Laying the Foundation for a Geographic Data Clearinghouse

Minnesota Governor's Council
on Geographic Information

Data Clearinghouse Committee

September 1997

MN PLANNING

The Governor's Council on Geographic Information was created in 1991 by Governor Arne H. Carlson to provide leadership in the development, management and use of geographic information and related technology. With assistance from Minnesota Planning, the council provides policy advice to all levels of government and makes recommendations regarding investments, management practices, institutional arrangements, education, stewardship and standards.

Minnesota Planning is charged with developing a long-range plan for the state, stimulating public participation in Minnesota's future and coordinating activities with state agencies, the Legislature and other units of government.

Upon request, *Laying the Foundation for a Geographic Data Clearinghouse* will be made available in alternate format, such as Braille, large print or audio tape. For TTY, contact Minnesota Relay Service at (800) 627-3529 and ask for Minnesota Planning.

MN PLANNING

658 Cedar St.
St. Paul, MN 55155
(612) 296-3985
www.mnplan.state.mn.us

September 1997

For additional information or copies of *Laying the Foundation for a Geographic Data Clearinghouse*, contact the council staff coordinator at (612) 296-1208 or via e-mail at gc@mnplan.state.mn.us. An electronic copy of this report can be found on the Governor's Council on Geographic Information's World Wide Web home page: www.lmic.state.mn.us/gc/gc.htm.

Cover Map — United States Geological Survey Topographic Quadrangle Digital Raster Graphic Map
Cover Inset — Data copied from a soils data attribute table

Data Clearinghouse Committee Members

Christopher Cialek, Land Management Information Center at Minnesota Planning
Tom Glancy, Minnesota Department of Transportation
Carl Hardzinski, Bureau of Indian Affairs (chair)
Roger Hirschman, Natural Resources Conservation Service
Susanne Maeder, Land Management Information Center at Minnesota Planning
Tanya Mayer, Metropolitan Council
Ken Saffert, City of Mankato
Wendy Treadwell, University of Minnesota
Ron Wencl, U.S. Geological Survey

Summary

Geography provides a commonly accepted framework for interpreting a variety of data. The use of geographic data in Minnesota continues to expand. More computerized data sets are being created all the time, and new geographic information system applications are being implemented at all levels of government, in education and in the private sector.

This expansion of geographic data producers and users spurred the Governor's Council on Geographic Information to form a Data Clearinghouse Committee to review data access issues in Minnesota and investigate the feasibility of establishing a clearinghouse for the state's computerized geographic information.

Clearinghouse components

The Data Clearinghouse Committee developed a conceptual framework for a statewide geographic information clearinghouse that would be a voluntary, cooperative effort of government agencies at all levels, public and private organizations, and geographic data users. The overriding goal of the clearinghouse would be to support the use of geographic data in planning, development and research by providing an access system for identifying existing data and determining its applicability. The clearinghouse would allow broad public access to search tools for acquiring information about the state's most valuable geographic data resources. It also would encourage the cooperation of data producers in collecting, sharing and preserving important geographic data sets.

The clearinghouse would include three components:

- **Documentation collected and maintained in an electronic format that could be searched through a single entry point.** Included would be metadata, or detailed descriptions of data, training in its use and a process to ensure continued updating of the metadata catalog.

- **Data storage and archiving to preserve data sets and supporting documentation that are determined to be useful to current and future data users.** This component would ensure the physical and intellectual integrity of data and supporting documentation over time by developing guidelines for the archive, providing storage and archiving alternatives for data producers, offering training in archive preservation processes and defining criteria for identifying data sets to be included in the archive.

- **Access to data and metadata through a search system responsive to the needs of data users.** Computerized access to metadata collections would be the initial focus of the clearinghouse. The access system would provide a range of information about available data, including basic summary information, detailed metadata and, where possible, data samples for inspection, preferably in a graphic display.

The search mechanism would be rapid, easy to use, reliable and secure. It would take into consideration the technical limitations of users and provide alternative, noncomputerized ways to access metadata. Methods for accessing the data itself would vary, depending on its source, but at a minimum, users would receive contact information for obtaining data directly from the producer.

Recommendations

Based on the committee's work, the council makes the following recommendations:

- Implement the geographic data clearinghouse model described in this report.

- Identify the Internet as the primary connection to clearinghouse resources but provide alternative means of access.

- Continue to develop policies on privacy, copyrights, liability, pricing practices, security, and documentation and preservation standards.

