topographic and hydrographic features. This list is expected to evolve into a Spatial Features Register (SFR) to provide a library of generic terms for all data themes that agencies can use for the transfer of their geographic data sets. The SDTS task force at the U.S. Geological Survey is working with the thematic data subcommittees of the FGDC to ensure that the SFR includes the geographic features needed for each data community.

At the conclusion of the forum, the participants acknowledged that the growing requirement for data sharing necessitates increased communication on assumptions and practices in collecting geographic data. At the GIS/LIS '93 Conference, November 2-4, 1993, in Minneapolis, several of the thematic data subcommittees are planning to conduct a joint session on content standards development.

To order a summary report of the Technical Forum on Spatial Features (available in late summer 1993), please use the FGDC Publications Request Form on the back page of this newsletter.

Testing Begins for the Revised Draft Spatial Metadata Standard

The revised draft of the "Content Standard for Spatial Metadata," recently released by the Standards Working Group of the Federal Geographic Data Committee (FGDC), provides a comprehensive specification for "data about data" to document the quality and other characteristics of geographic data sets. Participants in the prototype test of the Geospatial Data Clearinghouse (see related article, p. 1) are using the metadata standard to document the data sets included in the proposed clearinghouse.

The metadata standard addresses four activities that are common to all data users: (1) determining what data exist, (2) determining a "found" data set's fitness for a particular use, (3) providing information to access a data set, and (4) providing information needed during transfer so a data set can be processed and used.

The standard allows data producers to document data sets that locate information directly using coordinates and

addresses, as well as indirectly through features that have a known location, such as a county. Data producers may also document ancillary data sets that can be linked to a spatial data set, and they can document individual data sets or a series of data sets. The standard also permits documentation of data that are in work or planned.

From November 1992 to April 1993, the FGDC conducted a public review of the standard. Comments were received from over 50 individuals and organizations from Federal, State, and local government organizations; the private sector; and academia. Included in these comments were the perspectives of software vendors, data producers, users, and libraries.

Most of the original data elements in the metadata standard were retained as a result of the public review, although the information contained within them was reorganized in several instances. Data elements were added in some cases to better describe raster and remotely sensed data, and the documentation about map projections was reevaluated. The working group also applied techniques, such as entity-relationship diagrams, to refine and validate the data elements.

The working group recommends that data producers use existing standards whenever possible to implement the metadata standard. For transfer purposes, the working group is determining how to implement the metadata standard using the data elements available in the the Spatial Data Transfer Standard

(SDTS) (published as Federal Information Processing Standard 173). For cataloging purposes, the working group proposes the use of the U.S. Machine Readable Catalog, which is widely used by libraries.

As noted above, the Geospatial Data Clearinghouse prototype test constitutes the next phase of evaluation for the metadata standard. Individuals and organizations who are participating in the clearinghouse test will document their experiences in using the metadata standard and its use in enabling them to successfully search for other data sets.

Individuals and organizations not participating in the clearinghouse test are also encouraged to evaluate the metadata standard by using it to document their own data sets. Results and comments should be provided by mail to the FGDC Secretariat, U.S. Geological Survey, 590 National Center, Reston, Virginia 22092; or by Internet to gdc@usgs.gov. The envelope or Internet message should be marked to indicate, "Metadata Test Results."

The revised draft of the metadata standard is being sent to those individuals who commented on the draft or requested a copy. Other individuals may obtain a copy by using the FGDC Publications Request Form on the back page of this newsletter.

The National Spatial Data Infrastructure: A Technical Perspective

From February 18-21, 1993, 33 individuals with expertise in geographic data, GIS technology, and electronic networks were brought together in Charleston, South Carolina, to discuss the national spatial data infrastructure (NSDI). The NSDI, as envisioned in the spring 1993 *FGDC Newsletter*, consists of individuals, organizations, technology, and spatial data, integrated to facilitate development and dissemination of spatial data and the use of geographic information technologies.