- Designate a lead agency to act as the clearinghouse contact point. This agency would develop a process to ensure that clearinghouse data sets are current and usable, serve as the custodian for select data and develop partnership agreements and data exchange guidelines among Minnesota organizations. The Land Management Information Center at Minnesota Planning would serve as the lead agency until a permanent clearinghouse structure is established.

- Establish a Clearinghouse Node Working Group, Clearinghouse Advisory Steering Committee and Access Policy Committee through which major stakeholders would provide oversight and advice.

The Clearinghouse Node Working Group would bring together technical experts to plan and implement the technical aspects of the clearinghouse. Its long-term role would be to ensure the technological growth and development of the system.

The Clearinghouse Advisory Steering Committee would focus on facilitating interagency and external agreements, obtaining external funding and evaluating progress toward achieving clearinghouse goals. It would allow the lead agency to concentrate on day-to-day management issues and ensure input and representation from a range of stakeholders.

The Access Policy Committee would develop positions on broader access policy issues such as copyright, privacy, security and cost recovery. It would work closely with the executive committee of the Governor's council, the Minnesota Office of Technology and its Information Policy Council.

These recommendations are proposed to serve as the foundation of efforts by the Land Management Information Center at Minnesota Planning to develop the clearinghouse structure and identify the resources necessary to operate it.

Glossary

Below are generally accepted definitions for terms used in this report.

Archiving — Process for copying data onto electronic media for storage, preservation and retrieval.

Distributed data — Data stored on multiple computer systems.

Distributed network — A hardware and software environment in which computers are linked together so that software programs and data can be easily shared.

FGDC Clearinghouse Node — Any computer site registered with the Federal Geographic Data Committee containing metadata searchable as part of the National Spatial Data Clearinghouse.

Metadata — Detailed descriptions about data: geographic area covered, methods used to produce it, currency, accuracy and so on.

Node — Any computer site within a distributed network.

Z39.50 communications protocol — Internationally accepted set of procedures for computer-to-computer searching and retrieval of information in distributed network environments.

Introduction

eographic data integrates information from several disciplines into a common format and provides an efficient tool for organizing an array of details into a graphical, easy-to-comprehend format. Developers of computerized geographic data sets have long recognized that their data is expensive to create and could be used for multiple applications beyond those of their own organization. Providing access to such data sets through a statewide geographic data clearinghouse would result in a variety of benefits, including less duplication of effort in data creation and data being more readily useful to policy-makers and individuals in other fields.

A geographic data clearinghouse would allow broad public access to search tools for acquiring information about the state's most valuable geographic data resources. It would provide the structure, guidelines and access needed to derive the greatest benefit from the state's computerized geographic databases. Benefits of a well-designed clearinghouse would include:

- A catalog of data available from a variety of sources and instructions on how to access data in formats needed by the user

- Complete data documentation, including historical records about how the data was created and updated

- Improved access to data

- Fewer redundant data sets

- Efficiencies gained from the sharing and shared updating of data

- Improved data protection through a combination of preservation standards promoted by the clearinghouse and the option of a central archive for organizations unwilling or unable to provide long-term preservation on-site themselves

The Legislative Commission on Minnesota Resources, which has funded extensive geographic information system data development in Minnesota since the 1970s, recognizes the cumulative value of geographic information systems data. The commission endorsed the clearinghouse concept by specifying that all GIS projects it funds create data compatible with that maintained by the Land Management Information Center at Minnesota Planning.

Since the late 1980s, several initiatives have been proposed to develop a statewide clearinghouse for critical geographic data sets. LMIC has actively promoted the clearinghouse concept, as has the Federal Geographic Data Committee at the national level and MetroGIS at the regional level. Other agencies, including the U.S. Fish and Wildlife Service, the Minnesota Department of Transportation and several local governments, have made their geographic data sets freely available to outside users.

Meanwhile, for Minnesota GIS users, the problem of data access continues. A 1994 survey conducted by the Governor's Council on Geographic Information revealed the existence of hundreds of databases in transferable formats that data custodians were willing to distribute. Yet the perception of data users was that existing data was not adequate to fulfill their needs and that additional investment would be required to create the necessary data sets. With advice from the council, the Land Management Information Center used the survey results to construct an Internet directory of existing data sets. This improved the opportunity for data sharing but solved only a portion of the access problem identified by the survey.

Geographic data use continues to expand, with new GIS applications being implemented at all levels of government and by academic, private and research organizations. Along with increased use of geographic data, the number of data creators, providers and custodians also is growing.

Given this situation, the council established the Data Clearinghouse Committee in 1995 to be a forum for discussing issues of access to computerized geographic data and barriers to data sharing, exploring the concept of a geographic data clearinghouse, examining initiatives by other organizations offering public access to government information and developing a model for a geographic data access system in Minnesota.

Laying the Foundation for a Geographic Data Clearinghouse contains the findings and recommendations of the committee regarding establishing a statewide clearinghouse for geographic data.

Clearinghouse Model

The goal of the data clearinghouse is to encourage the use of geographic data for planning, development and research in the public and private sectors. The clearinghouse would provide a variety of tools to identify geographic data and determine its applicability and availability, encourage cooperation in collecting and disseminating data, preserve historically important data sets and ensure access to Minnesota's computerized data.

Three major groups have a stake in this goal:

- **Data users**, who are primarily concerned with access

- **Data producers**, who are concerned with metadata development and standards, options for data storage and archiving, and policy issues regarding levels of access

- **A lead agency** designated to handle the daily functioning of the clearinghouse, along with the coordination among its contributing members

The clearinghouse would be a distributed, Internet-based system using common standards and guidelines that would allow searching for data through a single access point. Links would provide direct access to data or information on how to obtain it. The clearinghouse would be a voluntary, cooperative venture of the state's geographic data producers and supporters. All data producers and users would be encouraged to become partners in the clearinghouse. State agency data would make up most of the early core data sets, but through partnerships the clearinghouse would expand to include regional and local data.

Data producers would determine their level of involvement in the clearinghouse. They would be encouraged to focus on their strengths and take advantage of the network of technical resources, support systems and training tools available through the clearinghouse. At the same time, they would control their own level of data security and access restrictions.

The clearinghouse would not be able to provide all the services needed by geographic data users. Value-added services such as attribute encoding and custom data formatting would need to be acquired outside the clearinghouse. This clearinghouse model also does not deal with such issues such as copyright, liability, pricing practices, security and privacy.

The clearinghouse model has three main components:

- **Documentation collected and maintained in an electronic format that can be searched through a single entry point.** Included would be metadata, or detailed descriptions of data, that meets the needs of both data producers and users; training in its use; and a process to ensure continued updating of the metadata catalog.

- **Data storage and archiving to preserve data sets and supporting documentation that are determined to be useful to current and future data users.** This component would ensure the physical and intellectual integrity of data and supporting documentation over time by developing guidelines for the archive, providing storage and archiving alternatives for data producers, offering training in archive preservation processes and defining criteria for identifying data sets to be included in the archive.

- **Access to data and metadata through a search system responsive to the needs of data users.** Computerized access to metadata collections would be the initial focus of the clearinghouse. The access system would provide a range of information about available data, including basic summary information, detailed metadata and, where possible, data samples for inspection, preferably in a graphic display.

In determining the clearinghouse components and structure, the committee considered critical issues and activities that would affect one or more components, identified stakeholders and determined what needs to be done to address the key issues.

Managing and maintaining documentation

Metadata needs to be created and maintained in a standardized, electronically searchable format to become the basis for a clearinghouse online catalog. The council's GIS Standards Committee has adopted metadata guidelines for geographic data sets that are consistent with federal standards and usable with federal electronic

search tools. These guidelines allow for varying levels of detail based on the needs and capabilities of data users and producers. Incentives for producing metadata based on the Minnesota guidelines need to be fully explored. The clearinghouse would be designed to provide support for complying with the guidelines, including training, software tools and technical assistance.

A major challenge for this component of the clearinghouse is the production and maintenance of metadata that can be searched in its entirety from a single World Wide Web site. Technical and structural limitations argue for a single storage site, but advances in search systems and communication protocols may eliminate this requirement. The lead agency also would need to maintain a traditional, hard-copy version of the metadata catalog for users without Internet access.

To provide a single metadata catalog of information about the availability, content and location of geographic data pertaining to Minnesota, the clearinghouse lead agency would:

- Establish a strong nucleus of data providers to start the clearinghouse and generate momentum so that prospective users could see the benefits of the clearinghouse

- Encourage the state's data producers to use the metadata guidelines adopted by the GIS Standards Committee

- Develop partnerships, agreements, memoranda of understanding and other mechanisms to encourage participation by data producers in the clearinghouse and use of documentation guidelines

- Work to achieve full compliance for important core data sets and base-level compliance for older, project-based, unique data sets considered valuable

- Identify important data sets that should be included in the clearinghouse but lack the necessary metadata. Develop policies on the level of metadata required before these older data sets can be included.