Participants for this meeting were chosen from different sectors (government, private and academic), representing both national and international interests, to provide a diversity of opinions to stimulate debate on different aspects of the NSDI. In preparation for the meeting, participants were asked to provide their thoughts on a number of key questions: (1) What is the NSDI? (2) What are the components of the NSDI? (3) Who needs the NSDI? (4) Who is responsible for developing the NSDI? For maintaining the NSDI? (5) Who pays for the NSDI? (6) What are the technical constraints to a national information infrastructure? (7) What are the institutional constraints? (8) What incentives are needed to facilitate development and use of the NSDI? (9) What role do standards play in the NSDI? (10) Who is responsible for creating standards? (11) Is there one data model or one representation of geographic reality to which we can all agree? If so, what is this model and who creates it? (12) If a variety of agencies and organizations are responsible for developing and maintain-

ing certain themes of geographic data (i.e., soils by the Soil Conservation Service, base cartography by the U.S. Geological Survey), how are these best integrated to optimize the use of GIS tools? (13) If there is not one model, and multiple representations of geography exist, how are they linked and integrated? (14) Is the trend of redundant copies of data on every user's system or server desirable, inevitable, or of concern? (15) If the NSDI is based on electronic communication, what are the networking issues that must be addressed? Who should address these? (16) What are the divisions of responsibility between public and private sectors in development, maintenance, and use of the NSDI?

The format for interaction at this meeting involved an initial presentation of issues and questions to all participants for general discussion. After general discussion, smaller groups were formed for more indepth dialogue. Results of the small group discussions were then presented to the collective group. Succeeding questions were sometimes formed on the basis of this discussion. There were five key issues addressed by each of the smaller groups: (1) What are the factors for success of the NSDI? (2) What are the impediments to the NSDI? (3) What is the mission statement of the NSDI? It's top 10 functions? Strategic directions or milestones? (4) What are core data? (5) What are the next critical steps for building the NSDI?

Participants were in accord on the nature of the NSDI and that some guidance and structure can contribute to a reduced cost and a more robust infrastructure. But there was no agreement on the steps that need to be taken or who needs to act. Unresolved issues include:

- understanding what is meant by a core data set, whether it is a national data set, the most requested data set, or the data set considered to be the most

fundamental to the efficiency and effectiveness of all other geographically based activities;

- determining the role of the Federal Government in producing core data;

- identifying the organizations that might maintain these data (such as local governments or utilities);

- resolving how centralized or decentralized the process of data collection and maintenance should be.

Synopses of the discussions will be provided in a report to be published in late summer 1993. Please use the FGDC Publications Request Form on the back page of this newsletter to order a copy.



The National Spatial Data Infrastructure: A Policy Perspective

In early May 1993, nearly 750 individuals gathered at the National Geo-Data Policy Forum in Tysons Corner, Virginia, to discuss policies affecting the development and use of the national spatial data infrastructure (NSDI). The main objective of the forum, which was accomplished in a series of plenary panel discussions, was to expose participants to the debates on existing and needed policies that affect the evolution of the NSDI.

Approximately 45 speakers shared their viewpoints on issues ranging from the development of standards and responsibilities for producing common national digital data sets, to questions of liability, privacy,

and cost recovery for digital spatial data. Vigorous debate occurred about public access to data, including roles, responsibilities, and fees. What roles do the private sector and libraries play in data dissemination? If data dissemination is not part of a public agency's mission, can fees for access be set higher? Does ultimate use of the data by entities other than the collecting agency affect how, to whom, and at what cost data are disseminated? In what form are data an asset of taxpayers?

The forum helped expose five major issues relating to spatial data that compound the challenges of both collecting and using data in a coordinated manner.

1. Numerous agencies with different missions, management objectives, and requirements for spatial accuracy share common geography.

2. Relatively few standards exist for producing geographic data, and the ones that do were often developed to apply to paper or nondigital spatial data; standards development is usually a long-term process.

3. Significant expense is involved in collecting spatial data accurately, including both positional and descriptive accuracy.

4. Detailed spatial data can be "generalized" into less detailed data to meet uses that are not dependent on precise spatial accuracy or specific attributes. But many questions surround this process, such as how to interpret the results of generalization and the concern that more detail means higher cost.

5. The recent spread of geographic information technologies has increased the demand for current spatial data, while giving more people the ability to create digital spatial data.