- Implement a process to ensure continued maintenance and updating of the metadata catalog

- Provide information, supporting software and training in the metadata guidelines for data producers and users

- Adopt the Z39.50 communications protocol — a standard Internet search protocol — for the clearinghouse

Storing and archiving data

The clearinghouse would allow for data to be distributed, that is, stored at more than one site. To ensure data quality, integrity, consistency, longevity and accessibility, data producers would be able to store their data on-site or with an external archiving service. Participating data producers could choose to retain custody of their data during its active lifetime and when it is no longer being actively used. Doing so would require them to commit to using a basic preservation protocol to ensure long-term access to the data. This protocol includes refreshing the data by periodically moving it onto new media and migrating it to new hardware platforms and software formats as necessary.

Other data producers could choose to use an archive custodian for some or all of their data. Materials would be deposited with the archive custodian on a schedule acceptable to the data producer and the clearinghouse's lead agency. Archiving schedules would take into consideration the point at which the data became inactive, the value of periodic snapshots of dynamic databases and the ability of data producers to provide ongoing support and maintenance of inactive data sets. The archive site need not be located within the lead agency, but the organization responsible for the archival function would need to fulfill a long-term preservation role for the clearinghouse.

The goal of data storage and archiving is to preserve useful data sets and their supporting documentation and ensure their physical and intellectual integrity over time. To accomplish this, the clearinghouse lead agency would need to:

- Develop a collection policy for the geographic data archive and build a statewide consensus on selection criteria for data valued by the geographic data community both now and in the future

- Set priorities for data acquisition and determine who will be responsible for dealing with incomplete documentation,

particularly for existing data sets; checking data quality; and determining conditions under which a data set should be improved before it is included in the clearinghouse

■ Establish guidelines for data archive criteria such as formatting and hardware and software independence, the schedule for migrating and refreshing data, and suitability for archival preservation, including content, documentation level and quality

■ Create a primary, centralized archive site and develop agreements with distributed archive sites to meet the needs and goals of data producers. The lead agency would offer archiving as a service if data producers choose not to do their own archiving. Agencies agreeing to act as data archive sites would have to agree to follow established data archive guidelines.

■ Seek out data that fits within the collection criteria of the clearinghouse and pursue agreements with data producers regarding long-term preservation and access

■ Identify agencies responsible for the various facets of data archiving, including ongoing data backups, backup storage and user access to metadata and data

■ Establish policies and procedures for data providers to follow in updating and modifying their metadata and data sets

■ Evaluate the quality of data submitted to the archive and assign quality assurance indicators based on agreed-upon criteria

■ Set up long-term depository schedules for data from each data producer

■ Develop training and support systems to help data producers establish in-house procedures that support the preservation goals and guidelines of the clearinghouse

Accessing metadata and data

Access goals differ for metadata and data. The goals for metadata would be within the domain of the clearinghouse, while those for data are more closely related to the priorities and restrictions of the individual data producers. The clearinghouse's lead agency would concentrate first on access to metadata, then on access to data.

The methods for obtaining data through the clearinghouse need to be responsive to users' needs as well as data producers' policies on security and cost recovery, and their ability to provide access. Metadata and data would become available through a single point of contact, though not necessarily a single computer site, as long as appropriate communication protocols were followed. Some sites might merely provide contact and ordering information, while others might have major data sets online.

The lead agency and data producers would not be able to provide the full range of data manipulation needed by GIS users. Value-added services such as reformatting data, adding attributes, linking to tabular data sets and performing custom queries would need to be made available outside the clearinghouse.

The clearinghouse ideally would provide ready access to both summary-level and detailed metadata and would help users gain direct access to Minnesota's geographic data. This would require the lead agency to:

■ Evaluate other clearinghouse efforts to determine the applicability of their designs to the Minnesota clearinghouse

■ Establish the identity of the clearinghouse as the central search site for Minnesota geographic data

■ Work with the Federal Geographic Data Committee, using the model of the National Spatial Data Infrastructure, to link to national data clearinghouse efforts

■ Develop a searchable data catalog that is quick, easy to use, reliable, up to date, secure and accessible. The search mechanism must take into consideration users' technical capabilities and provide a wide range of access and delivery options.

■ Identify or develop an effective search engine with a broad range of search protocols, sample data display and a logical protocol requiring access to metadata before data

■ Maintain and update the index and search system

■ Establish agreements with data providers for submitting metadata to the clearinghouse and clarifying the level of data access available

■ Set up links to individual data servers as appropriate

■ Provide sufficient information for the user to efficiently access data within the data producers' constraints and seek opportunities to provide direct electronic links to data

■ Determine the data and metadata formats most commonly needed by users

■ Develop processes to deliver data over the Internet and to produce less technical formats, such as hard copy for metadata. Streamline procedures for creating output formats such as CD-ROMs and tapes that can be shared by data producers.

■ Provide graphic display of data samples for inspection

■ Establish an automated system, along with a staffed help desk, to answer inquiries from data providers and users, provide training in a number of formats and promote efficient and effective use of clearinghouse materials

■ Identify organizations that would provide support services to users, including publicity about clearinghouse services and policies; guidance to convert data and metadata from nonstandard formats; and training and support for using the search system

Implementing the Clearinghouse

Developing and maintaining the clearinghouse would require the cooperation of a number of organizations. The initial goal of the lead agency would be to create a viable clearinghouse with a core base of information. A second phase would entail expanding that base to include a broad range of organizations and databases. Throughout this process, the Governor's Council on Geographic Information would provide guidance through its committees and advisory groups.

Based on its Data Clearinghouse Committee's work, the council makes the following recommendations:

■ Implement the geographic data clearinghouse model described in this report.

■ Identify the Internet as the primary connection to clearinghouse resources but provide alternative means of access.

■ Continue to develop policies on privacy, copyrights, liability, pricing practices, security, and documentation and preservation standards.

■ Designate a lead agency to act as the clearinghouse contact point, develop a process to ensure that clearinghouse data sets are current and usable, serve as the custodian for select data and develop partnership agreements and data exchange guidelines among Minnesota organizations. The Land Management Information Center at Minnesota Planning would serve as the lead agency until a permanent clearinghouse structure is established.

■ Establish a Clearinghouse Node Working Group, Clearinghouse Advisory Steering Committee and Access Policy Committee through which major stakeholders would provide oversight and advice.

These recommendations are proposed to serve as the foundation of efforts by the Land Management Information Center at Minnesota Planning to develop the clearinghouse structure and identify the resources needed to operate it.

Role of the lead agency

The lead agency would have ongoing responsibility for developing the clearinghouse through:

■ Identifying critical activities of the clearinghouse

■ Obtaining the participation of major data-producing organizations

■ Establishing a core set of databases and metadata and an access system based on the data provided by participating organizations

■ Promoting awareness and use of the clearinghouse

■ Providing training in clearinghouse operations

Role of the Clearinghouse Node Working Group

The Clearinghouse Node Working Group would be composed of members of the council's Metadata Guidelines Working Group and others with the technical expertise to implement a pilot clearinghouse and demonstrate its value.

The working group's short-term goal would be to use the model outlined by the Federal Geographic Data Committee in its National Geospatial Data Clearinghouse initiative to establish a direct link to the national clearinghouse. It's long-term goal would be to continually assess developments in technology and their impact on the clearinghouse, and to ensure that the clearinghouse's technical structure evolves effectively.

The working group would:

■ Investigate, test and implement technical developments that would help the clearinghouse effectively deliver geographic metadata and data over the Internet

■ Oversee technical implementation of a Federal Geographic Data Committee clearinghouse node in Minnesota and adapt the model to meet the needs of Minnesota citizens

■ Seek to ensure the clearinghouse's long-term value and responsiveness to changing technology and user needs

Role of the Clearinghouse Advisory Steering Committee

The Clearinghouse Advisory Steering Committee would focus on encouraging agreements among organizations, obtaining funding for the clearinghouse, reviewing how well participants' needs are being met and evaluating the lead agency's progress in meeting clearinghouse goals. The committee would be housed at the lead agency and communicate closely with the Governor's council.

The committee would:

■ Monitor interagency relationships to promote clearinghouse cooperation

■ Review data produced within Minnesota to identify sources of new data for the clearinghouse

■ Evaluate progress toward developing an effective clearinghouse structure

■ Provide policy direction to the lead agency and technical working groups addressing clearinghouse design issues, metadata and standards

The council recommends that representatives of the following stakeholders, among others, be invited to join the steering committee: Minnesota departments of Natural Resources, Transportation, and Agriculture; Minnesota Pollution Control Agency; Land Management Information Center; Minnesota Historical Society; MetroGIS; and library organizations.