Understanding and dealing with these issues is necessary for the NSDI to facilitate the most productive use of GIS tools. However, finding common solutions to these issues will involve more coordination than has previously occurred. The debates at the forum exposed some of the key concerns of many data providers and users. A follow-on discussion to the forum is being held on July 24-25, 1993, just prior to the URISA '93 Conference in Atlanta, Georgia. To order a summary report of the forum (available late summer 1993), please use the FGDC Publications Request Form on the back page of this newsletter.



Workshop Explores GPS Base Station Control Issues

A workshop on Federal GPS reference (base) station activities, held May 18-20, 1993, in Rockville, Maryland, explored how best to integrate the GPS reference station activities of Federal agencies to avoid duplication, improve efficiency, and reduce cost. The workshop, hosted by the Fixed Reference Station Working Group of the Federal Geodetic Control Subcommittee (FGCS) of FGDC, was attended by 25 participants representing 9 Federal agencies, and concentrated on 4 areas: (1) identification of user data requirements, (2) translation of requirements into reference station design, (3) current status of reference station implementation, and (4) identification of current barriers to reference station data exchange.

Four major ways in which reference station requirements varied from user to user were identified:

1. Use of pseudorange (meter-level accuracy) vs. carrier phase observables (centimeter-level accuracy).
2. Need for realtime vs. after-the-fact data access.
3. Fixed station vs. moving platform applications.
4. Level of reliability required.

The participants agreed that specifications were needed for several classes of reference stations at varying levels of complexity. Factors that increase complexity are: data links for realtime data transmission, dual receiver backup for high reliability, and high data recording rates (0.5 to 1 second sampling vs. 5 to 30 second sampling) for moving platform positioning.

Although agencies share reference station data at the local, State, and regional level, there are technical and procedural barriers to data sharing. A major technical barrier is the mixing of data from different types of receivers, even when using the RINEX (receiver

independent and exchange) data format. Procedural barriers include the closed nature of some agency data bases and differing agency policies on cost recovery. The agency representatives agreed to work through the Fixed Reference Station Working Group of the FGCS to resolve the technical problems, and through the FGCS and the FGDC to develop interagency agreements to eliminate procedural barriers.

Participants concluded that the problem of integrating GPS reference station activities is not because multiple agencies are involved. Indeed, many

agencies indicated they would be happy to use reliable GPS reference station data from another agency. The problem lies in the lack of assigned responsibility, exacerbated by inadequate funding, to put in place a multiuse GPS reference station network to meet multiple agency needs. The FGCS plans a follow-on meeting of the Fixed Reference Station Working Group to continue discussions on how to efficiently integrate Federal GPS reference station activities. This working group will also update its inventory of permanently operating base stations in July 1993.

Data Sharing Opportunities: A USGS Example

In the last issue, readers were offered an opportunity to document existing or planned data partnerships. The offer still stands to use this newsletter as a forum to communicate interests and requirements. Additionally, the FGDC Secretariat has established an Internet address — gdc@usgs.gov — where coordination efforts of individuals or organizations can be electronically registered. Please provide the following information: name of contact, name of organization, geographic area of interest, address, telephone and FAX numbers, and Internet address (if applicable). This information will be posted under the Internet Anonymous FTP account maintained by the FGDC (listed under the contact information for the FGDC Secretariat.)

As an example of the type of project that might be registered, the following is a brief summary of an initiative recently launched by the U.S. Geological Survey (USGS) to stimulate innovative partnerships:

The USGS is seeking to identify and to investigate sources of accurate base category map data from State and local governments, public utilities, the private sector, and others. The USGS is offering funds, source materials, public domain software, training, quality control or other technical assistance to take advantage of existing or planned map data. Program announcements are published in *Commerce Business Daily*. These announcements are not an inducement for speculative ventures by public or private map data producers. Instead, the USGS is willing to defray the additional cost to be incurred by the applicant in preparing the data to national standards. All data prepared under these agreements will be distributed by the USGS in a nonproprietary format as part of the public domain. Additionally, proposals to exchange accurate base category spatial data for USGS data of comparable value, or to donate such data, will be considered as part of this program.

The first program announcement sought digital line graph data only. The second announcement, due later this summer, will include digital elevation model data and digital orthophoto products.