Role of the Access Policy Committee

Several policy concerns were raised in developing the clearinghouse model that involve issues beyond the scope of geographic data users and need to be addressed in a broader context.

Among these issues are statewide goals for geographic data access and the State Data Practices Act, which addresses copyright, data privacy and security, and cost recovery concerns. The 1997 Minnesota Legislature authorized creation of an Information Policy Task Force, staffed by Minnesota Planning and the Minnesota Department of Administration, to study and make recommendations regarding public information policy, government data practices and information technology issues. The task force will submit a progress report to the Legislature in February 1998 and a final report in January 1999. These should be reviewed closely by the Access Policy Committee and used as a foundation for the committee's work.

To ensure that geographic data users have a strong, unified voice in discussions of these issues, the council recommends that an Access Policy Committee be formed to address policy issues related to geographic data access. The committee would refer its recommendations to the Minnesota Office of Technology and its Information Policy Council, which are charged with addressing statewide information and

technology issues. It also would develop links with the Federal Geographic Data Committee. Where this committee would be housed and who would provide staff have not been determined.

Committee members would include representatives of the Metropolitan Council, Land Management Information Center, Minnesota Department of Transportation, Minnesota Pollution Control Agency, library organizations, and federal and local governments.

Conclusion

The next step in creating a geographic data clearinghouse is for the Land Management Information Center at Minnesota Planning to be formally recognized as the interim lead agency. LMIC already has begun work on several initiatives described in this report, particularly those of the Clearinghouse Node Working Group. As the interim lead agency, LMIC would seek the resources needed to implement the clearinghouse in a way that ensures the participation and cooperation of a broad range of the state's geographic data producers and users.

Information Resources

Other organizations have developed standards and approaches from which Minnesota's clearinghouse developers can learn. Following are a few of the resources providing an overview of current activities in preservation, technical standards, search engines and presentation tools.

First, however, is a resource that tackles the problems of bringing together a large, diverse group of organizations to identify their needs for data within a geographic area:

■ **MetroGIS: Sharing Data Across Boundaries,** *http://www.state.mn.us/intergov/metrogis.* MetroGIS is designed to provide an ongoing, stakeholder-governed mechanism through which Twin Cities metropolitan area participants easily and equitably share geographically referenced graphic and associated attribute data that is accurate, current, of common benefit and readily usable. MetroGIS is notable for the size and scope of its participant base, which includes state, city and county agencies, school districts and a variety of other public and private organizations.

Preservation

■ **Center for Electronic Records, National Archives and Records Administration,** *http://www.nara.gov/nara/electronic.* This site provides information on the center's role in inventorying, scheduling, appraising and transferring electronic records from federal agencies to the archive. The center offers information on preservation practices and guides for the management of electronic records.

■ ***A National Archives Strategy for the Development and Implementation of Standards for the Creation, Transfer, Access, and Long-Term Storage of Electronic Records of the Federal Government.*** National Archives Technical Information Paper no. 8, Archival Research and Evaluation Staff, National Archives and Records Administration, June 1990, *gopher://gopher.nara.gov:70/00/managers/archival/papers/strategy.txt*.

■ ***Magnetic Tape Storage and Handling: A Guide for Libraries and Archives.*** Commission on Preservation and Access, National Media Laboratory, June 1995. This private, nonprofit commission acts on behalf of the nation's libraries, archives and universities to develop and encourage collaborative strategies for preserving and providing access to records.

■ **National Media Laboratory** (a host laboratory of the National Technology Alliance), *http://www.nta.org/AboutNTA/AboutNML.* This is an access point for information and resources about data recording systems and information technology.

■ **Inter-university Consortium for Political and Social Research,** ***Depositing Data with ICPSR,*** *http://www.icpsr.umich.edu/ICPSR/Archive/Deposit/deposit.html.* This

site provides access to documents on depository requirements, including *Guide to Social Science Data Preparation and Archiving* and *Data Deposit Form*.

Technical standards

■ **Z39.50-1992 document,** *http://www.cni.org:80/pub/NISO/docs/Z39.50-1992/www/50.brochure.toc.html.* Z39.50 is a computer-to-computer communications protocol designed to support searching and retrieval of information — full-text documents, bibliographic data, images and multimedia — in a distributed network environment. This site provides a hypertext edition of the documents. ANSI/NISO Z39.50-1992, also known as Z39.50, version 2, is also available for purchase from the National Information Standards Office.

■ **Federal Geographic Data Committee,** ***Content Standards for Digital Geospatial Metadata,*** *http://www.fgdc.gov/Metadata/metahome.html.* The standard is designed to provide a common set of terminology and definitions for documentation of digital geospatial data. Metadata is used to help organize an organization's spatial data; provide information about an organization's data to catalogs, clearinghouses and brokerages; and provide information to process and interpret data received through transfer from an external source. The Federal Geographic Data Committee metadata standard is the cornerstone of the National Spatial Data Clearinghouse organization and search tools described below. Minnesota has adopted a subset of the federal standard as a state metadata guideline.

■ **Minnesota Geographic Metadata Guidelines,** *http://www.lmic.state.mn.us/gc/stds/metadata.htm.* The council's GIS Standards Committee has developed formal guidelines and supporting tools for preparing metadata. The Minnesota guidelines are a streamlined version of the Federal Geographic Data Committee's *Content Standards for Geospatial Metadata.* Minnesota has supported development of DataLogr software for entering metadata and will develop a conversion program to make the Minnesota metadata fully compatible with the National Spatial Data Clearinghouse search tools.

■ **Data Documentation Initiative,** *http://www.icpsr.umich.edu/DDI/codebook.html.* This site provides background information on and describes current work of a committee established in 1995 by the Inter-university Consortium for Political and Social Research to develop a document type definition for an international codebook standard using standard generalized markup language. This committee includes representatives of major data archives and services in the United States, Canada and Europe.

Search engines

■ **Government Information Locator Service,** *http://www.usgs.gov/gils/.* GILS applies international standards within the federal government to allow for searching across GILS compliant databases. Compliant databases include the Government Printing Office, National Archives and Records Administration, and other federal agencies.

■ **National Geospatial Data Clearinghouse,** *http://fgdclearhs.er.usgs.gov.* The clearinghouse is a distributed, electronically connected network of geospatial data producers, managers and users through which users can determine who has what data. It provides access to data through (in descending order of preference) an online link, an order form or contact information for the organization that distributes the data. It is not a central repository where data sets are stored. The clearinghouse is sponsored by the Federal Geographic Data Committee, an organization spearheaded by the U.S. Geological Survey's National Mapping Division. The clearinghouse is based on the Z39.50 protocol for searching on the Internet. The Federal Geographic Data Committee provides the means for sites such as the Minnesota data clearinghouse or MetroGIS to register as searchable nodes on the National Geospatial Data Clearinghouse. All sites registered are indexed using the FGDC-provided Isite software and can be searched using the FGDC search tools. The FGDC metadata standard data elements provide the basis for the search, which can be geographic or theme-based.

■ **EDINA: Edinburgh Data and Information Access,** *http://edina.ed.ac.uk/.* This data center offers United Kingdom educators and researchers access to a library of data, information and research resources.

■ **Data Archive at the University of Essex,** *http://dawww.essex.ac.uk/.* This archive houses the largest collection of accessible computer-

readable data in the social sciences and humanities in the United Kingdom. It is a national resource center, disseminating data throughout the United Kingdom and, by arrangement with other national archives, internationally. *About the Data Archive* offers information on the data archive, sources and data holdings, confidentiality, bibliographic control, data processing, documentation, teaching materials and access to BIRON, HASSET and the Integrated Data Catalogue.

■ **Consortium for International Earth Science Information Network Gateway WWW Interface,** *http://wwwgateway.ciesin.org/*. The CIESIN gateway provides access to a distributed catalog that identifies data and information resources relevant to international earth science. The gateway conducts simultaneous search and retrieval across diverse databases worldwide.

■ **CIC Virtual Electronic Library,** *http:// cedar.cic.net/cic/velnew.html*. The VEL project provides a HTTP/Z39.50 gateway into 13 Committee on Institutional Cooperation university library public access catalogs using a single, customizable graphical user interface, allowing users to search across the CIC catalogs with a single search, search locally mounted databases and place a request using International Standards Organization protocols.

Presentation tools and web sites

■ **National Information Display Laboratory** (a host laboratory of the National Technology Alliance), *http://www.nta.org/AboutNTA/AboutNIDL/Introduction.* This site allows for the exchange of up-to-date information on hardware and software development for information display, image analysis, database tools and related information.