Base Cartographic

The Subcommittee on Base Cartographic Data is testing the feasibility of collecting base cartographic data through cooperative partnerships that pool the contributions of all organizations conducting cartographic activities in selected regions. Local agencies have the most knowledge of changes in many of the features represented on national series base maps and often maintain mapping systems to support their programs. Partnerships with national mapping programs can enable local agencies to upgrade their products through access to Federal base materials, aerial photographs, etc., and can improve national base map products through access to current feature data maintained by local partners.

The first mapping partnerships demonstration is being coordinated by the Tennessee Valley Authority (TVA) in cooperation with the North Carolina Department of Transportation (NCDOT), North Carolina Center for Geographic Information and Analysis, U.S. Forest Service, and U.S. Geological Survey (USGS). Using digital transportation feature data compiled by the Forest Service and photographs and digital orthophotographs provided

by the USGS and the TVA, the NCDOT will populate and maintain a current digital transportation data base for the State. This local data will be available to support the revision of national base series products by the TVA and the USGS. A pilot project is underway on the Asheville, North Carolina, 7.5-minute quadrangle to establish technical procedures. Following completion of the pilot project this summer a production demonstration on the 14 quadrangles covering Graham County, North Carolina, will be used to evaluate the benefits and cost to each of the partner agencies.

The subcommittee will also hold a standards meeting on July 26 at the URISA Conference in Atlanta, 1:30-4:30 p.m., Inforum, 250 Williams Street, Conference Room 12.

Cadastral

The Subcommittee on Cadastral Data has prepared draft cadastral data transfer standards, which are being reviewed by subcommittee members at this time. The review is scheduled for completion July 30, 1993, with transmittal to the National Institute of Standards and Technology for review by the end of September 1993. The subcommittee is planning the development of cadastral data content standards late in 1993.

The subcommittee will hold a standards meeting November 3 at the GIS/LIS '93 Conference in Minneapolis, 3:30-5:00 p.m., at the Minneapolis Convention Center. (For the specific room location for this meeting, please contact the FGDC Secretariat at the numbers listed in this newsletter.)

Cultural & Demographic

A meeting will be held on July 27 at the URISA '93 Conference in Atlanta (10:00 a.m.-2:00 p.m., Inforum, 250 Williams Street, Theater) to discuss how to achieve a nationwide, attribute-rich (including individual addresses), positionally accurate street centerline file that meets the needs of Federal, State, and local agencies; the business community; and other users of cultural and demographic data. Agreements on common requirements and partnerships also will be discussed. Those in attendance will include members of the Subcommittee on Cultural and Demographic Data; participants from the Transportation, Cadastral, and Base Cartographic Subcommittees; members of the National Academy of Sciences Mapping Science Committee; representatives of State and local governments; emergency service providers; and the private sector. A subsequent newsletter will describe the ideas presented at this meeting.

FGDC Publications Request Form

Please use this form to order publications from the FGDC Secretariat via FAX at **703-648-5755**, or write to:
FGDC Publications, c/o U.S. Geological Survey, 590 National Center, Reston, VA 22092.

Thank you for your interest!

Send me the publications checked below:

Newsletter

- FGDC Newsletter, no. 1, spring 1993.

Standards

- Content Standard for Spatial Metadata (revised draft), July 1993.
- Information about the Spatial Data Transfer Standard, FIPS 173.

Annual Reports

- First Annual Report to the Director, OMB, by the Federal Geographic Data Committee, December 1991.
- Second Annual Report to the Director, OMB, by the Federal Geographic Data Committee, 1992 (available late summer 1993).

Technical Reports (limited copies available)

- Information Exchange Forum on Spatial Metadata, Report of Meeting, June 1992.
- Application of Satellite Data for Mapping and Monitoring Wetlands, September 1992.

FGDC Forums

- Report of the Technical Meeting on the National Spatial Data Infrastructure, February 1993 (available late summer 1993).
- Report of the Technical Forum on Spatial Features, March 1993 (available late summer 1993).
- Report of the National Geo-Data Policy Forum, May 1993 (available late summer 1993).

Notice: Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

U.S. Department of the Interior
U.S. Geological Survey
590 National Center
Reston, VA 22092

Third Class Mail
Postage & Fees
PAID
Geological Survey
Permit No. G-23

Official Business
Penalty for Private use \$300