■ **U.S. Environmental Protection Agency Envirofacts Database,** *http://www.epa.gov/enviro/html/ef_home.html.* Envirofacts is a relational database that integrates data extracted monthly from five major facility-based EPA programs: Superfund data (CERCLIS), hazardous waste data (RCRIS), toxic release inventory (TRI), water discharge permits (PCS) and AIS facility subsystem (aerometric information). The database is available for public access through an online form that executes predefined queries for the types of facilities. The software required for accessing the database directly is described. This site also includes a spatial data library and a web-based mapping application that generates maps displaying database information.

■ **U.S. Geological Survey Water Resources Division — Water Data Queries,** *http://h20.usgs.gov.* This site offers public searches of water data stored on USGS Water Resources Division databases. Accessible data includes realtime streamflow, historical streamflow, water quality, water use, acid rain and geographic information.

■ **U.S. Census and TIGER data sources.** *http://www.census.gov/main/www/access.html.* This site offers access to 1990 U.S. census data tables, interactive TIGER map generation, state and county profiles, and U.S. Gazetteer. TIGER files are the GIS data layers created by the U.S. Census Bureau to help define and display census areas. They include census tracts, block groups and blocks; roads; surface water; and Public Land Survey boundaries.

■ **North Carolina Geographic Data Clearinghouse,** *http://cgia.cgia.state.nc.us/ncgdc/index.html.* North Carolina geographic data producers advertise their data on these pages, and users can perform many types of data searches. Using a web form, users can create a custom query through the North Carolina gateway or the Federal Geographic Data Committee world gateways. Users can also perform a static search of North Carolina data listings by data theme or agency.

■ **Montana State Library Natural Resources Information System/Geographic Information Systems,** *http://nris.mt.gov/gis/gis.html.* The GIS department acts as a clearinghouse for GIS databases within Montana and provides services to public and private organizations. Some data and all metadata are online. Metadata complies with Federal Geographic Data Committee standards.

■ **Montana State Library Natural Resources Information System/Water Information System,** *http://nris.mt.gov/wis/wis1.html.* This site offers access to a variety of water and climate information in many formats.

■ **Wisconsin Land Information Clearinghouse,** *http://badger.state.wi.us/agencies/wlib/sco/pages/wisclinc.html.* This clearinghouse is a gateway to geospatial data and metadata, related land and reference information, and the state agencies that produce or maintain these items. It also is a registered node of the National Spatial Data Clearinghouse. Geographic information available through this site is in the form of graphic interchange files and ARC/INFO shape files. This site also contains a searchable database of geospatial metadata or geographic data held by many state agencies.

■ **Kansas Geographic Information Systems Initiative's Data Access and Support Center,** *http://gisdasc.kgs.ukans.edu.* This site includes access to an online database catalog, search tools, online data request forms, and static and interactive image maps.

■ **Texas State Data Center,** *http://www-txsdc.tamu.edu/.* The section on data, maps, reports and services provides information on the electronic data catalog, online data and a general listing of available products and services.

■ **ForNet (University of Minnesota College of Natural Resources and Minnesota Department of Natural Resources),** *http://www.gis.umn.edu/fornet.* ForNet is one of 18 remote-sensing database programs funded as part of the National Aeronautics and Space Administration cooperative agreement notice, *Public Use of Earth and Space Science Data over the Internet*. ForNet includes an image delivery system and the ability to create maps interactively from Minnesota Department of Natural Resources forest stand inventory GIS data and supporting GIS data layers. Images for the Twin Cities metropolitan area are available online.

■ **NODIS,** *http://cua6.csuohio.edu/~ucweb/nodis/nodis.htm.* The Northern Ohio Data and Information Service is part of the Urban Center at the Maxine Goodman Levin College of Urban Affairs, Cleveland State University. NODIS is designated by the state of Ohio and the U.S. Census Bureau as the regional data center for northern Ohio. NODIS is a repository for census information, which is made available to the university community, public administrators, civic organizations and the general public on a cost-recovery basis.

■ **Michigan Information Center,** *http://www.michigan.state.mi.us/michome/mic.html-ssi.* Most information from this unit of the Michigan Department of Management and Budget is available in printed tabular reports and digital format on tape or diskette. Data extracts, geographic information systems and computer mapping are used extensively for creating standard products and custom projects.